CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	BEACON: below average cognitive ability in childhood - outcomes and prediction
Authors	Bowe, Andrea
Publication date	2025
Original Citation	Bowe, A. 2025. BEACON: below average cognitive ability in childhood - outcomes and prediction. PhD Thesis, University College Cork.
Type of publication	Doctoral thesis
Rights	© 2025, Andrea Bowe. - https://creativecommons.org/licenses/by/4.0/
Download date	2026-11-05 09:28:39
Item downloaded from	https://hdl.handle.net/10468/17920



UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Ollscoil na hÉireann, Corcaigh
National University of Ireland, Cork



BEACON: Below Average Cognitive Ability in Childhood - Outcomes and Prediction

Thesis submitted by
Dr Andrea Bowe
MB BCh BAO MPH MFPHMI
for the degree of
Doctor of Philosophy

University College Cork
Department of Paediatrics and Child Health and The Irish Centre for Maternal and Child Health Research (INFANT)

Head of Department
Professor Deirdre Murray

Supervisors
Professor Deirdre Murray
Dr Gordon Lightbody
Professor Anthony Staines

2025

Table of contents

LIST OF TABLES	9
LIST OF FIGURES	11
DECLARATION	13
LIST OF PUBLICATIONS	14
LIST OF CONFERENCE PRESENTATIONS	15
AWARDS	16
ABBREVIATIONS	17
ACKNOWLEDGMENTS	18
THESIS ABSTRACT	20
CHAPTER 1 INTRODUCTION	24
1.1 Cognitive ability	25
1.2 Psychometric assessment of cognitive ability	27
1.3 Terminology and categorisation of cognitive ability	29
1.3.1 Intellectual developmental disorders	29
1.3.2 Cognitive delay	30
1.3.3 Terms used in this thesis	31
1.4 The development of cognitive ability in early life	32
1.4.1 The first 1,000 days	32
1.4.2 Conceptual framework	35
1.4.2.1 Biological exposures	37
1.4.2.2 Psychological exposures	39
1.4.2.3 Social exposures	40
1.4.3 Preterm infants	41
1.5 The importance of childhood cognitive ability	42
1.5.1 Academic achievement	43
1.5.2 Physical health	44
1.5.3 Mental health	45
1.5.4 Financial well-being and social mobility	45

1.5.5 Outcomes and experiences in childhood	46
1.6 Evidence for early intervention	48
1.6.1 Experimental evidence: The Abecedarian Study	48
1.6.2 Experimental evidence in an Irish setting: The Preparing for Life Project	51
1.6.3 Real world evidence: The Head Start and Early Head Start Programmes	52
1.6.4 Early intervention for preterm infants	53
1.6.5 Eligibility for experimental and real world early intervention programmes	54
1.7 Current childhood screening for low cognitive ability in Ireland	56
1.8 Multivariable models for predicting cognitive outcomes in childhood	57
1.8.1 Published population-based multivariable models for predicting cognitive outcomes in childhood	57
1.8.2 Limitations of published population-based multivariable models for predicting cognitive outcomes in childhood	63
1.9 Machine learning for predicting cognitive outcomes in childhood	64
1.9.1 An introduction to machine learning	64
1.9.2 The Random Forest algorithm	67
1.9.3 The Gradient Boosting algorithm	71
1.9.4 The Support Vector Machine algorithm	72
1.9.5 Potential advantages of machine learning for predicting cognitive outcomes in childhood	75
1.9.6 Example applications of machine learning in neurodevelopmental paediatrics	77
1.9.7 Machine learning for population-based prediction	80
1.10 Summary	82
1.11 References	85
CHAPTER 2 AIMS AND OBJECTIVES	116
2.1 Aims and objectives	116
2.2 Data sources used in this thesis	118
2.2.1 Growing up in Ireland Infant Cohort	118

2.2.2 Cork Baseline Birth Cohort	122
2.2.3 Swedish Neonatal Quality Register	125
2.3 References	128
CHAPTER 3 BELOW AVERAGE COGNITIVE ABILITY – AN UNDER RESEARCHED RISK FACTOR FOR EMOTIONAL BEHAVIOURAL DIFFICULTIES IN CHILDHOOD	130
3.1 Abstract	131
3.2 Introduction	133
3.3 Methods	136
3.4 Results	141
3.5 Discussion	144
3.6 Conclusion	148
3.7 References	161
CHAPTER 4 THE RELATIONSHIP BETWEEN BELOW AVERAGE COGNITIVE ABILITY AT AGE 5 YEARS AND THE CHILD'S EXPERIENCE OF SCHOOL AT AGE 9	168
4.1 Abstract	169
4.2 Introduction	171
4.3 Methods	175
4.4 Results	182
4.5 Discussion	185
4.6 Conclusion	191
4.7 References	200
CHAPTER 5 THE PREDICTIVE VALUE OF THE AGES AND STAGES QUESTIONNAIRE IN LATE INFANCY FOR LOW AVERAGE COGNITIVE ABILITY AT AGE 5	215
5.1 Abstract	216
5.2 Introduction	217
5.3 Methods	219
5.4 Results	221
5.5 Discussion	223

5.6 Conclusion	226
5.7 References	233
CHAPTER 6 PREDICTING LOW COGNITIVE ABILITY AT AGE 5 – FEATURE SELECTION USING MACHINE LEARNING METHODS AND BIRTH COHORT DATA	238
6.1 Abstract	239
6.2 Introduction	240
6.3 Methods	242
6.4 Results	248
6.5 Discussion	251
6.6 Conclusion	254
6.7 References	263
CHAPTER 7 PREDICTING LOW COGNITIVE ABILITY AT AGE 5 USING PERINATAL DATA AND MACHINE LEARNING	270
7.1 Abstract	271
7.2 Introduction	272
7.3 Methods	273
7.4 Results	278
7.5 Discussion	281
7.6 Conclusion	287
7.7 References	297
CHAPTER 8 PREDICTING 2-YEAR COGNITIVE OUTCOMES IN VERY PRETERM INFANTS USING MACHINE LEARNING METHODS	307
8.1 Abstract	308
8.2 Introduction	310
8.3 Methods	311
8.4 Results	316
8.5 Discussion	319
8.6 Conclusion	322
8.7 References	330
CHAPTER 9 DISCUSSION	339

9.1 Introduction	339
9.2 Summary of data sources used	339
9.3 Summary of main findings	340
9.3.1 Summary of findings Chapter 3	342
9.3.2 Summary of findings Chapter 4	344
9.3.3 Summary of findings Chapter 5	346
9.3.4 Summary of findings Chapter 6	347
9.3.5 Summary of findings Chapter 7	349
9.3.6 Summary of findings Chapter 8	352
9.3.7 Application of conceptual framework	353
9.4 Interpretation of findings	354
9.4.1 Bias	355
9.4.1.1 Selection bias	355
9.4.1.2 Measurement bias	358
9.4.1.3 Recall bias	360
9.4.1.4 Social desirability bias	360
9.4.2 Confounding	361
9.4.3 Chance findings	362
9.5 Strengths and limitations	364
9.5.1 Strengths	364
9.5.2 Limitations	365
9.5.2.1 Measuring cognitive ability	365
9.5.2.2 Categorising cognitive ability	368
9.5.2.3 Use of existing data	369
9.5.2.4 Limitations of machine learning	370
9.6 Public health and policy implications of findings	371
9.6.1 Below average cognitive ability as a risk factor for adverse outcomes	372
9.6.2 Population screening in early life for below average cognitive ability in childhood	372

9.6.3 Machine learning in public health	374
9.6.4 Policy implications of findings	376
9.7 Ethical issues arising from thesis	380
9.7.1 Ethical issues related to measuring cognitive ability	380
9.7.2 Ethical issues related to early labelling	380
9.7.3 Ethical issues related to the use of machine learning in public health	382
9.8 Future Research	383
9.9 Conclusion	388
9.10 References	393
APPENDIX 1 DATA ACCESS GROWING UP IN IRELAND INFANT COHORT (CHAPTER 3)	410
APPENDIX 2 DATA ACCESS GROWING UP IN IRELAND INFANT COHORT (CHAPTER 4)	416
APPENDIX 3 DATA ACCESS GROWING UP IN IRELAND INFANT COHORT (CHAPTER 7)	422
APPENDIX 4 DATA ACCESS BASELINE BIRTH COHORT	428
APPENDIX 5 ETHICAL APPROVAL FROM THE CORK RESEARCH ETHICS COMMITTEE	437
APPENDIX 6 ETHICS APPLICATION TO THE SWEDISH ETHICAL REVIEW BOARD	438
APPENDIX 7 ETHICAL APPROVAL FROM THE SWEDISH ETHICAL REVIEW BOARD	456
APPENDIX 8 SUPPLEMENTARY MATERIAL FOR CHAPTER 3	457
APPENDIX 9 SUPPLEMENTARY MATERIAL FOR CHAPTER 4	464
APPENDIX 10 SUPPLEMENTARY MATERIAL FOR CHAPTER 5	494
APPENDIX 11 SUPPLEMENTARY MATERIAL FOR CHAPTER 6	498
APPENDIX 12 SUPPLEMENTARY MATERIAL FOR CHAPTER 7	511
APPENDIX 13 SUPPLEMENTARY MATERIAL FOR CHAPTER 8	529
APPENDIX 14 BIG DATA, MACHINE LEARNING, AND POPULATION HEALTH: PREDICTING COGNITIVE OUTCOMES IN CHILDHOOD	549
APPENDIX 15 MODULES AND COURSES COMPLETED DURING PHD	587

APPENDIX 16 ADDITIONAL PUBLICATIONS DURING PHD	588

List of Tables

Table 1.1 Published multivariable models for predicting cognitive outcomes in childhood at a population level	59
Table 1.2 Examples of machine learning applied in neurodevelopmental paediatrics	78
Table 3.1 Characteristics of study population according to cognitive ability at age 3 years	151
Table 3.2 Results of generalised linear mixed model examining cognitive ability as a predictor of longitudinal SDQ scores	152
Table 3.3 Results of generalised linear mixed model examining cognitive ability as a predictor of change in total SDQ score	154
Table 3.4 Logistic regression model examining association between cognitive ability at age 3 and clinically significant emotional-behavioural difficulties between age 3 and 9 years	156
Table 4.1 Characteristics of study population and comparison of characteristics between those who did and did not have below average cognitive ability at age 5 years	193
Table 4.2 Child self-concept at age 9 years	195
Table 4.3 Teacher-reported classroom engagement and behaviour at age 9 years	197
Table 4.4 Teacher-reported emotional behavioural function at age 9 years	199
Table 5.1 Characteristics of children in study sample compared to those excluded due to attrition or missing data	227
Table 5.2 Predictive value of ASQ at 24 or 27 months for identifying children with low cognitive ability at age 5 years	228
Table 5.3 Univariable and multivariable logistic regression model examining factors associated with low cognitive ability at age 5	229
Table 6.1 Characteristics of those with and without low cognitive ability at age 5 years	256
Table 6.2 Performance of Models A-D in repeated 10-fold cross validation	258
Table 6.3 Mean minimal depth of the six most frequently occurring interactions in random forest Model A	259

Table 7.1 Characteristics of study population and comparison of characteristics between those with and without low cognitive ability at age 5 years	289
Table 7.2 Performance of final models on unseen test set	292
Table 8.1 Features assessed as important or tentative for future cognitive function in very preterm infants (GA <32 weeks).	323
Table 8.2 External validation on unseen test data for Models A-D	326
Table 8.3 Contingency table for Model B at a decision threshold of 0.20	327
Table 9.1 Summary of data sources used	340
Table 9.2 Summary of original studies in thesis	341
Table 9.3 Ten most important features in random forest models predicting low cognitive ability at age 5 years using the Growing Up in Ireland and BASELINE cohorts.	351

List of Figures

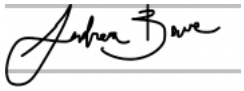
Figure 1.1 Conceptual framework for the perinatal and early life factors associated with cognitive ability in childhood according to Engel's Biopsychosocial Model	36
Figure 0.1 Schematic representation of the Random Forest algorithm	68
Figure 0.2 A hypothetical example of a linear Support Vector Machine classifying the data points into blue and pink circles	73
Figure 2.1 Overall thesis objectives with corresponding chapters and publications	117
Figure 2.2 Flowchart of participants in the Growing Up in Ireland Infant Cohort	120
Figure 2.3 Flowchart of participants in the BASELINE Birth Cohort	124
Figure 3.1 Distribution of total SDQ scores at age 3, 5, and 9 years according to cognitive ability	157
Figure 3.2 Alluvial plots of SDQ decile scores at age 3, 5, and 9 years according to cognitive ability	158
Figure 3.3 Parameter estimates for fully adjusted model predicting longitudinal total SDQ scores (Model 4)	159
Figure 3.4 Predicted change in total SDQ score between age 3 and 9 years according to cognitive ability group (Model 5)	160
Figure 5.1 Flowchart of study population from the BASELINE birth cohort	230
Figure 5.2 Receiver operating characteristic curve of the 24-month ASQ for predicting low cognitive ability at age 5 years	231
Figure 5.3 Receiver operating characteristic curve of the 27-month ASQ for predicting low cognitive ability at age 5 years	232
Figure 6.1 Flowchart of study population from the BASELINE birth cohort	260
Figure 6.2 Example decision tree from the random forest Model A predicting low cognitive ability at age 5	261
Figure 6.3 Feature importance plot showing the mean decrease in Gini and mean decrease in accuracy for the 15 most important features in random forest Model A	262
Figure 7.1 Overview of modelling process	293

Figure 7.2 Ten most important features in random forest Model 2	294
Figure 7.3 Partial dependence plots and histograms examining the relationship between the six most important features in random forest Model 2 and low cognitive ability	295
Figure 8.1 Overview of modelling process	328
Figure 8.2 Feature importance plots for Models A-D	329

Declaration

This is to certify that the work I am submitting is my own and has not been submitted for another degree, either at University College Cork or elsewhere. All external references and sources are clearly acknowledged and identified within the contents. I have read and understood the regulations of University College Cork concerning plagiarism and intellectual property.

Signature of the candidate:



Date: 03/03/2025

List of Publications

1. **Bowe AK**, Staines A, Murray DM. Below Average Cognitive Ability-An under Researched Risk Factor for Emotional-Behavioural Difficulties in Childhood. Int J Environ Res Public Health. 2021 Dec 8;18(24):12923. doi: 10.3390/ijerph182412923. PMID: 34948532; PMCID: PMC8702024.

2. **Bowe AK**, Urban M, Staines A, Murray DM. The relationship between below average cognitive ability at age 5 years and the child's experience of school at age 9. Front. Public Health. 2025 March;13. doi: 10.3389/fpubh.2025.1341797

3. **Bowe AK**, Hourihane J, Staines A, Murray DM. The predictive value of the ages and stages questionnaire in late infancy for low average cognitive ability at age 5. Acta Paediatr. 2022 Jun;111(6):1194-1200. doi: 10.1111/apa.16309. Epub 2022 Mar 3. PMID: 35202483; PMCID: PMC9314849

4. **Bowe AK**, Lightbody G, Staines A, Kiely ME, McCarthy FP, Murray DM. Predicting Low Cognitive Ability at Age 5-Feature Selection Using Machine Learning Methods and Birth Cohort Data. Int J Public Health. 2022 Nov 10;67:1605047. doi: 10.3389/ijph.2022.1605047. PMID: 36439276; PMCID: PMC9684182.

5. **Bowe AK**, Lightbody G, O'Boyle DS, Staines A, Murray DM. Predicting low cognitive ability at age 5 years using perinatal data and machine learning. Pediatr Res. 2024 May;95(6):1634-1643. doi: 10.1038/s41390-023-02914-6. Epub 2024 Jan 4. PMID: 38177251; PMCID: PMC11126385.

6. **Bowe AK**, Lightbody G, Staines A, Murray DM, Norman M. Prediction of 2-Year Cognitive Outcomes in Very Preterm Infants Using Machine Learning Methods. JAMA Netw Open. 2023 Dec 1;6(12):e2349111. doi: 10.1001/jamanetworkopen.2023.49111. PMID: 38147334; PMCID: PMC10751596.

7. **Bowe AK**, Lightbody G, Staines A, Murray DM. Big data, machine learning, and population health: predicting cognitive outcomes in childhood. Pediatr Res. 2023 Jan;93(2):300-307. doi: 10.1038/s41390-022-02137-1. Epub 2022 Jun 9. PMID: 35681091; PMCID: PMC7614199.

List of Conference Presentations

Oral presentations

1. **Bowe AK,** Lightbody G, O'Boyle D, Staines A, Murray DM
   Predicting low cognitive ability using machine learning methods: Findings from a population based cohort study.
   Faculty of Public Health Medicine Summer Scientific Meeting
   24th May 2023

2. **Bowe AK**, Staines A, Murray DM
   Below average cognitive ability – An under-researched risk factor for emotional-behavioural difficulties in childhood
   European Public Health Association Conference (virtual conference)
   Session: Child and Adolescent Mental Health
   10th – 12th November 2021

3. **Bowe AK**, Staines A, Murray DM
   Below average cognitive ability – An under-researched risk factor for emotional-behavioural difficulties in childhood
   Jaqueline Horgan Bronze Medal Meeting, Epidemiology and Public Health Medicine Section of the Royal Academy of Medicine in Ireland (virtual conference)
   28th October 2021

Poster presentations

1. Bowe AK, Lightbody G, Staines A, Murray DM, Norman M. Prediction of 2-Year Cognitive Outcomes in Very Preterm Infants Using Machine Learning Methods.
   15th International Newborn Brain Conference, Cork
   February 28 - March 2, 2024
   Presented by Professor Deirdre Murray (candidate on maternity leave)

2. Bowe AK, Lightbody G, Staines A, Kiely ME., McCarthy FP, Murray DM
   Predicting low cognitive ability at age 5 – feature selection using machine learning methods and birth cohort data
   European Public Health Association Conference, Berlin
   9th – 12th November 2022
   Presented by Professor Anthony Staines (candidate on maternity leave)

3. **Bowe AK**, Lightbody G, Staines A, Kiely ME., McCarthy FP, Murray DM
   Predicting low cognitive ability at age 5 – feature selection using machine learning methods and birth cohort data
   European Academy of Paediatric Societies (hybrid conference)
   7th – 11th October 2022

4. Bowe AK, Hourihane J, Staines A, Murray DM
Validity of the ages and stages questionnaire for detecting later below average cognitive function
European Public Health Association Conference (virtual conference)
10th - 12th November 2021

5. Bowe AK, Hourihane J, Staines A, Murray DM
Validity of the ages and stages questionnaire for detecting later below average cognitive function
Europaediatrics (hybrid conference)
7th – 9th October 2021

6. Bowe AK, Staines A, Murray DM
Emotional behavioural development in children with below average cognitive function
Europaediatrics (hybrid conference)
7th – 9th October 2021

Awards

University College Cork Postgraduate Research Students Awards 2024
Publication of the Year 2024 – awarded 19th June 2024 for "Prediction of 2-Year Cognitive Outcomes in Very Preterm Infants Using Machine Learning Methods". JAMA Netw Open. 2023 Dec 1;6(12):e2349111. doi: 10.1001/jamanetworkopen.2023.49111

Winner of first place prize for short oral presentation
Faculty of Public Health Medicine in Ireland Summer Scientific Meeting 2023
Bowe AK, Lightbody G, O'Boyle D, Staines A, Murray DM
Predicting low cognitive ability using machine learning methods: Findings from a population based cohort study.

Abbreviations

APA	American Psychological Association
AOR	Adjusted odds ratio
AI	Artificial Intelligence
ASQ	Ages and Stages Questionnaire
AUROC	Area under the receiver operating characteristic curve
BACA	Below average cognitive ability
BAS	British Ability Scales
BIF	Borderline intellectual function
BPD	Bronchopulmonary dysplasia
BSID	Bayley Scales of Infant and Toddler Development
CI	Confidence interval
DAG	Directed acyclic graph
DSM	Diagnostic and Statistical Manual of Mental Disorders
GBM	Gradient boosting machine
HIPAA	Health Insurance Portability and Accountability Act
HSE	Health Service Executive
ICD	International Statistical Classification of Diseases
IDD	Intellectual developmental disorder
IQ	Intelligence quotient
IVH	Intraventricular haemorrhage
KBIT	Kaufman Brief Intelligence Test
LCA	Low cognitive ability
NDD	Neurodevelopmental disorder
NEC	Necrotising enterocolitis
NICU	Neonatal intensive care unit
OR	Odds ratio
PVL	Periventricular leukomalacia
RF	Random forest
SD	Standard deviation
SVM	Support vector machine
TCD	Typical cognitive development
WHO	World Health Organisation
WGS	Whole genome sequencing
WISC III	Wechsler Intelligence Scale for Children – III
WPPSI-R	Wechsler Primary and Preschool Scales of Intelligence - Revised

Acknowledgements

I would like to sincerely thank my supervisors Professor Deirdre Murray, Professor Anthony Staines, and Dr Gordon Lightbody.

To Professor Deirdre Murray – This PhD would not have been possible without you. On top of providing your invaluable academic input, you helped me to navigate this PhD process around all of life's other commitments including getting married, becoming a mother (for the first and second time), and moving country (also for a first and second time) for Franke's fellowships. Myself and Franke are truly grateful for everything you have done to support me, and for enabling us to balance both our careers and our family. Meeting and working with you was to me one of the best things about doing this PhD. Thank you.

To Professor Anthony Staines – It has been a pleasure getting to know you over the past four years. I have learnt so much from you, both academically and otherwise. The scope of your knowledge and your ability to casually talk about such a vast array of topics, from child health to quantum physics, is truly astounding! I enjoyed every single one of our meetings and always left them having learnt something new. Thank you for your guidance in this PhD and always helping me to see the bigger picture.

To Dr Gordon Lightbody – Thank you for all of your exceptionally well-explained input throughout this thesis. It has been a pleasure working with you and exploring

the world of machine learning under your supervision. I always felt I had someone to turn to with my machine learning queries, no matter how big or small. Thank you.

Thank you to the Irish Clinical Academic Training Programme for giving me the opportunity to pursue this research. Thank you to all the other ICAT fellows who are a fountain of knowledge, tips, and support. Thank you to Dr Carol Stephens for being a great friend and colleague.

Sincere thanks to the Faculty of Public Health Medicine for their encouragement and support. I would particularly like to thank Professor Emer Shelley, Dr Regina Kiernan, Dr Emer O'Connell, Dr John Cuddihy, Dr Paul Kavanagh and Dr Anne Sheahan who have all provided fantastic support during both my clinical and academic training. I would like to pay a special thanks to Dr Peter Barrett who has been, and continues to be, an exceptional mentor and role model in public health medicine. Thank you Peter for your endless words of advice, encouragement, and support.

To my parents – Words cannot express how grateful I am for everything you have done and continue to do for me. I love you both. To my parents in law – Thank you for all your support and for all you do for our family. Thank you to all of my friends, without whom life would not be nearly as fun!

To Franke, Dáire, and Frankie – You are everything. I dedicate this thesis to you.

Abstract

Background

Approximately 16% of children have below average cognitive ability. Around 1-2% have an intellectual developmental disorder (IDD), but a larger proportion, approximately 14%, have cognitive ability between 1 and 2 standard deviations below the mean. The outcomes for this larger group have received little attention in scientific literature. For both groups, early intervention may improve cognitive outcomes, particularly if commenced in the first 1,000 days of life. The aim of this thesis was to explore the impact of below average cognitive ability on outcomes in childhood and to examine current and alternative methods, namely the application of machine learning to clinical and epidemiological data, for the early prediction of below average cognitive ability in childhood.

Structure and methods

The thesis begins with an introductory chapter (Chapter 1), followed by a statement of aims and objectives (Chapter 2). Six original studies are then presented (Chapters 3-8). The relationship between below average cognitive ability and emotional-behavioural function in childhood is examined using generalised linear mixed models and logistic regression (Chapter 3). The causal relationship between below average cognitive ability and the early experience of school is investigated using logistic regression and directed acyclic graphs (Chapter 4). The predictive validity of a widely used developmental screening tool, the Ages and Stages Questionnaire, is then investigated (Chapter 5), before a series of papers examining the application of a

range of machine learning techniques to clinical and epidemiological data for predicting below average cognitive ability in childhood among a general population (Chapters 6 & 7), and in infancy among a very preterm population (Chapter 8). The studies are based on data from the Growing up in Ireland Infant Cohort (Chapters 3, 4, & 7), the Cork BASELINE birth cohort (Chapters 5 & 6), and the Swedish Neonatal Quality Register (Chapter 8).

Results

Having cognitive ability between 1 and 2 standard deviations (SDs) below the mean at age 3, compared to having average cognitive ability, was independently associated with increased odds of experiencing a clinically significant emotional-behavioural problem between age 3 and 9 years (AOR 1.39, 95% CI 1.17-1.66, p<0.001). These children, on average, experienced a significant worsening in their emotional-behavioural function between age 5 and 9 years, the time coinciding with commencing formal education in Ireland (mean difference 0.48, 95% CI 0.06 to 0.90, p=0.016), compared to peers with average cognitive ability who had no significant change (mean difference -0.06, 95% CI -0.19 to 0.07, p=0.488). (Chapter 3)

Compared to their peers with average cognitive ability, children whose cognitive ability was more than 1 SD below the mean at age 5 had significantly higher odds, at age 9 years, of experiencing low self-concept (AOR 1.20, 95% CI 1.02-1.42); of hardly ever or never being interested, motivated, and excited to learn (AOR 2.29, 95% CI 1.70-3.10); and of experiencing teacher-reported emotional-behavioural difficulties (AOR 1.34, 95% CI 1.10-1.63). (Chapter 4)

The Ages and Stages questionnaire, a routinely used developmental screening tool, when performed at 24 or 27 months would only detect 21% of infants who would go on to have IQ more than 1 SD below the mean at age 5. (Chapter 5)

Using machine learning techniques to rebalance data and perform feature selection, on internal validation, a random forest model trained on the BASELINE cohort could correctly predict at birth, using 11 perinatal features, the cognitive outcome at age 5 of 95% of children. It achieved a sensitivity of 89% and a specificity of 98% for identifying those with an IQ less than 90 at age 5 years. External validation was not performed. (Chapter 6)

On external validation using unseen data, a random forest model trained on the Growing Up in Ireland cohort, using 15 participant-reported features in the first year of infant life, could correctly predict the cognitive outcome at age 5 for 87% of infants. However, the accuracy was largely driven by a high specificity of 90%, with a sensitivity of 40%. (Chapter 7)

In a high risk cohort of very preterm infants from the Swedish Neonatal Quality register, a logistic regression model containing 26 features could identify, at discharge from the neonatal unit, 93% of infants who experienced cognitive delay at 2-year follow up, with a specificity of 46%. (Chapter 8)

Conclusion

Children with below average cognitive ability, including those whose cognitive ability lies between 1 and 2 standard deviations below the mean, are more likely to experience emotional-behavioural difficulties, low self-concept, and poor classroom engagement in childhood. Early intervention could improve outcomes for these

children. Among a general paediatric population, applying machine learning to readily available clinical and epidemiological data, may be a step toward an early, scalable, risk stratification tool, to identify the children who would benefit most. The predictive performance of the model developed in this thesis is not yet good enough for implementation. Among a very preterm population, the model developed in this thesis shows great promise for early identification of those at risk of later cognitive delay.

Chapter 1 Introduction

"The question is not whether we can afford to invest in every child; it is whether we can afford not to"

Marian Wright Edelman

Cognitive ability in early life is both a predictor and a consequence of important factors in the life course including physical health, educational attainment, mental health, and occupational outcomes.[1-4] Early cognitive development is influenced by factors operating before, during, and after birth.[5,6] The 'First Thousand Days', the period from conception to a child's second birthday, represents a crucial window of opportunity in child development.[5-7] During this period, the infant brain is highly susceptible to influences, both positive and negative, that can have pervasive and far-reaching effects. [5-7]

Disparities in cognitive development which occur in early life tend to persist and increase over time.[8] Their association with later adverse outcomes has been well documented, but less has been published on how cognitive ability impacts outcomes in childhood. Providing the right intervention, at the right time, to the right people could help reduce disparities in cognitive development. From a biological, sociological, and economic perspective intervening in the first 1,000 days makes sense.[9,10] A challenge is identifying, at the earliest possible stage, which children and families are most likely to benefit. In this thesis we explore the impact of below average cognitive ability on outcomes in childhood and examine alternative methods of early identification of children at risk of below average cognitive ability.

1.1. Cognitive ability

Cognitive ability is a broad and complex construct, for which many definitions exist across different disciplines. The American Psychological Association (APA) define cognitive function as "the performance of the mental processes of perception, learning, memory, understanding, awareness, reasoning, judgement, intuition, and language".[11] Across definitions, there is largely consensus that cognitive ability can be broadly considered as one's ability to learn, reason, and problem-solve.[12,13]

The terms "intelligence" and "cognitive ability" are often used interchangeably in the literature, and while the two concepts have much in common, there are some distinguishing features.

The psychologist Heiner Rindermann described intelligence as the ability to think, free of specific prior knowledge. Intelligence, according to Rindermann, encompasses solving new problems by thinking, and not just recalling prior knowledge, by reasoning and drawing logical conclusions from new information.[14] He describes cognitive ability as a broader concept incorporating intelligence, memory, knowledge, and the intelligent use of this knowledge.[14] Rindermann's differentiation between intelligence and cognitive ability aligns with the concept of fluid and crystallised intelligence, described by American psychologist Raymond Cattell.[15] Cattell criticised existing intelligence tests at the time, such as the Binet intelligence test, over concerns that the tests focussed largely on acquired skills and verbal ability rather than intelligence.[16] Fluid intelligence refers to a general ability to reason, process information, perceive relationships and problem solve. Crystallised intelligence consists of acquired skills and knowledge, originally gained through fluid

ability, but no longer requiring fluid ability for use. It includes vocabulary, literacy, numeracy, and domain specific knowledge or skills.[15-17]

A contemporary of Cattell, Donald Hebb, developed a similar concept of intelligence based on his studies of brain injuries and performance on intelligence tests. Hebb also described two components to intelligence, one of which was an innate ability and the other the result of experience.[18] It was Hebb who proposed the use of a battery of tests including both verbal and non-verbal tests of intelligence to capture the different aspects of cognitive ability, and most contemporary tests of cognitive ability adopt this approach.

It makes intuitive sense that fluid and crystallised intelligence exhibit a degree of correlation. What one can learn and store as knowledge (crystallised intelligence) is determined, in part, by how well one can learn (fluid intelligence). People with high levels of one ability tend to have high levels of another.[17] However, what one learns is also determined by other factors such as motivation, personality, and opportunity.[17,19]

All cognitive abilities tend to follow a similar pattern – improvement with age until a peak point of performance is reached, followed by a slow decline. The timing, pattern, and speed of this process varies according to the specific ability.[20] Fluid ability increases rapidly throughout childhood until approximately age 20, after which time it will slowly decline. Conversely, crystallised ability will improve throughout adulthood as experience and knowledge increase, declining only at very old age or in the presence of certain disease processes.[20]

1.2. Psychometric assessment of cognitive ability

The study of within-person and between-person differences in cognitive ability is most commonly performed using standardised measures of cognitive ability, including intelligence quotient (IQ) tests, of which hundreds exist.[21] Some tests use a single item and measure one domain of ability. For example, the Peabody Picture Vocabulary test is a measure of verbal ability, a form of crystallised intelligence.[22] Other tests use a battery of subtests to measure different types of cognitive ability. For example, the Wechsler Intelligence Scale for Children is comprised of ten subtests measuring different non-verbal and verbal abilities.[23] Subtest scores are summed to give an overall ability or IQ score, and conventionally these scores are standardised to an age-neutral scale with a mean of 100 and a standard deviation of 15.[24] This means that the mean IQ of 5 year old children will be 100, as will the mean IQ of 30 year old adults. A cognitive ability or IQ score is therefore not an absolute measure of intelligence but is a relative measure, measuring an individual's cognitive ability compared to that of the general population.[25] Over decades of intelligence research many important observations on standardised measures of cognitive ability have been made.

Firstly, individual performance on different cognitive tasks tends to be positively correlated. If an individual performs above the population average on one cognitive task, they will likely perform above the population average on another cognitive task.[26] This observation was first made in 1904 by Charles Spearman, who noted that children's test scores, relative to others, across unrelated school subjects were positively correlated.[14] He termed this phenomenon "the positive manifold". To explain this observation he posited the existence of "the g-factor" which he

described as a general mental ability, unmeasurable with a single test, but which could be derived using factor analysis from performance on several cognitive tasks.[26,27] Multiple studies have replicated this finding, and have shown that among large samples who undertake multiple diverse cognitive tests, a "g-factor" accounts for approximately 25%-50% of the total test variance.[28]

While the existence of this statistical observation is beyond doubt, many have challenged Spearman's explanation. Shortly after Spearman published his work, he was challenged by Godfrey Thomson who proposed that the observation was due to overlap in the mental processes involved in each cognitive task. He suggested that the factor analysis results were due to overlap in the mental processes being used across different subtests, rather than a common factor 'g' measuring general mental ability.[29] This debate has not yet reached a conclusion in the literature.

Secondly, between individual differences in cognitive ability show substantial stability from childhood to adulthood.[30,31] One of the longest follow-up studies, among a cohort of Scottish children, reported a correlation of 0.73 between cognitive ability measured at age 11 and age 77.[30] Between age 18 and age 50 correlation coefficients for cognitive ability of 0.95 have been reported among a Swedish male cohort.[31] The correlation tends to decrease as the age at first measurement gets earlier, and the greatest variability in interindividual differences is seen in early childhood.[20,32]

1.3. Terminology and categorisation of cognitive ability

1.3.1. Intellectual developmental disorders

The International Statistical Classification of Diseases and Related Health Problems (ICD) is a World Health Organisation (WHO) classification system used globally for recording accurate and comparable diagnoses. It is used in research, epidemiology, in disease registries, death certification, and for resource allocation. The ICD-11 came into effect on January 1st 2022.[33] Disorders of intellectual development are defined in ICD-11 as "a group of etiologically diverse conditions originating during the developmental period characterised by significantly below average intellectual functioning and adaptive behaviour that are approximately two or more standard deviations (SDs) below the mean, based on appropriately normed, individually administered standardised tests".[34] Mild, moderate, and severe intellectual developmental disorders (IDDs) are defined by intellectual functioning and adaptive behaviour that are approximately 2-3 SDs below the mean, 3-4 SDs below the mean, and more than 4 SDs below the mean, respectively.[34] These cut-offs are equivalent to IQ scores of 55-70 for mild IDD, 40-55 for moderate IDD, and <40 for severe IDD.[35]

The distribution of IQ scores among a population tends to assume a normal distribution with approximately 68% of people having a score within plus or minus 1 SD of the mean, equivalent to an IQ score of 85-115.[13] This means approximately 14% of the population will have intellectual functioning that lies between 1 and 2 SD below the mean (IQ 70-85), on the border between typical intellectual functioning and intellectual disability.[36] In previous versions of both the ICD and the American Psychological Association Diagnostic and Statistical Manual of Mental Disorders (DSM), this level of intellectual function was clearly defined as borderline intellectual

function (BIF), and when combined with functional impairment came under the diagnostic category of what we now call IDDs. Around the year 1980, both the ICD and DSM removed borderline intellectual function from this diagnostic category, and in both classification systems it is now a condition contained in a non-specific "Other" category, lacking a clear definition and without a defined IQ criterion.[36] There were valid arguments for removing BIF from the IDD category, namely to avoid overdiagnosis of IDD in people without functional impairment, particularly in school settings, and to avoid what was deemed discrimination of socioeconomically disadvantaged people with this label. However, there were also adverse consequences.[37] Without a well-defined classification, people with borderline intellectual function have received little attention in scientific literature, and therefore little is known about the consequences and needs of this subgroup of the population.[38] What has been published confirms that people with borderline intellectual function have a higher prevalence of mental health difficulties, substance abuse disorders, homelessness, and poverty compared to the general population.[36,38]

1.3.2. Cognitive delay

The term cognitive delay comes under the general descriptor of developmental delay.[39] Unlike IDD and BIF, terms which can be used in childhood and adulthood, developmental delay is a term confined to use in childhood. Developmental delay occurs when a child fails to attain appropriate developmental milestones when compared to their typically developing peers.[40] In infants and young children, most standardised tests are not direct measures of cognitive ability, but instead use the trajectory of milestone attainment as a surrogate marker of brain development.[40]

Developmental delay includes cognitive delay, but also includes gross motor, fine motor, speech and language, and social delays.[39] It can be isolated to a single domain or can be a global developmental delay affecting two or more domains.[40] Most developmental delay is idiopathic and time-limited but for some children, it will persist and lead to a diagnosis of a developmental disorder, a large group of syndromes which includes IDDs.[39,40] Developmental delay refers to a transient state that either resolves or evolves into a developmental disorder.[40]

1.3.3. Terms used in this thesis

In this thesis, children with standardised cognitive ability scores more than 1 SD below the mean (equivalent to an IQ less than 85) are referred to as children with either "low cognitive ability", "below average cognitive ability", "low average cognitive ability" or "cognitive delay". These terms are used interchangeably and the choice of term in each paper was largely determined by the discipline of the academic journal in which the work was published and feedback from reviewers. The term "adverse cognitive outcome" is used in this thesis to describe a cognitive outcome which has been negatively impacted by a defined condition or event, or by a combination of defined conditions or events.

In this thesis, in the studies examining the association between below average cognitive ability and childhood outcomes (Chapters 3 and 4), separate analyses were conducted for those with cognitive ability scores between 1 and 2 standard deviations below the mean, and those with scores more than 2 standard deviations below the mean. This was done to ensure that any significant findings in the below average group as a whole were not driven only by those with more severe

cognitive deficits with general cognitive ability scores more than 2 standard deviations below the mean. In the studies predicting below average cognitive ability (Chapters 5-8) those with cognitive ability more than 2 standard deviations below the mean were combined with those scoring 1-2 standard deviations below the mean. The rationale for this was that any prediction tool aiming to enable early intervention should aim to identify all those who would potentially benefit.

1.4.The development of cognitive ability in early life

1.4.1. The first 1,000 days

The phrase "the first 1000 days", a period spanning conception to a child's second birthday, is one increasingly seen in policy documents relating to child health.[41,42] The phrase has its origins in the work of Professor David Barker, a physician and epidemiologist, whose trio of seminal articles published in the Lancet between 1986 and 1993 prompted the development of the 'fetal origins hypothesis', which posits that the events of early development have a profound impact on one's future risk of adult disease.[43–45] It is now widely accepted that the events in the first 1,000 days of life will have significant and life-long impacts on a child's health and well-being.[41,46]

A fundamental concept in Barker's work was that of "developmental plasticity", formally defined as "the ability of a single genotype to produce more than one alternative form of structure, physiological state, or behaviour in response to environmental conditions".[47] This complex interplay between genetics and the environment explains how genetically identical individuals, exposed to different environmental conditions in early life, can exhibit different traits and behaviours.[48]

With regard to the developing brain specifically, the first 1,000 days is considered a crucial window, of both vulnerability and opportunity.[49] During this window, neuroplasticity is at its greatest and the young brain has a variety of mechanisms, not evident in adults, that allow it to develop appropriately and adapt constantly to environmental stimuli.[9,50] The greater potentiality of the brain toward plasticity during this period is explained in part by the highly dynamic dendritic spines, which are predominant during early development, and display huge potential for adaptability but, which over time and in adulthood are much more stable.[9,50]

An individual's genetic code can be considered as the framework, upon which environmental factors will act to shape the structure and function of the developing brain.[46] The developing brain will change structurally and functionally in response to experience, and will adapt to the early environment in ways that have life-long consequences.[41] This adaptation occurs primarily through two mechanisms – epigenetics and synaptic pruning.[41]

Epigenetic change refers to the alteration in how genes function in response to the environment.[51] The DNA sequence making up the gene remains unchanged, but how the gene is expressed will vary depending on the environment to which it is exposed. The epigenetic system is most sensitive to environmental influences during the period of optimal neuroplasticity in the first 1,000 days of life, and changes can occur from just after conception throughout the prenatal and postnatal periods.[52] For example, early life stressors such as loss of a parent or poor quality parenting can produce epigenetic modifications to the genes controlling the neuroendocrine response to stress, resulting in higher basal stress hormone levels as an adult, which in turn can increase the risk of many noncommunicable diseases.[53] These epigenetic

modifications can be transmitted across generations, therefore the influence of this first 1,000 days may not just affect the long-term outcomes for one generation, but may impact outcomes in generations that follow resulting in an intergenerational transmission of adversity.[41,54]

The second mechanism critical to our understanding of the interplay between genetics and environment is synaptic pruning. A synapse is the point of contact between two brain cells, also known as neurons.[55] After birth, there is a surge in synaptogenesis, and billions of neurons will begin to communicate with each other during a period of rapid learning.[41] Different regions of the brain will reach peak synapse production at different time points. For example, areas of the prefrontal cortex, which largely control thoughts, actions and emotions, reach their peak of synaptogenesis at the 15th postnatal month.[55] The later the peak synapse production occurs, the longer the region remains plastic. This initial overproduction of synapses is followed by a gradual reduction, in a process called synaptic pruning, which is highly experience dependent and will occur at different time points depending on the area of the brain.[55] Connections which are used and activated through the child's interactions with their environment become stronger and endure, while those that are not may remain dormant.[55] These initial connections will form the foundations of life-long cognitive, social, and emotional skills.[41,55,56]

"Sensitive periods" represent key stages in a child's development, where their experiences can either potentiate or inhibit neural activity.[46] During these periods the brain is highly responsive to its environment, and the brains capabilities can be enhanced through quality positive experiences or adversely modified through absence of quality experiences or exposure to negative experiences.[46] A failure to

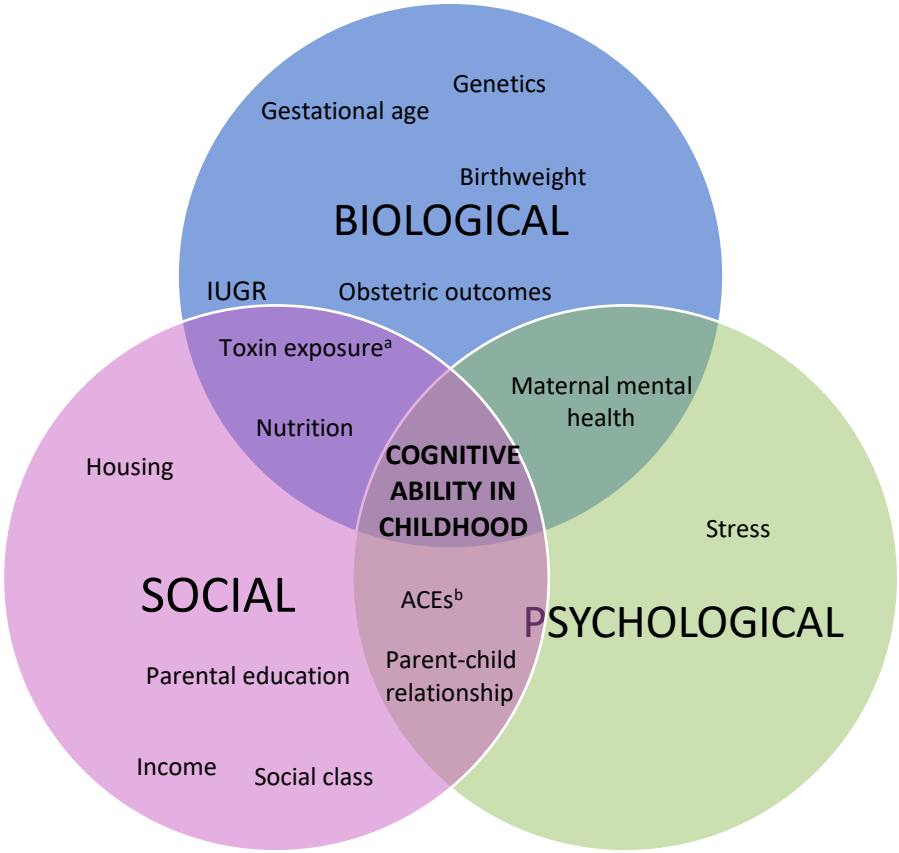
achieve early foundational cognitive skills during the first 1,000 days of life, a sensitive period in cognitive development, may result in a permanent loss of opportunity to reach full cognitive potential. Conversely, with appropriate intervention, it is a period during which there is opportunity to enrich the environment and enhance neural activity.[9]

1.4.2. Conceptual framework

During the first 1,000 days, there are a myriad of factors associated with cognitive outcomes in childhood, which can exert their effects through both epigenetic pathways and synaptic pruning. There are multiple ways to categorise these factors. They could be categorised according to the time period in which they predominantly act, for example whether they predominantly exert their effect in the perinatal or postnatal period, or they could be categorised according to whether they represent biological, social, or psychological risks. In Figure 1.1, Engel's Biopsychosocial model is used to visualise early life factors associated with childhood cognitive ability.[56] This model is helpful in that it enables us to acknowledge the intimately intertwined relationships between the biological, psychological, and social factors that contribute to the development of a child's cognition.[56]

A limitation of Engel's model and of Figure 1.1 is that it does not sufficiently capture the dynamic and constantly evolving relationships between these early life factors and cognitive ability, which itself is a not a static construct. There is a huge body of literature examining relationships between individual exposures and

Figure 1.1 Conceptual framework for the perinatal and early life factors associated with cognitive ability in childhood according to Engel's Biopsychosocial Model[56]



[a] IUGR, Intrauterine growth restriction; Toxin exposure includes alcohol, tobacco, cocaine, opioids, air pollutants including $PM_{2.5}$; [b] ACEs, Adverse childhood experiences.

cognitive ability in childhood, and to review each of these is beyond the scope of this thesis. For some exposures, the evidence is well-established and likely causal associations have been demonstrated consistently across studies, whilst for others there is conflicting or weak evidence based on single small studies. This section will focus on exposures which have demonstrated a consistent causal association with cognitive ability in childhood.

1.4.2.1 Biological exposures

Findings from across family, twin, and adoption studies consistently estimate that approximately 50% of the variance in general cognitive ability is accounted for by differences in DNA sequences inherited from the biological parents.[57] Recent advances in genetics research, including the ability to map whole genomic sequences, have resulted in the identification of genome sequences that account for 20% of this 50%.[58] However, the gene-environment interaction will determine how this genetic code is expressed, which genes will be amplified, and which will remain dormant.[59]

Toxin exposure in the prenatal period will have direct biological effects on the developing brain of the baby, but it is also reflective of the social environment of the mother-baby dyad. Exposure of the fetus to toxins consumed by the mother in the prenatal period, including alcohol[60], cocaine[61], and opioids[62] can increase the risk of cognitive difficulties in childhood, depending on the timing and concentration of exposure. There is also increasing evidence that maternal exposure to environmental toxins, such as air pollutants, during pregnancy is associated with lower IQ among offspring.[63]

With regard to nutrition, a suboptimal nutritional environment, both in the uterus and postnatally, can have detrimental and long lasting effects on the developing brain.[64] The period between the third trimester of pregnancy and two years of post-term life is characterised by rapid brain development, which is dependent on adequate macro- and micro-nutrients. The neonatal brain consumes approximately 60% of total caloric consumption. It is increasingly recognised that successful postnatal nutrition is highly dependent on the fetal nutritional status prior to birth, which is provided from the mother to the baby via the placenta.[64] Poor weight gain in the intrauterine period, or intrauterine growth restriction which refers to a rate of fetal growth that is less than normal considering the growth potential of the infant, are associated with worse neurodevelopmental outcomes.[64]

With regard to specific nutritional deficiencies, iodine deficiency is associated with adverse cognitive outcomes in childhood, with the mechanistic pathway being driven through the role of iodine in thyroid hormone production, which itself is crucial for cognitive development in the fetus and infant.[64] Iron deficiency, which results in maternal anaemia in early pregnancy and can subsequently lead to neonatal anaemia, is associated with a higher risk of neurodevelopmental disorders including intellectual disability.[65]

Maternal infections, both bacterial and viral, during pregnancy and particularly in the third trimester can be associated with lower IQ in childhood.[66] Specific infections such as rubella and herpes simplex virus which affect the central nervous system can result in profound intellectual deficits in offspring.[66] Other bacterial and viral infections in pregnancy, can result in vertical transmission, and can disrupt the developing immune system in the infant resulting in low grade

inflammation and a systemic immune response, which has been shown to have negative effects on childhood intelligence.[66]

Factors related to the delivery and birth of the infant, particularly preterm delivery which will be discussed in the following section, can result in adverse cognitive outcomes in childhood. Obstetric complications, such as perinatal asphyxia, can result in cerebral palsy and in hypoxic ischaemic encephalopathy. Children who suffer perinatal asphyxia, where therapeutic hypothermia is not applied, and who do not develop cerebral palsy, have lower cognitive assessment scores at early school-age compared to healthy controls.[67]

1.4.2.2 Psychological exposures

Maternal stress, anxiety, and depression in both the antenatal and postnatal period are associated with negative effects on the cognitive development of offspring.[68,69] These exposures affect cognitive development in offspring primarily through two mechanisms. Firstly, in the antenatal period, these exposures can negatively affect the developing fetal brain through direct biological effects.[68,69] Maternal stress, anxiety, and depression lead to increased fetal cortisol concentration, which can alter activity in the fetal hypothalamic-pituitary-adrenal axis, and can result in reduced placental blood flow with subsequent reductions in oxygen and nutrients to the developing fetus.[69] The timing of in-utero exposure to maternal stress is now thought to be important.[70] Evidence suggests that early gestational stress may be particularly relevant for cognitive outcomes, as it is during early gestation that important structures such as the hippocampus and amygdala, which will form later connections with the prefrontal cortex, are formed.[70]

Maternal stress, anxiety, and depression can also alter maternal behaviour and the ability of the mother to form a close, nurturing relationship with their baby.[68,69] It is well documented that stress, anxiety, and depression are intimately interlinked with socioeconomic circumstances, and people living in poverty will experiences greater levels of stress and stressful events in their lives.[68]

The early child environment may also be damaged through adverse childhood experiences (ACEs) and exposures such as domestic violence, drug-abuse, and homelessness. A population-based cohort study of 3,153 mother-child pairs in the United Kingdom found that IQ scores of children born to mothers exposed to physical violence remained significantly lower than those of unexposed children across full-scale, verbal, and performance IQ.[71] These exposures, which are inherently linked to social circumstances, damage parent-child relationships, result in stressful home environments, and can reduce quality parent-child interactions and cognitive stimulation, all of which adversely affect the cognitive development of the child through both epigenetic and synaptic pruning mechanisms.[72]

1.4.2.3 Social Exposures

There is a wealth of literature, both interventional and observational, which has consistently shown that the early socioeconomic environment of an infant is one of the strongest predictors of cognitive outcomes in childhood.[73–78.] Socioeconomic status, which may be reflected in parents education level, family income, housing, or social class, strongly influences both the prenatal environment and early-life environment of the infant. Low socioeconomic status is itself a risk factor for prenatal toxin exposure, nutritional deficiency, and maternal psychological distress.[75,76] In the

postnatal period, the effect of low SES on childhood cognitive outcomes may be mediated through many processes, which ultimately result in a suboptimal environment for a child's cognitive development.[75,76] For example, it may be mediated through increased maternal psychological distress resulting in insecure mother-infant attachment, through increased in-utero cortisol exposure resulting in direct biological effects on the developing brain, through a lack of physical resources such as books and toys, or through a lack of exposure to wide vocabulary and verbal interactions.[75-79]

There is now objective evidence, obtained through a randomised controlled trial involving 1,000 low-income mother-infant dyads and electroencephalograph monitoring of the infant at 1-year, that a reduction in poverty through an unconditional monthly cash payment can positively change infant brain activity, in ways linked to the development of subsequent cognitive skills.[80]

1.4.3. Preterm infants

Preterm infants are one particular subgroup of the population with increased biological risk for below average cognitive ability in childhood. The average IQ, across child and early adulthood, for infants born at <32 weeks gestation has been reported to be almost 1 SD lower than their full-term counterparts, equivalent to a mean difference of approximately 13 IQ points.[81] During weeks 24 to 40 of gestation, a multitude of processes critical for brain development occur, all of which are highly vulnerable to insults such as hypoxic events, inflammation, haemorrhage, and stressful environmental exposures.[81] Prematurity predisposes infants to conditions such as intraventricular haemorrhage (IVH), periventricular leukomalacia (PVL),

bronchopulmonary dysplasia (BPD), and necrotising enterocolitis (NEC) which have been associated with later cognitive impairment.[81] Early environmental stress in the neonatal unit may also contribute.[82]

There is also increasing evidence that children born moderate to late preterm (32-36 weeks gestation), even in the absence of IVH and PVL, are at increased risk of later poor cognitive outcomes.[83,84] This may suggest that during this later period of gestation, there is still considerable growth and maturation of the brain taking place, which remains vulnerable to interruption.[84] The factors which determine outcome in these children without objective brain injury on imaging are unclear.[83,84]

1.5. The importance of childhood cognitive ability

The cognitive skills achieved in early life form the foundation for more complex cognitive processes.[46] Cognitive skills are one of the most important determinants of academic achievement and educational attainment, which in turn can determine occupational and financial outcomes.[58,85] Within the workplace, it is cognitive ability that can determine how well one can adapt to novel tasks and challenges.[86] Cognitive ability will also influence physical and mental health. This relationship may be mediated through educational attainment, which may enable informed and healthy lifestyles choices, or through occupational and financial outcomes which will determine where and how one lives their life.[87,88]

Acknowledging that other non-cognitive skills, such as social, interpersonal, communication, and emotional skills, are also important determinants of these outcomes, there is well-established evidence to show that the level of cognitive function achieved in childhood will have a profound impact on outcomes throughout

the life course including academic achievement, physical health, mental health, financial well-being and social mobility.[87-96]

1.5.1. Academic achievement

Cognitive ability in childhood is not only an important predictor of later cognitive ability, it is also a predictor of academic and educational achievement.[89-91] Cognitive ability has been shown consistently, across both cross-sectional and longitudinal studies, to have a correlation coefficient of approximately 0.50 with educational achievement, indicating that as cognitive ability increases academic achievement will also increase in a positive manner.[85,89] A longitudinal study of 209 children in the Netherlands found that IQ measured at ages 5,7,10 and 12 had correlations of 0.41, 0.50, 0.60, and 0.63 respectively with educational achievement measured using a national test at age 12.[89] One of the largest longitudinal studies, examining the association between IQ at age 11 and educational attainment in national examinations at age 16, among more than 70,000 children in England found the correlation between a latent intelligence trait and a latent educational achievement trait was 0.81. For a one standard deviation increase in IQ the odds of attaining five or more honours grades increased 7-fold.[90]

Studies using causal models to examine the relationship have found that the positive relationship between IQ and academic achievement persists even after adjustment for early conduct problems, family, social, and childhood circumstances.[91] A longitudinal birth cohort of more than 1,200 infants in New Zealand which measured IQ at ages 8 and 9 found, after adjusting for a range of social, family, and childhood factors, intelligence had a direct relationship with later

educational attainment including attainment of post-school qualifications and university degrees.[91] Across the literature, children with higher IQ tend to have better educational outcomes, and this is likely due to a greater ability to learn. Other factors such as personality, motivation, quality of teaching, parental support may increase or decrease the correlation or association between IQ and educational achievement but none eliminate or mediate it completely.[24,92,93]

1.5.2. Physical health

A consistent inverse relationship has been reported in the literature between childhood cognitive ability and all-cause adult mortality. A meta-analysis of 16 prospective cohort studies from Sweden, Denmark, Australia, the United States, and United Kingdom found a 1 standard deviation advantage in childhood cognitive ability was associated with a 24% lower risk of mortality over a 17 to 69 year follow up period (Hazard ratio 0.76, 95% confidence interval (CI) 0.75-0.77). Adjusting for socioeconomic status in childhood did not attenuate this association.[2]

Calvin et al in a prospective cohort study of more than 65,000 individuals born in Scotland in 1936, with linked mortality data across 68 years of follow-up, found the strongest associations for coronary heart disease, stroke, smoking related cancers, digestive diseases and external causes such as accidents, where the risk of mortality from these diseases was halved for those in the highest 10% of intelligence scores in childhood compared to those in the lowest.[94] The causal pathway between early cognitive ability and the occurrence of disease or death may be mediated by health behaviours, health literacy, occupational class, or adult socioeconomic status.[2,94]

1.5.3. Mental health

A prospective cohort study of more than 1,000 participants in New Zealand found higher childhood IQ was associated with a reduction in the odds of lifetime schizophrenia, of adult depression, and of adult anxiety disorders, with findings remaining significant after adjustment for child socioeconomic status, adverse perinatal events, low birthweight, and childhood maltreatment.[88] A Danish study of 7,022 males also found childhood IQ scores were inversely related to the risk of total psychiatric illness, defined as any psychiatric discharge diagnosis from a psychiatric ward. Again, adjusting for paternal occupational social class and birth weight did not negate the effect of IQ.[3] In addition to the potential mediating factors outlined in the previous section, one hypothesis proposed to explain the causal pathway between childhood IQ and adult mental health outcomes is the "cognitive reserve" hypothesis which posits that individual differences in the structure and function of the brain can protect against or predispose to the effects of neuropathology.[88]

1.5.4. Financial well-being and social mobility

Childhood IQ has a positive association with upward social mobility.[95] In a prospective study of 1,142 children born to mothers in Newcastle, England in 1947 and followed up at age 49-51 years, after adjustment for gender, housing conditions, and adverse events in childhood, the level of education achieved was positively associated with intergenerational upward social mobility (defined as an increase in social class between parental social class in childhood and the individual's own social class in adulthood). Childhood IQ was positively associated with intragenerational upward

social mobility (defined as an increase in social class in adulthood between ages 25 and 49-51 years).[95] Cognitive ability in childhood is also an independent predictor of adult earning ability.[96] This relationship persists even after educational qualifications and occupational status are accounted for.[96]

1.5.5. Outcomes and experiences in childhood

Many studies in the literature have examined the relationships between childhood cognitive ability and adult or later-life outcomes. However, considerably less has been published on the role childhood cognitive ability plays in outcomes and experiences in childhood. Many of the studies in this area focus on children with intellectual developmental disorders and do not include children whose cognitive ability lies between 1 and 2 standard deviations below the mean and who do not meet the diagnostic criteria for an IDD. A small cross-sectional study of 104 children, of whom 25 had a standardised cognitive test score lower than the 15th percentile, found that these children had the lowest mean scores for self-esteem and highest mean scores for depression and anxiety, compared to children with typical cognitive ability and those with gifted ability. These results did not account for confounding by gender or by socioeconomic background.[97]

Emerson et al in a study of 4,337 Australian children found that 4-5 year old children with borderline intellectual functioning (BIF), defined as a standardised cognitive score between 1 and 2 SDs below the mean, and children with an IDD, defined as a score more than 2 SD below the mean, had significantly higher rates of possible mental health problems at age 6-7 years compared with their typically

developing peers. These differences persisted after accounting for gender and socioeconomic status.[98]

The relationship between low cognitive ability and poor mental health is complex and causality may be bidirectional. One hypothesis would be that poor mental health results in low cognitive ability, as a child with a mental health problem is less likely to engage in learning and education. An alternative hypothesis is that low cognitive ability is a cause of poor mental health in childhood. The findings of Emerson et al would support this hypothesis but their work does not account for the baseline mental health of the child, therefore reverse causality remains a possibility.[98]

It is highly plausible, that children with below average cognitive ability, would experience increasing cognitive difficulties upon starting school when the complexity of cognitive tasks increases, and would therefore experience a worsening in their mental health as they struggle with daily school tasks and endeavour to keep up with their typically developing peers. Evidence of a worsening in mental health function at the time of transitioning to formal education would support this alternative hypothesis. However, this has not been investigated in the literature to date.

Many studies have examined the relationship between child IQ and academic achievement, but very few have examined how cognitive ability in childhood influences a child's early experience of school, an experience which will have life-long impacts on a child's relationship with learning and education. A small cohort study by McIntyre et al, comparing early school outcomes for 24 children with an IDD and 43 children without, found children with an IDD had significantly more teacher-reported problem behaviour, fewer self-regulation and social skills, and poorer

student-teacher relationships.[99] Children with below average cognitive ability, not meeting criteria for an IDD, were excluded from their study. The early school experience for these children, whose difficulties may be more likely to go unrecognised and unsupported, has not previously been investigated in the published literature.

1.6. Evidence for Early Intervention

Across epidemiological, animal, and human research there is consensus agreement that early life is a period of fundamental importance in cognitive development, and strategies for intervention should ideally target this unique period.[6,7,9,41,42,53,55]

1.6.1. Experimental evidence: The Abecedarian Study

One of the most widely cited examples supporting early cognitive intervention is the Abecedarian Study, a randomised controlled trial commenced in 1972.[100] Included in this study were infants and families, recruited predominantly before birth through prenatal clinics in North Carolina Memorial Hospital, who were considered high risk of poor cognitive outcomes based on an individual high risk index score. This was calculated using a weighted combination of 13 sociodemographic factors which included parental education, family income, welfare payments, and the school history of family members. Infants with a biological condition associated with mental, sensory, or motor disabilities were excluded.[100,101] The trial involved 111 infants, 98% of whom were African-American, who were matched on sex, maternal IQ, number of siblings, and high risk index scores before being randomly assigned to the intervention or control group. All families, both in the control and intervention

groups, received social support services, standard health advice, free medical care for the child, infant formula, free transport to the research centre, free diapers, and financial remuneration for the mother if they undertook psychological testing.[101]

The intervention was delivered from age 0-5 years and consisted of high quality early education, delivered with high teacher-to-pupil ratios five days per week for fifty weeks of the year. Alongside this, there was an emphasis on family involvement and a home visitor from the education centre would visit the infants family at regular intervals to support them and help them implement educational activities in the home.[100] Each child had an individualised curriculum based on specific learning objectives and evidence-based practice at the time. Each activity was designed to achieve a specified learning objective, and a parent and teacher guide was created to demonstrate each activity. An example of an activity that was delivered for infants at age 2-3 months was "Holding the baby for a better look at things". This was a simple intervention to adjust the height at which the baby was held with the objective of improving head lift and visual attention to immediate surroundings.[101]

A critique of many early intervention programs is that long-term follow up is not completed, however one of the major strengths of the Abecedarian project is the robust long-term follow up with low attrition. At age 21, n=104/111 children in the original sample completed follow up. Children who received the intervention at age 0-5 years had significantly higher full scale IQ scores (Mean (sd) Intervention 89.7 (10.1) v Control 85.2 (8.6), $p<0.05$) and had grade equivalent scores almost 2 years higher than controls for both reading and maths. A significantly higher proportion of the intervention group were enrolled in 4-year college or university courses (36% v

14%, p<0.01), and were in skilled employment (47% v 21%, p<0.01). A significantly lower proportion reported marijuana use over the past 30 days (18% v 39%, p<0.05), regular smoking (39% v 55%, p<0.020), and teenage parenthood (26% v 45%, p<0.05). While lower proportions also reported incarceration (14% v 21%) this difference did not reach statistical significance but may be due to the small sample size in the groups.[102]

The 30-year follow up, which included n=101/111 of the original sample, again showed persistent effects on educational attainment with a significantly higher proportion of the intervention group graduating college compared to the control group (23% v 6%, p<0.05). Two unique aspects of the Abecedarian project, hypothesised to be part of the reason for the long-lasting effects seen well into adulthood, are firstly that it commenced very early in infancy with the mean age of commencement being 4.4 months, and secondly that it was a very intensive programme run five days per week for five years. [103]

The majority of published follow-up studies of the original Abecedarian project focus on the outcomes for the child, but one qualitative study involving 18 parents from a social housing complex in Winnipeg, where the government funded an Abecedarian inspired project, reported many positive parental outcomes. These included an increased awareness of their role in supporting their child's development, a feeling of pride in their child's increasing abilities and in their own role in this, and increased opportunities for parents own personal growth such as furthering their own education.[104]

1.6.2. Experimental evidence in an Irish setting: The Preparing for Life Project

The Preparing for Life Project was a randomised controlled trial of an early developmental intervention conducted in an area of disadvantage in Dublin, Ireland. All pregnant women living in the community between 2008-2010 were eligible for inclusion, and participants were recruited either through two maternity hospitals or through self-referral. Participants assigned to the treatment group received a twice monthly home visit from a trained mentor for the first 5 years of the infants life, which focussed on child development and parenting practices. They also participated in a baby massage course during the first year of life, and a positive parenting programme during the second year. Both treatment (n=115) and control (n=118) groups received developmentally appropriate toys and books, support to participate in community events and public health workshops, and access to a support worker.

Follow up assessments were completed face to face in participant's homes at regular intervals by trained researchers blinded to the allocation group. At age 4 years, just prior to starting school, children who remained in the study in the treatment group (n=69) had significantly higher mean general cognitive ability scores compared to those in the control group (n=59) (Mean (SD) Treatment group 97.7 (14.3) v Control group 88.0 (12.6), $p<0.01$). Children in the treatment group were significantly less likely to have below average cognitive ability at age 4 compared to controls (Treatment 20% v Control 60%, $p<0.01$).[105,106] The magnitude of effect seen in this RCT was larger than those of other home-based programmes. The reasons for this are not clear, but may be attributable in part to bias introduced by relatively high attrition rates. Nonetheless, this study demonstrates the feasibility and efficacy of implementing early intervention services in the Irish setting.

1.6.3. Real world evidence: The Head Start and Early Head Start Programmes

The Head Start Programme began in America in 1965. The aim was to break the cycle of poverty by providing preschool children from low-income families with a programme to meet emotional, social, psychological, and nutritional needs.[97] Due to emerging research on the importance of the first three years of life and very early intervention, the programme was then expanded in 1995 to include the Early Head Start program which serves pregnant women and infants under 3 years old. Both programmes focus on child development and health, as well as family well-being and engagement.[107]

Eligible for both programmes are infants and children from families with income below the poverty guidelines (for example a household income < $30,000 for a family of four)[108], children from homeless families, and foster children.[107] The Early Head Start programme offers families an option of a home-based service consisting of weekly home visits, lasting 90 minutes, focused on helping parents to support their child's development. Alternatively there is a centre-based service, providing education and child development services in a classroom setting for 1,380 hours annually, with two annual home visits. Both programmes aim to promote the physical, cognitive, social and emotional development of infants by enriching their environment and caregiving.[107]

A multi-centre randomised controlled trial across 17 Early Head Start Programmes in the United States, compared 3 year outcomes for more than 1,500 children and families all eligible for Early Head Start, who were randomly assigned to the Early Head Start Programme or a control group.[109] The study found that children

in the programme had higher mean Bayley Mental Developmental Index scores compared to controls (91.4 v 89.9, $p<0.05$) and higher Peabody Picture Vocabulary Test scores (83.3 v 81.1, $p<0.05$). They displayed higher emotional engagement with parents, and lower aggression compared to controls. Parents of the programme group were significantly more likely to read to their children daily compared to parents in the control group (56.8% v 52.0%, $p<0.01$).[109]

While some of the effects were very modest in size, there were consistent favourable effects for both the child and the parents in the programme group, and these effects were larger when the programme was implemented fully in a mixed approach setting with both a home-based and centre-based component.[109] While there are limited long term follow-up studies of the Early Head Start Programme due to its more recent introduction, long term follow-up studies of the Head Start programme have reported mixed results with some studies reporting a ”fade out” of the positive effect on test scores over time, while others have shown persistent positive effects on adult outcomes including educational attainment and criminal involvement.[110] These findings may be explained in part by the advantages in other domains, such as socio-emotional function and parental relationships, conferred to children participating in the programmes, which may contribute to successful adult outcomes but not necessarily a persistent improvement in test scores.

1.6.4. Early intervention for preterm infants

A pre-emptive approach to early cognitive intervention, where interventions are provided on the basis of risk rather than signs of delay, has also been shown to benefit other high risk cohorts including preterm infants. A 2015 Cochrane review

examined the effect of early intervention in the first 12 months of life compared to standard follow-up for infants born at <37 weeks' gestation on cognitive outcomes at infancy (16 studies), preschool-age (8 studies), school-age (5 studies) and adulthood (1 study). Meta-analyses demonstrated a treatment effect at infancy of approximately one-third of a SD on standardised cognitive tests (equivalent to 5 IQ points), and at preschool age of approximately one-half of a SD (equivalent to 7.5 IQ points). No differences in cognitive outcomes were found at school-age or in adulthood, however there was significant heterogeneity noted between the studies at school-age limiting the conclusions which could be drawn, and there was only one study with adult follow-up.[111]

1.6.5. Eligibility for experimental and real world early intervention programmes

In Ireland at present, there is no population-based early intervention programme for infants at risk of low cognitive ability in childhood. The Irish National Health Service (Health Service Executive (HSE)) has area-based early intervention teams for children age 0-5 years, but generally referrals are accepted only for children who are already displaying overt signs of a developmental delay in at least two domains.[112] For children with disabilities and complex health needs, there are 91 area-based Children's Disability Network Teams.[113] However, access to these services is hampered by lengthy waiting lists and staff shortages, which ensure that the majority of children are beyond the window of peak neuroplasticity before they are seen by these teams.[113,114]

The examples of early intervention provided in the previous sections demonstrate that, both in the controlled and real-world setting, early pre-emptive

intervention commenced based on risk, prior to overt signs of cognitive difficulty, at the time of optimal neuroplasticity can alter neurodevelopmental outcomes.[100-106,109-111] Each intervention outlined used different criteria to identify "at risk" children – The Abecedarian project used a combination of sociodemographic variables; The Preparing for Life project used area based disadvantage as their sole criterion; The Early Head Start programme primarily uses the poverty threshold; and the preterm interventions used a gestational age of <37 weeks.[100,105,107,111] All of these children, are children within a general paediatric population, who are at higher risk of adverse cognitive outcomes in childhood than those without these risk factors.

Many children will be exposed to a single risk factor during early life, be it biological, environmental, or otherwise, and endure few or no adverse effects.[115] It is well documented that children exposed to multiple early risk factors represent a more vulnerable subgroup for adverse childhood outcomes. The more risk factors a child is exposed to the worse these outcomes tend to be.[115-117] In the literature to date, the most utilised statistical approach to combining these risk factors for the prediction of cognitive outcomes has been to combine them in an additive fashion, with the main effect of each factor accounted for.[115]

However, we know that risk factor effects are not necessarily the same for all children and effect modification and interaction does exist.[115] The effect of one risk factor may be accentuated or diminished by another exposure. For example, the co-occurrence of preterm birth, a biological risk factor, and socioeconomic disadvantage, a social risk factor, may have synergistic adverse effects on neurodevelopmental outcomes.[118] Equally, the association between early-life brain injury due to prematurity and poor cognitive outcomes in childhood can be mitigated

almost entirely in the presence of higher maternal education.[119] Therefore, to implement a population-based intervention programme for all children at risk of adverse cognitive outcomes in childhood it is imperative that consideration is given to the spectrum of risk and mitigating factors, and both their cumulative and interactive effects.

1.7. Current childhood screening for low cognitive ability in Ireland

Developmental checks are provided, free of charge, by the HSE in Ireland typically at ages 3 months, 9-11 months, 18-24 months, and 3-3.5 years.[120] The Ages and Stages Questionnaire Third Edition (ASQ-3) was introduced nationally between 2018-2019 in Ireland as a recommended developmental screening tool. The initial roll-out began with its introduction at the 24 month developmental check.[121] The ASQ-3 is a parent-reported questionnaire, that can be administered between 1 and 66 months, and is designed to discriminate between children with developmental delay and those with typical development. Studies examining the validity of the ASQ-3 report different results depending on the age of administration. A recent systematic review, based on 43 studies, found that an ASQ score more than 2SDs below the mean in 1 or more domain had a pooled sensitivity of 0.77 (95% CI 0.64-0.86) and a specificity of 0.81 (0.75-0.86) for identifying developmental delay in children aged 12-60 months. However, for cognitive delay specifically the pooled sensitivity was much lower at 0.44 (95% CI 0.24-0.65), with the specificity remaining high at 0.93 (95% CI 0.81-0.95).[122] In this systematic review, the ASQ was compared against concurrently performed formal developmental assessment.

Charkaluk et al[123], using the French EDEN birth cohort, examined the predictive value of a total ASQ score at 36-months for predicting an IQ <85 at age 5-6 years. The area under the receiver operating characteristic curve (AUROC) was 0.78 (95% CI 0.72-0.84), and at an optimal cut-off value of 270 the 36-month ASQ achieved a sensitivity of 0.77 (95% CI 0.66-0.85) and a specificity of 0.68 (0.65-0.71), with a low positive predictive value of 0.16 (95% CI 0.12-0.20) and a high negative predictive value of 0.97 (0.96-0.98). This study used a total ASQ score for prediction, which is not routinely calculated or used in administration of the ASQ, and risks missing domain-specific severe delays which the instrument was designed to detect.[123] The predictive value of the ASQ, as it is currently used in practice, for predicting later low cognitive ability has not been fully investigated, and its utility as a predictive tool for later low cognitive ability in an Irish cohort has not previously been examined.

1.8. Multivariable models for predicting cognitive outcomes in childhood

1.8.1. Published population-based multivariable models for predicting cognitive outcomes in childhood

Generally two classes of statistical model are recognised – those where the aim is to understand the causal pathways leading to a certain outcome, and those where the aim is to predict as accurately as possible a certain outcome.[124] Statistical models intending to investigate the causal pathways between early life determinants and cognitive function in childhood estimate the effect size between a preceding factor and the outcome, adjusted for the confounding factors in this relationship, based on a formal modelling process.[125] The inclusion of additional variables in the model is based on previous evidence of confounding, changes in the size of the effect between

the causative factor and the outcome, and goodness of fit.[125,126] Directed acyclic graphs (DAGS) can be used in causal modelling to transparently document assumptions about the causal relationships between the variables, and to identify the confounders requiring adjustment to estimate the causal effect of a particular variable on the outcome of interest.[127] There are a wealth of published studies which investigate associations between single variables and intelligence or cognitive ability in childhood.[128–131]

Prediction models, however, combine multiple predictors based on their ability to predict the occurrence of a particular event in the future (prognostic model) or their ability to predict the presence of a particular health state (diagnostic model).[132] Prediction model studies aim to develop and validate multivariable prediction models which combine individual predictors to predict an individual's probability of either having or developing a particular outcome.[133] Prediction models designed specifically to predict the occurrence of cognitive difficulties in childhood in a general paediatric population are limited. A summary of relevant published models is contained in Table 1.1.

Eriksen et al[134] combined pregnancy, birth, parental, and postnatal factors in a series of linear regression models to predict the IQ score at age 5 among more than 1,700 children who participated in the Lifestyle During Pregnancy Study in Denmark. IQ was treated as a continuous variable, and was measured using the Wechsler Primary and Preschool Scales of Intelligence Revised (WPPSI-R). Non-linear terms were investigated and a significant quadratic association between head

Table 1.1 Published multivariable models for predicting cognitive outcomes in childhood at a population level

Authors	Eriksen et al[134]	Camargo-Figuera et al [135]	Lawlor et al [136]	Lawlor et al [137]	Camacho et al[138]
Year Published	2013	2014	2005	2006	2019
Participants	1,782	3,523	11,679 with follow up (f/u) at age 7 11,386 with f/u at age 9 11,324 with f/u at age 11	3,794[a]	9,487
Outcome	Continuous IQ score	Low IQ defined as a z-score >1 SD below the mean	Continuous IQ score	Continuous IQ score	Delayed school readiness defined as a composite score >1 SD below the mean
Age at outcome	5	6	7, 9, 11	5	3
Outcome measure	Wechsler Primary and Preschool Scales of Intelligence - Revised (WPPSI-R)	Wechsler Intelligence Scale for Children – III (WISC III)	Age 7 – Moray House picture intelligence test Age 9 – Schonell and Adams essential intelligence test Age 11 Moray House verbal reasoning tests.	Peabody Picture Vocabulary Test	Bracken School Readiness Assessment
Predictors in final model	Child's age, child's sex, tester, parental education, maternal IQ, maternal age,	Child's sex, parental skin colour, number of siblings, father and mother employment status, household income, maternal education,	Social class, gravidity, illegitimacy, maternal age, maternal height, maternal physical condition, pregnancy induced hypertension,	Maternal age, pregnancy complications, gravidity, fetal distress during labour, duration of	Child's sex, child's ethnicity, mother's age at birth of first child, birth weight, maternal mental health, child

	paternal age, parity, maternal prenatal marital status, birth weight, breastfeeding, irregularly breakfast, head circumference, height at 5 years	number of persons per room, duration of breastfeeding, height for age deficit, head circumference for age deficit, parental smoking during pregnancy, maternal perception of child's health status	antepartum haemorrhage, artificial rupture of membranes, method of delivery, child's sex, gestational age, standardised birth weight z-scores, number of siblings, maternal education	labour, mode of delivery, birthweight, gestational age, Apgar scores at l and 5 mins, intrauterine growth, duration of breast feeding (at 6-month follow up), family income, parental education	development milestones, duration of breastfeeding, number of children in family, maternal education, parents' workforce status, housing tenure, social class, annual family income.
Study reported investigation of non-linear terms	Yes	No	No	No	No
Study reported investigation of Interaction terms	No	No	Yes	Yes	No
Statistical approach	Linear regression	Logistic regression	Linear regression	Linear regression	Stepwise logistic regression
Evaluation results of best performing model	29% of variance in total IQ score explained	AUROC 0.80 on internal validation. Sensitivity 72%, specificity of 74%. AUROC 0.75 on external validation using an independent test set.	16% of variance in total IQ score explained at age 7, 9, 11	12.7% of variance in total IQ score explained	AUROC 0.80 (95% CI 0.78 to 0.81). At optimal cut-off sensitivity 72% (95% CI 69.0% - 74.3%), specificity 74% (95% CI 73.5% - 75.3%)
Validation performed	No	Yes	No	No	Yes

[a] This study followed children up at age 5 and age 14. The number included in the final model was specified for the model at age 14 only.

circumference and IQ was reported. Investigation of interaction effects was not reported. The final model which included 14 variables predicted 29% of the variance in IQ at age 5. Parental education and maternal IQ accounted for almost 20% of the variance in the model. The model did not undergo any validation procedures. Predictors from both early life and later childhood were included in the model limiting its usefulness for early prediction.[134]

Camargo-Figuera et al[135] used a logistic regression model to predict low IQ, defined as a z-score more than 1 SD below the mean on the Wechsler Intelligence Scale for Children-III, among a cohort of 3,500 children in Brazil. Their final model, which achieved an AUROC of 0.80 on internal validation and 0.75 on external validation, used 12 variables - child gender, parental skin colour, number of siblings, father and mother employment status, household income, maternal education, number of persons per room, duration of breastfeeding, height for age deficit, head circumference for age deficit, parental smoking during pregnancy, and maternal perception of child's health status. The paper did not report investigation of non-linear or interaction terms. The model was internally validated using bootstrapping, and externally validated on a similar birth cohort in Brazil. The predictors were measured at 3 and 12 months.[135]

Lawlor et al[136] used a multivariable linear regression model to predict intelligence at age 7,9, and 11 among a cohort of more than 10,000 children born in Aberdeen, Scotland between 1950 and 1956. The model, using a combination of intrauterine, pregnancy, parental, and childhood variables explained 16% of the variance in IQ at all three ages. All variables were measured at birth except child height and weight which were measured at school entry. Interaction terms were

investigated and significant interactions were reported between childhood height and gender for predicting IQ at age 7 years, and childhood weight and gender for predicting IQ at all ages.[136]

Lawlor et al[137] combined 14 variables in a linear regression model to predict intelligence, measured using the Peabody Picture Vocabulary Test, at age 5 among 3,999 children born to women in South Brisbane between 1981 and 1984. Using data on maternal age, pregnancy complications, gravidity, fetal distress during labour, duration of labour, mode of delivery, birthweight, gestational age, Apgar scores at 1 and 5 minutes, intrauterine growth, duration of breast feeding (at 6-month follow up), family income, and parental education, their model predicted 12.7% of the variance in IQ at age 5. The study reported investigating the interaction between birthweight and both height and BMI. The paper did not report investigation of non-linear terms and the model did not undergo external validation procedures.[137]

Camacho et al[138], using data from 9,487 children in the United Kingdom Millennium Cohort Study, developed a prediction model to predict school-readiness at age 3 years. A child's school readiness is not just based upon their cognitive ability, but also on their social and emotional readiness. However the scale used to measure school readiness in this study, the Bracken School Readiness Assessment, is largely focussed on the cognitive aspects of readiness. A child with a score more than 1 SD below the mean was categorised as not being school ready. Based on a combination of 14 predictors, measured at 9-months, the model achieved an AUROC of 0.80 when internally validated using bootstrapping. At the optimal cut-off it achieved a sensitivity of 72% and a specificity of 74%. The study did not report investigation of non-linear or interaction terms in their model.[138]

1.8.2. Limitations of published population-based multivariable models for predicting cognitive outcomes in childhood

Adverse cognitive outcomes in childhood are complex, and may result from highly interactive relationships between biological, psychological, and social factors, the mechanisms of which are often poorly understood.[115] This real-world complexity may be difficult to model, particularly with more traditional statistical methods.

A limitation in some of the statistical approaches used in the models outlined in Section 1.8.1, is their reliance on strong assumptions, many of which are unrealistic in complex real word data. Traditional regression methods are often used in ways which assume that the effect of a predictor on the outcome increases or decreases uniformly throughout the range of the prediction (assumption of linearity), however in truth many relationships are non-linear.[139] It is often assumed that the residuals (the difference between the observed value of the dependent variable and the model predicted value of the dependent variable) have a constant variance at each level of the predictor (assumption of homoscedasticity), and that the residuals have a normal distribution.[140] Methods exist to deal with all of these issues, but may not be properly or commonly used. In the five models outlined in the previous section, only one reported investigating non-linear terms.[134]

Further limitations are the constraints around interaction terms in traditional statistical models. Traditional statistical models assume predictors affect the outcome in an additive manner unless otherwise specified. Interaction terms must be determined a-priori and entered into the model by the user, a process not commonly performed despite evidence demonstrating interactions between

predictors in the models, such as that between the complications of prematurity and maternal education.[119] Only two of the published prediction models discussed in the Section 1.8.1 reported investigating interaction terms in the modelling process.

A further limitation noted in two of these models is the inclusion of risk factors which were measured in later infancy or childhood. Interventions for improving cognitive outcomes appear to be more effective when initiated in very early life, but including risk factors measured in later infancy or childhood precludes the possibility for very early intervention.[134,136] Finally, only two of the models performed validation procedures.[135,138] Model validation is a crucial step in predictive modelling. A predictive model, based on a finite sample, that maximises its predictive validity within that sample, will almost always have poorer prediction when used on the overall population from which the sample was drawn, or on a different population.[141] To truly determine a model's predictive ability, it must be tested on data that was not used in the development of the model, and which is representative of the population in which the model is intended for use.[142]

1.9. Machine learning for predicting cognitive outcomes in childhood

1.9.1. An introduction to machine learning

There is considerable debate regarding the true differences between traditional statistical approaches and machine learning (ML), often compounded by the fact that many ML techniques, such as linear and logistic regression, are drawn from traditional statistical methods.[143] While traditional statistical models and ML models may use similar means to reach their goal, the key difference tends to lie between the objectives of the models. Traditional statistical methods, while they can be used

for prediction, often focus on inferring relationships between variables, while following a defined set of assumptions and rules based on the data to which the model is fitted.[143,144] The purpose of ML is to find generalisable predictive patterns within the data, without prior assumptions or rules governing the process.[143,145] ML broadly refers to a collection of completely data-driven algorithms that attempt to extract patterns and learn from the data they are shown.[146]

There are two main approaches to ML – supervised and unsupervised ML. In unsupervised learning, there is no outcome or label to predict. The focus is purely on the discovery of patterns within the data. Eosinophilic subtype asthma was discovered through the application of an unsupervised ML algorithm to genomic data, which recognised a pattern within the data that led to its discovery.[147] In supervised learning, an outcome or label exists to supervise or guide the learning process. In a population-based dataset for example, this label could refer to whether a person is diabetic or not, or alternatively the label could refer to a blood glucose value. For both prediction problems, the algorithm will learn to associate the values of a set of predictors, referred to as features in ML, with the label. After the model is trained, it will be able to use the values of these features in a new population to make a prediction about the label.[146]

Supervised learning may be used for two families of problems, a classification problem or a regression problem. Classification problems involve a categorical label, for example, whether a patient is diabetic or not. Regression problems use a numeric label, for example, a blood glucose value. A numeric label can also be converted into a categorical label. For example the accepted blood glucose cut-off could be applied

to convert the blood glucose value into a categorical label indicating whether the patient is diabetic or not, and this could then be treated as a classification problem.[146]

The approach to a supervised learning problem generally involves a number of common steps. The dataset is often divided into a training dataset and a testing dataset. A frequently used data splitting ratio for training:testing data is a 70:30 split.[148] This approach, however, if performed as a single split can produce a lucky or unlucky estimate of model performance. Employing multiple data splits to estimate average performance can help overcome this.[149]

During model training, the labels are shown to the algorithm and it learns to associate patterns in the feature values with a specific label value. After training, the model can attempt to apply the set of rules it has learned about the feature values in the training dataset, to the unseen testing dataset, in order to make a prediction about the label values. The ability of the model to predict accurately on the new dataset can be measured using a variety of metrics, such as AUROC, sensitivity, specificity, and accuracy.[145-146]

During the training process prediction error is an important parameter used to help select the optimal model. Prediction error has two components – error due to bias and error due to variance, and the balance between these two errors is referred to as the bias-variance trade-off.[150] A high bias model simplifies the assumptions made by the model and may miss relevant patterns in the data (underfitting). A high variance model has learned very specific patterns from the training data and is highly sensitive to small fluctuations in the data, limiting the model's ability to make accurate predictions on unseen data (overfitting).[150] A careful

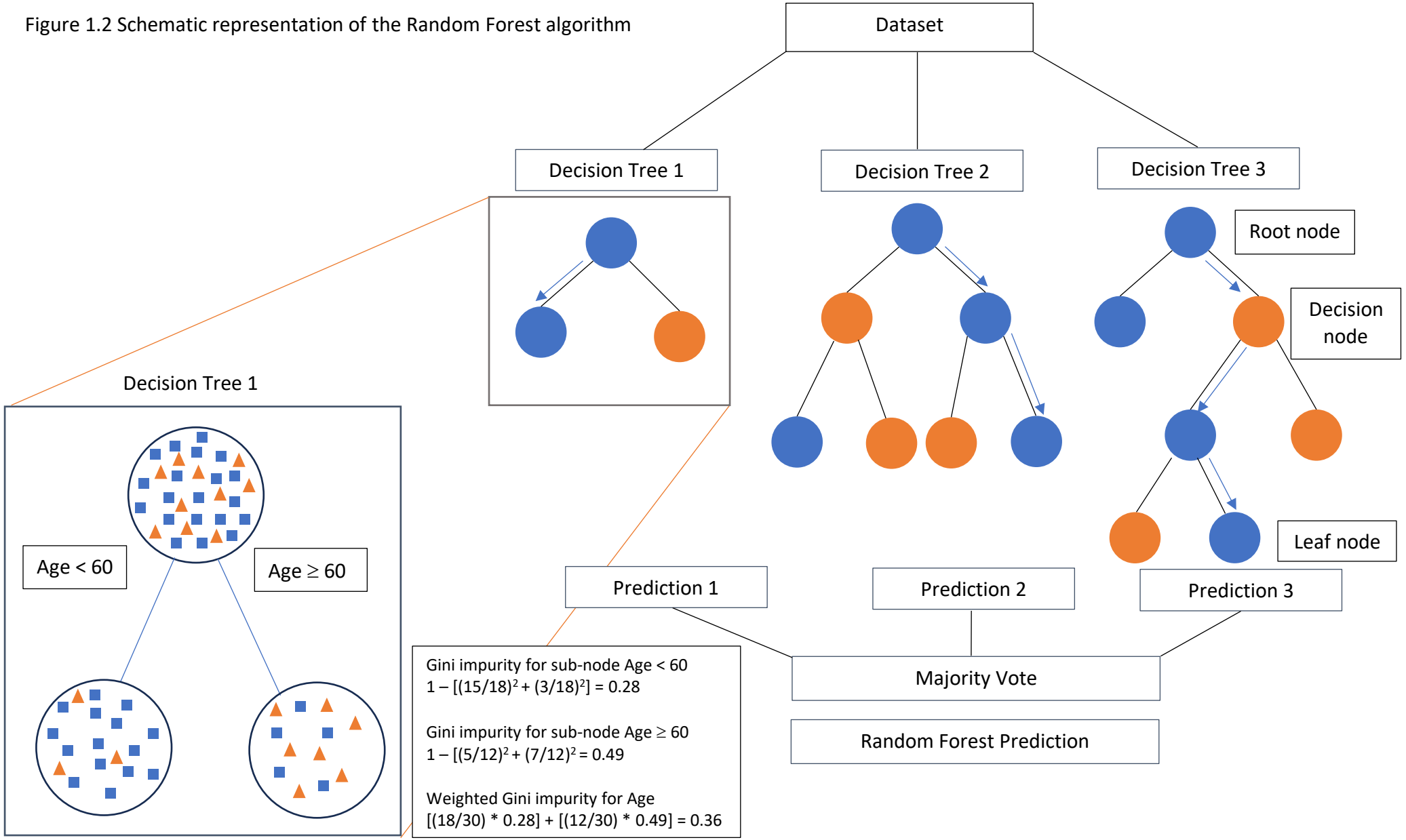
balance must be struck between the two error types in order to achieve a model which can produce predictable performance on unseen data.

Cross-validation is a technique that can be used in the training process to improve the generalisability of the model, by reducing the risk of overfitting and underfitting.[151] In *k*-fold cross-validation, the dataset is randomly divided into *k* subsets of equal size. At each iteration of cross-validation, *k*-1 subsets are used to train the model and one subset is held out for testing. Model performance is estimated over *k* iterations.[151] Repeated *k*-fold cross validation involves repeating this process a defined number of times. For example, using 5-fold cross-validation repeated 10 times would result in 50 different test subsets being used to estimate model performance. The estimate of model performance would be reported using a measure of central tendency, such as the mean or median, of the performance values across the test subsets. The algorithm with the best average performance would generally be selected as the final model. Examining the spread of results, using standard deviation, interquartile range, or maximum and minimum values can provide a good indication of whether the model will generalise well to unseen data. Models with a tighter spread of results across the repeated k-folds tend are likely to generalise better to unseen data.[151]

1.9.2. The Random Forest algorithm

The Random forest (RF) algorithm is a commonly used approach in supervised machine learning and is based on decision tree analysis. A decision tree, as shown in Figure 1.2, consists of a root node, branches, decision nodes, and leaf nodes. The root of the tree contains all samples in the dataset. The aim is then to split the samples

Figure 1.2 Schematic representation of the Random Forest algorithm



into branches so that after each split, the subset of samples remaining is more homogenous than its precursor.[151-152] Each decision node represents a feature within the dataset that is being used to split the sample down two or more branches, until a decision is made about the outcome at the leaf node.[146]

Measures of impurity, such as Gini impurity, are typically used to decide which feature to split the data on. For each node, Gini impurity measures the probability of misclassifying a randomly drawn sample if it was randomly labelled according to the class distribution of that node.[153] A node with zero impurity, i.e. all observations in node belong to the same class, has a Gini impurity score of zero and is considered a leaf node. A node with equal class distributions obtains the maximum Gini impurity score of 0.5.[154] Gini impurity is calculated as the sum of squares of count proportions for all target categories per node subtracted from one, with the result of this calculation multiplied by the proportion of observations in that given node.[153]

RF is referred to as an "ensemble learning technique" meaning multiple classifiers, or decision trees, are aggregated to solve the classification or regression problem. The random forest can consist of hundreds or even thousands of decision trees, that will vary in size and shape, and are built from random variants of the same dataset. The final decision or classification of the random forest is based on the majority vote of all trees in the random forest.[151,152] Majority voting or averaging the results across multiple decision trees can reduce the variance (overfitting) which may result when a single decision tree is used.[146]

There are two key concepts crucial for understanding how the random forest algorithm operates. The first key concept is bootstrap aggregation. Bootstrapping is a resampling technique that involves random sampling with replacement. Each

decision tree in the forest will be fitted on a different bootstrap sample. The random forest will combine the predictions from multiple decision trees all fitted on different bootstrap samples and generate a decision based on the majority vote. [156,157]

The second key concept is a parameter called *mtry*. At each decision node in each tree of the random forest, the algorithm will choose a feature on which to split the data. Instead of evaluating all features at each node, the *mtry* parameter is used to randomly select a subset of features for consideration at each node. *Mtry* refers to the number of candidate features that will be randomly selected for evaluation at each decision node. The feature which provides the greatest improvement in ordering the data is chosen as the decision attribute for that particular node. For example, *mtry*=3 will result in three randomly selected features being evaluated at each decision node in the tree.[151] By using a random selection of candidate features, the model avoids creating heavily correlated trees which use the same strong predictive features. This provides an opportunity for features with smaller effects to contribute. A low mtry value can ensure heterogenous, less correlated trees, and may increase the selection of features with small or moderate effects that can contribute to predictive performance, but which otherwise may be masked by features with stronger effects.[151] However, there is a risk that a low mtry value will result in poor predictive ability if only non-informative features are used. A large mtry reduces the risk of having only non-informative features and may ensure features with stronger predictive ability are selected.[151] This can improve predictive performance but may result in a model with poor generalisability.

1.9.3. The Gradient Boosting algorithm

Gradient boosting Machine (GBM) is a machine learning algorithm that builds an ensemble of shallow decision trees sequentially. Boosting is based on the principle that each decision tree will improve upon the mistakes of its predecessor.[158] Unlike the random forest algorithm where multiple trees are built independently and simultaneously, the trees in gradient boosting are built consecutively, in an additive manner.[158]

A key concept in GBM is the loss function. The loss function evaluates how good a model is at making predictions by tracking the amount of error ("loss") in the models predictions. It quantifies the difference between the model predicted value and the actual value. A model with accurate predictions will have a small loss, and conversely a model with inaccurate predictions will have a large loss.[159] The exact loss function used will vary depending on the type and distribution of the outcome. A classification problem with a binary outcome will use a different loss function than a regression problem with a continuous outcome.[160,161]

The aim in GBM is to minimise loss in a process called loss function optimisation. Loss function optimisation is performed using gradient descent, hence the name "Gradient Boosting". The gradient of the loss function refers to a vector of partial derivatives for each parameter in the model. Partial derivatives describe the rate of change of a function (or slope of a function) with respect to one parameter, while the other parameters are held constant.[162] The partial derivative of each parameter indicates how changing the value of that parameter impacts the value of the function. In other words, the slope of the function tells us the direction to move the value of the parameter in, in order to lower the cost function (error) on the next

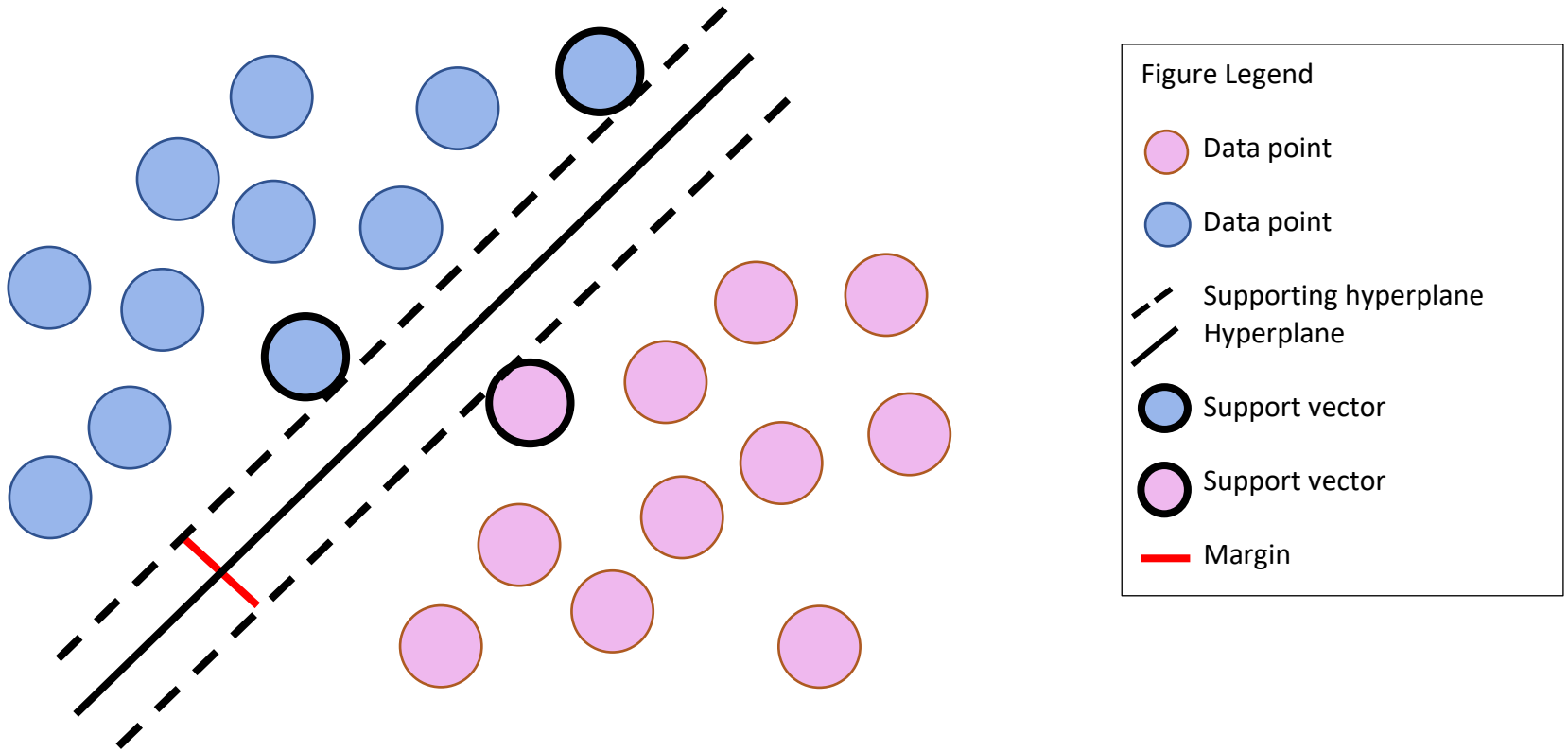
iteration The size of the change made at each iteration is controlled by the learning rate.[160-162] This hyperparameter, referred to as shrinkage, takes a value between 0 and 1. Smaller values result in a slower descent down the gradient. This can produce a more accurate model but is computationally more demanding as it requires more trees.[158]

The total number of trees in the ensemble can also be tuned, as can the depth of the trees. Shallower trees are efficient, whereas deeper trees allow for more complex trees to be fitted, at the risk of overfitting. Another hyperparameter controlling the complexity of the tree is the minimum number of observations in terminal nodes, which typically ranges from 5-15. Higher values can prevent the model from learning complex patterns which are very specific to a particular sample of data, but smaller values can be useful where there are imbalanced classes in classification problems.[158] At each iteration, a new weak learner is trained, predictions are computed, the loss is evaluated, the gradients are recalculated, and the parameters are adjusted according to the gradient and the learning rate. By repeating this process and iteratively adjusting the model parameters, gradient boosting gradually adjusts the model toward the optimal parameters that will minimise the loss function.[160-162]

1.9.4. The Support Vector Machine algorithm

A support vector machine (SVM) is a supervised machine learning algorithm frequently used for classification purposes.[163] The simple aim of a linear SVM, as shown in Figure 1.3, is to find the hyperplane, that best divides the data into two classes. The hyperplane represents a decision boundary, whereby data points on one

Figure 1.1 A hypothetical example of a linear Support Vector Machine classifying the data points into blue and pink circles



side of the decision boundary are classified as belonging to one class and data points on the other side are classified as belonging to the other class.[163] In a two-dimensional feature space, the hyperplane is a line, whereas in a three-dimensional feature space or more the hyperplane is a plane. Multiple hyperplanes can be found to separate classes, but the goal of the SVM is to find the hyperplane that maximises the distance between the closest data points of opposite classes.[164] The data points that determine the maximum distance between the classes are referred to as support vectors. The planes which run through these points are called the supporting hyperplanes, and the margin is the shortest perpendicular distance between the supporting hyperplanes. Maximising the margin will result in better generalisation performance on new data.[158] In reality, perfect separation of datapoints is not usually achievable and a degree of misclassification must be allowed. In this case, the SVM aims to find the hyperplane that provides the largest margin between the classes, but allows some datapoints to cross the margin boundaries. The amount of misclassification allowed is defined by the hyperparameter *C*, which is used to control the bias-variance trade-off.[158]

Most population-based datasets are complex and heterogenous and could not be easily separated by a simple linear hyperplane. When the data are not linearly separable, kernel functions can be used to map data to a higher-dimensional feature space where a hyperplane may be able to separate the data. This hyperplane, when projected back to the original feature space, will result in a non-linear decision boundary. There are multiple different kernel functions and the choice will depend on the data and type of problem. The gaussian radial based function kernel (RBF), a commonly used kernel function, measures the similarity between pairs of data points

based on the distance between them in the feature space. The RBF kernel will have a value closer to 1 when there is a high similarity between data points (i.e. they are close together in the feature space), and a value close to 0 when there is a low similarity (i.e. they are far apart in the feature space).[165] The distance determining which points are considered similar can be varied using the width parameter of the Gaussian kernel. The inverse-width parameter of the Gaussian kernel, referred to as γ, will control the flexibility of the resulting SVM in fitting the data. Large values of γ will result in a complex, overfitted model that will not generalise well to new data. Small values of γ can result in a model that does not capture the complexity of the data and will not perform well.[163,165]

1.9.5. Potential advantages of machine learning for predicting cognitive outcomes in childhood

From the preconceptual period, a myriad of factors including biological, psychological and social factors will influence and shape a child's cognition.[166] The 'exposome' is a term used for the totality of life course environmental exposures from the prenatal period onwards.[167] Adopting this lens enables a more complete understanding of the vast exposures involved in childhood cognitive development.[168,169] The complex interplay between this multitude of factors may be difficult to model with traditional statistical methods, and may be one of the reasons for the relatively poor predictive performance of models developed to date. The advantages conferred by machine learning may overcome some of the limitations of more traditional techniques and help to improve the accuracy of prediction.

A key advantage of machine learning lies in its ability to handle large numbers of predictive features. In recent years, there has been an increase in both the number and the size of birth cohort studies, and in Europe alone more than 110 birth cohorts representing 27 different countries have been formed.[170] Prospective birth cohorts and linked population-based registries have potential to follow an individual from the prenatal period, through intrauterine, and early life.[171] The capacity to collect large volumes of data, to biobank serum samples, to perform genomics, and to use technology such as wearable devices has resulted in an explosion in the quantity, complexity, and depth of data now collected in such cohorts. This provides us with an opportunity to use data to advance both our understanding of the aetiology of disease, and our ability to predict outcomes.

ML techniques can be used to simultaneously examine large numbers of multimodal predictive features and to model complex, non-linear, and interactive relationships between these features.[143] This can be done without the need for the user to pre-specify non-linear and interaction terms in the model, which algorithms such as logistic regression require the user to do.[151] Based on previously published literature these non-linear and interactive relationships, which generally have not been included in previously published models, may be important in the prediction of cognitive outcomes in childhood, and their inclusion may help to improve the predictive accuracy of these models.[115,139]

Model validation is a crucial step in the ML process, and best practice guidelines for ML modelling recommend validation is performed for all models.[172] This can be done using a split-sample approach, which involves training the model on a portion of the dataset (e.g. 70% of the dataset) and testing the model on the remaining

unseen data (e.g. 30% of the dataset). Alternatively, an entire-sample approach such as *k*-fold cross-validation can be used.[173]

1.9.6. Example applications of machine learning in neurodevelopmental paediatrics

The past decade has seen a rapid expansion in the application of ML and Artificial Intelligence (AI) for the purposes of diagnosis and prognosis in clinical medicine.[174] Significant progress has been made specifically in the areas of radiology, pathology, and cancer.[175] There is growing interest in the application of ML in the field of paediatric neurodevelopmental disorders (NDDs), a heterogenous group of disorders impacting different domains of function including cognition, language, social, motor and behavioural function. However, progress has been slower.[176]

Diverse data, ranging from MRI data to motor skills data collected via drag-and-drop gaming devices, have been combined with machine learning techniques for both diagnostic and prognostic purposes in the area of NDDs. In Table 1.2 a selection of published models, which highlight the diverse range of features being used to predict neurodevelopmental outcomes are shown. Accurate, objective, prognostic or diagnostic tools could be a particularly valuable aid to clinicians in this field, where many of the conditions currently rely on subjective, labour intensive, expert clinical assessments. In many countries, this results in lengthy waiting lists for assessment in developmental paediatric clinics and delays in diagnosis for those affected. One example is ADHD, the diagnosis of which relies upon a detailed clinical assessment based upon observed behaviour and reported symptoms.[177]

Table 1.2 Examples of machine learning applied in neurodevelopmental paediatrics

Authors	Das and Khanna[178]	Sen et al[179]	Emerson et al[180]	Girault et al[181]	Kim et al[182]
Year Published	2019	2018	2017	2019	2021
Purpose	Diagnostic	Diagnostic	Prognostic	Prognostic	Diagnostic
Outcome	ADHD	ADHD	ASD	Cognitive ability	Developmental disability
Outcome Measure	Diagnosed by child neurologist using DSM-IV criteria	Diagnostic criteria differed across study sites	DSM-IV Criteria	Mullens Scales of Early Ability. Outcome categorised as above or below the median	Diagnosis made by paediatric psychologist
Features	Pupillometric data	Structural and functional MRI	MRI at 6-months	MRI at birth	Drag and drop data from computer based game
Participants	50 (28 cases, 22 controls)	558 (279 cases, 279 controls)	59 infants with high familial risk of ASD (11 cases, 48 controls)	115 infants	370 children (147 cases, 233 controls)
Algorithms trained	LR, SVM, KNN, NB, DT, RF	SVM	SVM	DL	DL
Evaluation method	Nested 10-fold cross validation	Independent test set (77 cases, 94 controls)	Cross-validation with nested leave one out procedure	Independent test set (22 cases, 15 controls)	10-fold cross validation
Evaluation Result (best performing models)	NB: AUROC 0.87, sensitivity 0.66, specificity 0.86	Accuracy 0.67, sensitivity 0.45, specificity 0.85	Accuracy 0.97, sensitivity 0.82, specificity 1.00	Accuracy 0.84, sensitivity 0.86, specificity 0.80	AUROC 0.82, sensitivity 0.76, specificity 0.74

Abbreviations: LR, Logistic Regression; SVM, Support Vector Machine; KNN, K-Nearest Neighbours; NB, Naïve Bayes; DT, Decision Tree; RF. Random Forest; ADHD, Attention Deficit Hyperactivity Disorder; ASD, Autism Spectrum Disorder; DSM-IV, Diagnostic and Statistics Manual of Mental Disorders.

Pupillometric data (data based on dilation of the pupil) from 50 participants age 10 - 12 years used in an SVM algorithm achieved a sensitivity of 0.77 and a specificity of 0.75 in the detection of ADHD, based on the results of nested 10-fold cross validation.[178] This study, however, is limited by the small sample size of 50, and the absence of model validation on an independent test set raises questions over the generalisability of their model to wider populations.

MRI data has also been used in ML algorithms for the purpose of diagnosing ADHD. Sen et al[179] trained a SVM using structural and functional MRI data from 558 participants. The final model, which produced an accuracy of 0.67, sensitivity of 0.45, and specificity of 0.85 on an independent test set of 171 participants, is unlikely to be clinically useful in its current form, but suggests there may be signals within such data that could assist in the diagnosis of ADHD.[179]

For ASD, another neurodevelopmental disorder which benefits from early intervention but for which diagnosis is often delayed, ML has been increasingly studied as a potential tool for early detection.[180,183,184] A 2019 systematic review and meta-analysis of ML algorithms using structural brain MRI data to detect ASD reported a pooled AUROC of 0.90, a sensitivity of 0.83, and a specificity of 0.84.[184]

ML methods have also been deployed for prediction of later neurodevelopmental disorders. A deep learning model, trained on white matter connectomes from brain MRI data obtained at time of birth, and tested on an independent test set, correctly classified 84% of infants as having below or above the median level of cognitive ability at 2 years of age.[181] The results, which are based on a relatively small sample size of 37 in the independent test set, are promising but the clinical utility at a population-level may be limited.

Novel methods of neurodevelopmental assessment, outside of traditional medical investigations, are also being analysed using ML algorithms. For example, a deep learning model trained on data from a touch-based computer game, aimed to harness this measure of motor skill to detect developmental disabilities in children. An AUROC of 0.82 based on internal cross-validation provides a basis for further research in this area.[182]

Compared to the general paediatric population, there is comparatively more literature examining the use of ML for predicting neurodevelopmental outcomes in preterm infants.[185] The majority of ML models in this area to date have been based on MRI data. A systematic review published in 2023 reported 10 out of 11 eligible studies, which sought to predict neurodevelopmental outcomes in preterm infants, used imaging data as part of their input features.[185] The results of many models were promising, with one study reporting an AUROC of 1.0 for predicting cognitive impairment at 18-22 months based on MRI data obtained at a mean age of 37 weeks postmenstrual age. However the small number of infants (n=60) included in the study increases the risk of overfitting and producing a high variance model which may not generalise well to new populations. [186]

1.9.7. Machine learning for population-based prediction

Acknowledging the limitations of the studies to date in the field of NDDs, the results overall indicate that AI and ML may have an important role in the early prediction of certain NDDs. However, many of the prediction studies to date are based on data, such as MRI and EEG, which are not routinely collected at birth at a population level. Therefore, in clinical practice these studies may provide methods of prediction for

higher risk populations who undergo these investigations, but are likely to be largely unsuitable for implementation at a population-based level due to resource constraints and cost. At a population level, ML would be most advantageous if it could be used with data that is readily available, cost effective to collect, and available early in life. This may come at the cost of reduced predictive ability compared with models based on direct physiological and anatomical measures.

In the context of population health, the application of ML to address population health issues and mitigate health disparities has received much less attention than its clinical applications. Population-based data sources have the potential to incorporate social, environmental, and other structural determinants in prediction models, potentially improving accuracy.[187] Social determinants of health have long been studied in epidemiological research, but findings of these studies emphasise the need for methods that are better able to capture the complex, interactive, and flexible relationships between these predictors and health outcomes.[187]

The majority of well-established predictors of cognitive function in childhood, including income, gestational age, toxin exposure during pregnancy, birthweight, obstetric complications, parental education, and socioeconomic status, are predictors that are readily known and available in the perinatal period. To date, models aimed at predicting which children are at risk of low cognitive ability in childhood have shown mixed results with regards to their predictive ability, with none achieving an accuracy that would demand their imminent implementation into practice. The question of how we can best harness early life data for identifying high risk children remains. Application of machine learning methods to large population-

based datasets, using readily available, routinely collected information may assist in this objective. To date, this remains relatively unexplored in the literature.

1.10. Summary

In summary, approximately 16% of children will have cognitive ability below that of most of their peers at school-age, and will score more than 1 SD below the average score on a standardised assessment.[14,24] The level of cognitive ability achieved in childhood will impact academic, health, and socioeconomic outcomes in adulthood.[89-96]

Starting school is one of the most important and formative experiences of early childhood.[99] For children with below average cognitive ability, including those who do not meet criteria for an IDD, this transition may be particularly challenging. The impact of this transition on the mental health of children with below average cognitive ability has not been thoroughly investigated. Similarly, the early school experience for children with below average cognitive ability has not been fully described in the literature. If a child's cognitive ability, at the time of starting school, adversely impacts these outcomes this provides further impetus for early intervention prior to this important transition.

Cognitive ability is not fixed. Early intensive interventions, which enrich the environment of the infant and focus on the family unit, can improve cognitive outcomes in childhood for high risk children, particularly when started in the period of optimal neuroplasticity in the first 1,000 days of life.[100-107,109,110] High risk children may be those with biological risk factors such as prematurity, or those exposed to socioenvironmental risk factors such as poverty.[60-79] Population-based

developmental screening tools performed in infancy, such as the parent-reported Ages and Stages Questionnaire, have been reported to have a low sensitivity but high specificity for identifying cognitive delay.[122] The ASQ is now a recommended part of population-based developmental screening in Ireland.[121] The ability of the ASQ, performed in infancy, to predict later below average cognitive ability has not been extensively studied, and has not been investigated in an Irish cohort to date. This should be established prior to investigation of alternative methods of early identification.

A limitation of the ASQ and other developmental screening tools is that they tend to rely on the presence of a delay in a developmental domain to trigger action.[121,122] Evidence to date suggests that intervention may be more effective when implemented prior to the manifestation of delay. Multivariable models, with the specific aim of predicting cognitive ability in childhood, are limited.[134-138] Published population-based models have used linear and logistic regression methods and in general have not reported investigation of interaction terms or non-linear terms, which literature suggests may be important in predicting cognitive outcomes. ML methods, which can combine predictors in interactive and non-linear relationships without the need for pre-specification, may improve the prediction of below average cognitive ability at a population-level. This has not been investigated in the literature to date.

While this thesis will focus on cognitive ability, its relationship with childhood outcomes, and its prediction, it is important to acknowledge that cognitive ability alone does not determine one's destiny. It is only one of a number of skills that should be fostered and nourished in childhood. Undue emphasis should not be placed on

cognitive ability alone in early childhood, and a broad range of skills, talents and abilities should be recognised and celebrated. However, every child should have the opportunity to reach their individual cognitive potential and to benefit from their right to education. Achieving this is at the core of this research.

1.11. References

1. Hofer SM, Clouston S. Commentary: On the Importance of Early Life Cognitive Abilities in Shaping Later Life Outcomes. Res Hum Dev. 2014 Jan 1;11(3):241-246. doi: 10.1080/15427609.2014.936173. PMID: 25309140; PMCID: PMC4188394.
2. Calvin CM, Deary IJ, Fenton C, Roberts BA, Der G, Leckenby N, Batty GD. Intelligence in youth and all-cause-mortality: systematic review with meta-analysis. Int J Epidemiol. 2011 Jun;40(3):626-44. doi: 10.1093/ije/dyq190. Epub 2010 Oct 29. PMID: 21037248; PMCID: PMC3147066.
3. Batty GD, Mortensen EL, Osler M. Childhood IQ in relation to later psychiatric disorder: evidence from a Danish birth cohort study. Br J Psychiatry. 2005 Aug;187:180-1. doi: 10.1192/bjp.187.2.180. PMID: 16055831.
4. Hegelund ER, Flensborg-Madsen T, Dammeyer J, Mortensen, EL. Low IQ as a predictor of unsuccessful educational and occupational achievement: A register-based study of 1,098,742 men in Denmark 1968–2016. Intelligence. 2018 Nov; 71: 46–53. doi:10.1016/j.intell.2018.10.002
5. Gordon N. Some influences on cognition in early life: a short review of recent opinions. Eur J Paediatr Neurol. 1998;2(1):1-5. doi: 10.1016/1090-3798(98)01000-1. PMID: 10726840.
6. Flavia I, Pietrobelli A, Dargenio VN, Marchese F, Grillo A, Vural M, Giardino I, Mantovani M. The key 1000 life-changing days. Global Pediatrics. 2023; 4. doi: 10.1016/j.gpeds.2023.100049

7. Darling JC, Bamidis PD, Burberry J, Rudolf MCJ. The First Thousand Days: early, integrated and evidence-based approaches to improving child health: coming to a population near you? Arch Dis Child. 2020 Sep;105(9):837-841. doi: 10.1136/archdischild-2019-316929. Epub 2020 Feb 28. PMID: 32111596.
8. Han TH, Chae KY, Han B, Kim JH, Ha EK, Rhie S, Han MY. Early onset and increasing disparities in neurodevelopmental delays from birth to age 6 in children from low socioeconomic backgrounds. J Neurodev Disord. 2024 Nov 5;16(1):60. doi: 10.1186/s11689-024-09577-2. PMID: 39501156; PMCID: PMC11536651.
9. Cioni G, Inguaggiato E, Sgandurra G. Early intervention in neurodevelopmental disorders: underlying neural mechanisms. Dev Med Child Neurol. 2016 Mar;58 Suppl 4:61-6. doi: 10.1111/dmcn.13050. PMID: 27027609.
10. Doyle O, Harmon CP, Heckman JJ, Tremblay RE. Investing in early human development: timing and economic efficiency. Econ Hum Biol. 2009 Mar;7(1):1-6. doi: 10.1016/j.ehb.2009.01.002. Epub 2009 Jan 21. PMID: 19213617; PMCID: PMC2929559.
11. American Psychological Association. American Psychological Association APA Dictionary of Psychology [Internet]. 2020 [cited 2023 Sep 13]. Available from: https://dictionary.apa.org/
12. Colom R, Karama S, Jung RE, Haier RJ. Human intelligence and brain networks. Dialogues Clin Neurosci. 2010;12(4):489-501. doi: 10.31887/DCNS.2010.12.4/rcolom.

13. Plomin R, Deary IJ. Genetics and intelligence differences: five special findings. Mol Psychiatry. 2015 Feb;20(1):98-108. doi: 10.1038/mp.2014.105. Epub 2014 Sep 16. PMID: 25224258; PMCID: PMC4270739.

14. Rindermann H. Cognitive capitalism: Human capital and the wellbeing of nations. New York, NY, US: Cambridge University Press; 2018.

15. Raymond C. Intelligence: Its Structure, Growth and Action. New York, NY, US: North Holland; 1987.

16. Brown RE. Hebb and Cattell: The Genesis of the Theory of Fluid and Crystallized Intelligence. Front Hum Neurosci. 2016 Dec 15;10:606. doi: 10.3389/fnhum.2016.00606. PMID: 28018191; PMCID: PMC5156710.

17. Lövdén M, Fratiglioni L, Glymour MM, Lindenberger U, Tucker-Drob EM. Education and Cognitive Functioning Across the Life Span. Psychol Sci Public Interest. 2020 Aug;21(1):6-41. doi: 10.1177/1529100620920576. PMID: 32772803; PMCID: PMC7425377.

18. Hebb DO. The effect of early and late brain injury upon test scores, and the nature of normal adult intelligence. Proc Am Philos Soc. 1942;275–92.

19. Furnham A. Intelligence and Cognitive Abilities at Work [Internet]. In: Cartwright S, Cooper CL, editors. The Oxford Handbook of Personnel Psychology. Oxford University Press; 2008. Available from: https://doi.org/10.1093/oxfordhb/9780199234738.003.0002

20. Guez A, Peyre H, Williams C, Labouret G, Ramus F. The epidemiology of cognitive development. Cognition. 2021 Aug;213:104690. doi: 10.1016/j.cognition.2021.104690

21. Ganuthula VRR, Sinha S. The Looking Glass for Intelligence Quotient Tests: The Interplay of Motivation, Cognitive Functioning, and Affect. Front Psychol. 2019 Dec 17;10:2857. doi: 10.3389/fpsyg.2019.02857.

22. Eigsti IM. Peabody Picture Vocabulary Test. In: Volkmar FR. (eds) Encyclopedia of Autism Spectrum Disorders [Internet]. New York, NY: Springer; 2013 [cited 2024 Jan 02] Available from: https://doi.org/10.1007/978-1-4419-1698-3_531

23. Na SD, Burns TG. Wechsler Intelligence Scale for Children-V: Test Review. Appl Neuropsychol Child. 2016;5(2):156-60. doi: 10.1080/21622965.2015.1015337.

24. Neisser U, Boodoo G, Bouchard Jr. TJ, Boykin AW, Brody N, Ceci SJ, et al. Intelligence: Knowns and unknowns. American Psychologist. 1996;51:77–101.

25. Haier RJ. Increased intelligence is a myth (so far). Front Syst Neurosci. 2014 Mar 12;8:34. doi: 10.3389/fnsys.2014.00034. PMID: 24659957; PMCID: PMC3950413.

26. Spearman C. 'General intelligence,' objectively determined and measured. Am J Psychol. 1904;15:201–93.

27. Jensen AR. The g factor: The science of mental ability [Internet]. Westport, CT, US: Praeger Publishers/Greenwood Publishing Group; 1998. Available from: https://arthurjensen.net/wp-content/uploads/2020/04/The-g-factor-the-science-of-mental-ability-Arthur-R.-Jensen.pdf

28. de la Fuente J, Davies G, Grotzinger AD, Tucker-Drob EM, Deary IJ. A general dimension of genetic sharing across diverse cognitive traits inferred from molecular data. Nat Hum Behav. 2021 Jan;5(1):49-58. doi: 10.1038/s41562-020-00936-2. Epub 2020 Sep 7. PMID: 32895543; PMCID: PMC9346507.

29. Thomson GH. A Hierarch without a general factor. British Journal of Psychology. 1916 Sep 1;8(3):271–81. https://doi.org/10.1111/j.2044-8295.1916.tb00133.x

30. Deary IJ, Whalley LJ, Lemmon H, Crawford JR, Starr JM. The stability of individual differences in mental ability from childhood to old age: Follow-up of the 1932 Scottish Mental Survey. Intelligence. 2000;28:49–55. https://doi.org/10.1016/S0160-2896(99)00031-8

31. Rönnlund M, Sundström A, Nilsson LG. Interindividual differences in general cognitive ability from age 18 to age 65 years are extremely stable and strongly associated with working memory capacity. Intelligence. 2015;53:59–64. doi:10.1016/j.intell.2015.08.011

32. Bayley N. Consistency and variability in the growth of intelligence from birth to 18 years. J Genet Psychol. 1949 Dec;75(2):165-96. doi: 10.1080/08856559.1949.10533516. PMID: 15403674.

33. World Health Organisation. International Statistical Classification of Diseases and Related Health Problems (ICD). [Internet]. [cited 2023 Sep 10]. Available at: https://www.who.int/standards/classifications/classification-of-diseases

34. World Health Organisation. ICD-11 for Mortality and Morbidity Statistics: 6A00 Disorders of intellectual development.[Internet] 2023 [cited 2023 Sep 05]. Available at: https://icd.who.int/browse11/l-m/en#/http%253A%252F%252Fid.who.int%252Ficd%252Fentity%252F605267007

35. American Psychological Association. APA Dictionary of Psychology: IQ [Internet]. [cited 2023 Aug 10]. Available from: https://dictionary.apa.org/iq

36. Wieland J, Zitman FG. It is time to bring borderline intellectual functioning back into the main fold of classification systems. BJPsych Bull. 2016 Aug;40(4):204-6. doi: 10.1192/pb.bp.115.051490. PMID: 27512590; PMCID: PMC4967780.

37. Greenspan S. Borderline intellectual functioning: an update. Curr Opin Psychiatry. 2017 Mar;30(2):113-122. doi: 10.1097/YCO.0000000000000317. PMID: 28079674.

38. Nouwens PJG, Lucas R, Embregts PJCM, van Nieuwenhuizen C. In plain sight but still invisible: A structured case analysis of people with mild intellectual disability or borderline intellectual functioning. J Intellect Dev Disabil. 2017;42(1):36–44. https://doi.org/10.3109/13668250.2016.1178220

39. Khan I, Leventhal BL. Developmental Delay. 2023 Jul 17. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan–. PMID: 32965902.

40. Choo YY, Agarwal P, How CH, Yeleswarapu SP. Developmental delay: identification and management at primary care level. Singapore Med J. 2019 Mar;60(3):119-123. doi: 10.11622/smedj.2019025. PMID: 30997518; PMCID: PMC6441684.

41. Moore T, Arefadib N, Deery A, West S. The First Thousand Days: An Evidence Paper [Internet]. Melbourne: Centre for Community Child Health, Murdoch Children's Research Institute; 2017. Available from: https://www.rch.org.au/uploadedFiles/Main/Content/ccchdev/CCCH-The-First-Thousand-Days-An-Evidence-Paper-September-2017.pdf

42. Cusick S, Georgieff M. The first 1,000 days of life: The brain's window of opportunity. [Internet]. Unicef; 2013 [cited 2023 Sep 11]. Available at: https://www.unicef-irc.org/article/958-the-first-1000-days-of-life-the-brains-window-of-opportunity.html#:~:text=The%20first%201%2C000%20days%20of%20life%20%2D%20the%20time%20spanning%20roughly,across%20the%20lifespan%20are%20established.

43. Barker DJ, Gluckman PD, Godfrey KM, Harding JE, Owens JA, Robinson JS. Fetal nutrition and cardiovascular disease in adult life. Lancet. 1993 Apr 10;341(8850):938-41. doi: 10.1016/0140-6736(93)91224-a. PMID: 8096277.

44. Barker DJ, Osmond C. Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales. Lancet. 1986 May 10;1(8489):1077-81. doi: 10.1016/s0140-6736(86)91340-1. PMID: 2871345.

45. Barker DJ, Winter PD, Osmond C, Margetts B, Simmonds SJ. Weight in infancy and death from ischaemic heart disease. Lancet. 1989 Sep 9;2(8663):577-80. doi: 10.1016/s0140-6736(89)90710-1. PMID: 2570282.

46. Fox SE, Levitt P, Nelson CA 3rd. How the timing and quality of early experiences influence the development of brain architecture. Child Dev. 2010 Jan-Feb;81(1):28-40. doi: 10.1111/j.1467-8624.2009.01380.x. PMID: 20331653; PMCID: PMC2846084.

47. Barker DJP. Developmental origins of adult health and disease. J Epidemiol Community Health (1978) [Internet]. 2004 Feb 1;58(2):114 LP – 115. Available from: http://jech.bmj.com/content/58/2/114.abstract

48. Sahu M, Prasuna JG. Twin Studies: A Unique Epidemiological Tool. Indian J Community Med. 2016 Jul-Sep;41(3):177-82. doi: 10.4103/0970-0218.183593. PMID: 27385869; PMCID: PMC4919929.

49. Spencer N, Raman S, O'Hare B, Tamburlini G. Addressing inequities in child health and development: towards social justice. BMJ Paediatr Open. 2019 Aug 1;3(1):e000503. doi: 10.1136/bmjpo-2019-000503.

50. Fu M, Zuo Y. Experience-dependent structural plasticity in the cortex. Trends Neurosci. 2011 Apr;34(4):177-87. doi: 10.1016/j.tins.2011.02.001.

51. Duncan EJ, Gluckman PD, Dearden PK. Epigenetics, plasticity, and evolution: How do we link epigenetic change to phenotype? J Exp Zool B Mol Dev Evol. 2014 Jun;322(4):208–20. https://doi.org/10.1002/jez.b.22571

52. Barouki R, Gluckman PD, Grandjean P, Hanson M, Heindel JJ. Developmental origins of non-communicable disease: implications for research and public health. Environ Health. 2012 Jun 27;11:42. doi: 10.1186/1476-069X-11-42. PMID: 22715989; PMCID: PMC3384466.

53. Weaver IC. Shaping adult phenotypes through early life environments. Birth Defects Res C Embryo Today. 2009 Dec;87(4):314-26. doi: 10.1002/bdrc.20164.

54. Wang Y, Liu H, Sun Z. Lamarck rises from his grave: parental environment-induced epigenetic inheritance in model organisms and humans. Biol Rev Camb Philos Soc. 2017 Nov;92(4):2084-2111. doi: 10.1111/brv.12322.

55. Tierney AL, Nelson CA 3rd. Brain Development and the Role of Experience in the Early Years. Zero Three. 2009 Nov 1;30(2):9-13. PMID: 23894221; PMCID: PMC3722610.

56. Smith RC. The Biopsychosocial Revolution: Interviewing and Provider-patient Relationships Becoming Key Issues for Primary Care. J Gen Intern Med. 2002 Apr;17(4):309–10. doi: 10.1046/j.1525-1497.2002.20210.x. PMCID: PMC1495036.

57. Chipuer HM, Rovine MJ, Plomin R. LISREL modeling: Genetic and environmental influences on IQ revisited. Intelligence. 1990;14(1):11–29. https://doi.org/10.1016/0160-2896(90)90011-H

58. Plomin R, von Stumm S. The new genetics of intelligence. Nat Rev Genet. 2018 Mar;19(3):148-159. doi: 10.1038/nrg.2017.104. Epub 2018 Jan 8. PMID: 29335645; PMCID: PMC5985927.

59. Briley DA, Tucker-Drob EM. Explaining the increasing heritability of cognitive ability across development: a meta-analysis of longitudinal twin and adoption studies. Psychol Sci. 2013 Sep;24(9):1704-13. doi: 10.1177/0956797613478618. Epub 2013 Jul 1. Erratum in: Psychol Sci. 2021 Mar;32(3):466. doi: 10.1177/0956797621994569. PMID: 23818655; PMCID: PMC3954471.

60. Mamluk L, Edwards HB, Savović J, Leach V, Jones T, Moore THM, et al. Prenatal alcohol exposure and pregnancy and childhood outcomes: a systematic review of alternative analytical approaches. Lancet. 2016;388:S73. https://doi.org/10.1016/S0140-6736(16)32309-1

61. Singer LT, Minnes S, Short E, Arendt R, Farkas K, Lewis B, Klein N, Russ S, Min MO, Kirchner HL. Cognitive outcomes of preschool children with prenatal cocaine exposure. JAMA. 2004 May 26;291(20):2448-56. doi: 10.1001/jama.291.20.2448. PMID: 15161895; PMCID: PMC10249064.

62. Yeoh SL, Eastwood J, Wright IM, Morton R, Melhuish E, Ward M, Oei JL. Cognitive and Motor Outcomes of Children With Prenatal Opioid Exposure: A Systematic Review and Meta-analysis. JAMA Netw Open. 2019 Jul 3;2(7):e197025. doi: 10.1001/jamanetworkopen.2019.7025. PMID: 31298718; PMCID: PMC6628595.

63. Lopuszanska U, Samardakiewicz M. The Relationship Between Air Pollution and Cognitive Functions in Children and Adolescents: A Systematic Review. Cogn Behav Neurol. 2020 Sep;33(3):157-178. doi: 10.1097/WNN.0000000000000235. PMID: 32889949.

64. Georgieff MK, Ramel SE, Cusick SE. Nutritional influences on brain development. Acta Paediatr. 2018 Aug;107(8):1310-1321. doi: 10.1111/apa.14287. Epub 2018 Mar 22. PMID: 29468731; PMCID: PMC6045434.

65. Wiegersma AM, Dalman C, Lee BK, Karlsson H, Gardner RM. Association of Prenatal Maternal Anemia With Neurodevelopmental Disorders. JAMA Psychiatry. 2019 Dec 1;76(12):1294-1304. doi: 10.1001/jamapsychiatry.2019.2309.

66. Kwok J, Hall HA, Murray AL, Lombardo MV, Auyeung B. Maternal infections during pregnancy and child cognitive outcomes. BMC Pregnancy Childbirth. 2022 Nov 17;22(1):848. doi: 10.1186/s12884-022-05188-8. PMID: 36397016; PMCID: PMC9670450.

67. Dzikienė R, Lukoševičius S, Laurynaitienė J, Marmienė V, Nedzelskienė I, Tamelienė R, et al. Long-Term Outcomes of Perinatal Hypoxia and Asphyxia at

an Early School Age. Medicina (Kaunas). 2021 Sep 18;57(9):988. doi: 10.3390/medicina57090988.

68. Aizer A, Stroud L, Buka S. Maternal Stress and Child Outcomes: Evidence from Siblings. J Hum Resour. 2016;51(3):523-555. doi: 10.3386/w18422. PMID: 29118458; PMCID: PMC5673140.

69. Rogers A, Obst S, Teague SJ, Rossen L, Spry EA, Macdonald JA, et al. Association Between Maternal Perinatal Depression and Anxiety and Child and Adolescent Development: A Meta-analysis. JAMA Pediatr. 2020 Nov 1;174(11):1082-1092. doi: 10.1001/jamapediatrics.2020.2910.

70. Collins JM, Keane JM, Deady C, Khashan AS, McCarthy FP, O'Keeffe GW, Clarke G, Cryan JF, Caputi V, O'Mahony SM. Prenatal stress impacts foetal neurodevelopment: Temporal windows of gestational vulnerability. Neurosci Biobehav Rev. 2024 Sep;164:105793. doi: 10.1016/j.neubiorev.2024.105793. Epub 2024 Jul 5. PMID: 38971516.

71. Abel KM, Heuvelman H, Rai D, Timpson NJ, Sarginson J, Shallcross R, Mitchell H, Hope H, Emsley R. Intelligence in offspring born to women exposed to intimate partner violence: a population-based cohort study. Wellcome Open Res. 2019 Jul 10;4:107. doi: 10.12688/wellcomeopenres.15270.1. PMID: 31681855; PMCID: PMC6820818.

72. Parks RW, Stevens RJ, Spence SA. A systematic review of cognition in homeless children and adolescents. J R Soc Med. 2007 Jan;100(1):46-50. doi: 10.1177/014107680710000116. PMID: 17197688; PMCID: PMC1761678.

73. Bradley RH, Corwyn RF. Socioeconomic status and child development. Annu Rev Psychol. 2002;53:371-99. doi: 10.1146/annurev.psych.53.100901.135233. PMID: 11752490.

74. Larson K, Russ SA, Nelson BB, Olson LM, Halfon N. Cognitive ability at kindergarten entry and socioeconomic status. Pediatrics. 2015 Feb;135(2):e440-8. doi: 10.1542/peds.2014-0434. Epub 2015 Jan 19. PMID: 25601983.

75. von Stumm S, Plomin R. Socioeconomic status and the growth of intelligence from infancy through adolescence. Intelligence 2015; 48: 30–36. https://doi.org/10.1016/j.intell.2014.10.002

76. Hackman DA, Farah MJ, Meaney MJ. Socioeconomic status and the brain: mechanistic insights from human and animal research. Nat Rev Neurosci. 2010 Sep;11(9):651-9. doi: 10.1038/nrn2897. PMID: 20725096; PMCID: PMC2950073.

77. Kobrosly RW, van Wijngaarden E, Galea S, Cory-Slechta DA, Love T, Hong C, et al. Socioeconomic position and cognitive function in the Seychelles: a life course analysis. Neuroepidemiology. 2011;36(3):162-8. doi: 10.1159/000325779.

78. Repetti RL, Taylor SE, Seeman TE. Risky families: family social environments and the mental and physical health of offspring. Psychol Bull. 2002 Mar;128(2):330–66. PMID: 11931522.

79. McLoyd VC. Socioeconomic disadvantage and child development. Am Psychol. 1998 Feb;53(2):185-204. doi: 10.1037//0003-066x.53.2.185.

80. Troller-Renfree SV, Costanzo MA, Duncan GJ, Magnuson K, Gennetian LA, Yoshikawa H, et al. The impact of a poverty reduction intervention on infant brain activity. Proc Natl Acad Sci U S A. 2022 Feb 1;119(5):e2115649119. doi: 10.1073/pnas.2115649119.

81. Twilhaar ES, Wade RM, de Kieviet JF, van Goudoever JB, van Elburg RM, Oosterlaan J. Cognitive Outcomes of Children Born Extremely or Very Preterm Since the 1990s and Associated Risk Factors: A Meta-analysis and Meta-regression. JAMA Pediatr. 2018 Apr 1;172(4):361-367. doi: 10.1001/jamapediatrics.2017.5323. PMID: 29459939; PMCID: PMC5875339.

82. Soleimani F, Azari N, Ghiasvand H, Shahrokhi A, Rahmani N, Fatollahierad S. Do NICU developmental care improve cognitive and motor outcomes for preterm infants? A systematic review and meta-analysis. BMC Pediatr. 2020 Feb 13;20(1):67. doi: 10.1186/s12887-020-1953-1. PMID: 32054469; PMCID: PMC7017495.

83. Talge NM, Holzman C, Wang J, Lucia V, Gardiner J, Breslau N. Late-preterm birth and its association with cognitive and socioemotional outcomes at 6 years of age. Pediatrics. 2010 Dec;126(6):1124-31. doi: 10.1542/peds.2010-1536. Epub 2010 Nov 22. PMID: 21098151.

84. Cheong JL, Doyle LW, Burnett AC, Lee KJ, Walsh JM, Potter CR, et al. Association Between Moderate and Late Preterm Birth and Neurodevelopment and Social-Emotional Development at Age 2 Years. JAMA Pediatr. 2017 Apr 3;171(4):e164805. doi: 10.1001/jamapediatrics.2016.4805.

85. Rindermann H, Neubauer AC. Processing speed, intelligence, creativity, and school performance: Testing of causal hypotheses using structural equation models. Intelligence. 2004;32(6):573–89. doi:10.1016/j.intell.2004.06.005

86. Stasielowicz, Lukasz. How important is cognitive ability when adapting to changes? A meta-analysis of the performance adaptation literature. Personality and Individual Differences. 2020; 166. doi: 10.1016/j.paid.2020.110178.

87. Batty GD, Deary IJ, Macintyre S. Childhood IQ in relation to risk factors for premature mortality in middle-aged persons: the Aberdeen Children of the 1950s study. J Epidemiol Community Health. 2007 Mar;61(3):241-7. doi: 10.1136/jech.2006.048215. PMID: 17325403; PMCID: PMC2652919.

88. Koenen KC, Moffitt TE, Roberts AL, Martin LT, Kubzansky L, Harrington H, Poulton R, Caspi A. Childhood IQ and adult mental disorders: a test of the cognitive reserve hypothesis. Am J Psychiatry. 2009 Jan;166(1):50-7. doi: 10.1176/appi.ajp.2008.08030343. Epub 2008 Dec 1. PMID: 19047325; PMCID: PMC2705657.

89. Bartels M, Rietveld MJ, Van Baal GC, Boomsma DI. Heritability of educational achievement in 12-year-olds and the overlap with cognitive ability. Twin Res. 2002 Dec;5(6):544-53. doi: 10.1375/136905202762342017. PMID: 12573186.

90. Deary IJ, Strand S, Smith P, Fernandes C. Intelligence and educational achievement. Intelligence. 2007;35(1):13–21. https://doi.org/10.1016/j.intell.2006.02.001

91. Fergusson DM, Horwood LJ, Ridder EM. Show me the child at seven II: Childhood intelligence and later outcomes in adolescence and young

adulthood. J Child Psychol Psychiatry. 2005 Aug;46(8):850-8. doi: 10.1111/j.1469-7610.2005.01472.x. PMID: 16033633.

92. Furnham A, Monsen J. Personality traits and intelligence predict academic school grades. Learn Individ Differ. 2009;19(1):28–33. https://doi.org/10.1016/j.lindif.2008.02.001

93. Bee GM, Duke RL. Relationship between Intelligence, Scholastic Motivation, and Academic Achievement. Psychol Rep. 1960 Jan 1;6(1):3–8. https://doi.org/10.2466/pr0.1960.6.1.3

94. Calvin CM, Batty GD, Der G, Brett CE, Taylor A, Pattie A, et al. Childhood intelligence in relation to major causes of death in 68 year follow-up: prospective population study. BMJ. 2017 Jun 28;357:j2708. doi: 10.1136/bmj.j2708.

95. Forrest LF, Hodgson S, Parker L, Pearce MS. The influence of childhood IQ and education on social mobility in the Newcastle Thousand Families birth cohort. BMC Public Health. 2011 Nov 25;11:895. doi: 10.1186/1471-2458-11-895. PMID: 22117779; PMCID: PMC3248886

96. Furnham A, Cheng H. Childhood Cognitive Ability Predicts Adult Financial Well-Being. J Intell. 2016 Dec 27;5(1):3. doi: 10.3390/jintelligence5010003.

97. Alesi M, Rappo G, Pepi A. Emotional Profile and Intellectual Functioning: A Comparison Among Children With Borderline Intellectual Functioning, Average Intellectual Functioning, and Gifted Intellectual Functioning. Sage Open. 2015 Jul 1;5(3) https://doi.org/10.1177/2158244015589995

98. Emerson E, Einfeld S, Stancliffe RJ. The mental health of young children with intellectual disabilities or borderline intellectual functioning. Soc Psychiatry

Psychiatr Epidemiol. 2010 May;45(5):579-87. doi: 10.1007/s00127-009-0100-y. Epub 2009 Jul 19. PMID: 19618095.

99. McIntyre LL, Blacher J, Baker BL. The transition to school: adaptation in young children with and without intellectual disability. J Intellect Disabil Res. 2006 May;50(Pt 5):349-61. doi: 10.1111/j.1365-2788.2006.00783.x. PMID: 16629928.

100. Ramey CT, Ramey SL. Prevention of intellectual disabilities: early interventions to improve cognitive development. Prev Med. 1998 Mar-Apr;27(2):224-32. doi: 10.1006/pmed.1998.0279. PMID: 9579000.

101. Raney CT, Collier AM, Sparling JJ, Loda FA, Campbell FA, Finkelstein I, et al. The Carolina Abecedarian Project: A longitudinal and multidisciplinary approach to the prevention of developmental retardation [Internet]. Bethesda, Maryland: North Carolina University; 1974 [cited 2024 Jan 02]. Available from: https://eric.ed.gov/?id=ED104548

102. Campbell FA, Ramey CT, Pungello E, Sparling J, Miller-Johnson S. Early Childhood Education: Young Adult Outcomes From the Abecedarian Project. Appl Dev Sci. 2002 Jan 1;6(1):42–57. https://doi.org/10.1207/S1532480XADS0601_05

103. Muennig P, Robertson D, Johnson G, Campbell F, Pungello EP, Neidell M. The effect of an early education program on adult health: the Carolina Abecedarian Project randomized controlled trial. Am J Public Health. 2011 Mar;101(3):512-6. doi: 10.2105/AJPH.2010.200063. Epub 2011 Jan 13. PMID: 21233425; PMCID: PMC3036683.

104. Koshyk J, Wilson T, Stewart-Tufescu A, D'Souza M, Chase RM, Mignone J. The ripple effect: Examining the impact on parents of an Abecedarian early child care intervention in an urban social housing development. J Early Child Res. 2021 Mar;19(1):40-54. doi: 10.1177/1476718X20966696. Epub 2020 Dec 6. PMID: 34012334; PMCID: PMC8107451.

105. Orri M, Côté SM, Tremblay RE, Doyle O. Impact of an early childhood intervention on the home environment, and subsequent effects on child cognitive and emotional development: A secondary analysis. PLoS One. 2019 Jul 3;14(7):e0219133. doi: 10.1371/journal.pone.0219133. PMID: 31269050; PMCID: PMC6608972.

106. Planning for Life Evaluation Team. Final Report: Did Preparing for Life Improve Children's School Readiness? [Internet] Dublin: University College Dublin Geary Institute for Public Policy; 2016 August.[cited 2024 Dec 10]. Available from: https://www.preparingforlife.ie/wp-content/uploads/2017/12/PreparingforLife-Final-Report-LongVersion.pdf

107. U.S. Department of Health and Human Services. Head Start: Early Learning and Knowledge Center [Internet]. 2018 [cited 2023 Aug 23]. Available at: https://eclkc.ohs.acf.hhs.gov/

108. Office of the Assistant Secretary for Planning and Evaluation. Poverty Guidelines [Internet]. 2023 [cited 2023 Sep 01]. Available at https://aspe.hhs.gov/topics/poverty-economic-mobility/poverty-guidelines

109. Love JM, Kisker EE, Ross C, Raikes H, Constantine J, Boller K, Brooks-Gunn J, Chazan-Cohen R, Tarullo LB, Brady-Smith C, Fuligni AS, Schochet PZ, Paulsell D, Vogel C. The effectiveness of early head start for 3-year-old children and

their parents: lessons for policy and programs. Dev Psychol. 2005 Nov;41(6):885-901. doi: 10.1037/0012-1649.41.6.88. PMID: 16351335.

110. Garces E, Thomas D, Currie J. Longer-Term Effects of Head Start. The American Economic Review. 2002; 92(4), 999–1012.

111. Spittle A, Orton J, Anderson PJ, Boyd R, Doyle LW. Early developmental intervention programmes provided post hospital discharge to prevent motor and cognitive impairment in preterm infants. Cochrane Database Syst Rev. 2015 Nov 24;2015(11):CD005495. doi: 10.1002/14651858.CD005495.pub4.

112. Health Service Executive. Health Service Executive Early Intervention Team - Dun Laoghaire/Glasthule [Internet]. HSE [cited 2023 Aug 31]. Available from: https://www.hse.ie/eng/services/list/2/primarycare/pcteams/dublinsouthpcts/dunlaoghaireglasthulepct/eit.html

113. Health Service Executive Childrens Disability Network Team. Roadmap for Service Improvement 2023 – 2026 Disability Services for Children and Young People. [Internet]. HSE; 2024 [cited 2025 Jan 15]. Available at: https://www.hse.ie/eng/services/publications/disability/roadmap-for-service-improvement-2023-2026.pdf

114. Health Service Executive. HSE publishes 2023 Children's Disability Staff Census and Workforce Review - showing an increase in staff numbers. [Internet]. HSE ; 2024 [updated 2024 October 10; cited 2025 Jan 15]. Available at: https://about.hse.ie/news/hse-publishes-2023-childrens-disability-staff-census-and-workforce-review/]

115. Evans GW, Li D, Whipple SS. Cumulative risk and child development. Psychol Bull. 2013 Nov;139(6):1342-96. doi: 10.1037/a0031808.

116. Evans GW. A multimethodological analysis of cumulative risk and allostatic load among rural children. Dev Psychol. 2003 Sep;39(5):924-33. doi: 10.1037/0012-1649.39.5.924.

117. Luster T, McAdoo HP. Factors related to the achievement and adjustment of young African American children. Child Dev. 1994 Aug;65(4):1080-94. doi: 10.1111/j.1467-8624.1994.tb00804.x. PMID: 7956466

118. Ekeus C, Lindström K, Lindblad F, Rasmussen F, Hjern A. Preterm birth, social disadvantage, and cognitive competence in Swedish 18- to 19-year-old men. Pediatrics. 2010 Jan;125(1):e67-73. doi: 10.1542/peds.2008-3329. Epub 2009 Dec 7. PMID: 19969613.

119. Benavente-Fernández I, Synnes A, Grunau RE, Chau V, Ramraj C, Glass T, et al. Association of Socioeconomic Status and Brain Injury With Neurodevelopmental Outcomes of Very Preterm Children. JAMA Netw Open. 2019 May 3;2(5):e192914. doi: 10.1001/jamanetworkopen.2019.2914.

120. Citizens Information. Maternity care and the public health nurse [Internet]. [updated 2022 Dec 15; cited 2023 Aug 30]. Available at: https://www.citizensinformation.ie/en/health/health-services/reproductive_health/maternity-and-infant-welfare-services/#:~:text=Client%20Registration%20Unit.-,The%20public%20health%20nurse,of%20returning%20home%20from%20hospital.

121. The Nurture Programme Health Service Executive. National Guideline on the Use of the Ages and Stages Questionnaire for Developmental Screening of Children Between 1 month and 66 months of age. 2018 [cited 2023 Aug 05].

Available at: https://www.hse.ie/eng/health/child/nurture/asq3national-guideline-parta.pdf

122. Muthusamy S, Wagh D, Tan J, Bulsara M, Rao S. Utility of the Ages and Stages Questionnaire to Identify Developmental Delay in Children Aged 12 to 60 Months: A Systematic Review and Meta-analysis. JAMA Pediatr. 2022 Oct 1;176(10):980-989. doi: 10.1001/jamapediatrics.2022.3079. Erratum in: JAMA Pediatr. 2022 Dec 1;176(12):1274. doi: 10.1001/jamapediatrics.2022.4317. PMID: 36036913; PMCID: PMC9425289.

123. Charkaluk ML, Rousseau J, Calderon J, Bernard JY, Forhan A, Heude B, Kaminski M; EDEN Mother–Child Cohort Study Group. Ages and Stages Questionnaire at 3 Years for Predicting IQ at 5-6 Years. Pediatrics. 2017 Apr;139(4):e20162798. doi: 10.1542/peds.2016-2798. PMID: 28360034.

124. Ramspek CL, Steyerberg EW, Riley RD, Rosendaal FR, Dekkers OM, Dekker FW, van Diepen M. Prediction or causality? A scoping review of their conflation within current observational research. Eur J Epidemiol. 2021 Sep;36(9):889-898. doi: 10.1007/s10654-021-00794-w. Epub 2021 Aug 15. PMID: 34392488; PMCID: PMC8502741.

125. Linsell L, Malouf R, Morris J, Kurinczuk JJ, Marlow N. Risk Factor Models for Neurodevelopmental Outcomes in Children Born Very Preterm or With Very Low Birth Weight: A Systematic Review of Methodology and Reporting. Am J Epidemiol. 2017 Apr 1;185(7):601-612. doi: 10.1093/aje/kww135. PMID: 28338817.

126. Blakely T, Lynch J, Simons K, Bentley R, Rose S. Reflection on modern methods: when worlds collide-prediction, machine learning and causal

inference. Int J Epidemiol. 2021 Jan 23;49(6):2058-2064. doi: 10.1093/ije/dyz132. PMID: 31298274; PMCID: PMC8453374.

127. Tennant PWG, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. Int J Epidemiol. 2021 May 17;50(2):620-632. doi: 10.1093/ije/dyaa213.

128. Camprubi Robles M, Campoy C, Garcia Fernandez L, Lopez-Pedrosa JM, Rueda R, Martin MJ. Maternal Diabetes and Cognitive Performance in the Offspring: A Systematic Review and Meta-Analysis. PLoS One. 2015 Nov 13;10(11):e0142583. doi: 10.1371/journal.pone.0142583. PMID: 26566144; PMCID: PMC4643884.

129. Nelson S, Lerner E, Needlman R, Salvator A, Singer LT. Cocaine, anemia, and neurodevelopmental outcomes in children: a longitudinal study. J Dev Behav Pediatr. 2004 Feb;25(1):1-9. doi: 10.1097/00004703-200402000-00001. PMID: 14767350; PMCID: PMC2596987.

130. Derakhshan A, Korevaar TIM, Taylor PN, Levie D, Guxens M, Jaddoe VWV, Nelson SM, Tiemeier H, Peeters RP. The Association of Maternal Thyroid Autoimmunity During Pregnancy With Child IQ. J Clin Endocrinol Metab. 2018 Oct 1;103(10):3729-3736. doi: 10.1210/jc.2018-00743. PMID: 30020468.

131. Taylor CM, Kordas K, Golding J, Emond AM. Effects of low-level prenatal lead exposure on child IQ at 4 and 8 years in a UK birth cohort study. Neurotoxicology. 2017 Sep;62:162-169. doi: 10.1016/j.neuro.2017.07.003. Epub 2017 Jul 4. PMID: 28687448; PMCID: PMC5630203.

132. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, Reitsma JB, Kleijnen J, Mallett S; PROBAST Group†. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med. 2019 Jan 1;170(1):51-58. doi: 10.7326/M18-1376. PMID: 30596875.

133. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, Riley RD, Hemingway H, Altman DG; PROGRESS Group. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med. 2013;10(2):e1001381. doi: 10.1371/journal.pmed.1001381. Epub 2013 Feb 5. PMID: 23393430; PMCID: PMC3564751.

134. Eriksen HL, Kesmodel US, Underbjerg M, Kilburn TR, Bertrand J, Mortensen EL. Predictors of intelligence at the age of 5: family, pregnancy and birth characteristics, postnatal influences, and postnatal growth. PLoS One. 2013 Nov 13;8(11):e79200. doi: 10.1371/journal.pone.0079200.

135. Camargo-Figuera FA, Barros AJ, Santos IS, Matijasevich A, Barros FC. Early life determinants of low IQ at age 6 in children from the 2004 Pelotas Birth Cohort: a predictive approach. BMC Pediatr. 2014 Dec 16;14:308. doi: 10.1186/s12887-014-0308-1. PMID: 25510879; PMCID: PMC4272809.

136. Lawlor DA, Batty GD, Morton SM, Deary IJ, Macintyre S, Ronalds G, Leon DA. Early life predictors of childhood intelligence: evidence from the Aberdeen children of the 1950s study. J Epidemiol Community Health. 2005 Aug;59(8):656-63. doi: 10.1136/jech.2004.030205. PMID: 16020642; PMCID: PMC1733112.

137. Lawlor DA, Najman JM, Batty GD, O'Callaghan MJ, Williams GM, Bor W. Early life predictors of childhood intelligence: findings from the Mater-University

study of pregnancy and its outcomes. Paediatr Perinat Epidemiol. 2006 Mar;20(2):148-62. doi: 10.1111/j.1365-3016.2006.00704.x. PMID: 16466433.

138. Camacho C, Straatmann VS, Day JC, Taylor-Robinson D. Development of a predictive risk model for school readiness at age 3 years using the UK Millennium Cohort Study. BMJ Open. 2019 Jun;9(6):e024851. doi: 10.1136/bmjopen-2018-024851

139. Juul SE, Wood TR, German K, Law JB, Kolnik SE, Puia-Dumitrescu M, Mietzsch U, Gogcu S, Comstock BA, Li S, Mayock DE, Heagerty PJ; PENUT Consortium. Predicting 2-year neurodevelopmental outcomes in extremely preterm infants using graphical network and machine learning approaches. EClinicalMedicine. 2022 Dec 26;56:101782. doi: 10.1016/j.eclinm.2022.101782. PMID: 36618896; PMCID: PMC9813758

140. Barker LE, Shaw KM. Best (but oft-forgotten) practices: checking assumptions concerning regression residuals. Am J Clin Nutr. 2015 Sep;102(3):533-9. doi: 10.3945/ajcn.115.113498. Epub 2015 Jul 22. PMID: 26201816; PMCID: PMC4558311.

141. Copas JB. Regression, Prediction and Shrinkage. Journal of the Royal Statistical Society: Series B (Methodological). 1983 Jul 1;45(3):311–35. https://doi.org/10.1111/j.2517-6161.1983.tb01258.x

142. Ivanescu AE, Li P, George B, Brown AW, Keith SW, Raju D, Allison DB. The importance of prediction model validation and assessment in obesity and nutrition research. Int J Obes (Lond). 2016 Jun;40(6):887-94. doi: 10.1038/ijo.2015.214. Epub 2015 Oct 9. PMID: 26449421; PMCID: PMC4826636.

143. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. Nat Methods. 2018;15(4):233–4. https://doi.org/10.1038/nmeth.4642

144. Azzolina D, Baldi I, Barbati G, Berchialla P, Bottigliengo D, Bucci A, et al. Machine learning in clinical and epidemiological research: Isn't it time for biostatisticians to work on it? Epidemiol Biostat Public Health. 2019;16(4). doi:10.2427/13245

145. Rajula HSR, Verlato G, Manchia M, Antonucci N, Fanos V. Comparison of Conventional Statistical Methods with Machine Learning in Medicine: Diagnosis, Drug Development, and Treatment. Medicina (Kaunas). 2020 Sep 8;56(9):455. doi: 10.3390/medicina56090455. PMID: 32911665; PMCID: PMC7560135.

146. Jovel J, Greiner R. An Introduction to Machine Learning Approaches for Biomedical Research. Front Med (Lausanne). 2021 Dec 16;8:771607. doi: 10.3389/fmed.2021.771607. PMID: 34977072; PMCID: PMC8716730.

147. Woodruff PG, Modrek B, Choy DF, Jia G, Abbas AR, Ellwanger A, Koth LL, Arron JR, Fahy JV. T-helper type 2-driven inflammation defines major subphenotypes of asthma. Am J Respir Crit Care Med. 2009 Sep 1;180(5):388-95. doi: 10.1164/rccm.200903-0392OC. Epub 2009 May 29. Erratum in: Am J Respir Crit Care Med. 2009 Oct 15;180(8):796. PMID: 19483109; PMCID: PMC2742757.

148. Joseph VR. Optimal ratio for data splitting. Statistical Analysis and Data Mining: The ASA Data Science Journal. 2022 Aug 1;15(4):531–8. doi:10.1002/sam.11583

149. Topçuoğlu BD, Lesniak NA, Ruffin MT 4th, Wiens J, Schloss PD. A Framework for Effective Application of Machine Learning to Microbiome-Based Classification Problems. mBio. 2020 Jun 9;11(3):e00434-20. doi: 10.1128/mBio.00434-20. PMID: 32518182; PMCID: PMC7373189.

150. Belkin M, Hsu D, Ma S, Mandal S. Reconciling modern machine-learning practice and the classical bias-variance trade-off. Proc Natl Acad Sci U S A. 2019 Aug;116(32):15849–54. https://doi.org/10.1073/pnas.1903070116

151. Couronné R, Probst P, Boulesteix AL. Random forest versus logistic regression: a large-scale benchmark experiment. BMC Bioinformatics. 2018 Jul 17;19(1):270. https://doi.org/10.1186/s12859-018-2264-5

152. Chen X, Wang M, Zhang H. The use of classification trees for bioinformatics. Wiley Interdiscip Rev Data Min Knowl Discov. 2011 Jan;1(1):55-63. doi: 10.1002/widm.14. Epub 2011 Jan 6. PMID: 22523608; PMCID: PMC3329156.

153. Khozeimeh F, Sharifrazi D, Izadi NH, Jouloudari JH, Shoeibi A, Alizadehsani R et al . RF-CNN-F: random forest with convolutional neural network features for coronary artery disease diagnosis based on cardiac magnetic resonance. Sci Rep. 2022; 12:11178 https://doi.org/10.1038/s41598-022-15374-5

154. IBM. Gini Impurity Measure [Internet]. 2024 [updated 2024 Jan 18; cited 2025 Jan 16]. Available from: https://www.ibm.com/docs/en/cognos-analytics/12.0.0?topic=terms-gini-impurity-measure

155. Singh H. Splitting Decision Trees with Gini Impurity [Internet]. Analytics Vidhya; 2024 [updated 2024 Nov 18; cited 2025 Jan 16]. Available from: https://www.analyticsvidhya.com/articles/gini-impurity/#h-what-is-gini-impurity

156. Petkovic D, Altman R, Wong M, Vigil A. Improving the explainability of Random Forest classifier - user centered approach. Pac Symp Biocomput. 2018;23:204–15. PMID: 29218882; PMCID: PMC5728671.

157. Oshiro TM, Perez PS, Baranauskas JA. How Many Trees in a Random Forest? BT - Machine Learning and Data Mining in Pattern Recognition. In: Perner P (eds) Machine Learning and Data Mining in Pattern Recognition. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer https://doi.org/10.1007/978-3-642-31537-4_13

158. Boehmke B, Greenwell B. Hands-On Machine Learning with R [Internet]. 2020 [cited 2025 Jan 14]. Available at: https://bradleyboehmke.github.io/HOML/

159. IBM. What is a loss function? [Internet]. 2024 [cited 2025 Jan 14]. Available at: https://www.ibm.com/think/topics/loss-function

160. Medium. Deep Learning(Part 2). Loss Function and Gradient Function [Internet]. 2024 [cited 2025 Jan 10]. Available at: https://medium.com/@sumbatilinda/deep-learning-part-2-loss-function-and-gradient-function-2f64c566a1d6

161. Natekin A, Knoll A. Gradient boosting machines, a tutorial. Front Neurorobot. 2013 Dec 4;7:21. doi: 10.3389/fnbot.2013.00021. PMID: 24409142; PMCID: PMC3885826.

162. Liu C. 5 Concepts You Should Know About Gradient Descent and Cost Function [Internet]. 2022 [cited 2025 Jan 12]. Available at: https://www.kdnuggets.com/2020/05/5-concepts-gradient-descent-cost-function.html#:~:text=The%20loss%20function%20describes%20how,the%20parameters%20of%20our%20model.

163. Ben-Hur A, Weston J. A user's guide to support vector machines. Methods Mol Biol. 2010;609:223-39. doi: 10.1007/978-1-60327-241-4_13. PMID: 20221922.

164. IBM. What are Support Vector Machines? [Internet]. 2023 [cited 2025 Jan 14]. Available at: https://www.ibm.com/think/topics/support-vector-machine

165. Singh N. Unlocking the Power of RBF Kernel: Advantages, Local Boundaries, Gamma's Influence [Internet] 2023 [cited 2025 Jan 14]. Available at: https://pub.aimind.so/unlocking-the-power-of-rbf-kernel-advantages-local-boundaries-gammas-influence-2aebafce336a

166. Drago F, Scharf RJ, Maphula A, Nyathi E, Mahopo TC, Svensen E, Mduma E, Bessong P, Rogawski McQuade ET. Psychosocial and environmental determinants of child cognitive development in rural south africa and tanzania: findings from the mal-ed cohort. BMC Public Health. 2020 Apr 16;20(1):505. doi: 10.1186/s12889-020-08598-5. PMID: 32299410; PMCID: PMC7164138.

167. Santos S, Maitre L, Warembourg C, Agier L, Richiardi L, Basagaña X, Vrijheid M. Applying the exposome concept in birth cohort research: a review of statistical approaches. Eur J Epidemiol. 2020 Mar;35(3):193-204. doi: 10.1007/s10654-020-00625-4. Epub 2020 Mar 27. PMID: 32221742; PMCID: PMC7154018.

168. Vrijheid M, Fossati S, Maitre L, Márquez S, Roumeliotaki T, Agier L et al. Early-Life Environmental Exposures and Childhood Obesity: An Exposome-Wide Approach. Environ Health Perspect. 2020 Jun;128(6):67009. doi: 10.1289/EHP5975. Epub 2020 Jun 24. PMID: 32579081; PMCID: PMC7313401.

169. Steer CD, Bolton P, Golding J. Preconception and prenatal environmental factors associated with communication impairments in 9 year old children using an exposome-wide approach. PLoS One. 2015 Mar 4;10(3):e0118701. doi: 10.1371/journal.pone.0118701. PMID: 25739097; PMCID: PMC4349447.

170. Pansieri C, Pandolfini C, Clavenna A, Choonara I, Bonati M. An Inventory of European Birth Cohorts. Int J Environ Res Public Health. 2020 Apr 28;17(9):3071. doi: 10.3390/ijerph17093071. PMID: 32354101; PMCID: PMC7246583.

171. Canova C, Cantarutti A. Population-Based Birth Cohort Studies in Epidemiology. Int J Environ Res Public Health. 2020 Jul 23;17(15):5276. doi: 10.3390/ijerph17155276. PMID: 32717778; PMCID: PMC7432312.

172. Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, Shilton A, Yearwood J, Dimitrova N, Ho TB, Venkatesh S, Berk M. Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View. J Med Internet Res. 2016 Dec 16;18(12):e323. doi: 10.2196/jmir.5870. PMID: 27986644; PMCID: PMC5238707.

173. Coley RY, Liao Q, Simon N, Shortreed SM. Empirical evaluation of internal validation methods for prediction in large-scale clinical data with rare-event outcomes: a case study in suicide risk prediction. BMC Med Res Methodol. 2023 Feb 1;23(1):33. doi: 10.1186/s12874-023-01844-5. PMID: 36721082; PMCID: PMC9890785.

174. Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. N Engl J Med. 2016 Sep 29;375(13):1216-9. doi: 10.1056/NEJMp1606181. PMID: 27682033; PMCID: PMC5070532.

175. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nat Med. 2022 Jan;28(1):31-38. doi: 10.1038/s41591-021-01614-0.

176. Uddin M, Wang Y, Woodbury-Smith M. Artificial intelligence for precision medicine in neurodevelopmental disorders. NPJ Digit Med. 2019 Nov 21;2:112. doi: 10.1038/s41746-019-0191-0. PMID: 31799421; PMCID: PMC6872596.

177. The National Institute for Health and Care Excellence. Attention deficit hyperactivity disorder: diagnosis and management [Internet]. NICE guideline. 2018 [cited 2022 Apr 12]. Available from: https://www.nice.org.uk/guidance/ng87/chapter/Recommendations#diagnosis

178. Das W, Khanna S. A Robust Machine Learning Based Framework for the Automated Detection of ADHD Using Pupillometric Biomarkers and Time Series Analysis. Sci Rep. 2021 Aug 12;11(1):16370. doi: 10.1038/s41598-021-95673-5. PMID: 34385511; PMCID: PMC8361128.

179. Sen B, Borle NC, Greiner R, Brown MRG. A general prediction model for the detection of ADHD and Autism using structural and functional MRI. PLoS One. 2018 Apr 17;13(4):e0194856. doi: 10.1371/journal.pone.0194856. PMID: 29664902; PMCID: PMC5903601.

180. Emerson RW, Adams C, Nishino T, Hazlett HC, Wolff JJ, Zwaigenbaum L et al. Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age. Sci Transl Med. 2017 Jun 7;9(393):eaag2882. doi: 10.1126/scitranslmed.aag2882. PMID: 28592562; PMCID: PMC5819345.

181. Girault JB, Munsell BC, Puechmaille D, Goldman BD, Prieto JC, Styner M, Gilmore JH. White matter connectomes at birth accurately predict cognitive abilities at age 2. Neuroimage. 2019 May 15;192:145-155. doi: 10.1016/j.neuroimage.2019.02.060. Epub 2019 Feb 27. PMID: 30825656; PMCID: PMC6453706.

182. Kim HH, An JI, Park YR. A Prediction Model for Detecting Developmental Disabilities in Preschool-Age Children Through Digital Biomarker-Driven Deep Learning in Serious Games: Development Study. JMIR Serious Games. 2021 Jun 4;9(2):e23130. doi: 10.2196/23130.

183. Zwaigenbaum L, Penner M. Autism spectrum disorder: advances in diagnosis and evaluation. BMJ. 2018 May 21;361:k1674. doi: 10.1136/bmj.k1674. PMID: 29784657.

184. Moon SJ, Hwang J, Kana R, Torous J, Kim JW. Accuracy of Machine Learning Algorithms for the Diagnosis of Autism Spectrum Disorder: Systematic Review and Meta-Analysis of Brain Magnetic Resonance Imaging Studies. JMIR Ment Health. 2019 Dec 20;6(12):e14108. doi: 10.2196/14108. PMID: 31562756; PMCID: PMC6942187.

185. Baker S, Kandasamy Y. Machine learning for understanding and predicting neurodevelopmental outcomes in premature infants: a systematic review. Pediatr Res. 2023 Jan;93(2):293-299. doi: 10.1038/s41390-022-02120-w. Epub 2022 May 31. PMID: 35641551; PMCID: PMC9153218.

186. Schadl K, Vassar R, Cahill-Rowley K, Yeom KW, Stevenson DK, Rose J. Prediction of cognitive and motor development in preterm children using exhaustive feature selection and cross-validation of near-term white matter

microstructure. Neuroimage Clin. 2017 Nov 29;17:667-679. doi: 10.1016/j.nicl.2017.11.023. PMID: 29234600; PMCID: PMC5722472.

187. Mhasawade V, Zhao Y, Chunara R. Machine learning and algorithmic fairness in public and population health. Nat Mach Intell. 2021;3(8):659–66. https://doi.org/10.1038/s42256-021-00373-4

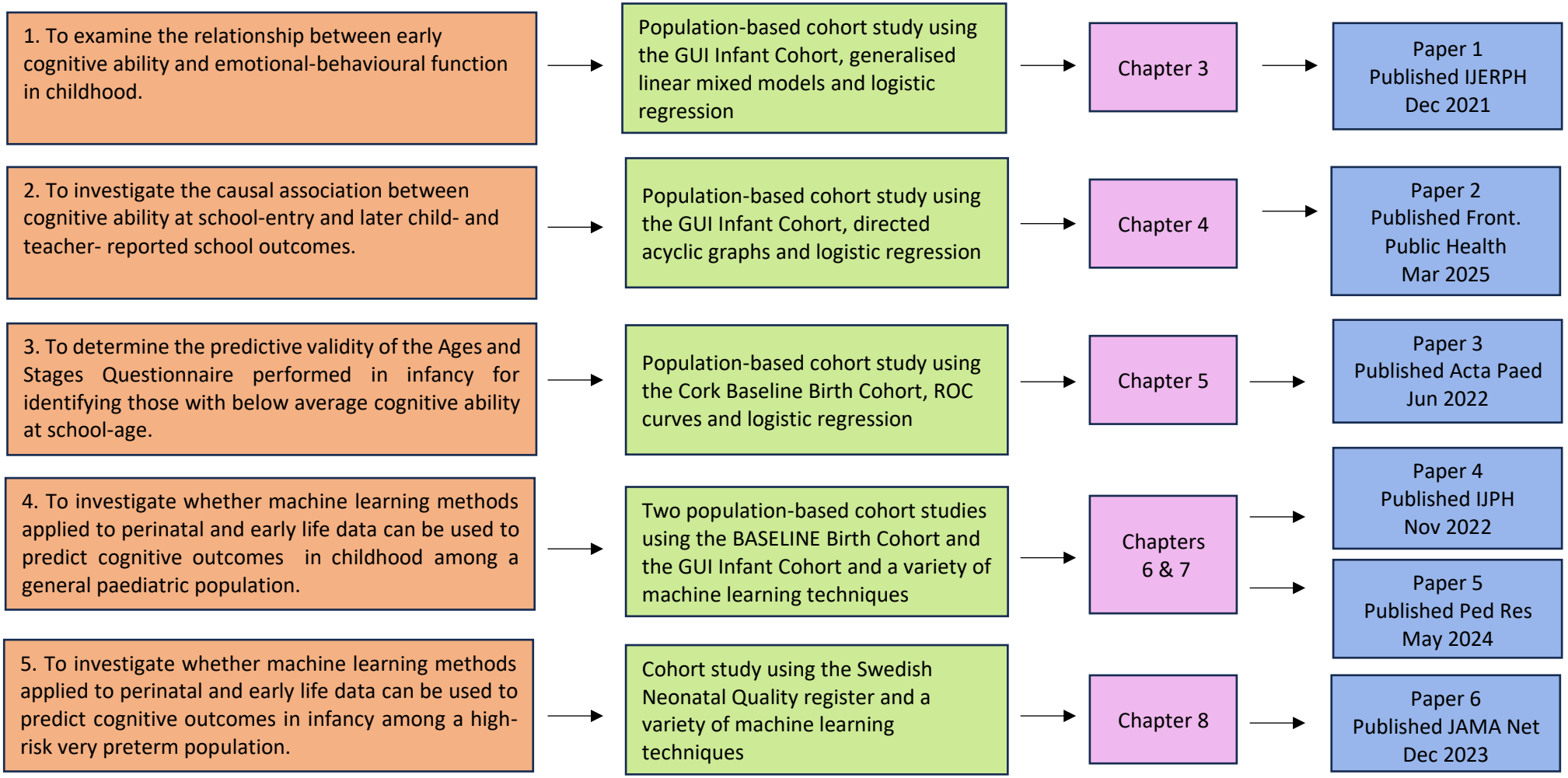
Chapter 2 Aims and Objectives

2.1.Aims and Objectives

The overall aim of this thesis was to explore the impact of below average cognitive ability on outcomes in childhood and to examine current and alternative methods of early identification of children at risk of below average cognitive ability in childhood. Specifically, the objectives were as follows:

1. To examine the relationship between early cognitive ability and emotional-behavioural function in childhood and to determine whether changes in emotional-behavioural function in childhood differ between those with and without below average cognitive ability.
2. To investigate the causal association between cognitive ability at school-entry and later child- and teacher- reported school outcomes.
3. To determine the predictive validity of the Ages and Stages Questionnaire performed in infancy for identifying those with below average cognitive ability at school-age.
4. To investigate whether machine learning methods applied to perinatal and early life data can be used to predict cognitive outcomes in childhood among a general paediatric population.
5. To investigate whether machine learning methods applied to perinatal and early life data can be used to predict cognitive outcomes in infancy among a high-risk very preterm population.

Figure 2.4 Overall thesis objectives with corresponding chapters and publications



Abbreviations: GUI, Growing Up in Ireland; ROC, Receiver operating characteristic curve; IJERPH, International Journal of Environmental Research and Public Health; Front. Public Health, Frontiers in Public Health; Acta Paed, Acta Paediatrica; IJPH, International Journal of Public Health; Ped Res, Pediatric Research; Jama Net, Journal of the American Medical Association (JAMA) Network Open.

2.2. Data sources used in this thesis

There were three data sources used in the original studies in this thesis, namely the Growing up in Ireland Infant Cohort, the Cork Baseline Birth Cohort, and the Swedish Neonatal Quality Register.

2.2.1. Growing Up in Ireland Infant Cohort

Overview: The Growing Up in Ireland (GUI) Infant Cohort was commissioned and funded by the Irish Government, with the principal objective of providing evidence-based research into childhood and the wellbeing of children in Ireland.[1]

Location: Republic of Ireland

Sampling strategy: The Irish Child Benefit Register is used to administer a monthly payment to the primary caregiver (usually the child's mother) of all children under the age of 16 years in Ireland. All children born in or resident in Ireland with a primary caregiver who lives in Ireland are eligible for child benefit. The Child Benefit Register was used as the sampling frame for the GUI Infant Cohort. Children born between 1st December 2007 and 30th June 2008 (n=41,185) were identified from the register, so they would be 9 months of age at the time of interview between September 2008 and April 2009. The sample was pre-stratified by marital status, county of residence, nationality, and number of children in the claim. Then, a systematic selection procedure with a random start and constant sampling fraction was used. Written information was sent to selected families before a face to face visit from the interviewer to arrange an appointment to conduct the interview at a convenient time

for the family. Data collection took place in person apart from at wave 4 (age 7/8) when a postal questionnaire only was used.[1]

Year data collection began: Wave 1 of data collection began in September 2008 when the infant was 9-months old.[1]

Follow up years: Wave 2 at age 3 years (December 2010 – July 2011), Wave 3 at age 5 years (March 2013 – September 2013), Wave 4 at age 7/8 years postal questionnaire only (Spring 2016), Wave 5 at age 9 years (June 2017 – February 2018).[2]

Survey Instruments: The primary and secondary caregivers were interviewed using Computer Assisted Personal Interviewing for non-sensitive questions. For sensitive questions Computer Assisted Self Interviewing was used. Paper versions of both were available in a number of different languages. All height and weight measurements were taken by the trained interviewer using a medically approved weighing scales and a Leicester measuring stick. Children were interviewed at age 9 years using Computer Assisted Personal Interviewing and pen and paper self-complete questionnaires for sensitive topics. Teachers were interviewed when the child was age 5 and 9 years using postal questionnaires. [1,2]

Number of participants: The number of participants at each wave of the survey is summarised in Figure 2.2. At wave 1, 11,134 infants (and their caregivers) participated. No additions were made to the sample after this time. These infants formed the target sample for all subsequent waves, except for those who moved outside of Ireland or who definitively opted out of the study.[2]

Figure 2.5 Flowchart of participants in the Growing Up in Ireland Infant Cohort

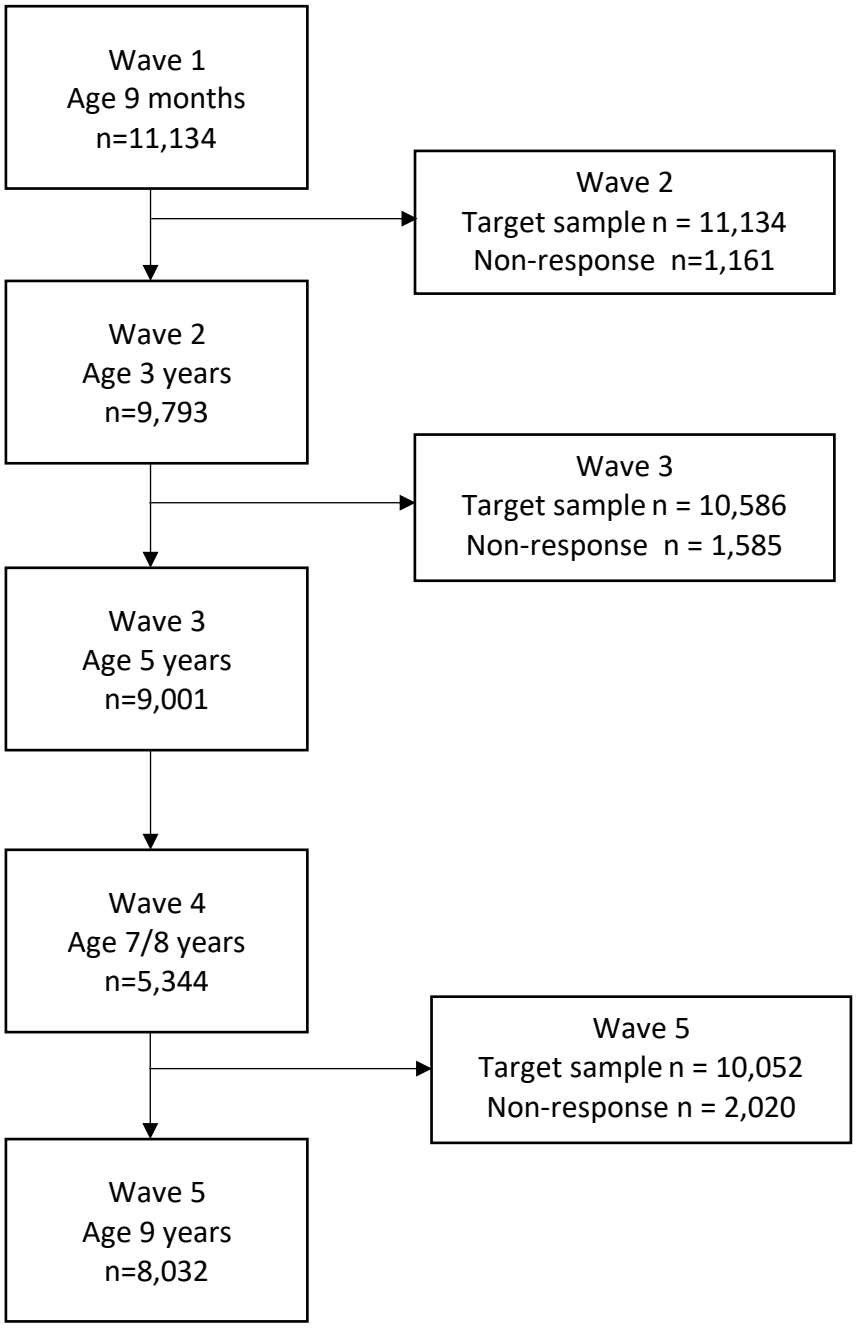


Figure 2.2 Legend
Shown in the main vertical boxes are the number of participants who completed each wave of the survey. Shown in the boxes to the right are the target sample at each wave and the number of non-responders. The target sample included those who participated in the previous round and/or those who had participated at one of the earlier rounds of the study. Families who moved abroad or definitively opted out of the study were excluded. Wave 4 of the study consisted of a postal questionnaire only.

Ethics and data access: The GUI study was approved by a dedicated and independent Research Ethics Committee convened by the Department of Health and Children. Secondary analysis of the anonymised microdata file does not require further ethical approval. Participants provided written informed consent. Data access was obtained through written request to the Irish Social Science Data Archive (ISSDA) with a study plan for each study using the dataset (Appendices 1-3).

Strengths: The GUI Infant cohort is a nationally representative study and a careful sampling strategy was used to select a representative sample of participants. Inter-wave attrition was low compared to many cohort studies, with 72% of the original sample participating at Wave 5.[3] The study team calculated a weighting variable which can be applied to the longitudinal samples to control for non-response and inter-wave attrition, which tended to be higher among younger mothers, those with lower levels of education or income, one-parent families, renters of social housing, those living in the Dublin region, those where the study child was never breastfed, and those where the primary caregiver smoked daily.[2]

Limitations: A limitation of this cohort, relevant to this thesis, is that the first wave of data collection began at infant age 9-months. Variables related to the antenatal period and the birth of the infant such as gestational age, birthweight, admission to the neonatal unit, and complications during pregnancy are therefore subject to recall bias.

2.2.2. Cork Baseline Birth Cohort

Overview: The Cork BASELINE Birth Cohort Study (Babies After SCOPE: Evaluating the Longitudinal Impact on Neurological and Nutritional Endpoints) is Ireland's first birth cohort study.[4] The overall objective of the cohort was to provide rich data on pregnancy and early infancy, particularly data related to the effect of biological, nutritional, and environmental factors on physical, neurological, and health outcomes. This was achieved through questionnaire data, clinical measurements, and blood samples.[4]

Location: Cork, Ireland. Cork is the largest county in Ireland and contains the second largest city.[4]

Sampling strategy: The study had two streams of recruitment. Stream 1 aimed to recruit participants from a separate pregnancy cohort study, the Screening for Pregnancy Endpoints (SCOPE) study, which began in 2008. All primiparous women in Cork were invited to participate in SCOPE. All SCOPE participants were informed about the BASELINE birth cohort and infants of mothers who consented to participate were registered with the cohort. In 2010, a second stream of recruitment was commenced and women in the postnatal wards at Cork University Maternity Hospital with a singleton pregnancy were invited to participate in the BASELINE cohort. Recruitment to BASELINE was conducted between August 2008 and November 2011 and 73% of participants were recruited through stream 1.

Year data collection began: Data collection began in 2008 during the antenatal period for women recruited through stream 1. For women recruited through stream 2 data collection began day 2 after birth of the baby.[4]

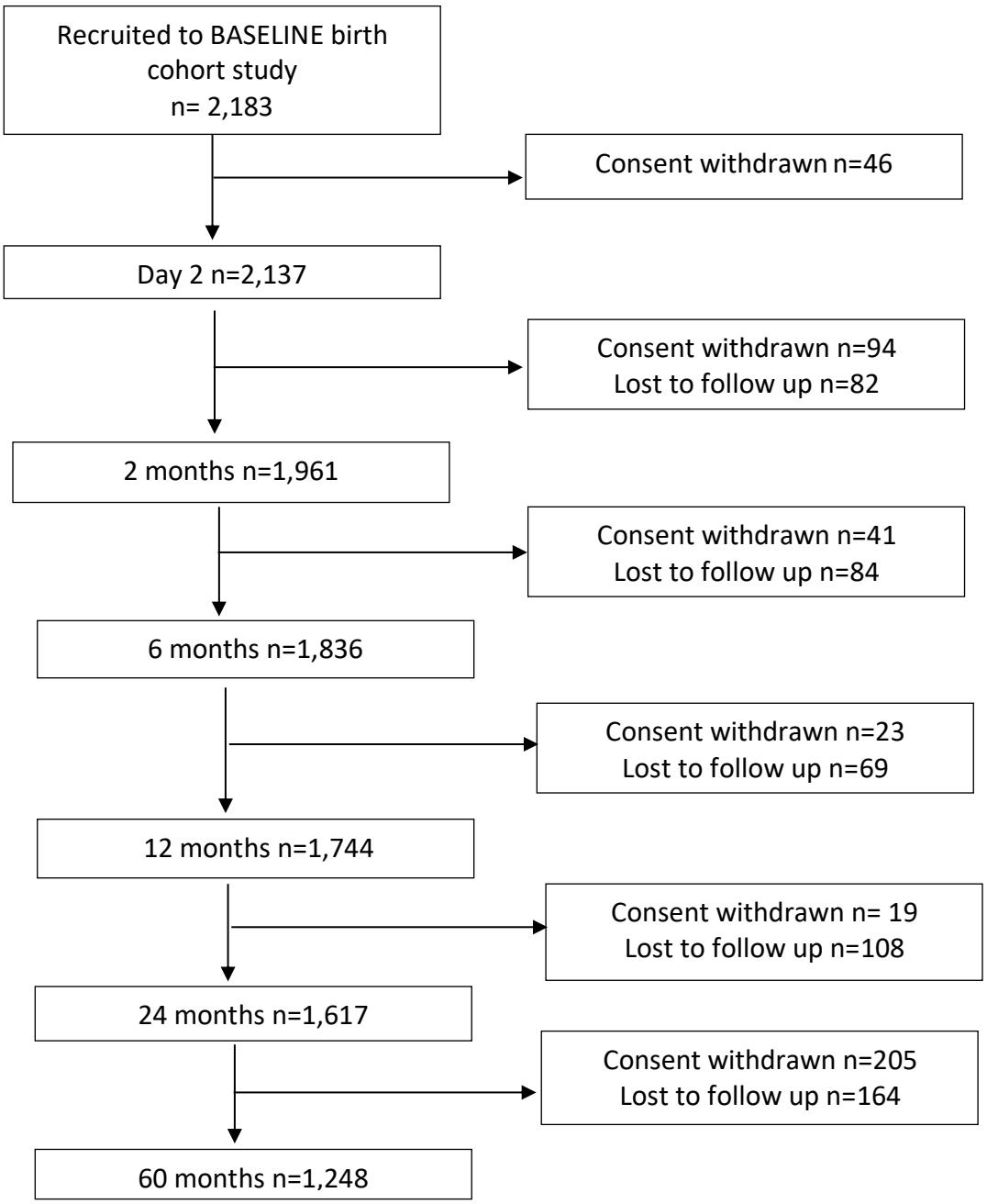
Follow up years: Postnatal visits were conducted at day 2, 2 months, 6 months, 12, months, 24 months, and age 5 years. The follow-up visits took place at a dedicated children's research centre in Cork University Hospital and were conducted by trained research nurses. Neurodevelopmental assessments were performed at age 24 months and age 5 years.[4,5]

Survey Instruments: Questionnaires administered to parents were self-completed paper-based questionnaires done in advance of the clinic visit, or done in person in clinic. All data were then transferred to a secure internet based online MedSciNet database that was Health Insurance Portability and Accountability Act (HIPAA) compliant. Clinical measurements including height, weight, and head circumference were measured by trained research nurses. Neurodevelopmental assessments were conducted by trained research nurses or research psychologists.[4,5] All data were subsequently cleaned, double checked for errors and compared to source data prior to locking of data for analysis.

Number of participants: The number of participants recruited to the cohort and who took part at each stage of follow up is outlined in Figure 2.3.

Ethics and data access: The BASELINE study was approved by the Clinical Research Ethics Committee of the Cork Teaching Hospitals. Participants provided written informed consent. Data were accessed following a written data access request submitted to the BASELINE study team and an ethical application to be added as co-investigator to the BASELINE study (Appendix 4). Ethical approval for the studies in this thesis using the BASELINE dataset was granted by the Clinical Research Ethics Committee of the Cork Teaching Hospitals (Appendix 5).

Figure 2.6 Flowchart of participants in the BASELINE Birth Cohort



Strengths: The majority of participants in the cohort (73%) were recruited in the antenatal period providing contemporaneous pregnancy data. Detailed assessment of a wide range of variables related to the pregnancy and birth of the infant was performed and standardised assessment of cognition using a well-established and validated IQ test was conducted.

Limitations: The study population was regional and may not be nationally representative. Of mothers recruited to the study, 55% had university level education, whereas census data from 2011 suggests this figure was closer to approximately 28% among women age 15-44 nationally, indicating a response bias.[6]

2.2.3. Swedish Neonatal Quality Register

Overview: The Swedish Neonatal Quality Register (SNQ) is a nationwide, population-based register, that captures data on all infants admitted for neonatal intensive care in Sweden. The register contains more than 300 variables for each admitted patient, providing detailed clinical information on conditions, investigations, and treatments initiated during pregnancy, delivery, and neonatal unit admission.[7]

Location: Sweden

Sampling strategy: All infants admitted to the 37 neonatal units in Sweden are captured in the opt-out register.[7]

Year data collection began: The SNQ began collecting data in 2001 and reached nationwide coverage in 2011. As part of a national follow-up programme launched in 2015, high-risk survivors of neonatal intensive care were recommended to receive comprehensive neurodevelopmental assessment at age 2 and 5.5 years of age.[7]

Follow up years: A comprehensive neurodevelopmental assessment is performed at age 2 years and 5.5 years of age for high-risk survivors of neonatal intensive care. High-risk survivors are defined as those with a gestational age <28 weeks, infants with a birthweight >3 SDs below the mean for gestational age, infants with morphological brain injury, infants with neonatal encephalopathy, infants with cerebral infection, and others with severe morbidity.[7]

Survey Instruments: Data related to the mother, the pregnancy, and the delivery are automatically extracted from medical records and transferred via the Swedish Pregnancy Register. Data on the neonatal unit admission are entered into the register by neonatal units, using an electronic platform which warns users of unreasonable values, blocks impossible values, and provides pop-ups with definitions and explanations. ICD codes are used, along with codes for surgeries and interventions.[7] Neurodevelopmental assessments at age 2 years are performed by psychologists in the vast majority of cases.

Number of participants: More than 207,000 infants have been registered in the SNQ since 2001.[8]

Ethics and data access: Parents to infants receiving neonatal care are informed that data are recorded in the SNQ with the possibility to opt out at any time. The SNQ project was approved by the Regional Ethics Review Board in Umea in 2001. To access SNQ data a principal investigator affiliated to a Swedish University must be part of the research team and apply for ethical approval with a detailed study plan (Appendix 6). Following approval from the Swedish Ethical Review Board (Appendix 7) de-identified data can be accessed and analysed without need for further consent from parents or caregivers. For the study conducted as part of this thesis Professor Mikael

Norman from the Department of Clinical Science, Intervention and Technology at the Karolinska Institutet was the principal investigator.

Strengths: As an opt-out population-based register, the SNQ is representative of the vast majority of infants who receive neonatal care in Sweden. The use of ICD, surgical, and intervention codes enhances data quality. The completeness of data has been studied and is excellent.[7] The SNQ captures very detailed clinical data, particularly that related to the birth and neonatal unit admission.

Limitations: Not all infants in the SNQ currently receive neurodevelopmental follow-up. This is restricted to infants deemed to have high-risk neonatal unit admissions, the majority of whom are very preterm infants born at a gestational age <28 weeks. There are a limited number of socioeconomic variables collected – namely parental education, the family language, and parental smoking status.[7,8]

2.3. References

1. Quail A, Williams J, McCrory C, Murray A, Thornton M. A Summary Guide to Wave 1 of the Infant Cohort of Growing Up in Ireland [Internet]. Dublin; 2011 [cited 2023 Aug 01] Available at: https://www.growingup.gov.ie/pubs/Summary-Guide_Infant-Cohort_Wave-1.pdf
2. Quail A, O'Reilly C, Watson D, McNamara E, O'Mahony D, Murray A. Growing Up in Ireland A Summary Guide to Cohort '08 at 9 Years [Internet]. Dublin; f2019 [cited 2023 Aug 01]. Available at: https://www.growingup.gov.ie/pubs/Summary-Guide-Infant-Cohort-Wave-5.pdf
3. Teague S, Youssef GJ, Macdonald JA, Sciberras E, Shatte A, Fuller-Tyszkiewicz M, et al. Retention strategies in longitudinal cohort studies: a systematic review and meta-analysis. BMC Med Res Methodol. 2018;18(1):151. https://doi.org/10.1186/s12874-018-0586-7
4. O'Donovan SM, Murray DM, Hourihane JO, Kenny LC, Irvine AD, Kiely M. Cohort profile: The Cork BASELINE Birth Cohort Study: Babies after SCOPE: Evaluating the Longitudinal Impact on Neurological and Nutritional Endpoints. Int J Epidemiol. 2015 Jun;44(3):764-75. doi: 10.1093/ije/dyu157. Epub 2014 Aug 7. PMID: 25102856.
5. The Cork BASELINE Birth Cohort [Internet]. 2013 [cited 2023 Aug 05]. Available at: http://www.baselinestudy.net/

6. Central Statistics Office. Central Statistics Office Census 2011. 2011 [cited 2023 Sep 07]. Available at: https://www.cso.ie/en/census/census2011reports/
7. Norman M, Källén K, Wahlström E, Håkansson S; SNQ Collaboration. The Swedish Neonatal Quality Register - contents, completeness and validity. Acta Paediatr. 2019 Aug;108(8):1411-1418. doi: 10.1111/apa.14823. Epub 2019 May 20. PMID: 31006126.
8. Norman M, Hakansson S, Swartling Schlinzig L. Swedish Neonatal Quality Register [Internet]. [cited 2023 Sep 12]. Available from: https://www.medscinet.com/PNQ/

Chapter 3 Below Average Cognitive Ability – An Under Researched Risk Factor For Emotional Behavioural Difficulties In Childhood

Andrea K. Bowe[1], Anthony Staines[2], Deirdre M. Murray[1,3]

[1] INFANT Research Centre, University College Cork, Ireland

[2] School of Nursing, Psychotherapy and Community Health, Dublin City University, Dublin, Ireland.

[3] Department of Paediatrics, School of Medicine, University College Cork, Cork, Ireland.

Published in The International Journal of Environmental Research and Public Health

December 2021

3.1.Abstract

Background

Children with below average cognitive ability represent a substantial yet under-researched population for whom cognitive and social demands, which increase in complexity year by year, may pose significant challenges. This observational study examines the longitudinal relationship between early cognitive ability and emotional-behavioural function between age 3-9 years.

Methods

Participants include 7,134 children from the population-based cohort study Growing Up In Ireland. Cognitive ability was measured at age 3 years using the Picture Similarities scale of the British Ability Scales. A t-score 1-2 standard deviations below the mean was defined as below average cognitive ability (n=767). Emotional-behavioural function was measured using the parent-reported Strengths and Difficulties Questionnaire (SDQ) at ages 3,5, and 9 years. Generalised linear mixed models and logistic regression were used to examine the relationship.

Results

Below average cognitive ability was an independent predictor of higher longitudinal SDQ scores (β 0.62, 95% confidence interval 0.34–0.91). The change in emotional-behavioural function between age 3 and 9 years differed between those with and without below average cognitive ability. After adjustment, the odds of experiencing clinically significant emotional-behavioural difficulties (EBDs) between age 3 and 9 years were 1.4 times higher for children with below average cognitive ability compared to those without (AOR 1.39, 95% CI 1.17-1.66, p<0.001).

Conclusions

Children with below average cognitive ability are at higher risk of EBDs in childhood. A scalable method of early identification should be a research priority for public health, enabling early intervention for cognitive and adaptive outcomes.

3.2. Introduction

Children with below average cognitive ability represent a substantial proportion of children, and often have unrecognised difficulties.[1,2] Intellectual function and adaptive behavior that are approximately 2 or more standard deviations (SDs) below the mean on a standardized test are diagnostic of an intellectual developmental disorder (IDD).[3] Those who perform 1-2 SDs below the mean fall into a "grey area", which is below that of the average population but not meeting the diagnostic criteria for an IDD.

Early deficits in cognitive ability are associated with later adverse outcomes in many domains of life including mental and physical health, educational attainment, socioeconomic status, and substance use.[4-7] Despite well-established adverse impacts, approximately 14% of children whose cognitive ability falls into this "grey area" receive little attention in the scientific literature and in ordinary life. Unlike children with more significant IDDs, whose marked deficits may be more likely to be detected on routine developmental screening, children with below average cognitive ability often reach their early motor and language milestones, and routinely commence formal education without recognition of the difficulties they may face and without additional educational supports.[8-10] This has potential to adversely affect both their transition into school, and their early years of experience in education, a critical period in the development of social and emotional well-being.[11-13]

The phrase "the first 1000 days", a period spanning conception until two years of age, is one increasingly seen in policy documents relating to child health. The phrase has its origins in the work of Professor David Barker, a physician and

epidemiologist, whose work prompted the development of the 'fetal origins hypothesis', also known as "Barker's hypothesis".[14,15] A fundamental concept in Barker's work was developmental plasticity, and he described critical periods 'when a system is plastic and sensitive to the environment, followed by loss of plasticity, and a fixed functional capacity'.[16] With regard to cognitive development specifically, this concept refers to periods during which neuronal connections are acutely susceptible to environment-dependent modification.[15] The "critical period" concept has since been challenged and while it is now broadly accepted that neuronal circuits can adapt throughout life, it is also accepted that early life is a unique period, differing from adulthood, during which the brain shows greater potentiality toward plasticity.[16] Since Barker, intervention in early life has been examined in lab-based studies, human intervention studies, and in epidemiological studies. There is now consensus agreement across the literature that early life is a period of fundamental importance in cognitive development, making it the ideal target for early intervention programs.[16]

The challenge for children with below average cognitive ability is that the early years of education may not unearth the cognitive difficulties experienced by children with below average ability. Thereafter, the increasing complexity of cognitive tasks and the accelerated academic demand may uncover underlying difficulties for these children, who may struggle to keep pace with their typically developing peers.[17] Unfortunately, a child may have to consistently fail the tasks asked of them for many years before they are considered for formal educational assessment.[10] The implications of this repeated failure on their psychological outcomes, particularly on emotional-behavioral development, is unclear.

Emotional behavioural difficulties (EBDs) are more prevalent in children with intellectual disabilities, though the literature examining the larger proportion of children with cognitive ability between 1-2 SD below the mean specifically is limited. Studies which have focused specifically on those with borderline intellectual ability suggest that these children are at increased risk of EBDs compared to their typically developing peers, though gaps exist in the literature in relation to the change in emotional-behavioral function over time, and the temporal relationship with commencing formal education.[17-20] Real-time recognition of below average cognitive ability occurs rarely.[8,10] For this reason, epidemiological studies of population-based data are ideally placed to identify the difficulties experienced by this substantial proportion of children, to inform policy and service planning, and to adequately meet their educational and psychological needs.

The aim of this study, which is based on a nationally representative sample of children in Ireland, is to inform childhood policy on early intervention through a detailed examination of the impacts of below average cognitive function on emotional-behavioral well-being. Our hypothesis is that having below average cognitive ability will have a negative effect on emotional-behavioural function in childhood, and this negative effect is likely to occur at the time of starting school when children are exposed to more complex cognitive demands.

The objectives of this study are: (1) to describe the epidemiology of below average cognitive ability at age three (2) to examine the relationship between early cognitive ability and emotional-behavioural function between the age of 3 and 9 years; and (3) to examine the independent association between early childhood

cognitive ability and clinically significant EBDs in childhood adjusting for relevant confounding factors.

3.3. Methods

Study Design and Setting

This study is based on secondary analysis of data from the Growing up in Ireland (GUI) Infant '08 Cohort, a nationally representative longitudinal cohort study of infants in Ireland, designed to inform government policy in relation to children, young people, and families.[21-25] The GUI dataset contains information collected from parents, teachers, school principals, and the children themselves. Information was collected on diverse aspects of the child's early life including birth characteristics, parental health, family income, dietary and exercise habits, as well as outcomes including physical health and development, social emotional and behavioral well-being, and educational achievement and intellectual ability.[22-25]

Participants

Families with infants born between December 2007 and June 2008 were identified from the Child Benefit Register, a register used to administer universal child benefit payments in Ireland. Potential participants were stratified according to marital status of claimant, county of residence, nationality, and number of children in the claim, before a systematic selection procedure with a random start and constant sampling fraction were applied. Full details of the GUI study are available elsewhere.[21] Data collection for Wave 1 began in September 2008 and recruited 11,134 nine-month-old infants and their primary caregivers, representing a 65% response rate. The participation and non-response rates for each wave of data collection are described

in Figure A8.1, Appendix 8. This study is based on data from 7,134 infants and their primary caregivers who completed Wave 1 (9 months), Wave 2 including cognitive assessment (age 3 years), Wave 3 (age 5 years) and Wave 5 (age 9 years). The study sample represents 67.4% of the original 11,134 participants in the cohort. Questionnaires were administered to the primary caregiver in their home using computer assisted personal interviewing.

Ethical Approval

Ethical approval for the GUI study was obtained from the Research Ethics Committee, Department of Children and Youth Affairs Ireland. Written consent was collected from parents/guardians prior to data collection. Secondary analysis of the anonymized microdata files of the GUI study does not require additional ethical approval. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines were followed in conducting and reporting this study.[26]

Measures

Cognitive Ability

Non-verbal cognitive ability was measured at age 3 using the Pictures Similarities Scale, a cognitive ability subtest from the British Abilities Scales (BAS) which measures non-verbal reasoning.[27] The BAS has demonstrated construct validity and reliability as a measure of intellectual ability.[28,29] In a time-restricted survey setting the BAS is particularly beneficial as each core sub-test is individually interpretable.[27] The test was administered by a trained interviewer at the home of the child. The child was shown four images and was given a card containing a single fifth image. The child must match the card to one of the four pictures based on a shared characteristic. Provided in the dataset were raw scores, ability scores, and standardised scores. The

standardised scores adjust for the performance of the child compared with other children of the same age.[27,29] The mean of the standardised score for the Picture Similarities Scale was 52.60 with a SD of 10.78. Those with scores more than 2 SD below the mean, consistent with a possible diagnosis of an intellectual disability, were excluded from the main analyses (*n* = 216), but a description of this cohort along with the results of models including this cohort are included in Appendix 8 (Figure A8.2, Figure A8.3, Table A8.1). Those with a score 1-2 SD below the mean were categorised as below average cognitive ability, and those with a score above this as average/above cognitive ability.

Emotional-behavioral function

Emotional-behavioral function was measured at age 3, age 5, and age 9 years using the parent-reported version of the Strengths and Difficulties Questionnaire (SDQ). The SDQ is comprised of four difficulties scales measuring emotional symptoms, conduct problems, hyperactivity/inattention, and peer relationship problems.

The parent-reported SDQ asks the parent to respond, using a 3-point Likert scale where 0=not true, 1=somewhat true, and 2=certainly true, to 20 statements about the child's behaviour over the last six months or school year. These 20 items make up the four difficulties subscales of 5 items each. The questionnaire provides a total difficulties score, ranging from 0-40, with a higher score indicating more difficulty.

Validation of the SDQ is well documented.[30,31]The total SDQ score was treated as both a continuous and categorical variable. In line with previous literature, scores above a 90th percentile cut-off were defined as "clinically significant" EBDs.[32]

This corresponded to a cut-off of ≥14 at age three, ≥14 at age five, and ≥16 at age nine.

Confounding Variables

Confounding variables were chosen a priori based on previously published literature demonstrating a relationship with both the predictor and outcome of interest. All were measured at the baseline survey at nine months by questionnaire completed by the primary caregiver. For 99.7% of participants the primary caregiver was the mother.

Child characteristics included sex (male, female), gestational age (weeks), birthweight (<2500g, 2500–4000g, >4000g), and admission to neonatal intensive care unit at birth (yes, no). Maternal characteristics included maternal age and maternal depression score (eight-item Center for Epidemiologic Studies Depression Scale). Sociodemographic characteristics included partner in household (yes, no); highest level of maternal education (lower secondary or less, upper secondary, sub degree (includes certificates and diplomas) and third level (primary degree and above); and income quintile (five equally sized groups each containing 20% of participants based on 'equivalized' disposable family income taking account of household size, composition, and disposable income after statutory deductions).

Where potential multicollinearity arose (maternal education & household income quintile; gestational age, birthweight, and admission to NICU) each variable was trialed in the model and chosen based on best model fit, assessed using the Bayesian Information Criterion (BIC).

Statistical Analysis

Statistical analyses were conducted using R Statistical Software Package.[33] A weighting variable calculated by the GUI study was applied to the population of 7,134 participants to adjust the structure of the sample to align with the population as a whole, allowing inferences to be made about the population from which participants were drawn.[34] Participants with complete data on the variables of interest were included in the analyses. Missing data was minimal and was ≤2% for all variables included in the linear and logistic analyses. Descriptive statistics were used to characterize the study population according to child, maternal, and sociodemographic characteristics. Pearson's chi-square tests and students t-tests were used to examine for significant differences in characteristics between cognitive ability groups. Correlations between the non-verbal cognitive ability score at age 3 years and total and subscale SDQ scores at age 3, 5, and 9 years were examined using Pearson's correlation coefficients (Table A8.2, Appendix 8)

The outcome of interest was total SDQ score at age 3, 5, and 9 years. Generalised linear mixed models, fitted using maximum likelihood, were used to examine the relationship between early cognitive ability and sequentially obtained total SDQ scores (main effect), and to examine the relationship between early cognitive ability and the change in SDQ between age 3 and 9 years (interaction effect).[35-38] Between participant variability in baseline SDQ was modelled using a random intercept and between participant variability in change of SDQ over time was modelled using a random slope.

The relationship was first examined using univariable models. Multivariable models were then built using a hierarchical approach, with the addition of infant characteristics, maternal characteristics, and sociodemographic characteristics.

Model fit was assessed following addition of each set of variables using the BIC, with lower values indicating better model fit.[38]

To evaluate the relationship between cognitive ability and the change in SDQ between age 3 and 9 years an interaction term between cognitive ability and age at SDQ measurement was investigated. Age was modelled as both a continuous and categorical variable, with the latter displaying better model fit. Each confounding variable was entered into an interaction term with age and if statistically significant on univariable analysis was then included in the multivariable model. Significant interaction terms with child, maternal, and sociodemographic characteristics were then added in blocks as before. Estimated marginal means with p-values and confidence intervals adjusted for multiple comparisons were used to examine the predicted mean emotional-behavioural function score for each cognitive ability group, adjusting for random effects and confounding. The independent association between early cognitive ability and the presence of a clinically significant EBD between age 3 and 9 years was examined using logistic regression.

3.4. Results

Description of the sample

There were 7,507 participants who completed the waves of interest in the GUI study. There were *n* = 157 participants with missing data on the Picture Similarities test at age 3 and *n* = 216 participants with a score more than 2 SD below the mean who were excluded from the main analysis. The population of 7,134 participants are described in Table 3.1 with respect to their cognitive ability at age 3 years. At age 3,

10.8% of the study population were categorised as below average cognitive ability. The remaining 89.2% were categorized as average/above.

Characteristics of study population

Mothers of children with below average cognitive ability at age 3 were more likely to report no partner present in household (21.4% v 14.7%, $p < 0.001$), and to experience higher mean depression scores (3.00 +/- 4.20 vs. 2.46 +/- 3.60, $p < 0.001$). There were statistically significant differences between the cognitive groups in relation to maternal education, income, and household social class. In the below average group 31.0% of mothers reported their highest level of education to be lower secondary or less compared with 17.4% in the average or above group. Similarly, 28.3% were in the lowest income quintile and 15.8% in the lowest social class group, compared to 19.0% and 8.7%, respectively, in the average/above group. Males comprised 61.8% of the below average cognitive ability group.

Total SDQ scores

The total SDQ scores for each cognitive ability group at age 3, 5, and 9 years are shown in Figure 3.1. There were statistically significant differences in mean total SDQ scores between cognitive ability groups at each age, with the below average cognitive group having a higher mean total SDQ at each time point (age 3 years 8.21 vs. 7.36 $p < 0.001$, age 5 years 7.80 vs. 6.88 $p < 0.001$, age 9 years 8.23 vs. 6.83 $p < 0.001$). To visualize the individual SDQ pathways, total SDQ scores at each age were divided into deciles with decile 1 representing the lowest 10% of scores and decile 10 the highest 10% of scores, and an alluvial plot, shown in Figure 3.2, was created for each cognitive ability group. At each age a higher proportion of the below average

cognitive ability group had total SDQ scores in the highest deciles, and a higher proportion remained in the highest deciles throughout childhood.

Cognitive ability as a predictor of longitudinal total SDQ scores

The predictive significance of early cognitive ability was examined using four models, described in Table 3.2. In Model 1, a univariable model, early cognitive ability was a significant predictor of higher longitudinal SDQ scores (BIC 120,291). Model 2 adjusts for child characteristics (BIC 120,231, $p < 0.001$). Model 3 adjusts for child and maternal characteristics (BIC of 119,628, $p < 0.001$), and Model 4 for child, maternal, and sociodemographic characteristics (BIC of 119,508, $p < 0.001$). Cognitive ability remained a significant predictor of higher longitudinal total SDQ scores after adjustment for child, maternal, and sociodemographic factors (β 0.62, 95% confidence interval 0.34–0.91). Parameter estimates for the fully adjusted Model 4 are summarized in Figure 3.3

Cognitive ability as a predictor of change in total SDQ score

Early cognitive ability was a significant predictor of the change in SDQ score between age 3 and 9 years (Table 3.3). The predicted change in total SDQ score between age 3 and 9 years for both groups, according to Model 5, with age treated as a categorical variable, is shown in Figure 3.4. The predicted change when age was treated as a continuous variable is shown in Figure A8.4, Appendix 8.

The estimated marginal means for Model 5 are shown in Table A8.3, Appendix 8. For those in the average/above cognitive ability group, the model predicted a mean decrease of 0.52 in total SDQ score between age 3-9 years (95% CI -0.67 to -0.37, p<0.001). For those in the below average group there was a predicted increase in total SDQ score between age 3 and 9 years, but this was not statistically

significant (mean difference 0.08, 95% CI -0.39 to 0.55, p=0.919). As shown in Figure 3.4, between age 3-5 years, both groups had a similar decrease in predicted SDQ score (average/above cognitive ability: mean difference -0.46, 95% CI -0.58 to -0.32, below average cognitive ability: mean difference -0.40, 95% CI -0.80 to -0.01). However, between age 5 and 9 years, the time coinciding with the start of school, there was a divergence between the groups. While the average/above group had a predicted decrease in total SDQ score (mean difference -0.06, 95% CI -0.19 to 0.07, p=0.488), the below average cognitive ability group had a significant increase in predicted score (mean difference 0.48, 95% CI 0.06 to 0.90, p=0.016).

Cognitive ability and clinically significant emotional behavioural difficulties

The proportion of those in the below average cognitive ability group who experienced a clinically significant EBD between age 3 and 9 years was 32.0% compared to 21.4% of those with average/above ability. The unadjusted and adjusted logistic regression models examining the relationship between below average cognitive ability and the presence of a clinically significant emotional-behavioral difficulty are described in Table 3.4. After adjusting for relevant confounding variables, compared to their peers with average/above ability, the odds of experiencing a clinically significant EBD between age 3-9 years was 1.4 times higher for those with below average cognitive ability (AOR 1.39 95% CI 1.17–1.66, $p < 0.001$).

3.5. Discussion

We have shown, in a large and well characterized prospective cohort study of more than 7,000 children, that having below average cognitive ability at age 3 years is a

significant predictor of worse emotional-behavioural function scores between age 3 and 9 years, and is also a significant predictor of having a clinically significant EBD between age 3 and 9 years. We have also shown that the trajectory of change in emotional-behavioral function between age 3 and 9 differs for those with and without below average cognitive ability.

The first objective of the study was to describe the epidemiology of below average cognitive ability in early childhood. At age three, 10.8% of children scored 1-2 SD below the mean on the Picture Similarities subtest of the British Ability Scales. There were no significant differences between cognitive groups in relation to potential biological risk factors including birthweight, gestational age, or admission to a neonatal intensive care unit at birth. The sociodemographic characteristics of the groups differed significantly, with the below average group significantly more likely to be characterized by low maternal education, low income, and single parent households. There is a substantial body of literature demonstrating the profound effect adverse socioeconomic environments have on infant and child neurodevelopment.[39-42] The exact mechanistic pathway is complex and likely mediated by multiple factors including prenatal influences, parental care, and cognitive stimulation.[42]

The second objective of the study was to model the change in emotional-behavioral function between age 3 and 9 years for those with and without below average cognitive ability. The a priori hypothesis was that children with below average cognitive ability would experience a different trajectory of change in emotional-behavioral function between age 3 and 9 years compared to their counterparts with average/above cognitive function. We postulated that at the time

of commencing formal education, which typically occurs at age five in Ireland, those with below average cognitive ability would experience increasing difficulties in emotional-behavioral function as they struggle to cope with increasingly complex cognitive demands, and may begin to experience failure at school. The results support this hypothesis. Unlike their counterparts who experienced an overall decrease in mean SDQ score between age 3 and 9 years, those with below average cognitive ability experienced a mean increase in total SDQ score, with the vast majority of this increase occurring between the age of 5 and 9 years. The adjusted multivariate logistic regression model confirmed that below average cognitive ability was associated with increased odds of having a clinically significant EBD between the age of 3 and 9 years.

There is limited published literature examining this specific cohort of children with cognitive ability between 1-2 SD below the mean, but the findings of our study appear consistent with what has been published to date. To our knowledge there is one published population-based study which has focused specifically on childhood mental health and borderline intellectual functioning, a level of intellectual function defined by an IQ between 70–85, a construct similar to that of below average cognitive ability used in this study. This 2010 study, which was based on more than 4,300 participants in the Longitudinal Study of Australian Children (LSAC), reported the prevalence of possible mental health problems at age 6/7 years among children with borderline intellectual function was 17% compared to 5% among typically developing peers. After controlling for potential confounding due to socioeconomic disadvantage using propensity score matching, significant differences in prevalence

remained between those with limited intellectual ability and those typically developing.[19]

A 2019 study using the same population cohort but at later waves of the study examined this association in the adolescent period and reported children with borderline intellectual functioning at age 8/9 years had significantly higher odds of having abnormal emotional-behavioral difficulties scores at age 14/15 years after adjustment for gender and a range of socioeconomic variables (adjusted odds ratio 4.33, 95% confidence interval 2.84–6.62).[18] Neither study examined the change in emotional-behavioral function over time for these groups.

Disentangling the relationship between cognitive ability and emotional-behavioral development is complex. Firstly, there is overlap between the constructs making it difficult to sharply distinguish between certain elements of both abilities. For example, attention is a component of both the broad construct of cognitive function, but also of behaviour.[31,43] Secondly, there are many confounding variables in the relationship between cognitive ability and emotional-behavioral development, as risk factors such as childhood adversity and markers of lower socioeconomic status predispose to both and it is therefore likely residual confounding exists in this relationship.[40,41,44]

Acknowledging these challenges in establishing a causal relationship between cognitive ability and emotional-behavioral development, the results of this study suggest a worsening in emotional-behavioral function for those with below average cognitive ability between the age of 5 and 9 years, a period coinciding with increasingly complex cognitive demands, with peer comparisons in the school setting, and with potential for school failure.

The findings of this study are subject to limitations, perhaps the most significant being the use of a single performance metric as a measure of cognitive ability in childhood. Use of this single measure, while accepted as being individually interpretable within the British Ability Scales, does not account for alternative, yet equally important abilities, such as musical or artistic abilities, nor does it account for a child's adaptive functioning in the presence of lower cognitive ability. Additionally, cognitive ability is not a static construct and may change over time. Further research using methods to account for this change may help to further decipher the temporal relationship between cognitive ability and emotional-behavioral function.

A further limitation of the study is that it is based on a longitudinal survey and is therefore subject to selection bias. Those who participate in population-based studies, particularly those who remain engaged in longitudinal studies, are likely to systematically differ from those who do not across key socio-demographic domains. To account for this, to the extent possible, the weighting variable calculated by the GUI study, to ensure the longitudinal cohort were representative of the population from which they were drawn, was applied.[34]

3.6.Conclusion

The relationship between early cognitive ability and emotional-behavioral function in childhood has not to-date been extensively investigated. While there are a number of systematic reviews and meta-analyses reporting a disproportionately high prevalence of mental health problems among children with intellectual developmental disabilities, no such reviews exist for the larger proportion of children with cognitive ability between 1-2 SD below the mean.[45]

In this study we examined the relationship between early cognitive ability and emotional-behavioral development using a variety of statistical models. We found that below average cognitive ability was associated with higher longitudinal SDQ scores in childhood, and that children with early below average cognitive ability are at higher risk of clinically significant emotional-behavioral difficulties in childhood. We also demonstrated that the change in emotional-behavioral function between age 3 and 9 years differs according to cognitive ability and that for those with below average cognitive ability, the period between the age of 5 and 9 years signals a worsening in emotional-behavioral function.

To our knowledge, this is the first study to examine in a longitudinal manner the change in mean SDQ scores over time for such a cohort of children. The findings of this study are suggestive of a temporal relationship between commencing formal education and worsening emotional behavioral function for this cohort of children. It is entirely plausible that this period of increasingly complex cognitive and social demand, during which one is routinely compared to typically developing peers, may culminate in the manifestation of emotional or behavioral problems. For many, it is only at this stage they will begin the process of assessment and support. Wieland and Zitman have outlined why taking this approach of waiting for adaptive difficulties to develop before recognising or intervening is problematic, and how for many it results in "problematic lives, functioning under high strain but unnoticed by the rest of society".[8]

This research raises the question of how early childhood policies can be adapted to identify high risk children and provide effective early intervention to address both cognitive and emotional-behavioral needs. The early identification of

children at risk of below average cognitive ability remains an important area for further research. The risk factors for a child failing to meet their full cognitive potential have been extensively researched and appear rooted in the social determinants of health. Factors such as parental education, family income, and socioeconomic status have shown consistent and strong associations with low cognitive ability. However, routinely this information is not utilised in early life to identify high risk children.

In addition, there is a paucity of research examining interventions to improve emotional-behavioral and cognitive outcomes specifically for those with below average cognitive ability, who may differ in many ways from those with more severe intellectual disabilities. The key question which remains, is how as a society, health, and social care system, we can recognize and support these children to help them to reach their full potential.

Author contributions

AB – Conceptualization, methodology, formal analysis, writing – original draft preparation, writing – review and editing

AS – Methodology, formal analysis, writing – review and editing

DM – Conceptualization, methodology, writing – review and editing

Table 3.5 Characteristics of study population according to cognitive ability at age 3 years

			Average/above cognitive ability n =6,362	Below average cognitive ability n=772	
	Valid	Total	n (%)	n (%)	p
Maternal					
Age[a], mean (SD), y	7,134	7,134	31.6 (5.3)	31.0 (5.7)	0.002[c]
No partner in household	7,134	1,099	934 (14.7)	165 (21.4)	<0.001[d]
Smoker - yes	7,134	1,917	1642 (25.8)	275 (35.6)	<0.001[d]
Depression, mean (SD)	7,030	2.3 (3.4)	2.5 (3.6)	3.0 (4.2)	<0.001[d]
Sociodemographic					
Maternal education	7,132				
Lower secondary		1,343	1104 (17.4)	239 (31.0)	
Upper secondary		1,758	1547 (24.3)	211 (27.3)	
Sub degree		2,053	1877 (29.5)	176 (22.8)	
Third level and higher		1,978	1832 (28.8)	146 (18.9)	<0.001[d]
Income Quintile	6,576				
Lowest		1,315	1113 (19.0)	202 (28.3)	
2nd		1,314	1150 (19.6)	164 (22.9)	
3rd		1,344	1195 (20.4)	149 (20.8)	
4th		1,398	1286 (21.9)	112 (15.7)	
Highest		1,205	1117 (19.1)	88 (12.3)	<0.001[d]
Household social class	6,431				
Professional/managerial		3,391	3107 (49.0)	284 (37.1)	
Non-manual/Skilled Manual		2,364	2108 (33.3)	256 (33.4)	
Semi-skilled/Unskilled Manual		676	571 (9.0)	105 (13.7)	
Never worked—no class defined		672	551 (8.7)	121 (15.8)	<0.001[d]
Infant					
Male gender	7,134	3,617	3140 (49.4)	477 (61.8)	<0.001[d]
Birth weight	7,062				
<2500 g		408	354 (5.6)	54 (7.0)	
2500–4000 g		5,655	5049 (80.2)	606 (79.0)	
>4000 g		999	892 (14.2)	107 (14.0)	0.283[d]
Gestational age, mean (SD), wk[b]	7,119	7,119	39.5 (2.1)	39.5 (2.3)	0.400[c]
Admitted to NICU at birth	7,127	976	857 (13.5)	119 (15.4)	0.141[d]

[a]Maternal age includes *n*=532 participants aged 40 years or above who were coded as 40 in original dataset, [b]Gestational age includes *n*=10 participants with a gestational age of 25 weeks or less which was recorded as 25, [c]Welch Two Sample t-test, [d]Pearson's Chi-squared test.

Table 3.6 Results of generalised linear mixed model examining cognitive ability as a predictor of longitudinal SDQ scores

	Model 1 Main effect cognitive ability	Model 2 Adjusted for child characteristics	Model 3 Adjusted for child and maternal characteristics	Model 4 Adjusted for child, maternal, and sociodemographic characteristics
	β Estimate (95% Confidence Interval)			
Fixed Effects				
Intercept	7.07 (6.98–7.17)	9.76 (8.04–11.47)	14.34 (12.57–16.12)	13.41 (11.62–15.19)
Below average cognitive ability	1.01 (0.70–1.31)	0.91 (0.61–1.21)	0.78 (0.49–1.07)	0.62 (0.34–0.91)
Infant Gender: Female		−0.77 (−0.94–−0.59)	−0.77 (−0.94–−0.60)	−0.79 (−0.96–−0.63)
Gestational age (weeks)		−0.06 (−0.10–−0.01)	−0.07 (−0.11–−0.03)	−0.06 (−0.10–−0.02)
Maternal age			−0.14 (−0.16–−0.13)	−0.10 (−0.12–−0.09)
Maternal depression score			0.23 (0.20–0.25)	0.21 (0.18–0.23)
Partner present				−1.09 (−1.42–−0.76)
Education: Non Degree				0.52 (0.32–0.72)
Education: Upper Secondary				0.67 (0.42–0.92)
Education: Lower Secondary				1.50 (1.18–−1.83)
Random Effects				
σ^2	1.67	1.68	1.64	1.56
$\tau_{00\ ID}$	17.11	17.03	16.01	15.88
$\tau_{11\ ID.Age5}$	15.10	15.09	15.16	15.29
$\tau_{11\ ID.Age9}$	22.79	22.77	22.84	22.99
ρ_{01}	-0.40	-0.40	-0.42	-0.43
	-0.38	-0.39	-0.41	-0.42
ICC	0.91	0.91	0.91	0.91

Participant number	7079	7079	7079	7079
Observations	21,237	21,237	21,237	21,237
Marginal/Conditional R2	0.005/0.911	0.013/0.911	0.077/0.914	0.092/0.919
BIC	120291.3	120231.4	119628.6	119508.3
p-value		<0.001	<0.001	<0.001

BIC, Bayesian Information Criterion; σ^2, Residual variance; $\tau_{00\ ID}$, Random intercept variance; $\tau_{11\ ID.Age5}$, Random slope variance; $\tau_{11\ ID.Age9,}$ Random slope variance; $\rho_{01,}$ Random intercept/slope correlation; ICC, Intraclass correlation coefficient.

Table 3.7 Results of generalised linear mixed model examining cognitive ability as a predictor of change in total SDQ score

	Model 5 Interaction effect Cognitive ability x Age	Model 6 Adjusted for child characteristics	Model 7 Adjusted for child and maternal characteristics	Model 8 Adjusted for child, maternal, and sociodemographic characteristics
	Estimate (95% Confidence Interval)			
Fixed effects				
Intercept	7.32 (7.21 – 7.42)	8.84 (6.88 – 10.80)	7.97 (6.03 – 9.90)	7.86 (7.44 – 8.28)
Below average cognitive ability	0.87 (0.52 - 1.21)	0.80 (0.45 - 1.14)	0.71 (0.37 – 1.05)	0.52 (0.18 – 0.85)
Age 5	-0.46 (-0.56 – -0.35)	1.39 (-0.54 – 3.32)	1.28 (-0.66 – 3.21)	0.12 (-0.31 – 0.55)
Age 9	-0.52 (-0.64 – -0.39)	2.32 (0.01 – 4.62)	2.11 (-0.19 – 4.42)	0.34 (-0.17 – 0.84)
Below average cognitive ability * Age 5	0.05 (-0.29 – 0.39)	0.02 (-0.32 – 0.36)	0.01 (-0.33 – 0.35)	0.01 (-0.33 – 0.35)
Below average cognitive ability * Age 9	0.60 (0.19 – 1.00)	0.52 (0.11 – 0.93)	0.50 (0.10 – 0.91)	0.47 (0.06 – 0.87)
Random effects				
σ^2	1.64	1.64	1.62	1.50
$\tau_{00\ ID}$	17.12	17.04	16.47	16.09
$\tau_{11\ ID.Age5}$	14.96	14.94	14.98	15.13
$\tau_{11\ ID.Age9}$	22.60	22.51	22.53	22.65
ρ_{01}	-0.40	-0.40	-0.41	-0.42
	-0.38	-0.38	-0.40	-0.41
ICC	0.93	0.93	0.92	0.93
Participants	7,079	7,079	7,079	7,079
Observations	21,237	21,237	21,237	21,237
Marginal/ Conditional R2	0.007 / 0.926	0.016 / 0.926	0.051 / 0.928	0.079 / 0.933
BIC	120230.1	120182.6	119841.0	119616.9

p-value	<0.001	<0.001	<0.001

Model 6 adjusted for interaction terms including child gender and gestational age
Model 7 adjusted for interaction terms including child gender, gestational age, maternal depression
Model 8 adjusted for interaction terms including child gender, gestational age, maternal depression, presence of a partner, maternal education
BIC, Bayesian Information Criterion; σ^2, Residual variance; $\tau_{00\ ID}$, Random intercept variance; $\tau_{11\ ID.Age5}$, Random slope variance; $\tau_{11\ ID.Age9}$, Random slope variance; ρ_{01}, Random intercept/slope correlation; ICC, Intraclass correlation coefficient.

Table 3.8 Logistic regression model examining association between cognitive ability at age 3 and clinically significant emotional-behavioural difficulties between age 3 and 9 years

	EBD		Unadjusted Models		Adjusted Model[a]		
	No n=5,425	Yes n=1,577	Crude OR	95% CI	AOR	95% CI	*p*-Value
Cognitive Ability							
Average and above	4914 (90.6)	1,337 (84.8)	Ref.		Ref.		
Below Average	511 (9.4)	240 (15.2)	1.72	1.46–2.03	1.39	1.17–1.66	<0.001
Gender:							
Male	2658 (49.0)	900 (57.1)	Ref.		Ref.		
Female	2767 (51.0)	677 (42.9)	0.69	0.62–0.77	0.70	0.62–0.78	<0.001
Depression, n, mean (sd)	5425, 2.13 (3.26)	1577, 3.81 (4.58)	1.11	1.10–1.13	1.09	1.07–1.10	<0.001
Education							
Degree or higher	1682 (31.0)	277 (17.6)	Ref.		Ref.		
Non-degree	1638 (30.2)	382 (24.2)	1.43	1.22-1.69	1.22	1.03-1.45	0.023
Upper secondary	1275 (23.5)	456 (28.9)	2.20	1.88-2.58	1.66	1.40-1.98	<0.001
Lower secondary or less	830 (15.3)	462 (29.3)	3.28	2.79-3.86	2.24	1.87-2.69	<0.001
Gestational age n, mean (sd)	5425, 39.57 (2.00)	1577, 39.37 (2.40)	0.96	0.94–0.98	0.96	0.94–0.99	0.004
Partner							
No	624 (11.5)	454 (28.8)	Ref.		Ref.		
Yes	4801 (88.5)	1123 (71.2)	0.33	0.29–0.38	0.54	0.46–0.64	<0.001
Maternal Age Category							
≤24 years	499 (9.2)	352 (22.3)	Ref.		Ref.		
25–34 years	3081 (56.8)	826 (52.4)	0.37	0.32–0.43	0.64	0.53–0.76	<0.001
≥35 years	1844 (34.0)	399 (25.3)	0.30	0.26–0.36	0.54	0.45–0.66	<0.001

Abbreviations: EBD, Emotional-behavioural difficulty; OR, odds ratio; CI, confidence interval; AOR, adjusted odds ratio. [a]Adjusted for partner, highest level education, maternal age category, maternal depression score, gender of child, and gestational age in weeks

Figure 3.5 Distribution of total SDQ scores at age 3, 5, and 9 years according to cognitive ability

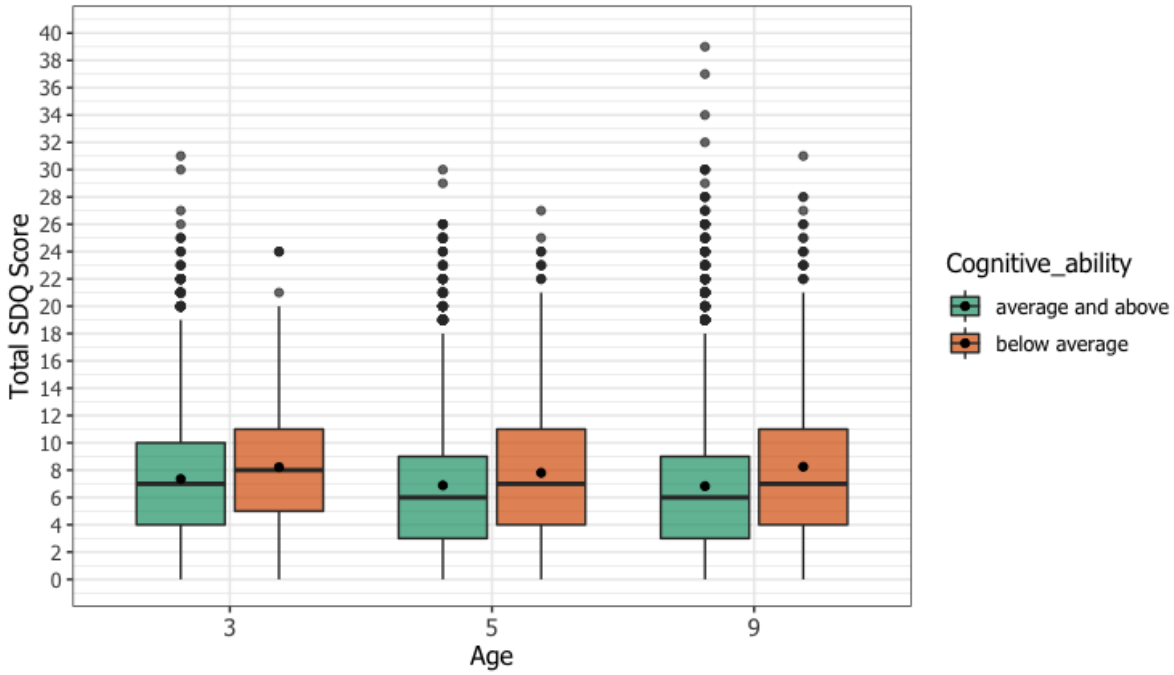


Figure 3.1 Legend

The distribution of total SDQ scores at each age is shown in a boxplot. Each box represents the 25th to 75th percentile. The length of the box is the interquartile range (IQR). The horizontal line inside the box represents the median. The whiskers represent the lowest and highest quartiles. The dot inside the box represents the weighted mean. Circles represent individual outliers. At each age, those with below average cognitive ability had higher mean and median total SDQ scores than those with average/above cognitive ability.

Figure 3.6 Alluvial plots of SDQ decile scores at age 3, 5, and 9 years according to cognitive ability

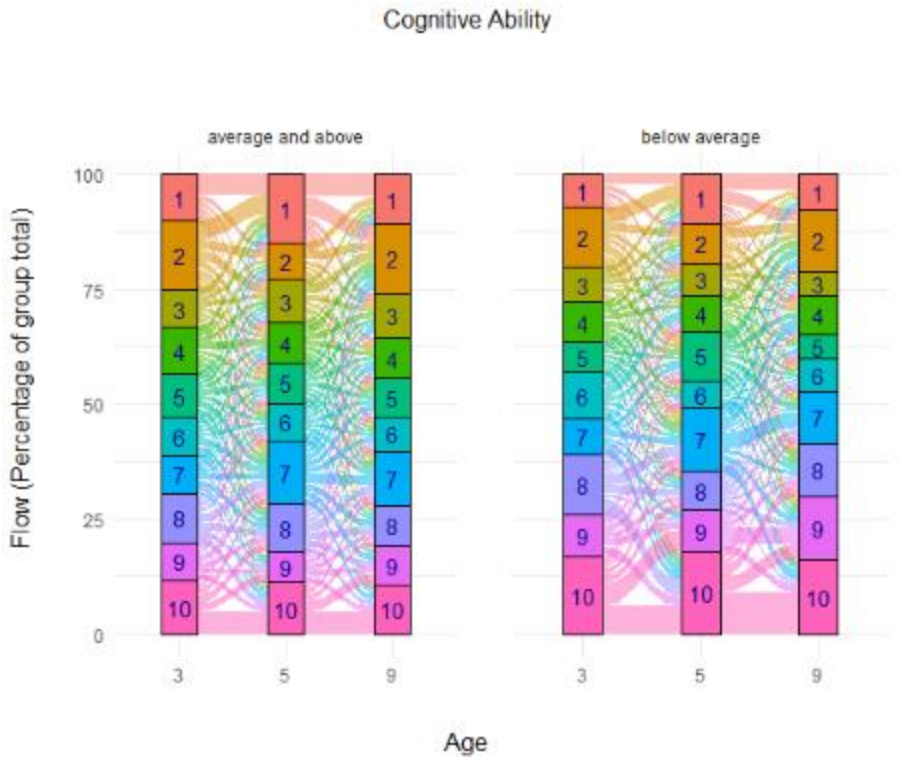


Figure 3.2 Legend

Total SDQ Scores were split into deciles with decile 10 representing the highest 10% of total SDQ scores at that time point and decile 1 representing the lowest 10% of total SDQ scores at that time point. Each flow represents the number of participants in that flow, with the width of the flow proportional to the number of participants. At age 3, 5, and 9 years a larger proportion of those in the below average cognitive ability group had SDQ scores in the highest two deciles.

Figure 3.7 Parameter estimates for fully adjusted model predicting longitudinal total SDQ scores (Model 4)

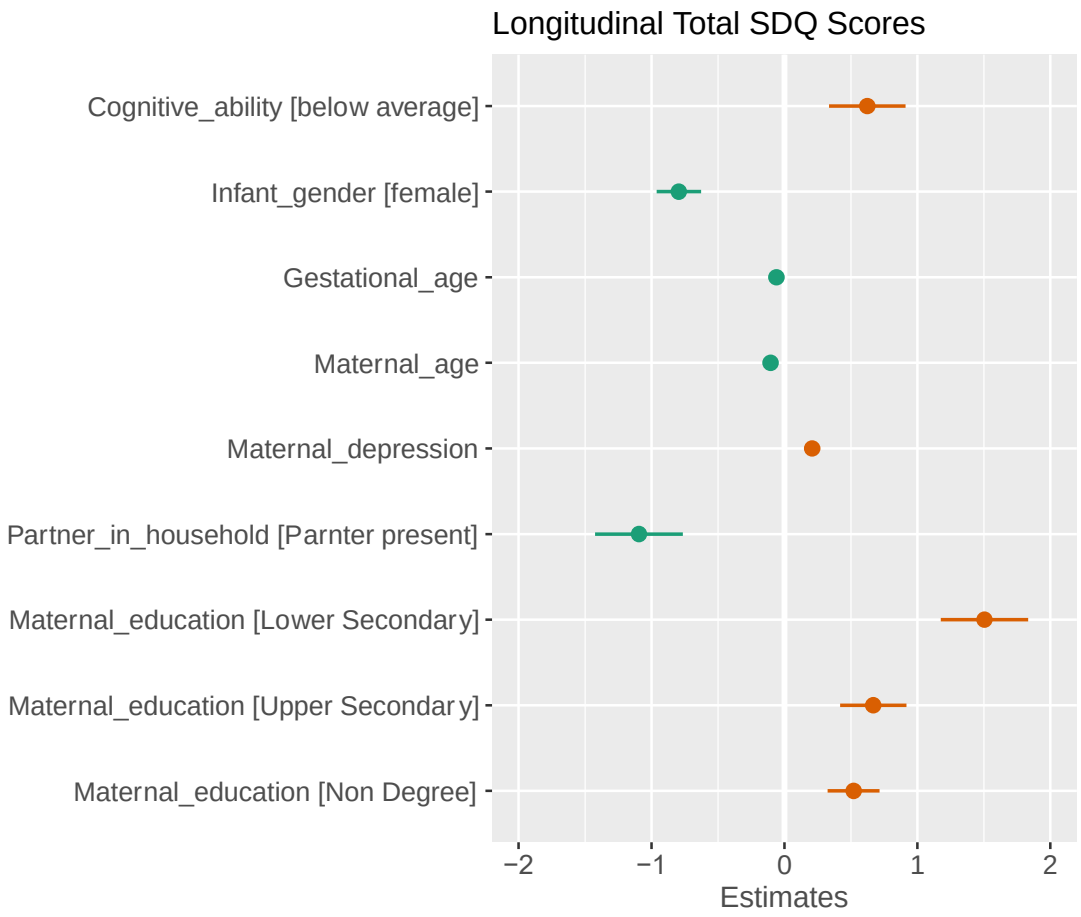


Figure 3.3 Legend

This figure demonstrates the parameter estimates for the variable of interest, cognitive ability, and the confounding variables included in the fully adjusted Model 4. Categorical variables are represented by circles in the centre of a line. The circles represent the point estimates and the lines represent 95% confidence intervals. Continuous variables are represented by a point only. Variables in orange are predictors of higher longitudinal total SDQ scores. Variables in green are predictors of lower longitudinal total SDQ scores. Reference group for cognitive ability – average/above. Reference group for gender – male. Reference group for partner in household – no partner present. Reference group for maternal education – degree or higher.

Figure 3.8 Predicted change in total SDQ score between age 3 and 9 years according to cognitive ability group (Model 5)

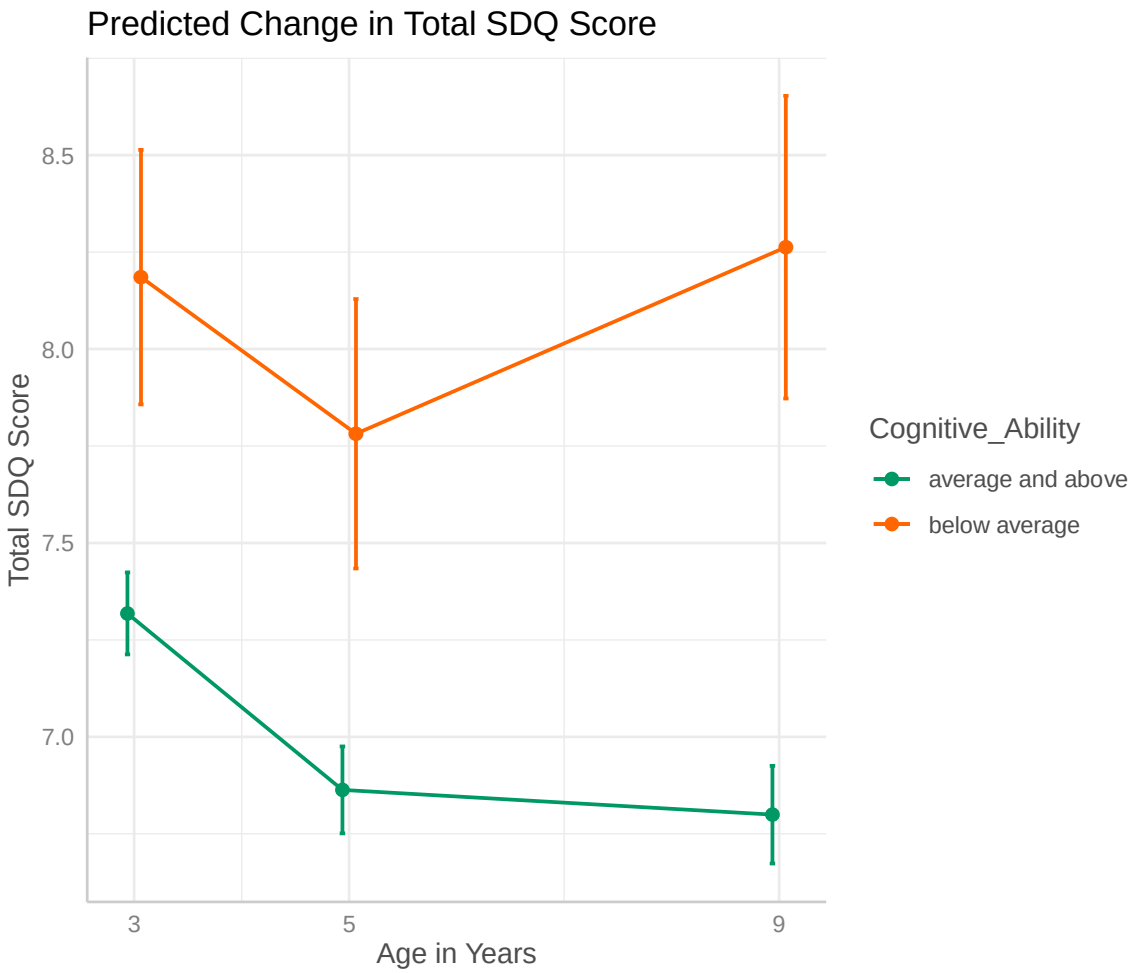


Figure 3.4 Legend

The predicted change in SDQ total score between age 3, 5, and 9 years is shown for those with below average cognitive ability and those with average/above ability. At each age point, those with below average cognitive ability had a significantly higher total SDQ score. Between age 5 and 9 years, there is a non-significant decrease in total SDQ score for those in the average/above group (estimated mean difference -0.06, p=0.488), whereas there is a statistically significant increase in total SDQ for those in the below average group (estimated mean difference 0.48, p=0.016). Between the age of 3-9 years, those in the average cognitive ability group had a statistically significant decrease in mean total SDQ score (-0.52, p<0.001) while those in the below average group had a non-significant increase (0.08, p=0.919).

3.7. References

1. Nouwens PJG, Lucas R, Embregts PJCM, van Nieuwenhuizen C. In plain sight but still invisible: A structured case analysis of people with mild intellectual disability or borderline intellectual functioning. J Intellect Dev Disabil. 2017;42(1):36–44. https://doi.org/10.3109/13668250.2016.1178220
2. Greenspan S. Borderline intellectual functioning: an update. Curr Opin Psychiatry. 2017 Mar;30(2):113-122. doi: 10.1097/YCO.0000000000000317.
3. World Health Organisation. ICD-10 for Mortality and Morbidity Statistics: Chapter V Mental and Behavioural Disorders, Mental retardation F70-F79.[Internet] 2019 [cited 2021 Feb 27]. Available at: https://icd.who.int/browse10/2019/en#/F70-F79
4. Deary IJ, Strand S, Smith P, Fernandes C. Intelligence and educational achievement. Intelligence. 2007;35(1):13–21. https://doi.org/10.1016/j.intell.2006.02.001
5. Wraw C, Deary IJ, Gale CR, Der G. Intelligence in youth and health at age 50. Intelligence. 2015 Nov-Dec;53:23-32. doi: 10.1016/j.intell.2015.08.001.
6. Wraw C, Deary IJ, Der G, Gale CR. Intelligence in youth and mental health at age 50. Intelligence. 2016 Sep-Oct;58:69-79. doi: 10.1016/j.intell.2016.06.005.
7. Batty GD, Mortensen EL, Osler M. Childhood IQ in relation to later psychiatric disorder: evidence from a Danish birth cohort study. Br J Psychiatry. 2005 Aug;187:180-1. doi: 10.1192/bjp.187.2.180. PMID: 16055831
8. Wieland J, Zitman FG. It is time to bring borderline intellectual functioning back into the main fold of classification systems. BJPsych Bull. 2016

Aug;40(4):204-6. doi: 10.1192/pb.bp.115.051490. PMID: 27512590; PMCID: PMC4967780.

9. Peltopuro M, Ahonen T, Kaartinen J, Seppälä H, Närhi V. Borderline intellectual functioning: a systematic literature review. Intellect Dev Disabil. 2014 Dec;52(6):419-43. doi: 10.1352/1934-9556-52.6.419.

10. Karande S, Kanchan S, Kulkarni M. Clinical and psychoeducational profile of children with borderline intellectual functioning. Indian J Pediatr. 2008 Aug;75(8):795-800. doi: 10.1007/s12098-008-0101-y.

11. Wylie C, Thompson J. The Long-term Contribution of Early Childhood Education to Children's Performance—Evidence from New Zealand. Int J Early Years Educ. 2003;11(1): 69–78. https://doi.org/10.1080/09669760304709

12. Kagan SL, Neuman MJ. Lessons from Three Decades of Transition Research. Elem. Sch. J. 1998;98(4):365–379. https://doi.org/10.1086/461902

13. Baxter C, Cummins RA, Polak S. A Longitudinal Study of Parental Stress and Support: From Diagnosis of Disability to Leaving School. Int J Disabil Dev Educ. 1995;42(2):125–136. doi: 10.1080/0156655950420204

14. Barker DJ, Gluckman PD, Godfrey KM, Harding JE, Owens JA, Robinson JS. Fetal nutrition and cardiovascular disease in adult life. Lancet. 1993 Apr 10;341(8850):938-41. doi: 10.1016/0140-6736(93)91224-a. PMID: 8096277.

15. Wadhwa PD, Buss C, Entringer S, Swanson JM. Developmental origins of health and disease: brief history of the approach and current focus on epigenetic mechanisms. Semin Reprod Med. 2009 Sep;27(5):358-68. doi: 10.1055/s-0029-1237424.

16. Cioni G, Inguaggiato E, Sgandurra G. Early intervention in neurodevelopmental disorders: underlying neural mechanisms. Dev Med Child Neurol. 2016 Mar;58 Suppl 4:61-6. doi: 10.1111/dmcn.13050. PMID: 27027609.

17. Pulina F, Lanfranchi S, Henry L, Vianello R. Intellectual profile in school-aged children with borderline intellectual functioning. Res Dev Disabil. 2019 Dec;95:103498. doi: 10.1016/j.ridd.2019.103498.

18. King TL, Milner A, Aitken Z, Karahalios A, Emerson E, Kavanagh AM. Mental health of adolescents: variations by borderline intellectual functioning and disability. Eur Child Adolesc Psychiatry. 2019 Sep;28(9):1231-1240. doi: 10.1007/s00787-019-01278-9.

19. Emerson E, Einfeld S, Stancliffe RJ. The mental health of young children with intellectual disabilities or borderline intellectual functioning. Soc Psychiatry Psychiatr Epidemiol. 2010 May;45(5):579-87. doi: 10.1007/s00127-009-0100-y. Epub 2009 Jul 19. PMID: 19618095.

20. Alesi M, Rappo G, Pepi A. Emotional Profile and Intellectual Functioning: A Comparison Among Children With Borderline Intellectual Functioning, Average Intellectual Functioning, and Gifted Intellectual Functioning. Sage Open. 2015 Jul 1;5(3) https://doi.org/10.1177/2158244015589995

21. Quail A, Williams J, McCrory C, Murray A, Thornton M. A Summary Guide to Wave 1 of the Infant Cohort of Growing Up in Ireland [Internet]. Dublin; 2011 [cited 2021 Mar 18] Available at: https://www.growingup.gov.ie/pubs/Summary-Guide_Infant-Cohort_Wave-1.pdf

22. Central Statistics Office (CSO). Growing up In Ireland Cohort ’08 (Infant Cohort) Wave 1 - 9 months, 2008. [dataset]. 2011. Version 1. Irish Social Science Data Archive. SN: 0019-01. URL http://www.ucd.ie/issda/data/GUIInfant/GUIInfantWave1

23. Central Statistics Office (CSO). Growing up In Ireland Cohort ’08 (Infant Cohort) Wave 2 - 3 years, 2011. [dataset]. 2013. Version 1. Irish Social Science Data Archive. SN: 0019-02. URL http://www.ucd.ie/issda/data/GUIInfant/GUIInfantWave2

24. Central Statistics Office (CSO). Growing up In Ireland Cohort ’08 (Infant Cohort) Wave 3 - 5 years, 2013. [dataset]. 2015. Version 1. Irish Social Science Data Archive. SN: 0019-03. URL http://www.ucd.ie/issda/data/GUIInfant/GUIInfantWave3

25. Central Statistics Office (CSO). Growing up in Ireland Cohort '08 (Infant Cohort) Wave 5 at 9 years, 2017/2018. [dataset]. 2024. Version 2. Irish Social Science Data Archive. SN: 0019-05. URL www.ucd.ie/issda/data/guiinfant/guiinfantwave5

26. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ. 2007 Oct 20;335(7624):806-8. doi: 10.1136/bmj.39335.541782.AD.

27. Elliot C, Smith P, McCullock K. British Ability Scales Second Edition (BAS II) Technical Manual. London: Nelson; 1997

28. Buckhalt JA, Denes GE, Stratton SP. Validity of the British Ability Scales Short-Form for a Sample of US Students. Sch Psychol Int. 1989;10(3):185–191. https://doi.org/10.1177/0143034389103003

29. Connelly R. Interpreting Test Scores Millennium Cohort Study Data Note: Interpreting Test Scores [Internet]. London; 2013. [cited 2021 Mar 18] Available at: https://cls.ucl.ac.uk/wp-content/uploads/2017/07/MCS-data-note-20131-Test-Scores-Roxanne-Connelly.pdf

30. Goodman A, Goodman R. Population mean scores predict child mental disorder rates: validating SDQ prevalence estimators in Britain. J Child Psychol Psychiatry. 2011 Jan;52(1):100-8. doi: 10.1111/j.1469-7610.2010.02278.x.

31. Stone LL, Janssens JM, Vermulst AA, Van Der Maten M, Engels RC, Otten R. The Strengths and Difficulties Questionnaire: psychometric properties of the parent and teacher version in children aged 4-7. BMC Psychol. 2015 Feb 20;3(1):4. doi: 10.1186/s40359-015-0061-8.

32. Spang KS, Thorup AAE, Ellersgaard D, Hemager N, Christiani C, Burton BK et al. The Strengths and Difficulties Questionnaire Is of Clinical Significance Regarding Emotional and Behavioral Problems in 7-Year-Old Children With Familial Risk of Schizophrenia or Bipolar Disorder and Population-Based Controls the Danish High Risk and Resilience Study-VIA 7; A Population-Based Cohort Study. Front Psychiatry. 2022 May 25;13:861219. doi: 10.3389/fpsyt.2022.861219. PMID: 35693960; PMCID: PMC9174569.

33. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna; 2021.

34. Quail A, O'Reilly C, Watson D, McNamara E, O'Mahony D, Murray A. Growing Up in Ireland A Summary Guide to Cohort '08 at 9 Years [Internet]. Dublin; f2019 [cited 2023 Aug 01]. Available at: https://www.growingup.gov.ie/pubs/Summary-Guide-Infant-Cohort-Wave-5.pdf

35. Bates D, Maechler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. J Stat Softw. 2015;67(1):1–48. https://doi.org/10.18637/jss.v067.i01

36. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. nlme: Linear and Nonlinear Mixed Effects Models. 2021. R package version 3.1-152. URL: https://CRAN.R-project.org/package=nlme>.

37. Dunn PK, Smyth GK. Generalized Linear Models with Examples in R [Internet]. New York: Springer; 2018 [cited 2023 Aug 6]. Available at: https://link.springer.com/book/10.1007/978-1-4419-0118-7

38. Assassi S, Leyva AL, Mayes MD, Sharif R, Nair DK, Fischbach M, Nguyen N, Reveille JD, Gonzalez EB, McNearney TA; GENISOS Study Group. Predictors of fatigue severity in early systemic sclerosis: a prospective longitudinal study of the GENISOS cohort. PLoS One. 2011;6(10):e26061. doi: 10.1371/journal.pone.0026061. Epub 2011 Oct 14. PMID: 22022507; PMCID: PMC3193535.

39. Huang J, Zhu T, Qu Y, Mu D. Prenatal, Perinatal and Neonatal Risk Factors for Intellectual Disability: A Systemic Review and Meta-Analysis. PLoS One. 2016 Apr 25;11(4):e0153655. doi: 10.1371/journal.pone.0153655.

40. Lawlor DA, Najman JM, Batty GD, O'Callaghan MJ, Williams GM, Bor W. Early life predictors of childhood intelligence: findings from the Mater-University study of pregnancy and its outcomes. Paediatr Perinat Epidemiol. 2006 Mar;20(2):148-62. doi: 10.1111/j.1365-3016.2006.00704.x.

41. Chin-Lun Hung G, Hahn J, Alamiri B, Buka SL, Goldstein JM, Laird N, et al. Socioeconomic disadvantage and neural development from infancy through early childhood. Int J Epidemiol. 2015 Dec;44(6):1889-99. doi: 10.1093/ije/dyv303.

42. Hackman DA, Farah MJ, Meaney MJ. Socioeconomic status and the brain: mechanistic insights from human and animal research. Nat Rev Neurosci. 2010 Sep;11(9):651-9. doi: 10.1038/nrn2897.

43. Dick BD, Pillai Riddell R. Cognitive and school functioning in children and adolescents with chronic pain: a critical review. Pain Res Manag. 2010 Jul-Aug;15(4):238-44. doi: 10.1155/2010/354812.

44. Bradley RH, Corwyn RF. Socioeconomic status and child *development.* Annual review of psychology. 2002 Feb; 53:371-99. https://doi.org/10.1146/annurev.psych.53.100901.135233

45. Buckley N, Glasson EJ, Chen W, Epstein A, Leonard H, Skoss R et al. Prevalence estimates of mental health problems in children and adolescents with intellectual disability: A systematic review and meta-analysis. Aust N Z J Psychiatry. 2020 Oct;54(10):970-984. doi: 10.1177/0004867420924101. Epub 2020 May 30. PMID: 32475125.

Chapter 4 The Relationship Between Below Average Cognitive Ability at Age 5 Years and the Child's Experience of School at Age 9

Andrea K. Bowe[1], Mathias Urban[2], Anthony Staines[3], Deirdre M. Murray[1,4]

[1] INFANT Research Centre, University College Cork, Ireland

[2] School of Language, Literacy & Early Childhood Education, Dublin City University, Dublin, Ireland

[3] School of Nursing, Psychotherapy, and Community Health, Dublin City University, Dublin, Ireland

[4] Department of Paediatrics, School of Medicine, University College Cork, Cork, Ireland.

Published in Frontiers in Public Health

March 2025

4.1.Abstract

Background

At age 5, while only embarking on their educational journey, substantial differences in children's cognitive ability will already exist. The aim of this study was to examine the causal association between below average cognitive ability at age 5 years and child-reported experience of school and self-concept, and teacher-reported class engagement and emotional-behavioural function at age 9 years.

Methods

This longitudinal cohort study used data from 7,392 children in the Growing Up in Ireland Infant Cohort, who had completed the Picture Similarities and Naming Vocabulary subtests of the British Abilities Scales at age 5. Principal components analysis was used to produce a composite general cognitive ability score for each child. Children with a general cognitive ability score more than 1 standard deviation (SD) below the mean at age 5 were categorised as "Below Average Cognitive Ability" (BACA), and those scoring above this as "Typical Cognitive Development" (TCD). The outcomes of interest, measured at age 9, were child-reported experience of school, child's self-concept, teacher-reported class engagement, and teacher-reported emotional behavioural function. Binary and multinomial logistic regression models were used to examine the association between BACA and these outcomes.

Results

Compared to those with TCD, those with BACA had significantly higher odds of never liking school (Adjusted odds ratio (AOR) 1.82, 95% CI 1.37-2.43, p<0.001), of being picked on (AOR 1.27, 95% CI 1.09-1.48) and of picking on others (AOR 1.53, 95% CI 1.27-1.84). They had significantly higher odds of experiencing low self-

concept (AOR 1.20, 95% CI 1.02-1.42) and emotional-behavioural difficulties (AOR 1.34, 95% CI 1.10-1.63, p=0.003). Compared to those with TCD, children with BACA had significantly higher odds of hardly ever or never being interested, motivated and excited to learn (AOR 2.29, 95% CI 1.70-3.10).

Conclusions

Children with BACA at school-entry had significantly higher odds of reporting a negative school experience and low self-concept at age 9. They had significantly higher odds of having teacher-reported poor class engagement and problematic emotional-behavioural function at age 9. The findings of this study suggest BACA has a causal role in these adverse outcomes. Early childhood policy and intervention design should be cognisant of the important role of cognitive ability in school and childhood outcomes.

4.2.Introduction

An important early milestone, for both children and parents, is the transition into primary school, which typically occurs around five years of age for children in Ireland.[1,2] The early experience of school lays the foundation for future academic and social development.[3] While only embarking on their educational journey, substantial inequalities in cognitive ability will already exist[4] Around 2% will have a standardised cognitive ability score more than two standard deviations (SDs) below the mean and may meet criteria for an intellectual developmental disorder (IDD).[5] A much larger proportion, approximately 14%, will have cognitive ability that lies between 1 and 2 SDs below the mean.[6,7]

Large epidemiological studies have demonstrated that cognitive ability in childhood contributes to important adult outcomes including educational attainment, social mobility, cardiovascular disease, and mental health.[8-11] Far less has been published on how cognitive ability affects early childhood experiences and outcomes, a formative one being a child's early experience of school. A small cohort study by McIntyre et al found children with an IDD had significantly more teacher-reported problem behaviour, fewer self-regulation and social skills, and poorer student-teacher relationships.[1] Children with below average cognitive ability, not meeting criteria for an IDD, were excluded from their study. The early school experience for these children, whose difficulties may be more likely to go unrecognised and unsupported, has not been extensively investigated.

In the early years of education, children with below average cognitive ability (BACA) may not display overt signs of difficulty. However, as the cognitive

complexity of academic tasks increases, the child may struggle to keep up with their peers and experience academic failure which may in turn influence, among other things, their experience of school, their self-concept, their engagement with learning, and their emotional and behaviour function.[12]

A child's self-concept refers to the child's personal perception of their self and is thought to play a critical role in how a child functions in almost all aspects of life, including how they cope with challenges.[13-15] It is generally viewed as a hierarchical construct, with "general self-concept", one's over-arching view of self, being comprised of multiple, correlated, domain-specific self-concepts, for example academic self-concept and physical self-concept.[14] Given the importance of self-concept for the adaptive functioning of the child it is essential factors which may contribute to negative self-concept are explored and understood.

Emotional functioning broadly describes the experience, expression, and regulation of both positive and negative emotions and is intimately intertwined with a child's behavioural function, one's actions and reactions to their environment.[16,17] Problematic emotional-behavioural function in childhood is common, with an estimated prevalence between 6-14%.[18-21] If persistent, frequent, distressing, and impacting function in other domains of life, it may constitute an emotional-behavioural disorder.[18,19] Common disorders in childhood include depression, anxiety, attention deficit hyperactivity disorder and conduct disorder.[18,19] Screening tools, of which the Strengths and Difficulties Questionnaire is one, can be completed by parents, the child, or teachers and aim to detect problematic emotional-behavioural functioning.[22]

Early subtle signs of academic difficulty in the early school years may go unnoticed or may be attributed to other causes such as personality or behavioural problems. Compulsory standardised testing in the Irish school system occurs when a child is in 2nd class (approximately 8 years of age), and until this point children with BACA may struggle unrecognised and often unsupported.[23] The Growing Up in Ireland (GUI) Infant '08 Cohort study, which directly measured cognitive ability at age 5 and interviewed both children and their teachers at age 9 provides an opportunity to explore the relationship between BACA and a child's experience of school, their self-concept, their classroom engagement, and their emotional-behavioural function. If the earliest years of formal education are adversely affected by cognitive deficits pre-dating school entry, this provides further rationale and impetus for very early intervention, prior to starting school, for at risk children.

The relationships between BACA and the aforementioned outcomes have multiple important confounders which must be addressed if a causal association between BACA and each outcome is to be established. Published literature was carefully examined to determine the appropriate confounding variables for inclusion and directed acyclic graphs were used to document our assumptions about the relationships between the potential cause (BACA), outcomes, and confounders.[24] Among the most important confounding factors is the sociodemographic background of the child, including their gender, socioeconomic position, and cultural background, which have been shown to be strongly associated with cognitive ability in childhood and the outcomes of interest.[25-31]

Child factors including the gestational age at which the child was born and the child's individual temperament have also been shown to be associated with

both cognitive ability and self-concept, emotional-behavioural function, and the formation of student-teacher relationships.[15,32-40] The relationship between BMI and cognitive ability in childhood has not been clearly established in the literature with some studies suggesting the association is completely mediated by socioeconomic factors, while others suggest BMI itself has a causal relationship with cognitive ability.[41,42] However, the respective associations between high BMI and low self-concept and problematic emotional behavioural function in childhood are well documented.[43-45]

Family factors such as the structure of the family, the parent-child relationship, and maternal mental health have been shown to be associated with both cognitive ability and self-concept, learning engagement, and emotional-behavioural function.[46-52] Multiple studies have shown that family size has an inverse association with childhood IQ, and many posit the 'resource dilution' model as an explanation, whereby the family resources including parental time and income reduce as family size increases.[46] There is consistent evidence that children in single parent households are at increased risk of lower academic achievement, lower self-concept and problematic emotional-behavioural function compared to those in two-parent households, with similar effects documented for the effect of maternal depression.[47,48,51]

The aim of this study was to examine the causal association between below average cognitive ability at age 5 and child-reported experience of school and self-concept, and teacher-reported class engagement and emotional-behavioural function at age 9. Our hypothesis was that early below average cognitive ability has an adverse relationship with self-concept, engagement with learning, and

emotional-behavioural function, as children with below average cognitive ability may struggle, often unrecognised, to cope with increasing academic demands.

4.3. Methods

Study design

This study is a secondary analysis of data collected in the GUI Infant Cohort, a nationally-representative longitudinal study of infants in Ireland, which contains information collected from primary caregivers (PCGs), teachers, and the children themselves.[53,54] The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines were followed in the conduct and reporting of this study.[55] Ethical approval for the GUI Infant Cohort was granted by a dedicated Research Ethics Committee convened by the Department of Children and Youth Affairs. Written consent was obtained from all participants in the study. PCGs were asked to provide consent to contact the child's school. Where a child had not yet started school at the time of the 5 year questionnaire, PCGs were asked for details of their intended school, and if this had not been decided they were recontacted prior to the start of school term. Secondary analysis of the GUI dataset does not require additional ethical approval.[53,54]

Study population

Families with infants born between 1st December 2007 and 30th June 2008 (approximately n=41,000 births) were identified from the Child Benefit Register, a register used to administer universal child benefit payments in Ireland, with the aim of carrying out the first wave of data collection between September 2008 and March/April 2009 when the infant would be 9 months old. Potential participants

were stratified according to marital status of claimant, county of residence, nationality, and number of children in the claim, before a systematic selection procedure with a random start and constant sampling fraction were applied. Full details of the GUI study are available.[56] Wave 1 recruited 11,134 nine-month-old infants, representing a 65% response rate. These 11,134 infants made up the target sample for subsequent waves (excluding those who moved outside of Ireland, who definitively opted out of the study, or who passed away during study period). A further three waves of in-person data collection occurred at age 3 (n=9,793 participants), age 5 (n=9,001), and age 9 years (n=8,032).[57] Included in this study are children (n=7,392) who completed standardised cognitive assessments at age 5 years and follow-up at age 9. Children with trisomy 21 (n=29) and cerebral palsy (n=11) were excluded.

Exposure

The exposure of interest was below average cognitive ability (BACA) at age 5. This was directly assessed using two core subtests of the British Ability Scales (BAS) Early Years Battery 2nd Edition, administered in the child's home by a trained interviewer.[58] The BAS has demonstrated construct validity as a measure of cognitive ability and has high reliability.[59]

The 36-item Naming Vocabulary subtest measured verbal ability in the English language. The child was requested to name everyday items, of increasing difficulty, displayed from a picture book. The Naming Vocabulary subtest primarily measures the broad ability of comprehension – knowledge or crystallised intelligence (Gc). The narrow abilities tested are language development, which refers to the ability to understand and communicate using language, and lexical knowledge, which

refers to the knowledge of words and their meanings.[60,61] The 33-item Picture Similarities subtest consisted of the child being shown four pictures and requested to match a fifth, based on a shared characteristic or construct. The Picture Similarities subtest primarily measures fluid intelligence (Gf). It tests the narrow ability of 'induction', which is the ability to discover underlying principles or rules governing a problem.[60,61] A standardised score for each scale, adjusted for both item difficulty and age (within a three month age band), was calculated.[62]

Multiple BAS subtest scores can be summed to produce a composite cognitive ability score. Due to time constraints, only two core subtests were administered in the GUI study which means this could not be done. To produce a composite score Principal Components Analysis (PCA) was used, as in previous research.[63] PCA of the two BAS subtests confirmed the presence of a general underlying cognitive ability factor. Principal component 1 (PC1) accounted for 64% of the total variance among the subtests. The Pearson correlation between this factor and the observed variable was 0.80 for Picture Similarities and 0.80 for Naming Vocabulary. PC1 was standardised to produce a general cognitive ability (GCA) score with a mean of 100 and a SD of 15. All children with a GCA score <85 (more than 1 SD below the mean) were categorised as "Below Average Cognitive Ability" (BACA) and those scoring ≥85 were categorised as typical cognitive development (TCD). Children with a GCA score <70 (more than 2 SD below the mean) were included in the BACA group in the main analysis. In supplementary analyses (Tables A9.1 – A9.6, Appendix 9) the BACA group was split into those with a GCA score <70 and those with a score between 70-85 (1-2 SD below the mean).

Outcomes

There were four outcomes of interest, all of which were measured at age 9

1) the child-reported experience of school

2) the child's self-concept

3) the teacher-reported class engagement

4) the teacher-reported emotional behavioural function of the child

Child-reported outcomes

The child-reported school experience was measured using the child's response to a number of questionnaire items, assessing whether the child liked school ("always like it", "sometimes like it", "never like it"), how well they thought they were doing in their school work ("well", "average/ok", "poorly"), whether they thought over the last year anyone (child or adult) had picked on them ("yes", "no"), or whether they thought over the last year they had picked on anyone else ("yes", "no").

The child's self-concept was measured using a modified version of the Piers-Harris Self-Concept Scale 2nd Edition, consisting of 32 items with a yes/no response.[64] This provided a measure of overall self-concept, as well as self-concept across six subscales of intellectual and school status, behavioural adjustment, physical appearance and attributes, freedom from anxiety, popularity, and happiness and satisfaction.[64] Accepted cut-offs are provided by the scale authors to categorise scores. For the purpose of this study these categories were collapsed into two categories - 'Low' which contained all those categorised as very low, low, or low average, and 'Average or High' which contained all those categorised as average, high average, or high.

Teacher-reported outcomes

Teachers were asked to complete a postal questionnaire when the child was aged 5 and 9. At both age 5 and age 9, the teacher was asked about learning limitations and the provision of additional resources. At age 9 the child's engagement in the classroom was investigated using teacher-reported responses ("always/almost always", "sometimes", "hardly ever/never") to a number of questions regarding the child's usual interest, behaviour, and involvement in class activities and learning. For example, "Would you say the study child is confident to try new activities, initiate ideas, and to speak in a familiar group?"

The teacher-reported emotional-behavioural function of the child was measured using the Strengths and Difficulties Questionnaire (SDQ), which has been shown to be a reliable and valid instrument.[65] The SDQ asks the teacher to respond, using a 3-point Likert where 0=not true, 1=somewhat true, and 2=certainly true, to 25 statements about the child's behaviour over the last six months or school year. These 25 items make up 5 subscales of 5 items each – Emotional subscale (Example item – "Often unhappy, downhearted, or tearful"), Conduct subscale (Example item. "Often has temper tantrums or hot tempers"), Hyperactivity subscale (Example item "Constantly fidgeting or squirming"), Peer problems subscale (Example item "Often fights with other children or bullies them"), and Pro-social subscales (Example item – "Considerate of other people's feelings"). The questionnaire provides a total difficulties score, with a higher score indicating more difficulty, along with scores for each subscale. In this study children with scores in the 'slightly raised', 'high', or 'very high' range were categorised as "High" and children scoring below this as

"Average".[66] Full details of the study outcomes are described in a data dictionary contained in Table A9.7 , Appendix 9.

Covariates

The existing literature and directed acyclic graphs, created using 'DAGitty'[24], were used to carefully choose relevant model-specific covariates for each outcome (Figures A9.1 – A9.4, Appendix 9). All covariates were measured when the child was aged 5 years, unless otherwise specified. Sociodemographic characteristics included child gender (male/female), siblings (yes/no), partner living in the household (yes/no), PCG highest education (lower secondary or less, upper secondary, technical or vocational, certificate or diploma, primary degree, post degree qualification), household social class (professional worker, managerial and technical, non-manual, skilled manual, semi-skilled, unskilled, all others gainfully occupied and unknown, validly no social class), and income quintile (1 lowest – 5 highest). The child's body mass index (BMI) was derived from the child's height and weight directly measured by the interviewer.

The parent-child relationship was measured using the 15-item Pianta Child Parent Relationship Scale (CPRS), which demonstrated acceptable reliability and validity in the pilot phase of the study.[62] The primary caregiver was asked to indicate the current applicability of 15 statements to their relationship with the child using a five-point Likert scale, where 1= definitely does not apply and 5= definitely applies. Seven items related to getting on with the child, for example– "I share an affectionate, warm relationship with my child", formed a positive aspects subscale. Scores could range from 7-35 with a higher score indicating a more positive relationship. Eight items related to the caregiver's perception of difficulties in the

relationship, for example – “My child and I always seem to be struggling with each other” formed the conflicts subscale. Scores could range from 8-40 with a higher score indicating more difficulties.[58,62]

Child temperament was measured using 12 adapted items from the parent-reported Short Temperament Scale for Children (STSC), which demonstrated reliability and validity in both the pilot phase of the study and in previous literature.[58,67] The parent was asked to indicate the answer that best described the current behaviour of the child using a six-point Likert scale where 1 = almost never and 6= almost always. Three subscale scores, ranging from 1-6, were produced by averaging the scores of the 4 items in each subscale - Persistence (4 items, example item – “When this child starts a project such as a puzzle or model, he/she works on it without stopping until it is completed, even if it takes a long time”), reactivity (4 items, example item – “When shopping together, if I do not buy what this child wants (e.g. sweets, clothing), he/she cries and yells”), and sociability (4 items, example item (reverse coded) – “This child is shy with strange adults”).

At the 9-month survey, the gestational age of the infant, the PCG cultural background (Irish, other white background, African or other black background, Chinese or other Asian background, Other), and PCG mental health were measured, the latter using The Center for Epidemiological Studies Depression Scale (CES-D), a validated screening tool for depression in the general population.[68] All covariates are described in Table A9.7, Appendix 9.

Statistical analysis

Analyses were conducted using SPSS software Version 28. A weighting variable, calculated by the GUI study team using a minimum information loss algorithm, to

adjust for non-response and attrition was applied to all analyses.[57] Missing data were managed using complete case analysis. The characteristics of the population were described using counts, percentages, and chi-square tests for categorical data; means, SDs, and t-tests for continuous normally distributed data; medians, interquartile ranges (IQRs), and Mann-Whitney U tests for continuous non-normally distributed data. Binary and multinomial logistic regression models were used to examine the association between BACA and the outcomes of interest. Models were adjusted using the DAG implied adjustment set for the estimand of interest. Unadjusted and adjusted odds ratios (AOR) with 95% confidence intervals (CI) were reported.

4.4. Results

Characteristics of study population

The characteristics of the 7,392 children included in the study are shown in Table 4.1. Children with BACA at age 5 (n=1,106, 15.0%) are compared to those without across key characteristics. There were significant differences between the groups across all socioeconomic variables. Compared to the TCD group, a higher proportion of PCGs of the BACA group had second level or lower education (40.0% v 27.8%), were in the lowest two income quintiles (60.3% v 38.5%), were in single parent households (21.2% v 14.9%), and a lower proportion were in social classes 1 or 2 (26.4% v 48.0%).

Among those with BACA, 44.9% (461/1,027) were reported by the teacher as having a limitation to the amount or kind of activity they could undertake in school at age 5 and 28.1% (289/1,027) were provided with additional support or resources (Table A9.2, Appendix 9). Among those with a GCA score <70, 30% were reported by

their teacher as having no limitation in the type or amount of school activities they could undertake. Of those with a recognised limitation, 21% were not provided with additional help or resources. Among those with a GCA score between 70-85, 41% were reported to have a limitation in the amount or kind of school activity they could undertake, of whom 42% were not provided with special help or resources.

Child-reported experience of school at age 9 years

Overall, 94.8% of children reported they always or sometimes liked school at age 9 (Table A9.8, Appendix 9). After adjusting for relevant confounding variables (Figure A9.1, Appendix 9), compared to those in the TCD group the odds of never liking school were 1.8 times higher for children with BACA (111 (10.2%) v 290 (4.7%), AOR 1.82, 95% CI 1.37-2.43, p<0.001). Children with BACA had significantly higher odds of being picked on (473 (46.8%) v 2,394 (40.1%), AOR 1.27, 95% CI 1.09-1.48) and of picking on others (231 (23.0%) v 822 (14.0%), AOR 1.53, 95% CI 1.27-1.84).

Child's self-concept

After appropriate adjustment for confounding (Figure A9.2, Appendix 9), children with BACA at age 5 had significantly higher odds of experiencing low overall self-concept at age 9 (AOR 1.20, 95% CI 1.02-1.42, p=0.028) (Table 4.2). The odds of experiencing low intellectual self-concept were 1.4 times higher (AOR 1.38, 95% CI 1.17-1.63, p<0.001), a measure which reflects both their assessment of intellectual ability and school performance, but also their future expectations about achievement. In both groups a substantial proportion of children reported low self-concept in happiness and satisfaction (TCD 32.0% v BACA 43.2%, p<0.001), but for children in the BACA, the odds were still 1.3 times higher than those in the TCD group (AOR 1.29, 95% CI 1.10-1.50, p<0.001)

Teacher-reported class engagement at age 9 years

As shown in Table 4.3, after adjustment for confounding (Figure A9.3, Appendix 9) children with BACA had less positive class engagement compared to their peers with TCD. They had significantly higher odds of hardly ever or never being interested, motivated and excited to learn (10.1% v 4.3%, AOR 2.29, 95% CI 1.70-3.10, p<0.001), of hardly ever or never being confident to try new activities, initiate ideas, and speak in a familiar group (16.0% v 6.0%, AOR 2.73, 95% CI 2.12-3.50, p<0.001), of hardly ever or never maintaining attention and concentration (13.3% v 6.2%, AOR 2.42, 95% CI 1.85-3.16, p<0.001,) and of hardly ever or never sustaining involvement and persevering particularly when trying to solve a problem (19.5% v 7.9%, AOR 2.75, 95% CI 2.17-3.50, p<0.001).

Teacher-reported emotional-behavioural function at age 9 years

As shown in Table 4.4, after adjustment for confounding variables children with BACA had significantly higher odds of having potentially problematic emotional-behavioural function (23.5% v 14.4%, AOR 1.34, 95% CI 1.10-1.63, p=0.003). When the BACA group was split into those with a GCA score of 70-85 and those with GCA score <70 (Table A9.6, Appendix 9), both groups remained significantly more likely to experience emotional-behavioural difficulties compared to their peers with TCD (GCA score 70-85 AOR 1.27, 95% CI 1.03-1.56; GCA score <70 AOR 1.95, 95% CI 1.20-3.17). Neither group had higher odds of conduct difficulties compared to their peers with TCD. After adjustment for confounding, children with BACA had significantly higher odds of having teacher-reported difficulties with hyperactivity compared to their peers with TCD (24.6% v 14.9%, AOR 1.41, 95% CI 1.16-1.71).

4.5. Discussion

The aim of this study was to examine the causal association between below average cognitive ability at age 5, the age most children begin primary school in Ireland, and child-reported experience of school, child-reported self-concept, and teacher-reported class engagement and emotional-behavioural function at age 9. We found in a large nationally-representative sample, after adjustment for confounding, that children beginning primary school with BACA had higher odds of never liking school, of being picked on, of picking on others, and of reporting low self-concept. They had significantly higher odds of poor class engagement and problematic emotional-behavioural function, as reported by their teacher. The modelling approach, which involved careful consideration and stringent adjustment for confounding variables, suggests that early below average cognitive ability has a causal association with each adverse outcome. The socioeconomic environment to which the child was exposed, likely the most important confounding factor, was carefully adjusted for. In keeping with recommendations in the literature, collinearity between potential measures of SES was examined prior to inclusion of a range of measures representing different domains of the socioeconomic environment – parental education, family income, presence of a partner in the household, and the cultural background of the family.[69]

There was evidence, after adjustment for confounding, of a causal relationship between BACA at age 5 and low overall self-concept at age 9. In particular, children with BACA had higher odds of low intellectual self-concept. Marsh's widely published reciprocal-effects model proposes that low academic performance will adversely impact a child's self-concept, which in turn will adversely affect a child's future academic achievement, as the child is more likely to be hesitant

to engage in learning, to fear failure, and to give up.[70] The findings of this study, while limited to only two time points, lend support to this hypothesis. They confirm that children beginning school with BACA were at higher risk of low self-concept four years later, and were less likely to be excited to learn, to persevere when problem-solving, and to be confident to try new things, initiate ideas, or speak in the classroom. These findings have important implications for educational policy and practice and for the development of interventions. Marsh and Craven maintain that in order to enhance performance "enhancing skills alone is not enough; people also need to hold positive self-concepts of their abilities in specific areas".[70]

Teachers were more likely to report poor classroom engagement for children with BACA, compared to their peers with typical cognitive development. Of note, they were no more likely to have conduct problems compared to their peers, suggesting disruptive behaviour did not contribute to their struggle to engage in the classroom. Children with BACA did however have difficulty with attention and concentration across a number of measures. They were significantly more likely to hardly ever or never maintain attention and concentration in the classroom compared to peers with TCD (13.3% v 6.2%), and were significantly more likely to have teacher-reported hyperactivity difficulties (24.6% v 14.9%).

While one could argue that these children may have other underlying diagnoses, such as attention deficit hyperactivity disorder (ADHD) or autism spectrum disorder (ASD), at age 9 years only 7% of all children with BACA had teacher-reported limitations in learning due to a diagnosis of an emotional-behavioural disorder such as ADHD, and only 5% due to a diagnosis of ASD. This does not rule out undiagnosed disorders, including specific learning disabilities, such as

dyslexia or dyscalculia, which may or may not be accompanied by lower cognitive ability. However, for some children with BACA the attention and concentration difficulties could reflect an inability to engage and keep up on a cognitive level with the activities taking place in the classroom. It is unsurprising that children with BACA had higher odds of never liking school, a sentiment which tends to persist over time, and which increases the risk of absenteeism and early school dropout.[71,72] Early school dropout itself is a risk factor for incarceration, substance use, teenage pregnancy and poor health.[73,74]

While cognitive ability alone does not determine a child's success, happiness, or achievement [75], the results of this study clearly show it has an association with a child's ability to engage in learning, with their appraisal of themselves, and with other aspects of their emotional and behavioural functioning. There is a risk of placing undue emphasis on a child's cognitive ability, and failing to recognise and foster other forms of ability and intelligence, such as musical, interpersonal, and bodily-kinaesthetic intelligence, however it cannot be denied that a child's cognitive ability is an important determinant of outcomes throughout the lifecourse.[76] If every child is to benefit equally from their right to education, a multi-faceted, inter-agency approach which looks at cognition from a primary, secondary, and tertiary prevention approach is required.

Every child should have the opportunity to reach their cognitive potential. For some children, genetic conditions at birth may result in a lower potential cognitive ability, but for many children it is the environment into which they are born which prevents them reaching their cognitive potential. In this study, a significantly higher proportion of children with BACA were from lower socioeconomic backgrounds, a

finding consistent across the literature.[77] Infants born into socioeconomic adversity are more likely to experience parental depression, homelessness, and drug or alcohol abuse. They may not receive adequate interaction and support, the basic prerequisites for cognitive development. They may not have access to the tangible resources, for example books and toys, and intangible resources, for example family social capital, that promote optimal cognitive development.[78] Political commitment, investment, and social policy reform are required to address homelessness, deprivation, unemployment, and addiction. Without this, achieving educational equality will remain an uphill struggle.

Early intensive interventions, which begin in infancy, focus on families as well as the infant, and provide developmentally enriching environments, can improve cognitive outcomes for at risk children.[79] The challenge lies in identifying, in a timely fashion, who these children are. Some countries, such as the United States, adopt a risk-based approach and provide population-based early intervention on the basis of socioeconomic risk factors.[79] Other countries have implemented standardised assessment at school-entry, however this has the potential to systematically delay school-starting age for children from disadvantaged backgrounds and risks further stigmatisation.[80] Additionally, this approach would only allow interventions to be started at the time of expected school entry, when significant differences in cognitive ability already exist. The ideal time for intervention is in very early life, when the brain is at its optimal stage of neuroplasticity. Current routinely used developmental screening tools, such as the Ages and Stages Questionnaire, will not detect the majority of children with BACA at school-age and cannot therefore be used to achieve this.[81,82] There is a need for alternative methods of early identification for children at

risk of poor cognitive outcomes at school-age. In Ireland, there is currently no population-based programme of early intervention for children at risk of poor cognitive outcomes. As alluded to previously, designing such an intervention should acknowledge the reciprocal relationship between cognitive ability and self-concept, and should aim to enhance both simultaneously.

Finally, from a tertiary prevention perspective, all children will begin their educational journey with different strengths and weaknesses, with different skills and abilities, and with different aspirations for what they wish to gain from their education. Every child should have equal opportunity to fulfil these aspirations in a supportive, bias-free system, which values both cognitive and non-cognitive skills. In Ireland, and in many countries internationally, this may mean shifting the focus from solely one of cognitive achievement, and adapting both teaching and assessment methods to incorporate other forms of ability and achievement.[83]

Major strengths of this study include the large sample size, the careful sampling strategy employed in recruitment of the cohort, and the application of a weighting variable to all analyses to account for non-response and attrition, all of which enable us to make relatively confident inferences about the population from which the sample was drawn.

Limitations

Cognitive ability was directly measured using standardised assessments. Children taking the naming vocabulary test, for whom English is not their native language, may be disadvantaged, with apparent poor performance. The British Ability Scales is a valid tool for measuring general cognitive ability, and is comparable to measures of IQ such as the Wechsler Intelligence Test.[84] Children in GUI did not complete all BAS

subtests and PCA was therefore used to calculate a GCA score. This approach would not be acceptable for a formal diagnosis of an IDD.

Efforts were made in GUI to minimise response bias through the use of anonymous self-complete questionnaires for sensitive topics. While the application of a weighting variable aimed to minimise the overall effect of non-response and attrition on the structure of the cohort, selective missing data remains a source of bias. For example, teacher-reported missing data for individual students may be systematically more likely in schools which are under-resourced. Missing data were treated with complete case analyses, a further source of bias.

Teacher-reported outcomes were used for assessing the child's classroom engagement and their emotional-behavioural function. It is well-documented that teacher-reported outcomes are subject to a teacher's implicit bias. A teacher's assessment of a child may be affected by the child's race, socioeconomic background, gender, or other characteristic.[85] Their treatment of the child can result in a self-fulfilling prophecy whereby the teacher's expectations of the student influence the actual performance and academic achievement of the student.[86]

The study did not adjust for the presence of underlying conditions at age 5 which may affect cognitive ability, such as ASD and ADHD, as this data was not available. However, data from the at 9 year survey would suggest that a minority of children in the BACA group were diagnosed with these conditions.

4.6. Conclusion

Acknowledging these limitations, this study shows that for children with BACA at the beginning of school, the odds of experiencing low child-reported self-concept,

poor teacher-reported classroom engagement, and problematic teacher-reported emotional-behavioural function are significantly higher. The findings support the hypothesis that cognitive ability itself has a causal relationship with these outcomes.

For families, the findings of this study suggest that if a child is disengaging from school; has a negative self-perception, particularly of their intellectual ability; is displaying, or is reported by their teacher to be displaying, problematic emotions or behaviours, then consideration should be given to the child's cognitive functioning as a potential contributing factor. Families and teachers, should together determine in these scenarios whether further exploration of the child's cognitive functioning is warranted. Teachers should have access, without undue delay or administrative barriers, to psychological assessments and to supports such as mental health services, speech and language therapists, psychological services, and learning support if they are necessary.

Every child should have the opportunity to reach their individual cognitive potential. Reducing socioeconomic inequality is crucial in this regard. Early intervention programmes can improve cognitive ability but further research is required on the optimal method to identify which children will benefit most. Finally, adapting our education system to provide support and resources for those with lower cognitive potential, and to foster and promote other types of non-cognitive abilities is essential.

Author contributions

AB – Conceptualisation, methodology, formal analysis, writing – original draft preparation, writing – review and editing

MU – Conceptualisation, writing – review and editing

AS – Methodology, formal analysis, writing – review and editing

DM – Conceptualisation, methodology, writing – review and editing

Table 4.1 Characteristics of study population and comparison of characteristics between those who did and did not have below average cognitive ability at age 5 years

	Valid	Total n=7,392	Typical cognitive development n=6,286	Below average cognitive ability n=1,106	p-value
Child characteristics					
Gender	7,392				
Male		3,783 (51.2)	3,131 (49.8)	652 (59.0)	<0.001[a]
Gestational age	7,375				
N, mean (sd)		39.5 (2.1)	6,277, 39.5 (2.1)	1,097, 39.5 (2.1)	0.802[b]
Child temperament					
Sociability – n, median (IQR)	7,389	4.0 (3.0-5.0)	6,285, 4.0 (3.0-5.0)	1,104, 3.8 (2.8-4.8)	0.014[c]
Persistence – n, median (IQR)	7,373	4.3 (3.5-5.0)	6,279, 4.3 (3.5-5.0)	1,095, 4.0 (3.0-4.8)	<0.001[c]
Reactivity – n, median (IQR)	7,381	2.3 (1.8-3.0)	6,282, 2.3 (1.8-3.0)	1,098, 2.3 (1.8-3.3)	<0.001[c]
Body mass index – kg/m^2	7,319	16.3 (1.7)	6,231, 16.3 (1.7)	1,088, 16.3 (1.7)	0.179[b]
General cognitive ability score	7,392				
N, mean (sd)		99.8 (14.9)	6,286, 103.8 (12.1)	1,105, 77.2 (6.7)	<0.001[b]
Socio-demographic characteristics					
Household location					
Urban	7,370	3,062 (41.5)	2,588 (41.2)	474 (43.3)	
Rural		4,308 (58.5)	3,688 (58.8)	620 (56.7)	0.207[a]
Partner in household	7,392				
Yes		6,224 (84.2)	5,352 (85.1)	872 (78.8)	<0.001[a]
Siblings	7,391				
Yes		6,479 (87.7)	5,514 (87.7)	965 (87.3)	0.755[a]
Cultural background	7,375				
Irish		6,398 (86.8)	5,598 (89.1)	800 (73.3)	
Other white		642 (8.7)	457 (7.3)	185 (17.0)	
African/Black		156 (2.1)	93 (1.5)	63 (5.8)	
Chinese/Asian		151 (2.0)	116 (1.8)	35 (3.2)	
Other		28 (0.4)	20 (0.3)	8 (0.7)	<0.001[a]
PCG Highest education	7,391				
Lower secondary		965 (13.1)	722 (11.5)	243 (22.0)	

or less					
Upper secondary		1,227 (16.6)	1,027 (16.3)	200 (18.0)	
Technical or vocational		1,430 (19.3)	1,169 (18.6)	261 (23.6)	
Certificate or diploma		1,605 (21.7)	1,421 (22.6)	184 (16.7)	
Primary degree		851 (11.5)	746 (11.9)	105 (9.5)	
Post degree qualification		1,313 (17.8)	1,201 (19.1)	112 (10.1)	<0.001[a]
Household social class	7,392				
Professional workers		934 (12.6)	858 (13.6)	76 (6.9)	
Managerial and technical		2,380 (32.2)	2,164 (34.4)	216 (19.5)	
Non-manual		1,334 (18.0)	1,126 (17.9)	208 (18.8)	
Skilled manual		1,115 (15.1)	868 (13.8)	247 (22.3)	
Semi-skilled		668 (9.0)	525 (8.4)	143 (12.9)	
Unskilled		110 (1.5)	95 (1.5)	15 (1.4)	
All others gainfully occupied		76 (1.0)	51 (0.8)	25 (2.3)	
Validly no social class[d]		775 (10.5)	599 (9.5)	176 (15.9)	<0.001[a]
Equivalised Income Quintile[e]	7,009				
Lowest		1,448 (20.7)	1,107 (18.5)	341 (33.3)	
2		1,472 (21.0)	1,194 (20.0)	278 (27.0)	
3		1,383 (19.7)	1,207 (20.2)	176 (17.1)	
4		1,360 (19.4)	1,231 (20.6)	129 (12.5)	
Highest		1,346 (19.2)	1,240 (20.7)	106 (10.3)	<0.001[a]
PCG depression n, median (IQR)	7,275	1.0 (0.0-3.0)	6,223, 1.0 (0.0-3.0)	1,052, 1.0 (0.0-4.0)	0.052[c]
Relationship characteristics					
Parent-child relationship					
Positive aspects[f] – n, median (IQR)	7,380	35.0 (33.0-35.0)	6,285, 35.0 (33.0-35.0)	1,095, 34.0 (33.0-35.0)	<0.001[c]
Conflict[g] – n, median (IQR)	7,384	14.0 (10.0-19.0)	6,284, 14.0 (10.0-18.0)	1,100, 15.0 (11.0-20.0)	<0.001[c]

Abbreviations: sd, standard deviation; IQR, interquartile range.

[a]Calculated using Pearson's Chi-squared test, [b]Independent samples t-test, [c]Independent samples Mann-Whitney U Test, [d]Validly no social class consists of households where both caregivers are currently economically inactive and have not held any previous employment in the past, [e]Equivalised Income Quintile is based on the disposable household income (total gross income less statutory deductions), divided by the equivalised household size which takes account of differences in size and composition of households, [f]Higher score indicates more positive relationship, [g]Higher score indicates more conflict.

Table 4.2 Child self-concept at age 9 years

	Valid	Total	Typical cognitive development	Below average cognitive ability	p-value[a]	Odds ratio (95% CI)	Adjusted odds ratio[b] (95% CI)	p-value
Overall self-concept	6,702							
Average or high		4,860 (72.5)	4,236 (73.8)	624 (64.7)		Ref		
Low		1,842 (27.5)	1,501 (26.2)	341 (35.3)	<0.001	1.54 (1.34-1.78)	1.20 (1.02-1.42)	0.028
Intellectual and school status	6,671							
Average or high		4,960 (74.4)	4,335 (75.9)	625 (65.2)		Ref		
Low		1,711 (25.6)	1,378 (24.1)	333 (34.8)	<0.001	1.68 (1.45-1.94)	1.38 (1.17-1.63)	<0.001
Behavioural adjustment	6,845							
Average[d]		4,388 (64.1)	3,841 (65.6)	547 (55.5)		Ref		
Low		2,457 (35.9)	2,018 (34.4)	439 (44.5)	<0.001	1.53 (1.33-1.75)	1.27 (1.09-1.48)	0.003
Physical appearance and attributes	6,799							
Average[c]		5,155 (75.8)	4,449 (76.6)	706 (71.3)		Ref		
Low		1,644 (24.2)	1,360 (23.4)	284 (28.7)	<0.001	1.32 (1.13-1.53)	1.16 (0.98-1.38)	0.081
Freedom from anxiety	6,878							
Average or high		5,317 (77.3)	4,584 (77.9)	733 (73.5)		Ref		
Low		1,561 (22.7)	1,297 (22.1)	264 (26.5)	0.002	1.27 (1.09-1.49)	1.03 (0.86-1.22)	0.781
Popularity	6,810							
Average or high		5,214 (76.6)	4,524 (77.7)	690 (69.8)		Ref		
Low		1,596 (23.4)	1,298 (22.3)	298 (30.2)	<0.001	1.51 (1.30-1.75)	1.21 (1.02-1.44)	0.028

Happiness and satisfaction	6,877							
Average[d]		4,562 (66.3)	3,997 (68.0)	565 (56.8)		Ref		
Low		2,315 (33.7)	1,885 (32.0)	430 (43.2)	<0.001	1.62 (1.41-1.85)	1.29 (1.10-1.50)	0.001

Abbreviations: CI, confidence interval. [a]Calculated using Pearson's Chi-squared test, [b]Adjusted for child gender, gestational age, siblings, primary caregiver education, household income, parent-child relationship, child BMI, cultural background, child temperament, partner in household, maternal mental health [c]Where 'Average' is listed alone no child score in the 'high average' or 'high' categories

Table 4.3 Teacher-reported classroom engagement and behaviour at age 9 years

	Valid	Total	Typical cognitive development	Below average cognitive ability	p-value[a]	Odds ratio (95% CI)	Adjusted odds ratio[b] (95% CI)	p-value
Shows an interest in classroom activities	6,729							
Always/almost always		5,175 (76.9)	4,542 (79.1)	633 (64.1)		Ref		
Sometimes		1,406 (20.9)	1,094 (19.1)	312 (31.6)		2.05 (1.76-2.38)	1.49 (1.25-1.78)	<0.001
Hardly ever/never		148 (2.2)	105 (1.8)	43 (4.4)	<0.001	2.90 (2.01-4.18)	2.01 (1.31-3.09)	0.001
High level of involvement in self-chosen activities	6,731							
Always/almost always		5,071 (75.3)	4,468 (77.8)	603 (61.2)		Ref		
Sometimes		1,476 (21.9)	1,159 (20.2)	317 (31.2)		2.03 (1.74-2.36)	1.59 (1.34-1.88)	<0.001
Hardly ever/never		184 (2.7)	118 (2.1)	66 (6.7)	<0.001	4.16 (3.04-5.69)	2.94 (2.05-4.21)	<0.001
Uses activities and resources independently	6,726							
Always/almost always		4,134 (61.5)	3,702 (64.5)	432 (43.7)		Ref		
Sometimes		2,254 (33.5)	1,797 (31.3)	457 (46.2)		2.18 (1.89-2.51)	1.74 (1.48-2.04)	<0.001
Hardly ever/never		338 (5.0)	238 (4.1)	100 (10.1)	<0.001	3.59 (2.78-4.63)	2.52 (1.87-3.39)	<0.001
Interested, motivated, and excited to learn	6,724							
Always/almost always		4,379 (65.1)	3,918 (68.3)	461 (46.8)		Ref		
Sometimes		2,000 (29.7)	1,575 (27.4)	425 (43.1)		2.29 (1.98-2.65)	1.87 (1.59-2.21)	<0.001
Hardly ever/never		345 (5.1)	246 (4.3)	99 (10.1)	<0.001	3.44 (2.67-4.42)	2.29 (1.70-3.10)	<0.001
Confident to try new things, initiate ideas and speak in familiar group	6,721							

Always/almost always		4,028 (59.9)	3,609 (62.9)	419 (42.5)		Ref		
Sometimes		2,193 (32.6)	1,785 (31.1)	408 (41.4)		1.97 (1.70-2.29)	1.71 (1.45-2.01)	<0.001
Hardly ever/never		500 (7.4)	342 (6.0)	158 (16.0)	<0.001	3.99 (3.22-4.94)	2.73 (2.12-3.50)	<0.001
Maintains attention and concentrates	6,720							
Always/almost always		3,813 (56.7)	3,426 (59.8)	387 (39.2)		Ref		
Sometimes		2,420 (36.0)	1,950 (34.0)	470 (47.6)		2.13 (1.84-2.47)	1.75 (1.48-2.07)	<0.001
Hardly ever/never		487 (7.2)	356 (6.2)	131 (13.3)	<0.001	3.25 (2.59-4.07)	2.42 (1.85-3.16)	<0.001
Perseveres when problem solving	6,714							
Always/almost always		3,676 (54.8)	3,324 (58.0)	352 (35.8)		Ref		
Sometimes		2,397 (35.7)	1,958 (34.2)	439 (44.7)		2.12 (1.82-2.46)	1.80 (1.52-2.14)	<0.001
Hardly ever/never		641 (9.5)	450 (7.9)	191 (19.5)	<0.001	4.03 (3.29-4.93)	2.75 (2.17-3.50)	<0.001

Abbreviations: CI, confidence interval. [a]Calulcated using Pearson's Chi-squared test, [b]Adjusted for child gender, gestational age, siblings, primary caregiver education, household income, parent-child relationship, child BMI, cultural background, child temperament, partner in household, maternal mental health.

Table 4.4 Teacher-reported emotional behavioural function at age 9 years

	Valid	Total	Typical cognitive development	Below average cognitive ability	p-value	Odds ratio (95% CI)	Adjusted odds ratio[b] (95% CI)	p-value
Total Difficulties	6,738							
Average		5,680 (84.3)	4,923 (85.6)	757 (76.5)		Ref		
High		1,058 (15.7)	826 (14.4)	232 (23.5)	<0.001[a]	1.83 (1.55-2.15)	1.34 (1.10-1.63)	0.003
Emotional	6,739							
Average		5,631 (83.6)	4,853 (84.4)	778 (78.7)		Ref		
High		1,108 (16.4)	897 (15.6)	211 (21.3)	<0.001[a]	1.46 (1.24-1.73)	1.23 (1.02-1.49)	0.031
Conduct	6,739							
Average		6,118 (90.8)	5,242 (91.2)	876 (88.6)				
High		621 (9.2)	508 (8.8)	113 (11.4)	0.011[a]	1.34 (1.08-1.66)	0.92 (0.71-1.18)	0.507
Hyperactivity	6,738							
Average		5,641 (83.7)	4,895 (85.1)	746 (75.4)				
High		1,097 (16.3)	854 (14.9)	243 (24.6)	<0.001[a]	1.87 (1.59-2.20)	1.41 (1.16-1.71)	<0.001
Peer problems	6,739							
Average		5,799 (86.1)	5,000 (87.0)	799 (80.8)				
High		940 (13.9)	750 (13.0)	190 (19.2)	<0.001[a]	1.58 (1.33-1.89)	1.18 (0.96-1.45)	0.112
Prosocial[c]	6,737							
Average		5,888 (87.4)	5,062 (88.1)	826 (83.6)				
High		849 (12.6)	687 (11.9)	162 (16.4)	<0.001[a]	1.45 (1.20-1.75)	0.99 (0.79-1.23)	0.905

Abbreviations: CI, confidence interval. [a] Pearson's Chi-squared test, [b] Adjusted for child gender, gestational age, siblings, primary caregiver education, household income, parent-child relationship, child BMI, cultural background, child temperament, partner in household, maternal mental health [c] The prosocial scale is a positive subscale but was reverse scored so that a higher score indicates potentially problematic function in this domain.

4.7. References

1. McIntyre LL, Blacher J, Baker BL. The transition to school: adaptation in young children with and without intellectual disability. J Intellect Disabil Res. 2006 May;50(Pt 5):349-61. doi: 10.1111/j.1365-2788.2006.00783.x. PMID: 16629928.
2. Government of Ireland. Information for parents of children aged 2-5 years on the early childhood care and education (ECCE) and primary school systems in Ireland [Internet]. [Cited 2025 Jan 9] Available from: https://assets.gov.ie/234516/cc0e413d-d507-451a-8d23-5cf155cfcc3f.pdf
3. Entwisle D R, Hayduk LA. Lasting Effects of Elementary School. Sociology of Education. 1988;61(3):147-159. doi:10.2307/2112624
4. Lee V, Burkam D. Inequality at the Starting Gate: Social Background Differences in Achievement As Children Begin School.[Internet]. Washington DC: Economic Policy Institute; 2002 [cited 2024 Jul 10]. Available from: https://www.epi.org/publication/books_starting_gate/
5. World Health Organisation. International Statistical Classification of Diseases and Related Health Problems p. A00 Disorders of intellectual development.[Internet]. World Health Organisation; 2022 [cited 2023 Jan 18]. Available from: https://icd.who.int/browse11/l-m/en#/http%3A%2F%2Fid.who.int%2Ficd%2Fentity%2F605267007
6. Wieland J, Zitman FG. It is time to bring borderline intellectual functioning back into the main fold of classification systems. BJPsych Bull.

2016 Aug;40(4):204-6. doi: 10.1192/pb.bp.115.051490. PMID: 27512590; PMCID: PMC4967780.

7. Nouwens PJG, Lucas R, Embregts PJCM, van Nieuwenhuizen C. In plain sight but still invisible: A structured case analysis of people with mild intellectual disability or borderline intellectual functioning. J Intellect Dev Disabil. 2017;42(1):36–44. https://doi.org/10.3109/13668250.2016.1178220

8. Forrest LF, Hodgson S, Parker L, Pearce MS. The influence of childhood IQ and education on social mobility in the Newcastle Thousand Families birth cohort. BMC Public Health. 2011 Nov 25;11:895. doi: 10.1186/1471-2458-11-895. PMID: 22117779; PMCID: PMC3248886.

9. Furnham A, Cheng H. Childhood Cognitive Ability Predicts Adult Financial Well-Being. J Intell. 2016;5(1):3. doi: 10.3390/jintelligence5010003.

10. Calvin CM, Deary IJ, Fenton C, Roberts BA, Der G, Leckenby N, Batty GD. Intelligence in youth and all-cause-mortality: systematic review with meta-analysis. Int J Epidemiol. 2011 Jun;40(3):626-44. doi: 10.1093/ije/dyq190. Epub 2010 Oct 29. PMID: 21037248; PMCID: PMC3147066.

11. Koenen KC, Moffitt TE, Roberts AL, Martin LT, Kubzansky L, Harrington H, Poulton R, Caspi A. Childhood IQ and adult mental disorders: a test of the cognitive reserve hypothesis. Am J Psychiatry. 2009 Jan;166(1):50-7. doi: 10.1176/appi.ajp.2008.08030343. Epub 2008 Dec 1. PMID: 19047325; PMCID: PMC2705657.

12. Karande S, Kanchan S, Kulkarni M. Clinical and psychoeducational profile

of children with borderline intellectual functioning. Indian J Pediatr. 2008;75(8):795-800. doi: 10.1007/s12098-008-0101-y. Epub 2008 Jun 25. PMID: 18581071.

13. Butler RJ, Gasson SL. Self Esteem/Self Concept Scales for Children and Adolescents: A Review. Child Adolesc Ment Health. 2005 Nov;10(4):190-201. doi: 10.1111/j.1475-3588.2005.00368.x. PMID: 32806865

14. Shavelson R.J., Bolus R. Self-concept: The interplay of theory and methods. Journal of Educational Psychology. 1982;74(1): 3–17. doi:10.1037/0022-0663.74.1.3

15. Brown GL, Mangelsdorf SC, Neff C, Schoppe-Sullivan SJ, Frosch CA. Young Children's Self-Concepts: Associations with Child Temperament, Mothers' and Fathers' Parenting, and Triadic Family Interaction. Merrill Palmer Q (Wayne State Univ Press). 2009;55(2):184-216. doi: 10.1353/mpq.0.0019. PMID: 25983365; PMCID: PMC4429799.

16. Milojevich HM, Lindquist KA, Sheridan MA. Adversity and Emotional Functioning. Affect Sci . 2021 Sep 18;2(3):324-344. doi: 10.1007/s42761-021-00054-w. PMID: 36059901; PMCID: PMC9382958.

17. Molteni M. Behavior, human. Clin Ter. 2010; 161(1):91-2. PMID: 20393688.

18. Ogundele MO. Behavioural and emotional disorders in childhood: A brief overview for paediatricians. World J Clin Pediatr .2018 Feb 8;7(1):9-26. doi: 10.5409/wjcp.v7.i1.9. PMID: 29456928; PMCID: PMC5803568.

19. Brauner CB, Stephens CB. Estimating the prevalence of early childhood serious emotional/behavioral disorders: challenges and

recommendations. Public Health Rep. 2006 May-Jun;121(3):303-10. doi: 10.1177/003335490612100314. PMID: 16640155; PMCID: PMC1525276.

20. Swift A, McConkey R, Curry P, Iriarte EG. Social-Emotional Difficulties in Irish Children Aged Five and Nine Years: A National, Longitudinal Study. Children (Basel). 2021 Jul 29;8(8):656. doi: 10.3390/children8080656. PMID: 34438549; PMCID: PMC8392290.

21. Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. J Child Psychol Psychiatry. 2015 Mar;56(3):345-65. doi: 10.1111/jcpp.12381. Epub 2015 Feb 3. PMID: 25649325.

22. Ortuño-Sierra J, Chocarro E, Fonseca-Pedrero E, Riba SSI, Muñiz J. The assessment of Emotional and Behavioural problems: Internal structure of The Strengths and Difficulties Questionnaire. Int J Clin Health Psychol. 2015 Sep-Dec;15(3):265-273. doi: 10.1016/j.ijchp.2015.05.005. Epub 2015 Jun 22. PMID: 30487843; PMCID: PMC6225034.

23. Department of Education. Standardised Testing in Primary Schools [Internet]. Ireland; 2020. [updated 2024 April 8; cited 2025 Jan 9]. Available from: https://www.gov.ie/en/service/af92e8-standardised-testing-in-primary-schools/#:~:text=Standardised%20tests%20are%20used%20to,to%20access%20the%20full%20curriculum.

24. Textor J, van der Zander B, Gilthorpe MS, Liskiewicz M, Ellison GT. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. Int

J Epidemiol. 2016 Dec 1;45(6):1887-1894. doi: 10.1093/ije/dyw341. PMID: 28089956.

25. Twenge J.M., Campbell W.K. Self-Esteem and Socioeconomic Status: A Meta-Analytic Review. Personality and Social Psychology Review. 2002;6(1):59-71. https://doi.org/10.1207/S15327957PSPR0601_3

26. Chaplin TM, Aldao A. Gender differences in emotion expression in children: a meta-analytic review. Psychol Bull. 2013 Jul;139(4):735-65. doi: 10.1037/a0030737. Epub 2012 Dec 10. PMID: 23231534; PMCID: PMC3597769.

27. von Stumm S, Plomin R. Socioeconomic status and the growth of intelligence from infancy through adolescence. Intelligence. 2015 Jan-Feb;48:30-36. doi: 10.1016/j.intell.2014.10.002. PMID: 26640306; PMCID: PMC4641149.

28. Ardila A, Rosselli M, Matute E, Inozemtseva O. Gender differences in cognitive development. Dev Psychol. 2011 Jul;47(4):984-90. doi: 10.1037/a0023819. PMID: 21744957.

29. Obregon-Cuesta AI, Mínguez-Mínguez LA, León-Del-Barco B, Mendo-Lázaro S, Fernández-Solana J, González-Bernal JJ, González-Santos J. Bullying in Adolescents: Differences between Gender and School Year and Relationship with Academic Performance. Int J Environ Res Public Health. 2022 Jul 29;19(15):9301. doi: 10.3390/ijerph19159301. PMID: 35954658; PMCID: PMC9368596.

30. Tran CD, Arredondo MM, Yoshida H. Early Executive Function: The Influence of Culture and Bilingualism. Biling (Camb Engl). 2019

Aug;22(4):714-732. doi: 10.1017/S1366728918000160. Epub 2018 Jul 5. PMID: 31558885; PMCID: PMC6762025.

31. Cole PM, Bruschi CJ, Tamang BL. Cultural differences in children's emotional reactions to difficult situations. Child Dev. 2002 May-Jun;73(3):983-96. doi: 10.1111/1467-8624.00451. PMID: 12038564.

32. Shenkin SD, Starr JM, Deary IJ. Birth weight and cognitive ability in childhood: a systematic review. Psychol Bull. 2004;130(6):989–1013. doi: 10.1037/0033-2909.130.6.989

33. Poulsen G, Wolke D, Kurinczuk JJ, Boyle EM, Field D, Alfirevic Z, Quigley MA. Gestational age and cognitive ability in early childhood: a population-based cohort study. Paediatr Perinat Epidemiol. 2013 Jul;27(4):371-9. doi: 10.1111/ppe.12058. Epub 2013 May 28. PMID: 23772939.

34. Alterman N, Johnson S, Carson C, Petrou S, Kurinzcuk JJ, Macfarlane A, Boyle E, Quigley MA. Gestational age at birth and academic attainment in primary and secondary school in England: Evidence from a national cohort study. PLoS One. 2022 Aug 17;17(8):e0271952. doi: 10.1371/journal.pone.0271952. PMID: 35976808; PMCID: PMC9385000.

35. Liu Y, Mendonça M, Bartmann P, Wolke D. Very preterm birth and trajectories of domain-specific self-concept from childhood into adulthood. Dev Psychopathol. 2022 Dec;34(5):1926-1937. doi: 10.1017/S0954579421000432. Epub 2021 Aug 20. PMID: 34414877.

36. Domellöf E, Johansson AM, Farooqi A, Domellöf M, Rönnqvist L. Risk for Behavioral Problems Independent of Cognitive Functioning in Children

Born at Low Gestational Ages. Front Pediatr. 2020 Jun 26;8:311. doi: 10.3389/fped.2020.00311. PMID: 32676487; PMCID: PMC7333186.

37. Coplan R.J., Barber A.M., Lagacé-Séguin D.G. The role of child temperament as a predictor of early literacy and numeracy skills in preschoolers. Early Childhood Research Quarterly. 1999; 14*:* 537-553*.* https://doi.org/10.1016/S0885-2006(99)00025-3

38. Hedges A, Corman H, Noonan K, Reichman NE. Gestational Age at Term and Educational Outcomes at Age Nine. Pediatrics. 2021 Aug;148(2):e2020021287. doi: 10.1542/peds.2020-021287. Epub 2021 Jul 9. PMID: 34244451; PMCID: PMC8344335.

39. Abulizi X, Pryor L, Michel G, Melchior M, van der Waerden J; EDEN Mother–Child Cohort Study Group. Temperament in infancy and behavioral and emotional problems at age 5.5: The EDEN mother-child cohort. PLoS One. 2017 Feb 15;12(2):e0171971. doi: 10.1371/journal.pone.0171971. PMID: 28199415; PMCID: PMC5310866.

40. Susa-Erdogan G, Benga O, Albu-Răduleț M, Macovei T. Child temperament and child-teacher relationship quality: Implications for children's emotional functioning during preschool period. Front Psychol. 2022 Nov 10;13:992292. doi: 10.3389/fpsyg.2022.992292. PMID: 36438346; PMCID: PMC9686403

41. Dennis E, Manza P, Volkow ND. Socioeconomic status, BMI, and brain development in children. Transl Psychiatry. 2022 Jan 24;12(1):33. doi: 10.1038/s41398-022-01779-3. PMID: 35075111; PMCID: PMC8786961.

42. Afzal AS, Gortmaker S. The Relationship between Obesity and Cognitive Performance in Children: A Longitudinal Study. Child Obes. 2015 Aug;11(4):466-74. doi: 10.1089/chi.2014.0129. PMID: 26258562; PMCID: PMC4528984.

43. Bacchini D, Licenziati MR, Affuso G, Garrasi A, Corciulo N, Driul D et al. The Interplay among BMI z-Score, Peer Victmization, and Self-Concept in Outpatient Children and Adolescents with Overweight or Obesity. Child Obes. 2017 Jun;13(3):242-249. doi: 10.1089/chi.2016.0139. Epub 2017 Mar 7. PMID: 28267352.

44. Griffiths LJ, Dezateux C, Hill A. Is obesity associated with emotional and behavioural problems in children? Findings from the Millennium Cohort Study. Int J Pediatr Obes. 2011 Jun;6(2-2):e423-32. doi: 10.3109/17477166.2010.526221. Epub 2010 Nov 30. PMID: 21114457; PMCID: PMC3465802.

45. Ali A, Al-Ani O, Al-Ani F. Children's behaviour and childhood obesity. Pediatr Endocrinol Diabetes Metab. 2024;30(3):148-158. doi: 10.5114/pedm.2024.142586. PMID: 39451187; PMCID: PMC11538919.

46. Yu W, Yan H.X. Effects of Siblings on Cognitive and Sociobehavioral Development: Ongoing Debates and New Theoretical Insights. American Sociological Review. 2023; 88(6):1002-1030. https://doi.org/10.1177/00031224231210258

47. Chavda K, Nisarga V. Single Parenting: Impact on Child's Development. Journal of Indian Association for Child and Adolescent Mental Health. 2023;19(1):14-20. doi:10.1177/09731342231179017

48. Daryanani I, Hamilton JL, Abramson LY, Alloy LB. Single Mother Parenting and Adolescent Psychopathology. J Abnorm Child Psychol. 2016 Oct;44(7):1411-23. doi: 10.1007/s10802-016-0128-x. PMID: 26767832; PMCID: PMC5226056.

49. McAdams TA, Rijsdijk FV, Narusyte J, Ganiban JM, Reiss D, Spotts E et al. Associations between the parent-child relationship and adolescent self-worth: a genetically informed study of twin parents and their adolescent children. J Child Psychol Psychiatry. 2017 Jan;58(1):46-54. doi: 10.1111/jcpp.12600. Epub 2016 Jul 18. PMID: 27426633; PMCID: PMC5215430.

50. McFarlane AC. The relationship between patterns of family interaction and psychiatric disorder in children. Aust N Z J Psychiatry. 1987 Sep;21(3):383-90. doi: 10.1080/00048678709160935. PMID: 3435378.

51. Maternal depression and child development. Paediatr Child Health. 2004 Oct;9(8):575-598. doi: 10.1093/pch/9.8.575. PMID: 19680490; PMCID: PMC2724169.

52. Shao Y, Kang S. The Link Between Parent–Child Relationship and Learning Engagement Among Adolescents: The Chain Mediating Roles of Learning Motivation and Academic Self-Efficacy. Frontiers in Education [Internet] 2022;7. Available from: https://www.frontiersin.org/articles/10.3389/feduc.2022.854549/pdf

53. Central Statistics Office (CSO). Data from: Growing up In Ireland Cohort ’08 (Infant Cohort) Wave 3 - 5 years, 2013. [data file] Irish Social Science Data Archive SN: 0019-03; 2015 [cited 2023 Jan 9]. Available from:

http://www.ucd.ie/issda/data/GUIInfant/GUIInfantWave3

54. Central Statistics Office (CSO). Data from: Growing up in Ireland Cohort '08 (Infant Cohort) Wave 5 at 9 years, 2017/2018. [data file]. Irish Social Science Data Archive SN: 0019-05; 2020 [cited 2023 Jan 9]. Available from: www.ucd.ie/issda/data/guiinfant/guiinfantwave5

55. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ et al. STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Int J Surg. 2014 Dec;12(12):1500-24. doi: 10.1016/j.ijsu.2014.07.014. Epub 2014 Jul 18. PMID: 25046751.

56. Quail A, Williams J, McCrory C, Murray A, Thornton M. A Summary Guide to Wave 1 of the Infant Cohort (At 9 months) of Growing Up in Ireland [Internet]. 2011 [cited 2025 Jan 10]. Available from: https://www.growingup.gov.ie/pubs/Summary-Guide_Infant-Cohort_Wave-1.pdf

57. Quail A, O'Reilly C, Watson D, McNamara E, O'Mahony D, Murray A. Growing Up in Ireland A Summary Guide to Cohort '08 at 9 Years [Internet]. 2019 [cited 2024 Jul 6]. Available from: https://www.growingup.gov.ie/pubs/Summary-Guide-Infant-Cohort-Wave-5.pdf

58. Murray A, Williams J, Quail A, Neary M, Thornton M. Growing Up in Ireland A Summary Guide to Cohort '08 at 5 Years [Internet]. 2015 [cited 2024 Jul 6]. Available from: https://www.growingup.gov.ie/pubs/Summary-Guide-_Infant-

Cohort_Wave-3.pdf

59. Elliott CD, Smith P, McCulloch K. British Ability Scales (2nd ed.). England: NFER-Nelson; 1996.

60. McGrew KS, Flanagan DP. The intelligence test desk reference (ITDR): Gf-Gc cross-battery assessment. America: Allyn & Bacon; 1998

61. Flanagan DP, Dixon SG. The Cattell-Horn-Carroll Theory of Cognitive Abilities. In Encyclopedia of Special Education [Internet]. 2014 [cited Jan 31 2025]. Available from https://doi.org/10.1002/9781118660584.ese0431

62. Thornton M, Williams J. Growing Up in Ireland Report on the Pilot Phase of Wave Three, Infant Cohort (at 5 Years of Age) [Internet]. 2016 [cited 2024 Jul 6] Available from: https://www.growingup.gov.ie/pubs/BKMNEXT331.pdf

63. Connelly R. Millennium Cohort Study Data Note: Interpreting Test Scores [Internet]. 2013 [cited 2023 Jan 4]. Available from: https://cls.ucl.ac.uk/wp-content/uploads/2017/07/MCS-data-note-20131-Test-Scores-Roxanne-Connelly.pdf

64. Murray A, Williams J, Gallagher S, Thornton M. Report on the Pilot Phase of Wave Three, Infant Cohort (at 9 Years of Age) [Internet]. 2020 [cited 2025 Jan 10]. Available from: https://www.growingup.gov.ie/pubs/Cohort08at9-Pilot-Report-2020-2.pdf

65. van den Heuvel M, Jansen DEMC, Stewart RE, Smits-Engelsman BCM, Reijneveld SA, Flapper BCT. How reliable and valid is the teacher version

of the Strengths and Difficulties Questionnaire in primary school children? PLoS One. 2017 Apr 28;12(4):e0176605. doi: 10.1371/journal.pone.0176605. PMID: 28453573; PMCID: PMC5409073.

66. YouthinMind. Scoring the Strengths and Difficulties Questionnaire for ages 4-17 [Internet]. 2016 [cited 2023 Jan 30]. Available from: https://www.sdqinfo.org/py/sdqinfo/c0.py

67. Sanson AV, Smart DF, Prior M, Oberklaid F, Pedlow R. The Structure of Temperament From Age 3 to 7 Years: Age, Sex, and Sociodemographic Influences [Internet]. Merrill Palmer Q. 1994 Aug 17 [cited 2025 Jan 5];40(2):233–52. Available from: http://www.jstor.org/stable/23087863

68. Radloff LS. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. Appl Psychol Meas. 1977 Jun 1;1(3):385–401. https://doi.org/10.1177/014662167700100306

69. Braveman PA, Cubbin C, Egerter S, Chideya S, Marchi KS, Metzler M, Posner S. Socioeconomic status in health research: one size does not fit all. JAMA. 2005 Dec 14;294(22):2879-88. doi: 10.1001/jama.294.22.2879. PMID: 16352796.

70. Marsh HW, Craven RG. Reciprocal Effects of Self-Concept and Performance From a Multidimensional Perspective: Beyond Seductive Pleasure and Unidimensional Perspectives. Perspect Psychol Sci. 2006 Jun;1(2):133-63. doi: 10.1111/j.1745-6916.2006.00010.x. PMID: 26151468.

71. Graham LJ, Gillett-Swan J, Killingly C, Van Bergen P. Does It Matter If Students (Dis)like School? Associations Between School Liking, Teacher

and School Connectedness, and Exclusionary Discipline. Front Psychol. 2022 Mar 3;13:825036. doi: 10.3389/fpsyg.2022.825036. PMID: 35310270; PMCID: PMC8927887.

72. Gubbels J, van der Put CE, Assink M. Risk Factors for School Absenteeism and Dropout: A Meta-Analytic Review. J Youth Adolesc. 2019 Sep;48(9):1637-1667. doi: 10.1007/s10964-019-01072-5. Epub 2019 Jul 15. PMID: 31312979; PMCID: PMC6732159.

73. Lansford JE, Dodge KA, Pettit GS, Bates JE. A Public Health Perspective on School Dropout and Adult Outcomes: A Prospective Study of Risk and Protective Factors From Age 5 to 27 Years. J Adolesc Health. 2016 Jun;58(6):652-8. doi: 10.1016/j.jadohealth.2016.01.014. Epub 2016 Mar 19. PMID: 27009741; PMCID: PMC4877222.

74. Freudenberg N, Ruglis J. Reframing school dropout as a public health issue. Prev Chronic Dis. 2007 Oct;4(4):A107. Epub 2007 Sep 15. PMID: 17875251; PMCID: PMC2099272.

75. Duckworth AL, Quirk A, Gallop R, Hoyle RH, Kelly DR, Matthews MD. Cognitive and noncognitive predictors of success. Proc Natl Acad Sci U S A. 2019 Nov 19;116(47):23499-23504. doi: 10.1073/pnas.1910510116.

76. Gardner H. Taking a multiple intelligences (MI) perspective. Behav Brain Sci. 2017 Jan;40:e203. doi: 10.1017/S0140525X16001631.

77. Larson K, Russ SA, Nelson BB, Olson LM, Halfon N. Cognitive ability at kindergarten entry and socioeconomic status. Pediatrics. 2015 Feb;135(2):e440-8. doi: 10.1542/peds.2014-0434. Epub 2015 Jan 19. PMID: 25601983.

78. Lynch K, Baker J. Equality in education: An equality of condition perspective. Theory and Research in Education. 2005;3(2):131-164. https://doi.org/10.1177/1477878505053298

79. Love JM, Kisker EE, Ross C, Raikes H, Constantine J, Boller K, Brooks-Gunn J, Chazan-Cohen R, Tarullo LB, Brady-Smith C, Fuligni AS, Schochet PZ, Paulsell D, Vogel C. The effectiveness of early head start for 3-year-old children and their parents: lessons for policy and programs. Dev Psychol. 2005 Nov;41(6):885-901. doi: 10.1037/0012-1649.41.6.88. PMID: 16351335.

80. Bradbury A, Jarvis P, Nutbrown C, Roberts-Holmes G, Stewart N, Whitebread D. Baseline assessment: Why it doesn't add up. More Than A Score [Internet]. 2018 [cited 2024 Jul; 7]. Available from: https://morethanascorecampaign.files.wordpress.com/2018/02/neu352-baseline-a4-16pp-crop.pdf

81. Duggan C, Irvine AD, O'B Hourihane J, Kiely ME, Murray DM. ASQ-3 and BSID-III's concurrent validity and predictive ability of cognitive outcome at 5 years. Pediatr Res. 2023 Oct;94(4):1465-1471. doi: 10.1038/s41390-023-02528-y. Epub 2023 Feb 25. PMID: 36841883; PMCID: PMC10589087.

82. Bowe AK, Hourihane J, Staines A, Murray DM. The predictive value of the ages and stages questionnaire in late infancy for low average cognitive ability at age 5. Acta Paediatr. 2022 Jun;111(6):1194-1200. doi: 10.1111/apa.16309. Epub 2022 Mar 3. PMID: 35202483; PMCID: PMC9314849.

83. Jeffers G, Lillis C. Responding to educational inequality in Ireland; Harnessing teachers' perspectives to develop a framework for professional conversations. Educational Studies. 2021; 50(2):139–165. doi.org/10.1080/03055698.2021.1931043

84. Cook J. An investigation of the validity of the British Ability Scales with respect to the Wechsler Intelligence Scale for Children-Revised and the wide range achievement test-revised on a group of Canadian children. Br J Educ Psychol. 1988 Jun;58 (Pt 2):212-6. doi: 10.1111/j.2044-8279.1988.tb00895.x. PMID: 3166936.

85. Breese AC, Nickerson AB, Lemke M, Mohr R, Heidelburg K, Fredrick S, Allen K. Examining Implicit Biases of Pre-Service Educators Within a Professional Development Context. Contemp Sch Psychol. 2023 Feb 23:1-16. doi: 10.1007/s40688-023-00456-6. Epub ahead of print. PMID: 36855339; PMCID: PMC9948799.

86. Jussim L, Harber KD. Teacher expectations and self-fulfilling prophecies: knowns and unknowns, resolved and unresolved controversies. Personal Soc Psychol Rev an Off J Soc Personal Soc Psychol Inc. 2005;9(2):131–55. https://doi.org/10.1207/s15327957pspr0902_3

Chapter 5 The Predictive Value of The Ages and Stages Questionnaire in Late Infancy for Low Average Cognitive Ability at Age 5

Andrea K. Bowe[1], Hourihan J[1,2], Anthony Staines[3], Deirdre M. Murray[1,4]

[1] INFANT Research Centre, University College Cork, Ireland.

[2] Paediatrics, Royal College of Surgeons in Ireland, Dublin, Ireland.

[3] School of Nursing, Psychotherapy and Community Health, Dublin City University, Dublin, Ireland.

[4] Department of Paediatrics, School of Medicine, University College Cork, Cork, Ireland.

Published in Acta Paediatrica

June 2022

5.1.Abstract

Aim

This retrospective, longitudinal study examined the predictive value of the Ages and Stages Questionnaire (ASQ) in late infancy for identifying children who progressed to have low cognitive ability at 5 years of age.

Methods

The ASQ was performed on 755 participants from the Irish BASELINE birth cohort at 24 or 27 months of age. Intelligence quotient (IQ) was measured at age 5 with the Kaufmann Brief Intelligence Test, Second Edition and low cognitive ability was defined as a score more than 1 standard deviation below the mean. The predictive value of the ASQ was examined, together with other factors associated with low cognitive ability at 5 years.

Results

When the ASQ was performed at 24 or 27 months, the overall sensitivity for identifying low cognitive ability at 5 years was 20.8% and the specificity was 91.1%. Using a total score cut-off point increased the sensitivity to 46.6% and 71.4% at 24 and 27 months, but specificity fell to 74.1% and 67.2%, respectively. After adjusting for ASQ performance, maternal education and family income were strongly associated with cognitive outcomes at 5 years.

Conclusion

The ASQ did not detect the majority of children with low cognitive ability at age 5. Alternative methods need investigation.

5.2. Introduction

Children with low average cognitive ability represent a substantial proportion of children with often unrecognised difficulties.[1,2] Early deficits in cognitive ability are associated with later adverse outcomes in many domains of life including mental and physical health, educational attainment, and substance use.[3-7] Despite the well-established adverse impacts, the approximately 14% of children whose cognitive ability falls into this lower range, receive little attention in scientific literature and often little recognition in ordinary life.[2]

Defining this level of cognitive ability can be difficult. The International Classification of Diseases, 11th Edition clearly defines a disorder of intellectual development as one that is characterised by significantly below average intellectual functioning and adaptive behaviour that are approximately two or more standard deviations (SD) below the mean on appropriately normed, individually administered standardised tests.[8] The intellectual functioning of some children is substantially below the population average, but they do not meet this criteria. They remain in a grey area between average intellectual functioning and disordered intellectual development. Children with more significant intellectual developmental disorders may be more likely to display marked deficits that can be detected during routine developmental screening. However, children in the grey area may not display overt signs during infancy. Their cognitive difficulties may go unrecognised until they experience repeated failures or behavioural problems at school.[1,7,9]

Early interventions are best focused on the early years of life, a fundamental period in cognitive development, and the aim of routine developmental screening is to identify infants who would benefit from such interventions.[10] The American

Academy of Pediatrics recommends that parent-reported measures should be considered during routine developmental screening and the Ages and Stages Questionnaire (ASQ) is one example.[11] The ASQ is a parent-reported questionnaire designed to discriminate between children with developmental delays and those with typical development.[12] The National Steering Group for the Revised Child Health Programme in Ireland has recommended its use during the 24-month public health nursing assessment.[13]

Studies have suggested that the ability of the ASQ to predict below average childhood IQ depends on the cohort and the cut-offs used.[14,15] The ASQ has high specificity at a general population level and may effectively identify children not at risk of neurodevelopmental disorders.[16] A significant shortcoming of many developmental screening tests is that they do not adequately account for the social environment to which the child will be exposed.[17] Charkulak et al examined the predictive value of the 36-month ASQ for IQ at age 5-6 years. They found that the risk of low IQ was similar for children with an ASQ score below the optimal cut-off and a high maternal education, and children with an ASQ score above the cut-off and a low maternal education level.[14]

This study examined the predictive value of the ASQ performed in late infancy for identifying those who progress to have low average cognitive ability at 5 years of age. We also assessed whether the 24-month or 27-month ASQs would have a better predictive value if an optimal cut-off , based on the total score, was used. Finally, we examined other important factors associated with lower than average cognitive ability in early childhood.

5.3. Methods

Participants

The study population consisted of mothers and infants from the Irish Cork BASELINE Birth Cohort Study, which started in 2008.[18] BASELINE recruited 1,583 of the 2,183 participants (73%) from the Screening for Pregnancy Endpoints study, a multi-centre prospective study of low-risk, primiparous women. The other 600 were recruited from the postnatal wards of Cork University Maternity Hospital. The subjects were followed up at 2, 6, 12 and 24 months, and 5 years. 1,102 children completed the ASQ Third Edition (ASQ-3) at either 24 or 27 months and 755 of these completed the Kaufmann Brief Intelligence test, Second Edition (KBIT-2) at the age of 5. These 755 participants formed the study cohort (Figure 5.1).

Ethical Approval

The Clinical Research Ethics Committee of Cork Teaching Hospitals provided ethical approval for the BASELINE study, which was registered with clinical trials.gov (ID NCT01498965) and was carried out in line with the Declaration of Helsinki. It also approved the secondary data analysis by the lead author. All the mothers provided informed consent for participation in the BASELINE study.

Measures

ASQ

The ASQ is a parent-completed questionnaire with 30 items covering 5 development domains: communication, gross motor, fine motor, problem solving and personal social skills.[12] The score is based on the parental response, which may be “yes” (10 points), “not yet” (0 points), or “sometimes” (5 points), to a statement about the child’s activity. Parents are asked to try the activities with their child, if

necessary, before answering. Total scores are calculated for each domain and missing answers may be replaced by the mean of the other answers, if there are no more than 2 missing items per domain. Age-specific statistically derived cut-off points have been established for onward referral. The ASQ has not been validated in Ireland and onward referral for further professional assessment is made based on an age-specific score, in any domain, that falls at least 2 SD below the mean in the reference population.[12] The BASELINE study mothers received an age-appropriate ASQ-3 and were asked to complete it before their next assessment at the research facility: 468 received the 24-month ASQ and 287 received the 27-month ASQ. The cut-offs for onward referral are outlined in Table A10.1, Appendix 10.

Cognitive ability

The KBIT-2 was administered by trained research nurses to measure cognitive ability at 5 years of age. It comprises three subtests: verbal knowledge and riddles, which measure verbal intelligence, and matrices, which measure non-verbal intelligence. The verbal and non-verbal scores are combined to provide a composite intelligent quotient (IQ) score. The Kaufmann test has been shown to be a reliable and valid measure of IQ in children and adults aged 4-90 years.[19] Children who scored more than 1 SD below the cohort mean were categorised as "Low Cognitive Ability" and those who scored above this as "Average or Above Cognitive Ability".

Child, Maternal, and Sociodemographic Variables

Child characteristics included gender and weeks of gestation. Maternal characteristics included age in years (at birth of child), smoking status in first trimester (smoker, non-smoker), weekly units of alcohol during the first trimester and their body mass index in kg/m^2 (measured at 15 weeks gestation).

Socioeconomic characteristics included total years of schooling and family income category (≥€106000, €85,000-€105,000, €64,000-€84,000, €43,000-€63,000, €21,000-€42,000, <€21,000).

Statistical analysis

The statistical analysis was conducted using R version 4.0.3 (R Foundation, Vienna, Austria).[20] The study population were described using counts and percentages for categorical data and means and standard deviations for continuous data. T-tests were used to compare the mean ASQ scores between the cognitive ability groups. The predictive value of the ASQ was examined by calculating its sensitivity, specificity, positive and negative predictive values and their 95% confidence intervals (CI). Previous studies have reported sensitivities ranging from 28-77% for predicting low cognitive performance, with results varying by the cohort tested and cut-offs used.[14,15]

The pROC package in R was used to generate receiver operating characteristic curves and calculate optimal cut-off points for the total ASQ Score using Youden's J statistic.[21] Univariable and multivariable logistic regression were used to examine child, maternal and socioeconomic predictors of low cognitive ability at 5 years. The findings are reported in line with the Strengthening the Reporting of Observational Studies in Epidemiology guidelines.[22]

5.4. Results

Description of Sample

The study sample consisted of 755 children who underwent the ASQ at either 24 or 27 months and the KBIT-2 at 5 years. Their characteristics are compared to those who

were not included due to missing data or attrition (Table 5.1). Children in the study sample had a higher mean birthweight and higher mean maternal age compared to those not included. A significantly lower proportion of those in the study sample had mothers who smoked in the first trimester, and a significantly higher proportion lived in a owner-occupied home. A higher proportion of children in the study sample were from the highest income backgrounds compared to those not included.

The ASQs were completed at a mean (sd) age of 25.1 (1.1) months, but the results are presented separately for 24 and 27 months, as the cut-off points and score distributions differ between the two questionnaires. The Kaufmann test was completed at a mean (sd) age of 60.8 (1.7) months. The mean composite IQ score was 104.3 and 13.4% scored more than 1 SD below this point.

Mean ASQ scores

We examined the mean ASQ scores for those with and without low cognitive ability at 5 years of age (Table A10.2, Appendix 10) and the predictive value of the ASQ (Table 5.2). Children with low cognitive ability had significantly lower total ASQ scores, with a mean difference of 16.6 points (95% CI 7.4-25.7) for the 24-month ASQ and 24.8 points (95% CI 8.1-41.5) for the 27-month ASQ.

Predictive value of the ASQ

The ASQ had a sensitivity of 20.8% (95% CI 13.6%-30.2%) for identifying children with low cognitive ability at 5 years and a high specificity of 91.1% (95% CI 88.6%-93.2%) for identifying those without low cognitive ability at 5 years. The smooth receiver operating characteristic curves for the 24 and 27 month ASQs are shown with their respective area under the curve values (Figure 5.2, Figure 5.3). The empirical ROC curves are provided in Figure A10.1 and Figure A10.2, Appendix 10. At 24 months, a

total ASQ cut-off of 249.5 increased the sensitivity to 46.6% (95% CI 35.6%-57.5%), but reduced the specificity to 74.1% (95% CI 69.8%-78.2%). At 27 months, a total ASQ cut-off of 244.5 increased the sensitivity to 71.4% (95% CI 53.6% - 89.3%) and reduced the specificity to 67.2% (95% CI 61.4% - 72.9%).

Association between ASQ failure and low cognitive ability

The results of the univariable and multivariable logistic regression analyses are shown in Table 5.3. After adjusting for maternal and socioeconomic factors, the odds of having low cognitive ability at 5 years were 2.9 times higher for those who failed at least one ASQ domain compared to those who did not (adjusted odds ratio (AOR) 2.94, 95% CI 1.51-5.56, p<0.001). Other significant predictors of low cognitive ability at age 5 were gestational age, weekly units of alcohol in the first trimester, total years of schooling, and family income. After adjustment, a 1-year increase in maternal schooling was associated with a 30% reduction in the odds of low cognitive ability at 5 years (AOR 0.69, 95% CI 0.49 – 0.96, p=0.026). Compared to those in the highest income category, the odds of experiencing low cognitive ability at age 5 were 5 times higher for children in the lowest income category (AOR 5.01, 95% CI 1.18-21.47, p=0.027).

5.5. Discussion

Our main aim was to examine the predictive value of the ASQ in late infancy for identifying children in the general population who progress to have low cognitive ability at 5 years of age. When the onward referral criterion of failing in a single domain was used, the overall sensitivity was 20.8%. The specificity was 91.1%, in keeping with previous literature[16], making it a useful tool for identifying those not at

risk. While the ASQ may identify children with significant domain specific developmental delay, for the purpose of identifying children with low cognitive ability in childhood, the ASQ at 24 or 27 months has an unacceptably low sensitivity, missing 80% of children who may benefit from earlier intervention.

The second aim was to determine whether an optimal cut-off point, based on the total ASQ score, could improve the tools predictive value. Total ASQ scores are not currently used for onward referral, but studies suggest this may be a useful additional measure.[15,23] Using optimal cut-off points of 249.5 and 244.5 on the 24-month and 27-month ASQs improved sensitivity to 46.6% and 71.4%, respectively, but reduced specificity. In population-based screening a careful balance must be struck between the sensitivity and specificity of a test. This improved sensitivity would mean that almost half of all children who progress to low cognitive ability at age 5 could be identified at 24-months. However, the accompanying reduction in specificity would mean that at 24-months approximately 1 in every 4 children not at risk of low cognitive ability would be falsely identified as positive. This has significant ethical and societal implications. Labelling a child incorrectly as at risk of low cognitive ability may result in stigmatisation, or may lead to lower expectations of ability which can adversely affect evolving capabilities.[24] Additionally, a criterion for implementing any population-based screening programme is that there must be an available course of action with regard to further diagnosis, treatment, or intervention.[25] Implementing this cut-off could threaten to overwhelm the finite resources dedicated to early intervention programmes for at risk children. Implications would be similar with implementation of the 27-month total score cut-off.

The third aim was to examine the child, maternal, and socioeconomic characteristics associated with low cognitive ability at 5 years. The findings were consistent with studies that have demonstrated strong, consistent associations between families' socioeconomic characteristics and children's cognitive outcomes.[17] The odds of low cognitive ability were significantly higher for children whose mothers had fewer years of schooling and who were from lower income backgrounds. This finding again highlights the pivotal role socioenvironmental factors have in the cognitive development of a child.

This study investigated one potential screening method, the ASQ, for the purpose of early identification of children who progress to have low cognitive ability. It was found that, as it is currently used, this tool will not identify the majority of children with low cognitive ability at school entry. Other tools exist for identifying children at risk of low cognitive ability, the most widely used being the Bayley Scales of Infant Development (Bayley). This developmental assessment tool is usually administered by psychologists, paediatricians, or physiotherapists and takes 30-70 minutes to complete, with extra time needed for scoring.[26] It is resource intensive and is generally used for high risk populations, such as preterm or low birth weight infants.[27] Studies of its use in the general population are limited, but evidence suggests that if it was performed in early life, particularly the first 24 months, it would have modest correlations with later cognitive ability.[28]

Our study has limitations. The research question focussed on the use of the ASQ at a single time point in late infancy, but the authors acknowledge that developmental assessment should be a continuous process performed at multiple timepoints throughout infancy and childhood. The outcome was cognitive ability at

age 5, but it is well recognised that the social and emotional competence of a child at this age are equally important determinants of outcomes in childhood.[29] Cohort studies are subject to selection bias, and those who participate and remain engaged tend to systematically differ from those who do not.

5.6.Conclusion

Studies have shown that the ASQ has acceptable performance characteristics for identifying significant neurodevelopmental disorders. However, for the purpose of identifying children in the general population who progress to have low average cognitive ability at 5 years of age, this study found it is not a useful tool. Alternative methods of early identification are required. Further research into a systematic and robust method of identifying high-risk children, based on environmental, genetic, biological, social, and demographic factors, as well as the interactions between these factors, should be a public health priority.

Author contributions

AB – Conceptualization, methodology, formal analysis, writing – original draft preparation, writing – review and editing

JH – Data acquisition, writing – review and editing

AS – Methodology, formal analysis, writing – review and editing

DM – Conceptualization, data acquisition, methodology, writing – review and editing

Table 5.4 Characteristics of children in study population compared to those excluded due to attrition or missing data

	Children included n=755 (34.6%)	Children excluded n=1,428 (65.4%)	p-value
Child characteristics			
Male	381 (50.5)	719 (50.4)	0.996[a]
Birthweight, mean (SD), g	3538 (488)	3454 (538)	<0.001[b]
Gestational age, mean (SD), wk	39.6 (1.4)	39.4 (1.8)	0.003[b]
Maternal characteristics			
Maternal age, mean (SD), y	30.8 (3.9)	29.5 (4.7)	<0.001[b]
Smoking status in first trimester			
Smoker	126 (21.4)	213 (23.0)	0.003[a]
Body Mass Index – n, mean (SD), kg/m^2	588, 25.0 (4.2)	995, 24.8 (4.1)	0.577[b]
Socioeconomic characteristics			
Total years of schooling, n, mean (SD)	755, 13.4 (0.7)	1,197, 13.2 (0.9)	<0.001[b]
University completed	428 (56.7)	631 (52.8)	0.099[a]
Accommodation type			
Own house/flat	601 (79.6)	795 (66.5)	
Private rental	124 (16.4)	329 (27.5)	
Council/government rental	12 (1.6)	36 (3.0)	
Other	18 (2.4)	36 (3.0)	<0.001[a]
Family income category			
<21k	27 (3.7)	129 (11.3)	
21-42k	126 (17.2)	235 (20.6)	
43-63k	169 (23.1)	267 (23.4)	
64-84k	189 (25.8)	226 (19.8)	
85-105k	112 (15.3)	162 (14.2)	
≥106k	110 (15.0)	123 (10.8)	<0.001[a]
Ages and Stages Questionnaire			
Completed 24 month	468 (62.0)	254 (62.3)	
Completed 27 month	287 (38.0)	154 (37.7)	1.000[a]
Total ASQ Score, n, mean (SD)	755, 255.0 (31.0)	412, 246.3 (44.9)	<0.001[b]
Fail in ASQ domain			
Communication	17 (2.5)	27 (6.7)	<0.001
Fine Motor	18 (2.4)	24 (6.2)	0.002
Gross Motor	20 (2.6)	15 (3.7)	0.412
Problem solving	24 (3.2)	23 (5.9)	0.040
Social Skills	25 (3.3)	33 (8.3)	<0.001
At least one domain	79 (10.5)	70 (18.6)	<0.001[a]
KBIT-2 IQ Test			
Composite IQ Score, n, mean (SD)	755, 104.6 (8.7)	315, 103.6 (9.0)	0.098[b]

[a]Pearson's Chi-squared test, [b]Welch Two Sample t-test

Table 5.5 Predictive value of ASQ at 24/27 months for identifying children with low cognitive ability at age 5 years

	Sensitivity	Specificity	PPV	NPV
	% (95% Confidence Interval)			
One fail criterion	20.8 (13.6-30.2)	91.1 (88.6-93.2)	26.6 (17.6-37.9)	88.2 (85.5-90.5)
Fail Communication	6.9 (3.1-14.2)	98.0 (96.5-98.9)	35.0 (16.3-59.1)	87.2 (84.5-89.5)
Fail Gross Motor	7.9 (3.7-15.5)	98.2 (96.7-99.0)	40.0 (20.0-63.6)	87.4 (84.5-89.6)
Fail Fine Motor	0.9 (0.5-6.2)	97.4 (95.8-98.4)	5.6 (0.2-29.4)	86.4 (83.7-88.8)
Fail Problem Solving	7.9 (3.7-15.5)	97.4 (95.8-98.4)	32.0 (15.7-53.6)	87.3 (84.6-89.6)
Fail Social Skills	10.9 (5.8-19.0)	97.9 (96.4-98.8)	44.0 (25.0-64.7)	87.7 (85.0-89.9)

Abbreviations: PPV, positive predictive value; NPV, negative predictive value.

Table 5.6 Univariable and multivariable logistic regression model examining factors associated with low cognitive ability at age 5

	Low n=101	Average/above n=653	Odds ratio (95% CI)	p	Adjusted odds ratio (95% CI)	p
Child characteristics						
ASQ One Fail	21 (20.8)	58 (8.9)	2.69 (1.53 – 4.61)	<0.001	2.94 (1.51 – 5.56)	0.001
Gender – female	44 (43.6)	330 (50.5)	0.76 (0.49 – 1.15)	0.193	1.08 (0.64-1.82)	0.781
Birthweight*, mean (SD), g	3,498 (529)	3,545 (482)	1.00 (0.99 – 1.00)	0.368		
Gestational age, mean (SD), wk	39.4 (1.8)	39.6 (1.3)	0.87 (0.76 – 1.01)	0.052	0.84 (0.71 -0.99)	0.036
Maternal characteristics						
Maternal age, mean (SD), y	30.5 (4.5)	30.8 (3.8)	0.98 (0.92 – 1.04)	0.457	1.02 (0.95-1.09)	0.593
Smoking status						
Non-Smoker	54 (71.1)	407 (79.6)	Ref		Ref	
Smoker	22 (28.9)	104 (20.4)	1.59 (0.91 – 2.71)	0.091	1.11 (0.59 – 2.01)	0.742
Weekly units of alcohol in first trimester, mean (SD)	6.1 (8.3)	4.1 (4.6)	1.06 (1.02 – 1.10)	0.003	1.06 (1.02 – 1.11)	0.004
Body Mass Index, mean (SD), kg/m^2	25.8 (4.9)	24.8 (4.0)	1.05 (1.00 – 1.11)	0.064	1.03 (0.97 – 1.09)	0.267
Socioeconomic characteristics						
Total years of schooling, mean (SD)	13.2 (0.9)	13.4 (0.7)	0.68 (0.52 – 0.89)	0.004	0.69 (0.49 – 0.96)	0.026
Income Category						
>106k	9 (9.2)	101 (15.9)	Ref		Ref	
85-105k	10 (10.2)	102 (16.1)	1.10 (0.43 – 2.88)	0.842	1.72 (0.59 – 5.41)	0.325
64-84k	25 (25.5)	164 (25.8)	1.71 (0.79 – 4.01)	0.189	1.93 (0.74 – 5.65)	0.198
43-63k	26 (26.5)	143 (22.5)	2.04 (0.95 – 4.78)	0.080	2.29 (0.89 – 6.72)	0.103
21-42k	21 (21.4)	105 (16.5)	2.24 (1.01 – 5.38)	0.055	2.49 (0.86 – 7.88)	0.100
<21k	7 (7.1)	20 (3.1)	3.93 (1.27 – 11.82)	0.015	5.01 (1.18 –21.47)	0.027

*In logistic regression model, a one unit increase corresponds to a 10g increase in birthweight. Abbreviations: CI, confidence interval; p, p-value.

Figure 5.4 Flowchart of study population from the BASELINE birth cohort

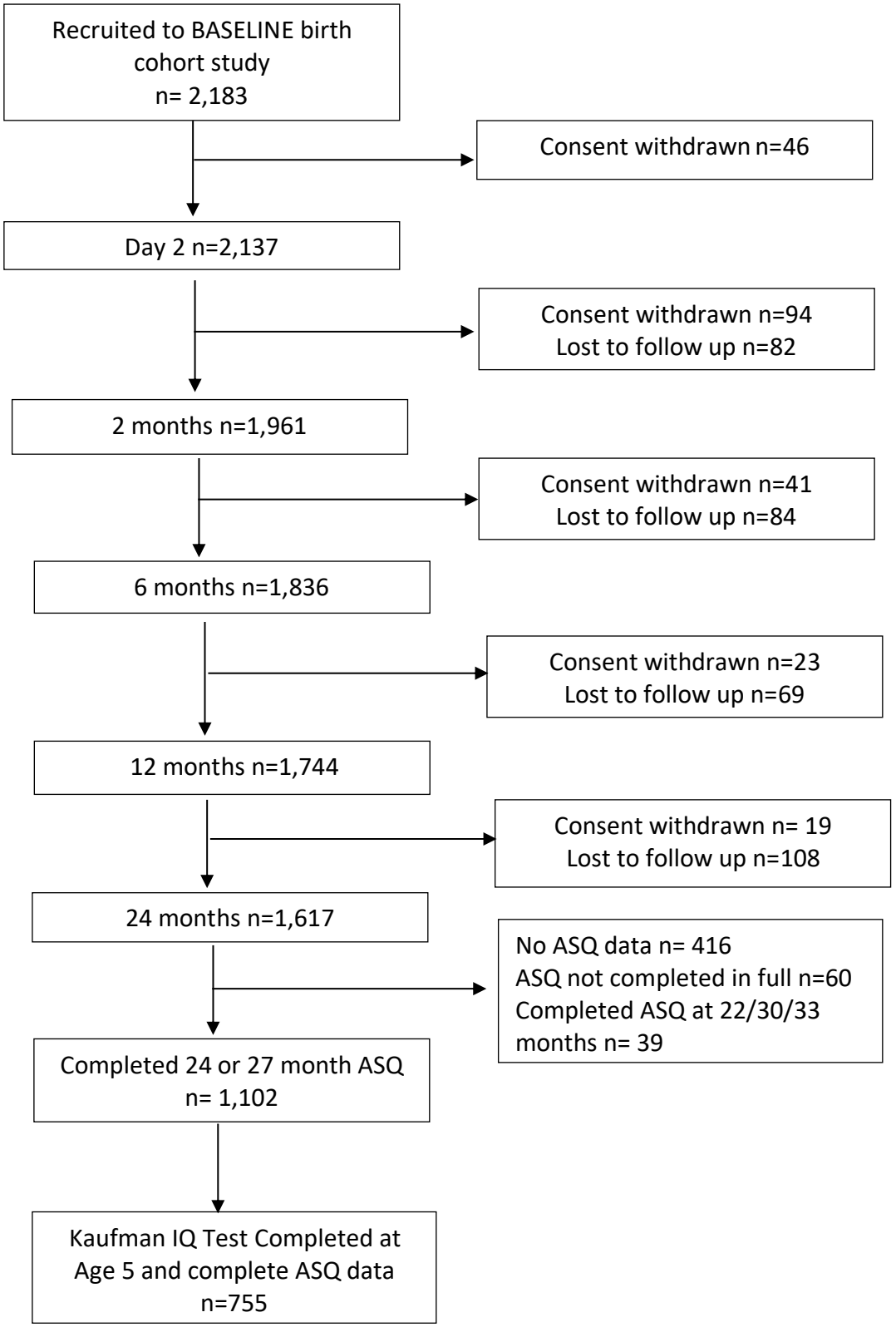


Figure 5.5 Receiver operating characteristic curve of the 24-month ASQ for predicting low cognitive ability at age 5 years

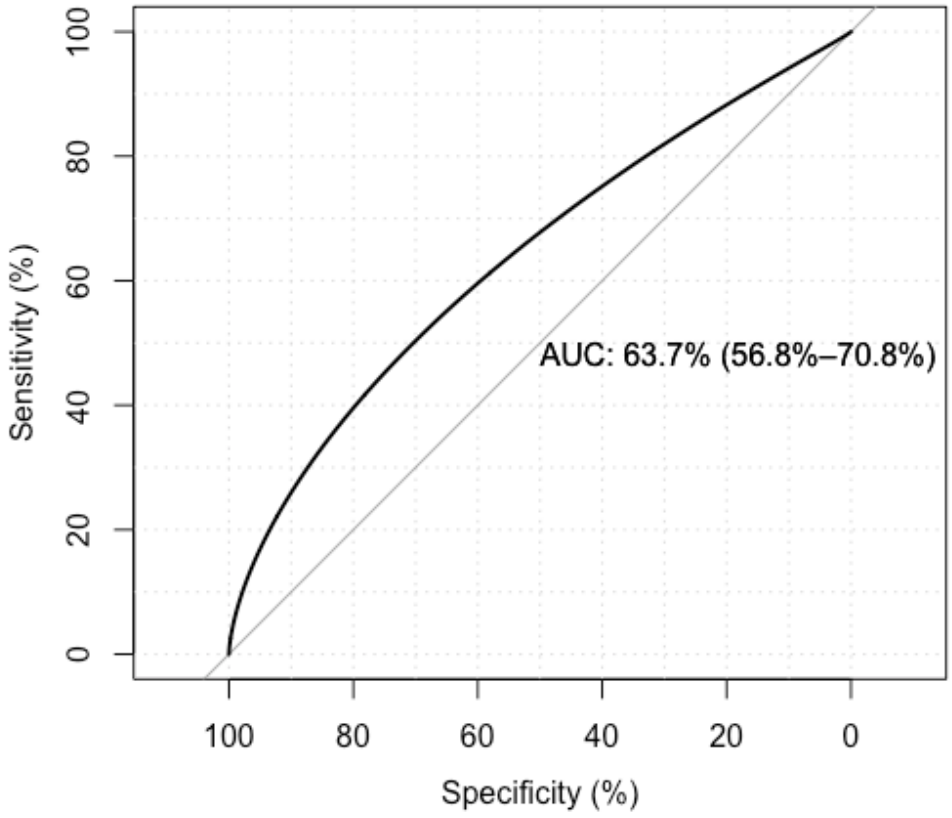


Figure 5.2 Legend

ROC curve performed using total ASQ score at 24-months for the prediction of a composite IQ score more than 1 standard deviation below the cohort mean at age 5 years. A smoothed ROC curve was generated using the "pROC" package in R. Analysis includes n=394 controls and n=73 cases. The Area Under the Curve was 0.64, 95% Confidence Interval 0.57-0.71.

Figure 5.6 Receiver operating characteristic curve of the 27-month ASQ for predicting low cognitive ability at age 5 years

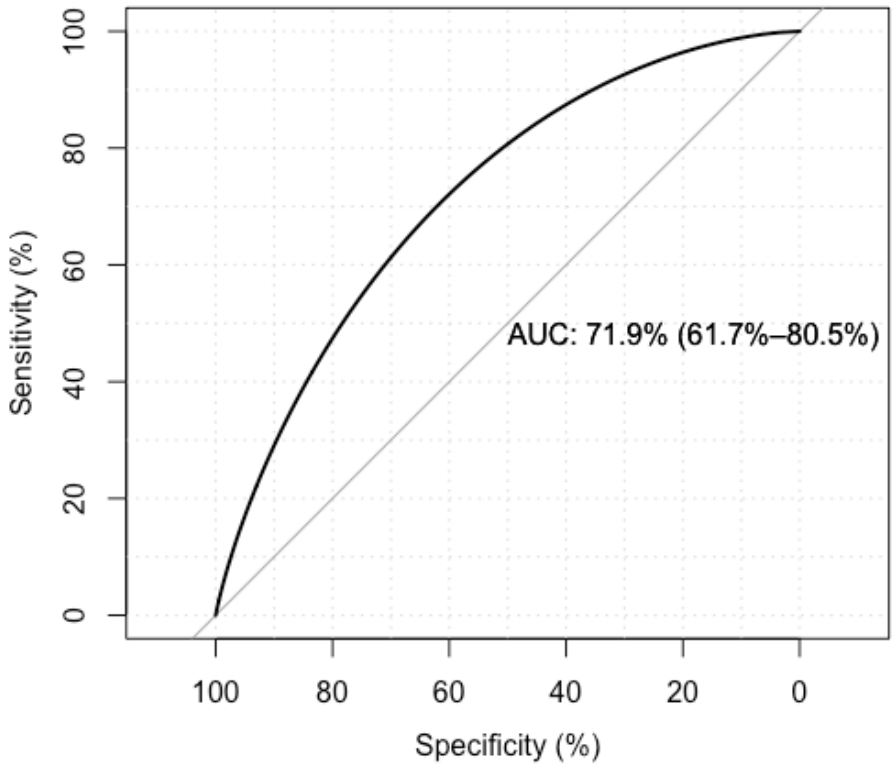


Figure 5.3 Legend
ROC curve performed using total ASQ score at 27-months for the prediction of a composite IQ score more than 1 standard deviation below the cohort mean at age 5 years. A smoothed ROC curve was generated using the "pROC" package in R. Analysis includes n=259 controls and n=28 cases. The Area Under the Curve was 0.72, 95% Confidence Interval 0.62-0.81.

5.7. References

1. Wieland J, Zitman FG. It is time to bring borderline intellectual functioning back into the main fold of classification systems. BJPsych Bull. 2016 Aug;40(4):204-6. doi: 10.1192/pb.bp.115.051490. PMID: 27512590; PMCID: PMC4967780.
2. Nouwens PJG, Lucas R, Embregts PJCM, van Nieuwenhuizen C. In plain sight but still invisible: A structured case analysis of people with mild intellectual disability or borderline intellectual functioning. J Intellect Dev Disabil. 2017 Jan;42(1):36–44. https://doi.org/10.3109/13668250.2016.1178220
3. Deary IJ, Strand S, Smith P, Fernandes C. Intelligence and educational achievement. Intelligence. 2007;35(1):13–21. https://doi.org/10.1016/j.intell.2006.02.001
4. Wraw C, Deary IJ, Gale CR, Der G. Intelligence in youth and health at age 50. Intelligence. 2015 Nov-Dec;53:23-32. doi: 10.1016/j.intell.2015.08.001.
5. Schijven EP, Didden R, Otten R, Poelen EAP. Substance use among individuals with mild intellectual disability or borderline intellectual functioning in residential care: Examining the relationship between drinking motives and substance use. J Appl Res Intellect Disabil. 2019 Jul;32(4):871-878. doi: 10.1111/jar.12578. Epub 2019 Mar 7. PMID: 30844128; PMCID: PMC6850364.
6. Alesi M, Rappo G, Pepi A. Emotional Profile and Intellectual Functioning: A Comparison Among Children With Borderline Intellectual Functioning, Average Intellectual Functioning, and Gifted Intellectual Functioning. Sage

Open. 2015 Jul 1;5(3). https://doi.org/10.1177/2158244015589995

7. Fernell E, Ek U. Borderline intellectual functioning in children and adolescents - insufficiently recognized difficulties. Acta Paediatr. 2010 May;99(5):748-753. doi: 10.1111/j.1651-2227.2010.01707.x. Epub 2010 Feb 9. PMID: 20146756.

8. World Health Organisation. ICD-11 - Mortality and Morbidity Statistics [Internet]. [cited 2021 Apr 18]. Available from: https://icd.who.int/browse11/l-m/en#/http%3A%2F%2Fid.who.int%2Ficd%2Fentity%2F605267007

9. Karande S, Kanchan S, Kulkarni M. Clinical and psychoeducational profile of children with borderline intellectual functioning. Indian J Pediatr. 2008 Aug;75(8):795-800. doi: 10.1007/s12098-008-0101-y. Epub 2008 Jun 25. PMID: 18581071.

10. Tierney AL, Nelson CA 3rd. Brain Development and the Role of Experience in the Early Years. Zero Three. 2009 Nov 1;30(2):9-13. PMID: 23894221; PMCID: PMC3722610.

11. Council on Children With Disabilities; Section on Developmental Behavioral Pediatrics; Bright Futures Steering Committee; Medical Home Initiatives for Children With Special Needs Project Advisory Committee. Identifying infants and young children with developmental disorders in the medical home: an algorithm for developmental surveillance and screening. Pediatrics. 2006 Jul;118(1):405-20. doi: 10.1542/peds.2006-1231.

12. Squires J, Twombly E, Bricker D, Potter L. ASQ-3 Technical Report [Internet]. 2009 [cited 2020 Aug 14]. Brookes Publishing. Available from:

https://agesandstages.com/wp-content/uploads/2019/08/ASQ-3-Technical-Appendix_web.pdf

13. Health Service Executive. The Nurture Programme. National Guideline on the Use of the Ages and Stages Questionnaire for Developmental Screening of Children Between 1 and 66 Months of Age [Internet].2018 [cited 2021 Oct 07]. Available from: https://www.hse.ie/eng/health/child/nurture/asq3national-guideline.pdf

14. Schonhaut B L, Pérez R M, Castilla F AM, Castro M S, Salinas A P, Armijo R I. Validez del Ages & Stages questionnaires para predecir el desempeño cognitivo en los primeros años de educación escolar [Predictive value of Ages & Stages Questionnaires for cognitive performance at early years of schooling]. Rev Chil Pediatr. 2017 Feb;88(1):28-34. Spanish. doi: 10.1016/j.rchipe.2016.08.008. PMID: 28288224.

15. Charkaluk ML, Rousseau J, Calderon J, Bernard JY, Forhan A, Heude B, et al; EDEN Mother–Child Cohort Study Group. Ages and Stages Questionnaire at 3 Years for Predicting IQ at 5-6 Years. Pediatrics. 2017 Apr;139(4):e20162798. doi: 10.1542/peds.2016-2798.

16. Lamsal R, Dutton DJ, Zwicker JD. Using the ages and stages questionnaire in the general population as a measure for identifying children not at risk of a neurodevelopmental disorder. BMC Pediatr. 2018 Apr 3;18(1):122. doi: 10.1186/s12887-018-1105-z.

17. Eriksen HL, Kesmodel US, Underbjerg M, Kilburn TR, Bertrand J, Mortensen EL. Predictors of intelligence at the age of 5: family, pregnancy and birth characteristics, postnatal influences, and postnatal growth. PLoS One. 2013

Nov 13;8(11):e79200. doi: 10.1371/journal.pone.0079200.

18. O'Donovan SM, Murray DM, Hourihane JO, Kenny LC, Irvine AD, Kiely M. Cohort profile: The Cork BASELINE Birth Cohort Study: Babies after SCOPE: Evaluating the Longitudinal Impact on Neurological and Nutritional Endpoints. Int J Epidemiol. 2015 Jun;44(3):764-75. doi: 10.1093/ije/dyu157. Epub 2014 Aug 7. PMID: 25102856.

19. Bain SK, Jaspers KE. Test Review: Review of Kaufman Brief Intelligence Test, Second Edition. J Psychoeduc Assess. 2010 Apr;28(2):167–74. https://doi.org/10.1177/0734282909348217

20. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria;2021. [cited 2021 Nov 10]. Available from https://www.R-project.org/.

21. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011 Mar 17;12:77. doi: 10.1186/1471-2105-12-77.

22. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ. 2007 Oct 20;335(7624):806-8. doi: 10.1136/bmj.39335.541782.AD.

23. Halbwachs M, Muller JB, Nguyen The Tich S, de La Rochebrochard E, Gascoin G, Branger B, Rouger V, Rozé JC, Flamant C. Usefulness of parent-completed ASQ for neurodevelopmental screening of preterm children at five years of age. PLoS One. 2013 Aug 27;8(8):e71925. doi:

10.1371/journal.pone.0071925. PMID: 24014166; PMCID: PMC3754941.

24. McKown C, Weinstein RS. Teacher expectations, classroom context, and the achievement gap. J Sch Psychol. 2008 Jun;46(3):235-61. doi: 10.1016/j.jsp.2007.05.001.

25. Dobrow MJ, Hagens V, Chafe R, Sullivan T, Rabeneck L. Consolidated principles for screening based on a systematic review and consensus process. CMAJ. 2018 Apr 9;190(14):E422-E429. doi: 10.1503/cmaj.171154.

26. Balasundaram P, Avulakunta ID. Bayley Scales Of Infant and Toddler Development In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan. [Updated 2021 Nov 24, cited 2021 Nov 30]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK567715/

27. Lin JY, Chen CH, Li ST, Chang HY. Predictive Validity of the Bayley-III Cognitive Scores at 6 Months for Cognitive Outcomes at 24 Months in Very-Low-Birth-Weight Infants. Front Pediatr. 2021 May 7;9:638449. doi: 10.3389/fped.2021.638449.

28. Månsson J, Stjernqvist K, Serenius F, Ådén U, Källén K. Agreement Between Bayley-III Measurements and WISC-IV Measurements in Typically Developing Children. Journal of Psychoeducational Assessment. 2019;37(5):603-616. doi:10.1177/0734282918781431

29. Miller MM, Goldsmith HH. Profiles of Social-Emotional Readiness for 4-Year-Old Kindergarten. Front Psychol. 2017 Jan 31;8:132. doi: 10.3389/fpsyg.2017.00132.

Chapter 6 Predicting Low Cognitive Ability at Age 5 – Feature Selection Using Machine Learning Methods and Birth Cohort Data

Andrea K. Bowe[1], Gordon Lightbody[1,2], Anthony Staines[3], Mairead E. Kiely[1,4], Fergus P. McCarthy [1,5], Deirdre M. Murray[1,6]

[1] INFANT Research Centre, University College Cork, Cork, Ireland

[2] Department of Electrical and Electronic Engineering, University College Cork, Cork, Ireland

[3] School of Nursing, Psychotherapy, and Community Health, Dublin City University, Ireland

[4] Cork Centre for Vitamin D and Nutrition Research, School of Food and Nutritional Sciences, University College Cork, Cork, Ireland

[5] Department of Obstetrics and Gynaecology, Cork University Maternity Hospital, Cork, Ireland

[6] Department of Paediatrics, Cork University Hospital, Cork, Ireland

Published in The International Journal of Public Health

Nov 2022

6.1.Abstract

Objectives

In this study, we applied the random forest (RF) algorithm to birth-cohort data to train a model to predict low cognitive ability at 5 years of age and to identify the important predictive features.

Methods

Data were from 1,070 participants in the Irish population-based BASELINE cohort. A RF model was trained to predict an intelligence quotient (IQ) score ≤90 at age 5 years using maternal, infant, and sociodemographic features. Feature importance was examined and internal validation performed using 10-fold cross validation repeated 5 times.

Results

The five most important predictive features were the total years of maternal schooling, infant Apgar score at 1 minute, socioeconomic index, maternal BMI, and alcohol consumption in the first trimester. On internal validation a parsimonious RF model based on 11 features showed excellent predictive ability, correctly classifying 95% of participants. This provides a foundation suitable for external validation in an unseen cohort.

Conclusions

Machine learning approaches to large existing datasets can provide accurate feature selection to improve risk prediction. Further validation of this model is required in cohorts representative of the general population.

6.2. Introduction

Cognitive function is a broad construct consisting of multiple domains including learning, understanding, reasoning, problem solving, memory, language, attention, and decision making.[1] Early life is a crucial period for shaping the developing brain, and represents a window of both vulnerability and opportunity.[2] A failure to achieve early foundational cognitive skills may result in a permanent loss of opportunity to achieve full cognitive potential. This has significant adverse implications for outcomes throughout the life course including educational attainment[3], mental health[4], social mobility[5] , financial well-being[6] , and physical health[7].

Unlike children with more severe neurodevelopmental disorders who may be identified in infancy through routine developmental screening programmes, children with low cognitive ability may not display such overt deficits in early infancy.[8] Risk prediction models for poor cognitive outcomes in childhood have been widely studied among certain high risk populations such as pre-term and low birthweight infants.[9] However, far less has been published on their potential use among a general paediatric population.

Adverse cognitive outcomes in childhood are complex, heterogenous, and result from highly interactive relationships between biological, environmental, and social factors. The causative mechanisms are often poorly understood. This real-world complexity may be difficult to model with traditional statistical methods that rely on strong assumptions, which are often unrealistic in complex real world data.[10] Traditional regression methods can handle only a small number of predictors and require interaction terms to be specified a priori by the investigator, a process not commonly performed despite evidence demonstrating interactions between many

of the relevant predictors, for example between prematurity and maternal education.[11–13]

In contrast, machine learning (ML) aims to make a repeatable prediction by learning from patterns within the data, without prior assumptions or rules governing the process.[14] ML methods have the ability to handle many potential variables and to statistically model highly complex, non-linear, interactive relationships, without the need for prior manual specification of the interactions. With recent advances in both the quantity and quality of population based data available through large birth cohorts and electronic health records, the application of artificial intelligence and machine learning may assist in finding the optimal predictive patterns to enable early interventions.[15] To date, the application of ML for predicting childhood cognitive outcomes at a population level has not been explored.

Decision trees and random forests are one type of supervised ML algorithm. A decision tree resembles a flow chart whereby data is successively divided according to decisions based on the predictor variables (features).[16] A random forest (RF) consists of multiple decision trees whose results are aggregated into a single result.[16,17] One advantage of the RF is the relative 'interpretability' when compared with other ML methods.[18] This means that it is possible to determine the relative importance of the features contained within the model and to interpret and describe important interactions. For many clinical applications the interpretability of the model for both clinicians and the public is important for successful implementation.

In this study we focus on the early identification of infants at risk of poor cognitive outcomes at 5 years of age in a general paediatric population. Acknowledging that interventions are more successful when implemented at the

earliest possible stage, we focus on features which do not require invasive testing, and are readily available in the perinatal period at a population level. The objectives of this study are to determine the most important of these features for predicting low cognitive ability at age 5; to examine the important interactions between these features; to train a RF classification model using these features; and to examine the accuracy of this model within our cohort.

6.3. Methods

Study population

The study population was mothers and infants from the Irish Cork BASELINE Study, the first longitudinal birth cohort in Ireland.[19] The BASELINE birth cohort was established to examine the effect of the in-utero and early life environment on health and neurodevelopmental outcomes in children.[19] Recruitment for BASELINE occurred through two streams. Stream 1 recruitment (n=1,583) was from the Screening for Pregnancy Endpoints (SCOPE) pregnancy cohort, a multi-centre prospective study of low-risk, primiparous women which began in Cork in 2008, with the aim of examining adverse outcomes in pregnancy. All women who participated in the SCOPE study were invited to participate in the BASELINE birth cohort. Stream 2 recruitment (n=600) began in 2010 in the postnatal wards of Cork University Maternity Hospital. Women who had a singleton pregnancy were invited to participate in BASELINE. Postnatal assessments were completed at day 2, and then at 2, 6, 12, 24 and 60 months. Infants with complete cognitive outcome data at age 5 years were eligible for inclusion in this study. A flowchart describing the BASELINE cohort and the participants included in this study is contained in Figure 6.1.

Data preprocessing

To create the training dataset, the data were preprocessed and transformed to a suitable format for model training. The main issues with the dataset were the imbalance and missing values.

The dataset was highly imbalanced with regard to the outcome of interest, meaning that the outcome classes of interest were not equally represented – specifically only a minority of the population fell into the outcome group of primary interest (low cognitive ability defined as an IQ $\leq$ 90). Learning from class-imbalanced data tends to produce a model bias in favour of the majority class (average/above cognitive ability), with poor predictive performance for the minority class (low cognitive ability).[20] There are many different techniques to address the issue of class imbalance including over-sampling the minority class, under-sampling the majority class, boosting, bagging, and repeated random sub-sampling. Synthetic Minority Oversampling Technique (SMOTE) was used in this study, and involves a combination of over-sampling the minority class and under-sampling the majority class.[21] Unlike other oversampling techniques which simply duplicate minority class cases, SMOTE utilises an over-sampling approach where the minority class is over-sampled by creating synthetic examples.[21] This is achieved by taking each minority class sample and introducing random synthetic examples along the line segments that join it to a randomly selected sample from its k nearest minority class neighbours. This random convex combination of feature vectors is repeated until the desired amount of over-sampling, set by the user, is achieved.[20,21] Under-sampling is achieved by randomly removing samples from the majority class. Through this combination the learning bias toward the majority class can be compensated for.[20,21] The use of SMOTE is not

without disadvantages – under-sampling may disregard potentially useful data points from the majority class in its attempt to rebalance the classes; the random selection of synthetic minority class data points used in over-sampling may produce distribution marginalisation.[22]

Due to the different times of recruitment (antenatal recruitment vs post-delivery recruitment), there were participants with missing data on features of interest measured in the antenatal period. As these data absences could not be determined to be at random, complete case analysis was performed. Imputation was also performed for missing predictor data using the standard random forest imputation method ('missForest package) and analysis of imputed data is contained in online only material.[23]

Data analysis

Outcome

The outcome of interest in this study was cognitive ability at 5 years of age. This was measured using the Kaufman Brief Intelligence Test Second Edition (KBIT-2), which was administered by trained research nurses at the 60-month follow up. It comprises three subtests: verbal knowledge and riddles, which measure verbal intelligence, and matrices, which measure non-verbal intelligence. The verbal knowledge subtest requires the child to point to a picture, from a selection of six, that best represents a vocabulary word spoken by the examiner. For the riddles subtest the child is asked to point to a picture or provide a word that solves a riddle spoken by the examiner. Together, these two subtests primarily measure crystalised verbal ability which comes from prior learning and experience.[24] The matrices subtest, which requires the child to choose a picture which follows a pattern or concept in pictures shown by

the examiner, measures fluid reasoning and visual processing and is not dependent on knowledge of vocabulary or language.[24]

The verbal and non-verbal scores are combined to provide a composite intelligent quotient (IQ). The Kaufmann test has been shown to be a reliable and valid measure of IQ in children and adults aged 4-90 years.[24] Composite IQ scores ≤ 90 were categorised 'low cognitive ability' and scores above this as 'average and above'. A cut-off of 90 corresponds to a cut-off of 1.5 standard deviations below the cohort mean IQ.

Features

A machine learning algorithm was trained to learn how to make this classification between low and average/above cognitive ability using the predictive features available in the dataset. The only inclusion criterion for predictive features was that the feature was easily measurable at a population level in the perinatal period. In total, there were 21 predictive features included in the study dataset. These were categorised into sociodemographic, lifestyle/behavioural, and birth/delivery and a description of measurement is provided (Table A11.1, Appendix 11). A correlation matrix (Figure A11.1, Appendix 11) was used to examine the correlation between features of interest and assess redundancy. All correlation coefficients were ≤0.70 and no features were deemed redundant.

The machine learning algorithm used for this classification problem was the Random Forest algorithm (using the 'caret' and 'randomForest' packages in R).[25,26] RF is an ensemble ML algorithm that is widely used in classification problems. A RF consists of multiple decision trees, each of which consists of a root node, split node(s), and terminal node(s). At each node a splitting step occurs whereby a

decision is taken on how to partition the data for the nodes below, with each partition of data representing a branch in the tree.[27] At each node of the tree, a random subset of a defined number of features (mtry) is selected and only these variables are considered for partitioning the data at that node.[11] Splitting at nodes will continue down each branch until a stopping rule is satisfied, with this node being referred to as the leaf or terminal node.[27] To classify a participant, each tree will determine which terminal node that participant belongs to, culminating in a vote for that particular class. The forest predicts the class that gets the majority vote across all decision trees in the forest.[26] The RF algorithm allows for tuning of hyperparameters which control the learning process in the RF and these are discussed in supplementary text (Appendix 11).

Feature importance, a measure of the degree of association between the feature and the classification result, was assessed using the Gini importance and the permutation accuracy importance. Gini impurity is a measure of how often a randomly chosen observation would be incorrectly labelled if it was randomly labelled according to the distribution of labels at that node.[28] A lower Gini impurity indicates a lower probability of misclassification. The Gini importance for a given node is the mean decrease in node impurity, weighted by the proportion of observations reaching that node in each decision tree in the forest. [28] A split which results in a large decrease of Gini impurity is important and the feature used for splitting will therefore have a high mean decrease in node impurity and a high Gini importance.

Bootstrap sampling of the training set for each tree means that each tree in the forest has its own ‘out of bag’ (OOB) data which was not used in the construction

of that particular tree.[17] The prediction accuracy is first measured on the OOB data. The values of the feature of interest are randomly shuffled and prediction accuracy is calculated again. The mean difference between the OOB errors before and after the random permutation of the feature is calculated. If a feature is important it would be expected that randomly changing its value would result in a stronger decrease in accuracy than the permutation of an unimportant variable.[11,29] An alternative measure is the minimal depth of the feature which is described in supplementary material. The relationship between each feature and the outcome was examined using partial dependence plots, which provide a graphical representation of the marginal effect of a feature on the class probability.

The method used to evaluate the RF was repeated k-fold cross-validation. Performing k-fold cross validation involves first specifying a value for k, and in this study we chose to set it to k=10, a value commonly used and supported by previous literature.[30] The dataset is shuffled randomly and split into 10 mutually exclusive folds. Each fold will function as a test set one time and will be used as part of the training set k-1 (or 9) times. This procedure is repeated a specified number of times, in this study 5 times, and the mean result across all folds and all repeats is reported.[31] A 10-fold cross validation repeated 5 times means that the accuracy estimates are performed across 50 different held-out test sets. The measures used in the evaluation were overall accuracy (proportion of correct predictions), sensitivity (proportion of those with low cognitive ability correctly identified), and specificity (proportion of those without low cognitive ability correctly identified),

Recursive feature elimination (RFE) with 10-fold cross validation repeated 5 times was then used to train a parsimonious RF model (Model B). [32] The performance

was compared to the original RF model (Model A). Two logistic regression models, one containing all 21 predictors (Model C) and the other containing the 11 most important features (Model D) were also evaluated.

Ethical Approval

The Clinical Research Ethics Committee of Cork Teaching Hospitals provided ethical approval for the BASELINE study (ref ECM5(9) 01/07/2008). It is registered at the United States National Institutes of Health Clinical Trials Registry (http://www.clinicaltrials.gov), ID: NCT01498965, and was carried out in line with the Declaration of Helsinki. The STROBE guidelines were followed in the conduct and reporting of this study.[33]

6.4. Results

Demographics

There were 1,070 children who completed the KBIT-2 at a mean age of 60.8 months (standard deviation (SD) 1.7 months), and who were eligible for inclusion in this study. Composite IQ values ranged from 76-143, and a histogram is contained in Figure A11.2, Appendix 11. The characteristics of those who did and did not complete the KBIT-2 at 60 months are compared in Table A11.2, Appendix 11. There were 66 participants in the low cognitive ability group which had a mean IQ of 87.0, and 1,004 participants in the average/above group which had a mean IQ of 105.5.

The characteristics of participants according to cognitive ability group are described in Table 6.1. The association between each feature and the outcome is described using unadjusted odds ratios. Male participants had a mean IQ of 103.3, compared with females who had a mean IQ of 105.5. In the low cognitive ability

group there was a significantly higher proportion of males compared with females (males 65.2% v females 34.8%, p=0.022). For males, the odds of experiencing low cognitive ability at age 5 were 1.9 times higher compared with females (unadjusted odds ratio 1.88, 95% confidence interval 1.13-3.22). Other features which were associated with a significant increase in the odds of a child experiencing low cognitive ability were lower socioeconomic index, being in a lower income category, being from a single parent family, being formula fed, smoking in the first trimester of pregnancy, and living in government housing.

Machine Learning

After applying complete case analysis there were 36 participants (5.0%) in the low cognitive ability group and 683 (95.0%) in the average/above group. SMOTE was applied with 600% over-sampling to the minority class and 200% under-sampling to the majority class. The rebalanced dataset contained 252 participants (36.8%) and 432 (63.2%) in the low and average/above groups respectively.

The results of the final RF model containing 21 features (Model A) are presented in Table 6.2, alongside the parsimonious model containing 11 features (Model B) and logistic regression models (Models C and D). Of note, the 11 features identified on RFE for the parsimonious model were contained within the 13 most important features consistently identified using the random forest importance measures. The accuracy of Model B was 0.95 and did not improve by more than 0.01 following inclusion of further features beyond these 11. The parsimonious Model B achieved a sensitivity of 0.89 and a specificity of 0.98 on 10-fold cross validation. To illustrate how a RF is constructed, an example decision tree from Model A is shown in Figure 6.2.

There were 13 features identified consistently across measures of feature importance (Figure 6.3 and Figure A11.3, Appendix 11).These were total years maternal schooling, socioeconomic index, family income, maternal age, maternal relationship status, maternal BMI, weekly units of alcohol in first trimester, maternal depression score, Apgar score at 1 minute, Apgar score at 5 minutes, infant birthweight, infant gestational age, and infant head circumference. The correlations between importance rankings for the 21 features are described in Figure A11.4, Appendix 11. The relationship between each of these features and the probability of low cognitive ability can be visualised in the partial dependence plots in Figure A11.5, Appendix 11.

Using these variables a dataframe of the most important interactions in the model, determined by occurrence was generated. Table 3 outlines the six interaction terms which occurred most frequently in Model A. The mean minimal depth refers to the mean minimal depth of the interacting feature conditional on the main feature being the root node. The unconditional mean minimal depth is the mean minimal depth of the feature without stipulating this root. An interaction term with a lower mean minimal depth indicates the interaction is a more important feature than either feature alone.[34] For example, the interaction between total years of maternal schooling and weekly units of alcohol in the first trimester was a more important predictive feature than weekly units of alcohol or total years of maternal schooling alone.

6.5. Discussion

In this study we identified the most important features for predicting low cognitive ability at age 5 from a range of features which could be easily measured in the perinatal period at a population level. There were 13 features consistently identified across the range of importance measures used - total years maternal schooling, socioeconomic index, family income, maternal age, maternal relationship status, maternal BMI, weekly units of alcohol in first trimester, maternal depression score, Apgar score at 1 minute, Apgar score at 5 minutes, infant birthweight, infant gestational age, and infant head circumference.

No previous studies using ML methods for prediction of cognitive outcomes in childhood among a general paediatric population have been identified in the literature for comparison. However, the multivariable prediction models that have been developed for similar outcomes using more traditional linear and logistic regression methods identified many of the same predictors including maternal education, socioeconomic status, income, and employment status.[35–37] An advantage of this study is that through use of feature importance metrics, partial dependence plots, and an exploration of the most important interactions, a model that is relatively interpretable was developed where one can understand the decision logic of the model. In addition, an explanation of the key components required to understand the development and interpretation of a ML model were provided.

The findings of the study are subject to limitations. Bias in ML models can arise in data collection, model development, and model evaluation.[38] First, the BASELINE cohort is subject to sampling bias as the majority of those recruited were low-risk, primiparous women who may not be representative of the pregnant

population as a whole. In addition, those who remain engaged in cohort studies systematically differ from those who do not. As shown in Table A11.2 participants whose child completed IQ assessment at age 5 were more likely to be older non-smokers, with higher education and lower depression scores, and were more likely to have infants with a higher birthweight, head circumference, and gestational age. The generalisability, equity, and utility of a predictive model at a population level is dependent on data that is representative of the population.[39]

Sampling bias, such as that outlined here, is a significant challenge in the application of ML methods to address population health problems. Often the most vulnerable populations, including homeless populations, intravenous drug users, and ethnic minorities, are absent from large cohort studies, but may be among the populations who would benefit most from optimised prediction models. As the potential to collect large quantities of data increases, through electronic health records, electronic devices, and wearable technology, researchers must ensure that minority groups are represented and barriers to participation which may be cultural, financial, linguistic, or time-bound must be removed.

The evaluation of a ML model should be performed across diverse patient groups, on data that was not used in the training process. In this study external validation on unseen data was not performed and therefore the performance metrics of the model should be interpreted with caution. As the dataset used in this study was imbalanced with regard to the outcome of interest the entire dataset was used in the training process and internal validation only was applied. The next step is to now validate this model using unseen data, ideally data that is more representative of the population as a whole.

A further limitation associated with the use of pre-existing birth cohort data is that features, such as genetic information and parental IQ, which have an important causal relationship with the outcome of interest and would likely improve predictive power, may not have been collected.[40] However, a model relying on genetic data would not currently be feasible to implement at a population level.

ML was used in this study to optimise the predictive power of the model. Risk prediction and early intervention is a form of secondary prevention, and should occur in conjunction with primary prevention strategies to reduce the causative factors. However, a limitation of ML is that it does not inform us of the causal relationship between the features and the outcome. While many of the predictive features identified in this study, such as parental education[41], have an existing literature base demonstrating a causal relationship with the outcome, predictors should not be assumed to be causal factors.

Finally, there are important ethical issues to be considered. Identifying children at risk of poor cognitive outcomes risks labelling a child early in life. It is well documented that early labelling can negatively impact both parent and teacher expectations, as well as the child's own self-concept.[42,43] These risks are especially pertinent for children and families who may be incorrectly identified as at risk on a screening or risk prediction tool and who are unlikely to benefit from early intervention. For those correctly identified, the question is whether these risks are outweighed by the benefit of providing early intervention.

6.6. Conclusion

In conclusion, the application of machine learning to address population health challenges has received much less attention than its application in the clinical setting. Disparities in cognitive development can be seen as early as 2 years of age, but unabated will amplify over time, making early intervention essential.[44] The socioenvironmental exposures of a child in early life are modifiable, and enrichment of the early environment has been shown to be both feasible, achievable, and effective in improving cognitive outcomes, if appropriately resourced.[45,46] However, the effectiveness of such intervention programmes is dependent upon the early and accurate identification of the children who are most likely to benefit.

Current population based strategies which rely on the presence of a delay in development prior to intervention miss the opportunity for pre-emptive intervention in the period of optimal neuroplasticity, and risk missing the most vulnerable who may not present to routine screening. The perinatal period is a unique window of opportunity where there is almost universal contact between the public and healthcare professionals, and is an ideal time for engagement with high-risk mother-infant dyads. As clinical disciplines strive toward a personalised approach to medicine, population health must not be left behind. Investigation into how machine learning can be used to assist in addressing disparities in cognitive development in early childhood - a significant population health challenge, rooted in the social determinants of health and exacerbated by inequity – requires further research. This targeted approach must be considered in the context of wider changes to educational policy, planning policy, and economic policy which have a central role to play in addressing disparities in childhood development.

Author contributions

AB – Conceptualization, methodology, formal analysis, writing – original draft preparation, writing – review and editing

GL – Methodology, formal analysis, writing – review and editing

AS - Methodology, writing – review and editing

MK – Data acquisition, writing – review and editing

FMcC - Data acquisition, writing – review and editing

DM – Conceptualization, data acquisition, methodology, writing – review and editing

Table 6.4 Characteristics of those with and without low cognitive ability at age 5 years

	Valid	Low cognitive ability (n= 66)	Average/above cognitive ability (n= 1,004)	p	Unadjusted odds of low cognitive ability OR (95% CI)
		n (%)	**n (%)**		
Sociodemographic					
Gender					
Female	1,070	23 (34.8)	504 (50.2)		Ref
Male		43 (65.2)	500 (49.8)	0.022[a]	1.88 (1.13-3.22)
Maternal Relationship Status	1,070				
In a relationship		60 (90.9)	973 (96.9)		Ref
Not in a relationship		6 (9.1)	31 (3.1)	0.025[a]	3.14 (1.15-7.32)
Maternal migration history	1,059				
Lived in country >2 generations		55 (84.6)	870 (87.5)		Ref
Mother immigrated		7 (10.8)	94 (9.5)		1.18 (0.48-2.50)
One or both parents immigrated		3 (4.6)	30 (3.0)	0.534[b]	1.58 (0.37-4.62)
Ethnicity					
European	1,070	65 (98.4)	993 (98.9)		Ref
Other		1 (1.6)	11 (1.1)	0.536[b]	0.08-7.31
Total years of schooling	1,070	13.2 (0.8)	13.4 (0.8)	0.109[c]	0.75 (0.56-1.03)
Maternal Employment Status	1,067				
Full time		49 (75.4)	822 (82.0)		Ref
Part time		9 (13.8)	89 (8.9)		1.70 (0.76-3.41)
Student		1 (1.5)	15 (1.5)		1.12 (0.06-5.69)
Homemaker		0 (0.0)	37 (3.7)		0.00 (0.00-3.15)
Unemployed/sickness benefit		6 (9.2)	39 (3.9)	0.061[b]	2.58 (0.95-5.98)
Socioeconomic index – mean (sd)	1,070	37.3 (15.1)	44.5 (15.5)	<0.001[c]	0.97 (0.95-0.98)
Accommodation type	1,070				
Own house/flat		46 (69.7)	797 (79.4)		Ref
Private rental		14 (21.2)	172 (17.1)		1.41 (0.73-2.56)
Government/council rental		4 (6.1)	13 (1.3)		5.33 (1.46-15.76)
Other		2 (3.0)	22 (2.2)	0.026[b]	1.58 (0.25-5.57)
Family Income	1,042				
<21k		8 (12.7)	45 (4.6)		Ref
21-42k		14 (22.2)	164 (16.7)		0.48 (0.19-1.27)

43-63k		16 (25.4)	217 (22.1)		0.41 (0.17-1.08)
64-84k		16 (25.4)	249 (25.4)		0.36 (0.15-0.94)
85-105		5 (7.9)	158 (16.1)		0.18 (0.05-0.56)
106k-140k		4 (6.3)	122 (12.4)		0.18 (0.05-0.62)
>140k		0 (0.0)	25 (2.6)	0.042[b]	0.0 (0.0-0.0)
Maternal age, mean (sd), y	1,070	29.7 (4.3)	30.7 (3.9)	0.123[c]	0.93 (0.86-1.01)
Behavioural/Lifestyle					
Maternal BMI, mean (sd), kg/m^2	806	26.1 (4.6)	24.9 (4.0)	0.115[c]	1.07 (0.99-1.14)
Depression score[e], median (IQR)	795	6.0 (7.5)	5.0 (7.0)	0.743[d]	1.02 (0.95-1.09)
Units of alcohol per week in 1st trimester, median (IQR)	806	4.0 (8.0)	2.9 (5.0)	0.127[d]	1.07 (1.02-1.11)
No. cigarettes per day in 1st trimester, median (IQR)	806	0 (5.0)	0 (0.0)	0.042[d]	1.08 (1.02-1.13)
Birth/Delivery					
Apgar score at 1 minute, median (IQR)	1,061	9.0 (0.0)	9.0 (0.0)	0.773[d]	0.94 (0.79-1.14)
Apgar score at 5 minute, median (IQR)	1,061	10.0 (1.0)	10 (1.0)	0.877[d]	1.08 (0.78-1.61)
Birth head circumference, mean (sd), cm	1,055	35.1 (1.4)	34.9 (1.5)	0.287[c]	1.09 (0.92-1.30)
Admission to NICU	1,070				
Not Admitted		62 (93.9)	917 (91.3)		Ref
Admitted		4 (6.1)	87 (8.7)	0.648[a]	0.68 (0.20-1.70)
Infant gestational age, mean (sd), wk	1,070	39.5 (1.8)	40.0 (2.0)	0.155[d]	0.90 (0.78-1.05)
Infant birthweight, mean (sd), g		3,546.2 (1.6)	3,526.8 (1.5)	0.183[c]	1.00 (0.99-1.00)
Infant feeding on discharge	1,061				
Predominantly breastfed		21 (31.8)	460 (46.2)		Ref
Partially breastfed		19 (28.8)	289 (29.0)		1.44 (0.76-2.73)
Formula fed		26 (39.4)	246 (24.7)	0.018[a]	2.32 (1.28-4.24)
High Risk Birth	833	1 (2.2)	67 (8.5)	0.167[b]	0.24 (0.01-1.15)
Cognitive					
IQ, mean (sd)	1,070	87.0 (3.4)	105.5 (7.8)	<0.001[c]	

Abbreviations: OR, odds ratio; CI, confidence interval.

[a]p-value calculated using Pearson's Chi-square test, [b]p-value calculated using Fishers' Exact test, [c]p-value calculated using Welch Two Sample t-test, [d]p-value calculated using Wilcoxon rank sum test, [e]Measured at 15 weeks' gestation using Edinburgh Depression Scale

Table 6.5 Performance of Models A-D in repeated 10-fold cross validation

	Method	**Features**	**Mtry**	**Accuracy**	**Sensitivity**	**Specificity**
Model A[a]	Random forest	21	4	0.95	0.89	0.99
Model B[b]	Random forest	11	5	0.95	0.89	0.98
Model C[c]	Logistic regression	21	NA	0.78	0.64	0.86
Model D[d]	Logistic regression	11	NA	0.77	0.48	0.83

[a]Model A: Random forest model using 21 features (total years schooling, Apgar score 1 minute, socioeconomic index, Apgar score 5 minutes, family income, gestational age, units of alcohol in first trimester, head circumference at birth, maternal age, maternal depression score, accommodation type, infant gender, maternal relationship status, migration history, maternal employment status, admission to NICU, accommodation type, infant feeding on discharge, high risk birth)
[b]Model B: Parsimonious random forest model using 11 features (total years schooling, Apgar score 1 minute, socioeconomic index, maternal BMI, Apgar score 5 minutes, family income, gestational age, units of alcohol in first trimester, head circumference at birth, maternal age, maternal depression score)
[c]Model C: Logistic regression model using 21 features (outlined in Model A)
[d]Model D: Logistic regression model using 11 features (outlined in Model B)

Table 6.6 Mean minimal depth of the six most frequently occurring interactions in random forest Model A

Main Feature	**Interacting Feature**	**Occurrences**	**Unconditional mean minimal depth of interacting feature**	**Mean minimal depth of interacting feature conditional on main feature**
Total Years Schooling	Weekly units alcohol first trimester	334	3.21	2.00
Total Years Schooling	Maternal BMI	333	2.71	2.14
Total Years Schooling	Infant Birthweight	329	3.73	2.43
Total Years Schooling	Maternal Depression Score	325	3.78	2.73
Apgar Score 1 Minute	Weekly units alcohol in first trimester	324	3.21	2.18
Total Years Schooling	Infant head circumference	321	3.68	2.72

Figure 6.4 Flowchart of study population from the BASELINE birth cohort

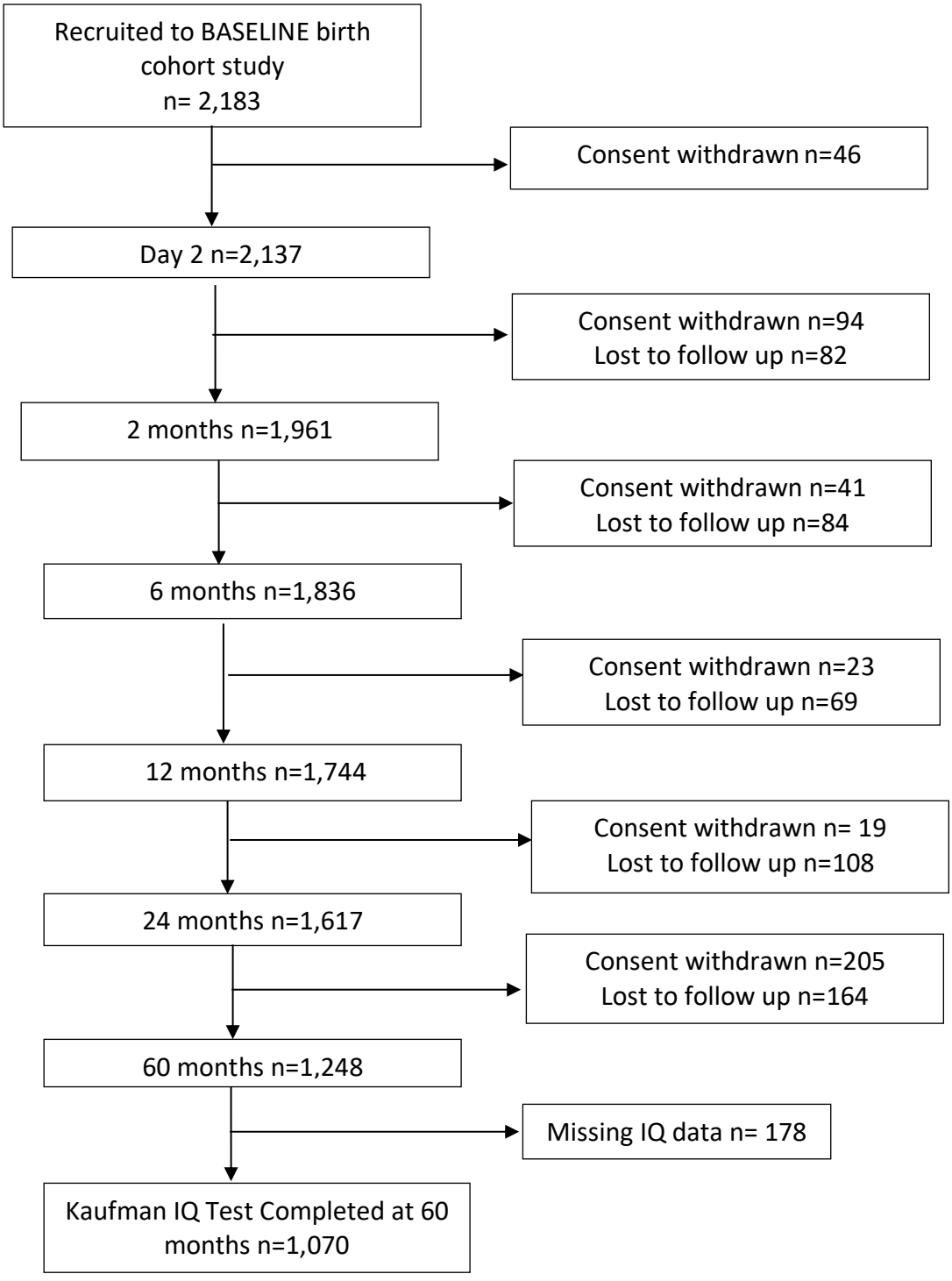


Figure 6.5 Example decision tree from the random forest Model A predicting low cognitive ability at age 5

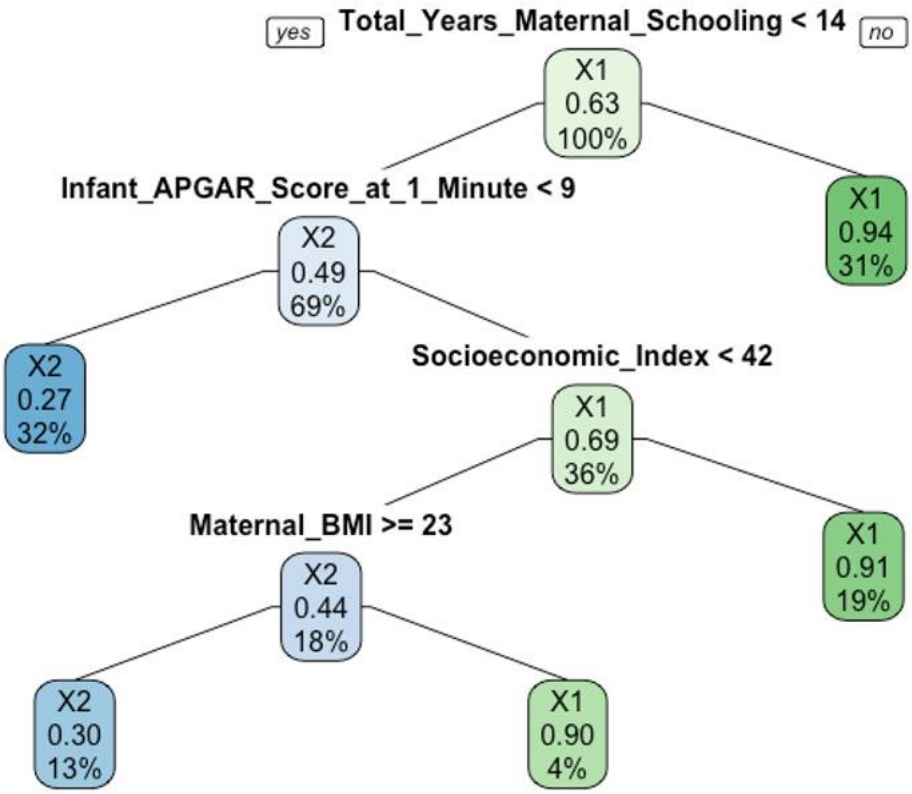


Figure 6.2 Legend

X1 refers to the 'average/above cognitive ability' group and X2 refers to the 'low cognitive ability' group. At the top of decision tree, the overall probability of 'average/above' cognitive ability is shown. 63% of participants had average/above cognitive ability (after SMOTE applied). The root node asks whether the total years of maternal schooling was < 14. If no, then you proceed down the right branch. 31% had ≥ 14 years of schooling and the probability of average/above cognitive ability was 0.94. If yes, then proceed down left branch. 69% had <14 years and a probability of average/above cognitive ability of 0.49. The next splitting node asks whether the Apgar score at 1 minute was <9. If yes, then proceed down left branch to the leaf or terminal node. 32% had an Apgar score <9 at 1 minute and a probability of average/above cognitive ability of 0.27. If no, then proceed down the right branch where the next splitting node asks whether the socioeconomic index was less than 42.

Figure 6.6 Feature importance plots showing the mean decrease in Gini and mean decrease in accuracy for the 15 most important features in random forest Model A

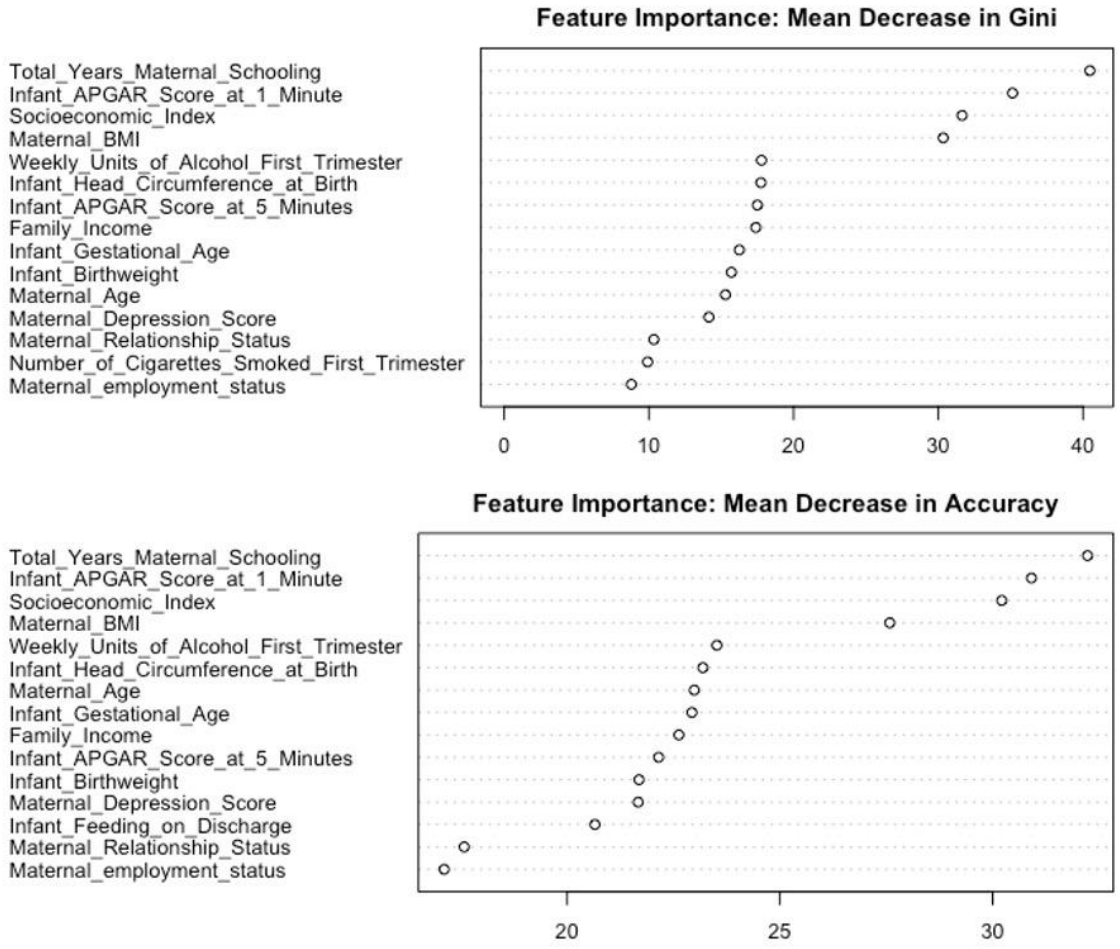


Figure 6.3 Legend

The top plot in the figure shows the mean decrease in Gini for 15 features ranked in descending order of importance. A feature which results in a larger decrease in Gini impurity is a more important feature. The bottom plot in the figure shows the mean decrease in accuracy for 15 features ranked in descending order of importance. A feature with a larger mean decrease in accuracy is a more important feature. Importance was measured in 10-fold cross validation repeated 5 times.

6.7. References

1. Nouchi R, Kawashima R. Improving Cognitive Function from Children to Old Age: A Systematic Review of Recent Smart Ageing Intervention Studies. Advances in Neuroscience 2014 Aug 11;(2):235479. https://doi.org/10.1155/2014/235479
2. Spencer N, Raman S, O'Hare B, Tamburlini G. Addressing inequities in child health and development: towards social justice. BMJ Paediatr Open. 2019 Aug 1;3(1):e000503. doi: 10.1136/bmjpo-2019-000503.
3. Lager A, Bremberg S, Vågerö D. The association of early IQ and education with mortality: 65 year longitudinal study in Malmö, Sweden. BMJ. 2009 Dec 10;339:b5282. doi: 10.1136/bmj.b5282.
4. Alesi M, Rappo G, Pepi A. Emotional Profile and Intellectual Functioning: A Comparison Among Children With Borderline Intellectual Functioning, Average Intellectual Functioning, and Gifted Intellectual Functioning. Sage Open. 2015 Jul 1;5(3) https://doi.org/10.1177/2158244015589995
5. Forrest LF, Hodgson S, Parker L, Pearce MS. The influence of childhood IQ and education on social mobility in the Newcastle Thousand Families birth cohort. BMC Public Health. 2011 Nov 25;11:895. doi: 10.1186/1471-2458-11-895. PMID: 22117779; PMCID: PMC3248886.
6. Furnham A, Cheng H. Childhood Cognitive Ability Predicts Adult Financial Well-Being. J Intell. 2016 Dec 27;5(1):3. doi: 10.3390/jintelligence5010003.
7. Whalley LJ, Deary IJ. Longitudinal cohort study of childhood IQ and survival up to age 76. BMJ. 2001 Apr 7;322(7290):819. doi: 10.1136/bmj.322.7290.819.

8. Fernell E, Ek U. Borderline intellectual functioning in children and adolescents - insufficiently recognized difficulties. Acta Paediatr. 2010 May;99(5):748-753. doi: 10.1111/j.1651-2227.2010.01707.x. Epub 2010 Feb 9. PMID: 20146756.

9. Linsell L, Malouf R, Morris J, Kurinczuk JJ, Marlow N. Risk Factor Models for Neurodevelopmental Outcomes in Children Born Very Preterm or With Very Low Birth Weight: A Systematic Review of Methodology and Reporting. Am J Epidemiol. 2017 Apr 1;185(7):601-612. doi: 10.1093/aje/kww135. PMID: 28338817.

10. Fan J, Han F, Liu H. Challenges of Big Data Analysis. National science review 2014 Jun; 1(2): 293–314. https://doi.org/10.1093/nsr/nwt032

11. Couronné R, Probst P, Boulesteix AL. Random forest versus logistic regression: a large-scale benchmark experiment. BMC Bioinformatics. 2018 Jul 17;19(1):270. doi: 10.1186/s12859-018-2264-5. PMID: 30016950; PMCID: PMC6050737.

12. Morgenstern JD, Buajitti E, O'Neill M, Piggott T, Goel V, Fridman D, et al. Predicting population health with machine learning: a scoping review BMJ Open 2020;10:e037860. doi: 10.1136/bmjopen-2020-037860

13. Benavente-Fernández I, Synnes A, Grunau RE, Chau V, Ramraj C, Glass T, et al. Association of Socioeconomic Status and Brain Injury With Neurodevelopmental Outcomes of Very Preterm Children. JAMA Netw Open. 2019 May 3;2(5):e192914. doi: 10.1001/jamanetworkopen.2019.2914.

14. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. Nat Methods. 2018 Apr;15(4):233-234. doi: 10.1038/nmeth.4642.

15. Dwyer D, Koutsouleris N. Annual Research Review: Translational machine learning for child and adolescent psychiatry. J Child Psychol Psychiatry. 2022 Apr;63(4):421-443. doi: 10.1111/jcpp.13545.

16. Speiser JL, Callahan KE, Houston DK, Fanning J, Gill TM, Guralnik JM, et al. Machine Learning in Aging: An Example of Developing Prediction Models for Serious Fall Injury in Older Adults. J Gerontol A Biol Sci Med Sci. 2021 Mar 31;76(4):647-654. doi: 10.1093/gerona/glaa138.

17. Altman N, Krzywinski M. Ensemble methods: bagging and random forests. Nat Methods 2017 Sep29; 14: 933–934. https://doi.org/10.1038/nmeth.4438

18. Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable AI: A Review of Machine Learning Interpretability Methods. Entropy (Basel). 2020 Dec 25;23(1):18. doi: 10.3390/e23010018.

19. O'Donovan SM, Murray DM, Hourihane JO, Kenny LC, Irvine AD, Kiely M. Cohort profile: The Cork BASELINE Birth Cohort Study: Babies after SCOPE: Evaluating the Longitudinal Impact on Neurological and Nutritional Endpoints. Int J Epidemiol. 2015 Jun;44(3):764-75. doi: 10.1093/ije/dyu157. Epub 2014 Aug 7. PMID: 25102856.

20. Blagus R, Lusa L. SMOTE for high-dimensional class-imbalanced data. BMC Bioinformatics. 2013 Mar 22;14:106. doi: 10.1186/1471-2105-14-106.

21. Chawla N, Bowyer K, Hall L, Kegelmeyer W. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 2002; 16: 341–348. https://doi.org/10.1613/jair.953

22. Wang S, Dai Y, Shen J, Xuan J. Research on expansion and classification of imbalanced data based on SMOTE algorithm. Sci Rep. 2021 Dec 15;11(1):24039. doi: 10.1038/s41598-021-03430-5.

23. Stekhoven DJ, Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012 Jan 1;28(1):112-8. doi: 10.1093/bioinformatics/btr597.

24. Bain SK, Jaspers KE. Test Review: Review of Kaufman Brief Intelligence Test, Second Edition. J Psychoedul Asses. 2010; 28 (2): 167–174. https://doi.org/10.1177/0734282909348217

25. Kuhn M. Building Predictive Models in R Using the caret Package. Journal of Statistical Software. 2008 Nov; 28(5): 1–26. https://doi.org/10.18637/jss.v028.i05

26 Breiman L. Random Forests. Machine Learning. 2001 Oct; 45: 5–32. https://doi.org/10.1023/A:1010933404324

27 Venkatasubramaniam A, Wolfson J, Mitchell N, Barnes T, JaKa M, French S. Decision trees in epidemiological research. Emerg Themes Epidemiol. 2017 Sep 20;14:11. doi: 10.1186/s12982-017-0064-4.

28 Nembrini S, König IR, Wright MN. The revival of the Gini importance? Bioinformatics. 2018 Nov 1;34(21):3711-3718. doi: 10.1093/bioinformatics/bty373.

29 Khalilia M, Chakraborty S, Popescu M. Predicting disease risks from highly imbalanced data using random forest. BMC Med Inform Decis Mak. 2011 Jul 29;11:51. doi: 10.1186/1472-6947-11-51.

30 Marcot BG, Hanea AM. What is an optimal value of k in k-fold cross-validation in discrete Bayesian network analysis? Computational Statistics 2021 Sep; 36: 2009–2031. https://doi.org/10.1007/s00180-020-00999-9

31 de Rooij M, Weeda W. Cross-Validation: A Method Every Psychologist Should Know. Advances in Methods and Practices in Psychological Science. 2020;3(2):248-263. doi:10.1177/2515245919898466

32 Kuhn M, Johnson K. Feature Engineering and Selection: A Practical Approach for Predictive Models [Internet]. Taylor and Francis Group; 2019. [cited 2022 Mar 17]. Available at http://www.feat.engineering/

33 von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ. 2007 Oct 20;335(7624):806-8. doi: 10.1136/bmj.39335.541782.AD.

34 Paluszynska A. Understanding random forests with randomForestExplainer [Internet].[cited 2022 Jun 11]. Available at: https://cran.rstudio.com/web/packages/randomForestExplainer/vignettes/randomForestExplainer.html

35 Eriksen HL, Kesmodel US, Underbjerg M, Kilburn TR, Bertrand J, Mortensen EL. Predictors of intelligence at the age of 5: family, pregnancy and birth

characteristics, postnatal influences, and postnatal growth. PLoS One. 2013 Nov 13;8(11):e79200. doi: 10.1371/journal.pone.0079200.

36 Camargo-Figuera FA, Barros AJ, Santos IS, Matijasevich A, Barros FC. Early life determinants of low IQ at age 6 in children from the 2004 Pelotas Birth Cohort: a predictive approach. BMC Pediatr. 2014 Dec 16;14:308. doi: 10.1186/s12887-014-0308-1. PMID: 25510879; PMCID: PMC4272809.

37 Camacho C, Straatmann VS, Day JC, Taylor-Robinson D. Development of a predictive risk model for school readiness at age 3 years using the UK Millennium Cohort Study. BMJ Open. 2019 Jun;9(6):e024851. doi: 10.1136/bmjopen-2018-024851

38 Vokinger KN, Feuerriegel S, Kesselheim AS. Mitigating bias in machine learning for medicine. Commun Med (Lond). 2021 Aug 23;1:25. doi: 10.1038/s43856-021-00028-w.

39 Cahan EM, Hernandez-Boussard T, Thadaney-Israni S, Rubin DL. Putting the data before the algorithm in big data addressing personalized healthcare. NPJ Digit Med. 2019 Aug 19;2:78. doi: 10.1038/s41746-019-0157-2.

40 Allegrini AG, Selzam S, Rimfeld K, von Stumm S, Pingault JB, Plomin R. Genomic prediction of cognitive traits in childhood and adolescence. Mol Psychiatry. 2019 Jun;24(6):819-827. doi: 10.1038/s41380-019-0394-4.

41 Rindermann H, Ceci SJ. Parents' education is more important than their wealth in shaping their children's intelligence: Results of 19 samples in seven countries at different developmental levels. J Educ Gift. 2018; 41(4), 298–326. https://doi.org/10.1177/0162353218799481

42 Jussim L, Harber KD. Teacher expectations and self-fulfilling prophecies: knowns and unknowns, resolved and unresolved controversies. Personality and Social Psychology Review. 2005 May; 9: 131–155. https://doi.org/10.1207/s15327957pspr0902_3

43 Taylor LM, Hume IR, Welsh N. Labelling and self-esteem: the impact of using specific vs. generic labels. Educational Psychology. 2010 Jan; 30: 191–202. doi:10.1080/01443410903494478

44 von Stumm S, Plomin R. Socioeconomic status and the growth of intelligence from infancy through adolescence. Intelligence 2015; 48: 30–36. doi:10.1016/j.intell.2014.10.002

45 Love JM, Chazan-Cohen R, Raikes H, Brooks-Gunn J. What makes a difference: Early Head Start evaluation findings in a developmental context. Monogr Soc Res Child Dev. 2013 Feb;78(1):vii-viii, 1-173. doi: 10.1111/j.1540-5834.2012.00699.x.

46 Troller-Renfree SV, Costanzo MA, Duncan GJ, Magnuson K, Gennetian LA, Yoshikawa H, et al. The impact of a poverty reduction intervention on infant brain activity. Proc Natl Acad Sci U S A. 2022 Feb 1;119(5):e2115649119. doi: 10.1073/pnas.2115649119.

Chapter 7 Predicting Low Cognitive Ability at Age 5 Years using Perinatal data and Machine Learning

Andrea K. Bowe[1], Gordon Lightbody[1,2], Daragh S. O'Boyle[1], Anthony Staines[3], Deirdre M. Murray[1,4]

[1] INFANT Research Centre, University College Cork, Cork, Ireland

[2] Department of Electrical and Electronic Engineering, University College Cork, Cork, Ireland

[3] School of Nursing, Psychotherapy, and Community Health, Dublin City University, Ireland

[4] Department of Paediatrics, Cork University Hospital, Cork, Ireland

Published in Paediatric Research

Jan 2024

7.1. Abstract

Background

There are no early, accurate, scalable methods for identifying infants at high risk of poor cognitive outcomes in childhood. We aim to develop an explainable predictive model, using machine learning and population-based cohort data, for this purpose.

Methods

Data were from 8,858 participants in the Growing Up in Ireland cohort, a nationally representative study of infants and their primary caregivers (PCGs). Maternal, infant, and socioeconomic characteristics were collected at 9-months and cognitive ability measured at age 5 years. Data preprocessing, synthetic minority oversampling, and feature selection were performed prior to training a variety of machine learning models using ten-fold cross validated grid search to tune hyperparameters. Final models were tested on an unseen test set.

Results

A random forest (RF) model containing 15 participant-reported features in the first year of infant life, achieved an area under the receiver operating characteristic curve (AUROC) of 0.77 for predicting low cognitive ability at age 5. This model could detect 72% of infants with low cognitive ability, with a specificity of 66%.

Conclusions

Model performance would need to be improved before consideration as a population-level screening tool. However, this is a first step towards early, individual, risk stratification to allow targeted childhood screening.

7.2.Introduction

Early life is a unique period where the developing brain has great plasticity and huge potential for adaptability.[1] There is consensus agreement that individual interventions to improve cognitive development should be initiated early.[2,3] A failure to achieve early foundational cognitive skills may result in a permanent loss of opportunity to reach full academic potential.[4] This, in turn, may adversely affect outcomes throughout the life course including educational attainment[5], mental health[6], social mobility[7], financial well-being[8], and physical health.[9]

Internationally, many countries rely on universal screening programmes to identify children who may benefit from early intervention. The majority of developmental screening assessments are based on the presence of a delay in developmental milestones.[10] A limitation of this approach is that opportunities for intervention in the period of optimal neuroplasticity are lost, and intervention begins when an infant is already substantially behind their typically developing peers. There is increasing evidence that early pre-emptive interventions, initiated prior to overt signs of delay or difficulty, can alter neurodevelopmental outcomes.[11-13] The challenge we face is predicting at an individual level, using accurate and scalable methods, who the highest risk infants are.

In the United States, the population-based early intervention programmes Head Start and Early Head Start, base eligibility primarily on a family income at or below the poverty level.[14] Adverse socioeconomic conditions are among the strongest predictors of poor cognitive outcomes in childhood, but there are other important psychosocial, biological, genetic, and environmental influences, which often have complex and interactive relationships both with each other and with

cognitive outcomes.[15,16] There is now increasing potential to statistically model complex interactive relationships using machine learning techniques.[17,18] This has not been sufficiently explored for prediction of poor cognitive outcomes in childhood at a population level.[19]

In this study we aim to develop an explainable predictive algorithm, using population-based cohort data, for identifying infants at risk of low cognitive ability (LCA) at school-age. The objectives of the study are to train a variety of machine learning models to predict LCA at age 5; to test these models on an independent unseen test-set; to compare model performance using a range of measures; and to identify the most important features for prediction.

7.3. Methods

Data

Data are from the Growing Up in Ireland (GUI) Infant Cohort, a nationally representative survey of infants and their primary caregivers (PCG). Wave 1 of data collection commenced in 2008 at infant age 9-months, Wave 2 occurred at age 3 years, and Wave 3 at age 5 years. The sample was drawn from the National Child Benefits Register, a universal welfare entitlement in the Republic of Ireland. It was selected on a systematic basis with a random start and constant sampling fraction, and was pre-stratified by marital status, county of residence, nationality, and number of children in the family. There were 11,134 families who participated at Wave 1, of whom 9,001 completed Wave 3, representing 80.8% of the original sample. Full details of the sample design, response, and survey instruments are available.[20] Eligible for inclusion in this study were the 8,858 infants who completed

cognitive assessments at age 5 and their PCGs, who in 99.7% of cases were the infant's mother. A flow chart of the study population is contained in Figure A12.1, Appendix 12.

Outcome

Cognitive ability at age 5 years was directly assessed using two core subtests of the British Ability Scales (BAS) Early Years Battery Second Edition, administered in the child's home by a trained interviewer. The BAS consists of a battery of individually administered subtests (detailed in Table A12.1, Appendix 12) and has demonstrated construct validity as a measure of cognitive ability and high test-retest reliability.[21] To minimise participant burden in the GUI study, there was an upper limit of 90 minutes contact time in the home. Therefore, it was not feasible to administer the full battery of tests, and two core subtests which most closely align with measures of crystallised and fluid cognitive ability were chosen.[22]

The Naming Vocabulary test measures verbal ability in the English language and consists of the child naming everyday items displayed from a picture book. The Picture Similarities test measures non-verbal ability and consists of the child being shown four pictures and requested to match a fifth picture, based on a shared characteristic or construct. A standardised score for each scale is provided in the dataset and is adjusted for both item difficulty and age (within a three month age band).

Multiple BAS subtest scores can be combined with summation to produce a General Conceptual Ability Score. To produce composite scores using fewer subtests Principal Components Analysis (PCA) was used, as in previous research.[23] PCA of the two BAS subtests confirmed the presence of a general underlying

cognitive ability factor. Principal component 1 (PC1) accounted for 64% of the total variance among the subtests. The Pearson correlation between this factor and the observed variable was 0.80 for Picture Similarities and 0.80 for Naming Vocabulary subtests. PC1 was then standardised to produce a general cognitive ability (GCA) score with a mean of 100 and a standard deviation (sd) of 15.

There is no consensus agreement on a cut-off that defines LCA in childhood. The International Classification of Disease 11th Revision (ICD-11) use a standardised test score that is ≥ 2 standard deviations (SD) below the mean to define a disorder of intellectual development, while other research in this area has used cut-offs of ≥1 or 1.5 SDs below the mean.[16,24,25] As this study was exploratory in nature, we examined both a 1 and 1.5 SD cut-off. For clarity of presentation the 1.5 SD cut off is used in the main study and children scoring below this cut-off are referred to as having low cognitive ability (LCA). Those with cognitive ability score above this cut-off are referred to as having average or above cognitive ability. The results using a 1 SD cut off are included in Appendix 12 and this group are referred to as below average cognitive ability (BACA).

Data preparation

No feature had more than 12% missing values. Missing values were imputed using the 'missForest' package, a random forest imputation method.[26] The post-imputation dataset was stratified by the outcome and then randomly split into a training set, containing 70% (n=6,202) of participants, and a testing set, containing 30% (n=2,656).

Feature Selection

The GUI dataset contains more than 600 variables measured at wave 1 (9-months) across domains of health, education, cognitive development, social class, and neighbourhood characteristics. The framework used in the selection of relevant features is shown in Figure A12.2, Appendix 12. These features capture pregnancy and birth, maternal and infant characteristics, the socioeconomic circumstances of family, and the infant's early environment. Only features based on information that could be easily obtained at a population level without the need for invasive testing were eligible for inclusion. Features which could be collected in the perinatal period were preferable to enable prediction soon after birth. It is well established that the early learning environment is important in cognitive development, and a set of features intended to measure this were included.[27] Previous published work by our group has focussed on feature selection to predict low IQ in a separate cohort and this was also considered.[28] All features considered are described in Table A12.2, Appendix 12.

To remove redundant features, Pearson correlation coefficients were calculated and plotted. Features with a correlation greater than 0.6 were examined for redundancy. In choosing which features to remove, the potential timing of collection, objectivity, and clinical opinion were considered. Three feature sets were then created (Table A12.3, Appendix 12). Set 1 contained the features identified as most important in our previous work.[28] Set 2 contained only features that could potentially be collected in the perinatal period. Set 3 included those representing the child's early environment that would require later measurement. Recursive feature elimination (RFE), with five-fold cross validation repeated five times, was performed for each feature set to determine the optimal combination

for the final models. The features included in the final models are indicated by an asterisk in Table A12.3, Appendix 12.

Modelling

The dataset was imbalanced with regard to the outcome of interest. Training a model on class imbalanced data risks producing a model bias in favour of the majority class, with poor predictive performance for the minority class.[29] To address this Synthetic Minority Oversampling Technique (SMOTE) was applied to the training set only.[30] Unlike other oversampling techniques which simply duplicate minority class cases, SMOTE utilises an over-sampling approach where the minority class is over-sampled by creating synthetic examples.[29,30] Random forest (RF), logistic regression (LR), and support vector machine (SVM) algorithms were trained and optimal hyperparameters selected using the rebalanced training dataset and 10-fold cross validated grid search.

Evaluation

To select the most appropriate machine learning algorithm, accuracy across 10-fold cross-validation was compared. After selection of the algorithm, the final models were tested on an independent unseen test set. Area under the receiver operating curve (AUROC) was used to evaluate overall model performance. Explanation of performance metrics is provided in Supplemental Text Appendix 12. A summary of the modelling process is contained in Figure 7.1.

Explainability

Feature importance plots were created using the permutation method in the 'vip' package.[31] First, baseline model performance is measured using a measure set by the user (in this study – AUROC). The feature of interest is then randomly shuffled

and model performance is measured again. The difference between the two measures is used as a measure of feature importance. For each feature shuffling was simulated ten times and importance was averaged across the simulations. It would be expected that randomly shuffling the values of an important feature would degrade model performance.[31]

The relationships between the most important features and the outcome were examined using partial dependence plots (PDPs), which provide a visualisation of the relationship between a feature and the response while accounting for the average effect of the other features in the model.[32] A lower value on the y-axis of the PDP suggests that the positive class (low cognitive ability) is less likely at that value of the feature on the x-axis, according to the model.

7.4. Results

Characteristics of study population

There were 8,858 infants included, of whom n=573 (6.5%) had LCA. A summary of characteristics are described in Table 7.1, with a complete description provided in Table A12.4, Appendix 12. The LCA group, which was comprised of 60.6% boys, had a lower mean maternal age (30.6 v 32.1 years, p <0.001), a higher proportion of mothers who smoked (30.4% v 22.4%, p <0.001), and a lower proportion of mothers reporting English as their native language (53.9% v 86.9%, p<0.001). In the LCA group, 43.8% of mothers reported their highest education level as being primary or secondary only, compared with 27.5% of those without LCA. The LCA group had a lower median family income (€31,200 v €48,000, p<0.001), a lower

proportion of families living in owner occupied accommodation (38.9% v 73.5%, p<0.001), and a higher proportion in the lower social classes.

Rebalanced training dataset

In the study dataset, n=573/8,858 (6.5%) children had LCA. As detailed in Figure 1, in the training dataset, which contained 70% of participants, n=402/6,202 had LCA. After application of SMOTE, the rebalanced training dataset consisted of n=1,206 (42.9%) children with LCA and n=1,608 (57.1%) without. This rebalanced dataset was used to train the models. The testing dataset was not rebalanced and contained the original 30% of participants, of whom n=171/2,656 (6.5%) had LCA. This was the dataset used to evaluate the models and this data was not used in the training process.

Feature and Algorithm Selection

Following RFE 8 features were retained for Model 1, 15 for Model 2 and 23 features for Model 3. As shown in Table A12.5 Appendix 12, the random forest algorithm was the best performing algorithm for all three models in repeated 10-fold cross validation, and was selected for the final models.

Final Model Evaluation

The independent test set contained n=2,656 participants, of whom n=171 had LCA. Models 2 and 3 achieved the highest AUROC of 0.77 and 0.78 respectively (Figure A12.3, Appendix 12). At a decision threshold of 0.5 the models correctly predicted the cognitive outcome of 87% of participants at age 5 (Table 7.2). The odds of having low cognitive ability were 5.9 times higher in the group with a 'low' prediction. Model 2 was deemed to be the best performing model overall, achieving similar performance to Model 3 but using only 15 features, all of which

had potential to be collected in the perinatal period. To further explore its performance, the decision threshold was altered in increments of 0.05 and the corresponding sensitivities, specificities, positive and negative predictive values were examined using the independent test set. The optimal threshold of 0.28, using Youden's index, yielded a sensitivity of 72% with a specificity of 66%. (Table A12.6, Appendix 12).

A worked example of how the model, at a decision threshold of 0.5, would function in the real world is provided in Table A12.7, Appendix 12. In 2021 there were almost 60,000 births in Ireland. Assuming a prevalence of 6.5%, 3,900 infants will have LCA at age 5, of whom 1,560 would be detected in infancy (sensitivity 0.40). Of the 56,100 infants with average or above cognitive ability at age 5, 50,490 would be correctly identified as not at risk (specificity 0.90). There would be 5,610 false positives and 2,340 false negatives.

The modelling and results for predicting LCA using an alternative cut-off of a GCA score more than 1 SD below the mean are shown in Figure A12.4, Table A12.8, and Table A12.9, Appendix 12.

Explainability

The five most important features in Model 2 were PCG alcohol intake (measured when infant was 9-months old), household social class, PCG highest education, number of bedrooms in the home, and household equivalised income (Figure 7.2). PDPs are shown in Figure 7.3, alongside a histogram showing the distribution of each feature in the dataset.

The five most important features were all markers of socioeconomic status (SES). To determine whether any of these markers of SES could achieve similar

predictive performance alone, five separate random forest models were trained and tested using each feature alone as a predictor. As shown in Supplementary Table A12.10 Appendix 12, none of the features alone achieved as high an AUROC as the final model. Family income achieved an AUROC of 0.64 (95% CI 0.59-0.68). Each feature alone could achieve a high accuracy, but this was driven by high specificity with relatively poor sensitivity for detecting those with low cognitive ability.

7.5. Discussion

We have shown by evaluating a variety of machine learning algorithms in a large population-based cohort that a RF model based on 15 features, all of which have potential to be collected in the perinatal period, achieved an AUROC of 0.77 for predicting low cognitive ability at age 5. At a decision threshold of 0.5, the model could correctly predict the cognitive outcome of 87% of infants, however this was largely driven by a high specificity and its ability to detect the majority of children with normal cognitive ability. When the alternative cut off of 1 SD was used, an AUROC of 0.70 was achieved using 31 features in a RF model. Model performance was similar to that reported by Camargo et al, who developed a LR model with an AUROC of 0.75, to predict low IQ (defined by a z-score greater than 1 standard deviation below the mean) at age 6. Their model, however, included predictors that could not be measured until the infant was 12-months old.[16] No other similar predictive models were identified in the literature for comparison, and to our knowledge, this is the first study to examine the potential of machine learning for the prediction of low cognitive ability in childhood.

It is difficult to make direct comparisons of model performance with other screening tools due to differences in cohort characteristics, cognitive tests, and cut-off scores used to determine LCA. The available literature examining the performance of the Ages and Stages Questionnaire (ASQ) suggests very similar performance. The ASQ is one of the most widely used parent-reported screening tools and is currently recommended for early screening of developmental delay at two years of age in many countries, including by the American Academy of Pediatrics. The ASQ at 24 and 36 months have reported AUROCs of 0.64 and 0.78 respectively, for predicting low IQ at age 5, defined as an IQ more than 1SD below the cohort mean and an IQ less than 85 points in the respective studies.[33,34] A study examining the performance of the ASQ, performed between 8-30 months, for predicting low IQ in the early years of schooling reported similarly low sensitivities ranging from 28-50%, with specificities ranging from 79-96% across a range of cut-offs.[35] Our data suggest that similar prediction can be made at birth allowing early intervention in the most high risk cases.

The statistical approach used in this study was designed to optimise prediction, not investigate causal relationships. However, it is notable that among the ten most important predictors, six (alcohol intake, social class, education, income, bedrooms in the home, and maternal age) are inherently associated with socioeconomic status.[36,37] This is in keeping with a wealth of literature, both interventional and observational, which has consistently shown that the socioeconomic environment an infant is born into is one of the strongest predictors of cognitive outcomes in childhood.[38-42] The findings of this study would suggest that these features, while all representative of socioeconomic status, may

contribute to risk in different ways as their cumulative effects better predicted the outcome than the effect of any single feature.

Previous research has shown that children exposed to multiple early risk factors represent a more vulnerable subgroup for adverse childhood outcomes. The more risk factors a child is exposed to the worse these outcomes tend to be.[43] The most commonly used statistical approach to combining risk factors for the prediction of cognitive outcomes has been to combine them in an additive fashion, with the main effect of each factor accounted for.[15,16] However, we know that risk factor effects are not necessarily the same for all children and effect modification and interaction does exist.[44] The effect of one risk factor may be accentuated or diminished by another exposure. For example, the co-occurrence of brain injury due to preterm birth, a biological risk factor, and socioeconomic disadvantage, a social risk factor, have synergistic adverse effects on neurodevelopmental outcomes.[45] An advantage of the ML approach used in this study is that the algorithms can combine features in interactive relationships without the need for prior specification.

Our study has many limitations to consider. There is much debate about the merits and limitations of standardised cognitive assessments. Both the ICD-11 and the Diagnostic and Statistical Manual of Mental Disorders 5th Edition (DSM-V) now include the impact on adaptive functioning as a diagnostic criteria for IDDs, reflecting a move away from using standardised testing alone.[24,46] The purpose of this study was not to develop a diagnostic model, but a prognostic model that could identify high-risk children whose cognitive outcome could adversely impact other aspects of life, such as educational attainment and mental health.[47,48] In this

regard, the use of a binary cut-off is justified, although there is no consensus on the most appropriate one. In this study cut-offs of 1 and 1.5 SD were examined.

Cognitive ability was directly measured using standardised assessments, however children taking the naming vocabulary test, for whom English is not their native language, may be disadvantaged, with apparent poor performance. A more appropriate measure of cognitive ability in children from multilingual backgrounds may include total vocabulary across their languages, or novel language-independent cognitive assessments.[49,50] Children in GUI did not complete all BAS subtests and PCA was therefore used to calculate a GCA score. This may not be acceptable for a formal diagnosis of an IDD, however, a cited advantage of the BAS is that not all tests in the battery are required to assess performance and tests are individually interpretable.[21] While there is overlap in the mental processes used in different subtests, the two subtests chosen in the GUI study largely measure verbal and non-verbal ability, and are not primarily intended to measure numerical ability, spatial ability, perceptual speed, or memory. Performing the full battery of BAS subtests would provide a more robust assessment of cognitive ability, but for large population-based cohort studies may not be feasible.[21,51]

There is now a substantial body of evidence to suggest that early interventions are most effective when started in the first year of life.[11-13] For this reason, our study focussed on features which could be collected in the perinatal period at a population level. However, this approach may come at the cost of reduced predictive performance as detailed information on the child's early home and school environment were not included in the model. Information on the number of books in the home and the time spent on learning activities at age 3

were included in feature set C but predictive performance was not significantly better than with perinatal features alone. This likely reflects the fact that the early learning environment is intimately intertwined with the socioeconomic background of the family.

While all 15 features in the final model could be collected in the perinatal period, in the GUI study they were collected at infant age 9-months. Validation of the model in a cohort with data collection in the perinatal period is required, but this is challenging. There are more than 110 birth cohorts in Europe, however, there are no consensus guidelines on measurements, scales, data sources, or timing.[52] This leads to significant challenges with merging and harmonising datasets which could provide large pools of data for validation. Work in this area is ongoing.[53]

This study included a large sample size, and a careful sampling strategy was employed in recruitment of the cohort.[20] Attrition in the GUI study was relatively low and the sample included in this study who completed cognitive assessments at age 5 represented 79.6% of the original cohort recruited at 9-months. However, attrition was higher among those from disadvantaged backgrounds, who are also more likely to have poor cognitive outcomes.[54] Ideally, the model should be validated using population-based registry data which is not currently collected in Ireland.

Extensive validation is required to ensure wider generalisability of predictive models. If the relationship between predictor and outcome changes, model performance will be affected. For example, in Ireland only 35% of babies receive any breastmilk at 3 months , and this is significantly associated with socioeconomic status (SES).[55] In our model breastfeeding may be a surrogate marker of SES, a

relationship which may not exist in other countries where there are very high rates of breastfeeding.

An important challenge with ML models is explainability, the concept that the prediction a model makes can be explained in an acceptable way on a human level. In this study PDP plots were used to help understand how the model was making predictions. However, these must be interpreted with caution. For example, the PDP plot examining the relationship between PCG alcohol intake and risk of low cognitive ability in the child would suggest that those with moderate alcohol consumption have the lowest risk, while those with no alcohol or very high alcohol consumption have the highest risk. This U-shaped relationship curve between alcohol consumption and health outcomes has been described previously in the literature.[56] In our study, among those who reported English was not their native language 41.4% reported no alcohol intake, compared to only 12% of those for whom English was the native language. Therefore, it is plausible that the increased risk seen for those who reported no alcohol consumption is due to confounding, and the relationship is actually being driven by immigration, cultural, religious, language, or socioeconomic factors.

Finally the individual, health-system, and resource implications of a risk-based approach must be considered. In its current iteration this model may not be suitable for use at a population level. The sensitivity is too low and the resource implications of the high false positive rate, at a lower decision threshold, are too great. However, it is a first step toward enabling an early personalised, prediction, which could assist decisions on further intervention. This is not without consequence. Labelling a child so early in life could have detrimental effects on

both child and family. In addition to adverse impacts on the child's self-concept, it can perpetuate a self-fulfilling prophecy due to effects on parent and teacher expectations.[57,58] Using a label based largely on social factors, often rooted in inequities generated by social and political policy, may shift the blame of societal failures onto the individual child and family. Any risk stratification must come from a position of support, environmental enrichment, and education.

7.6. Conclusion

In conclusion, the current practice of waiting for overt signs of developmental delay before intervention goes against a large body of developmental literature. In this study, it was possible to develop a model, using 15 simple features available at birth and readily incorporated into an electronic health record, that correctly predicted the cognitive ability of 87% of children at age 5. Whilst further improvements in predictive performance would be required to use this as a population level screening tool, it provides a strong basis for further research. New methods of direct assessment of early cognitive function are lacking. Neurophysiological measures such as electroencephalography and eye-tracking show some promise.[42,59] Combined with targeted direct assessment, this model could, in the future, form the foundation for an early targeted screening protocol similar to that now being widely implemented for the early diagnosis of cerebral palsy.[60] Improved capacity to collect large volumes of rich data and to apply novel statistical methods, particularly suited to prediction, provides an opportunity for researchers and clinicians to investigate alternative approaches.

Author contributions

AB – Conceptualization, methodology, formal analysis, writing – original draft preparation, writing – review and editing

GL – Methodology, formal analysis, writing – review and editing

DO'B - Methodology, formal analysis, writing – review and editing

AS - Methodology, writing – review and editing

DM – Conceptualization, methodology, writing – review and editing

Table 7.3 Characteristics of study population and comparison of characteristics between those with and without low cognitive ability at age 5 years

Characteristics	Total	Low cognitive ability n=573 (6.5%)	Average or above cognitive ability n=8,285 (93.5%)	p-value
Pregnancy and Birth				
Mode of delivery				
Normal vaginal	5,239 (59.1)	367 (64.0)	4,875 (58.8)	
Forceps/suction assisted	1,299 (14.7)	60 (10.5)	1,238 (14.9)	
Planned c-section	1,151 (13.0)	61 (10.6)	1,090 (13.2)	
Emergency c-section	1,169 (13.2)	85 (14.8)	1,082 (13.1)	0.003[a]
Admission to NICU/SCBU	1,217 (13.7)	88 (15.4)	1,129 (13.6)	0.271[a]
Gestational age, mean (sd), wk	39.5 (2.1)	39.3 (2.4)	39.5 (2.0)	0.035[b]
Singleton pregnancy	8,546 (96.5)	290 (92.9)	7,995 (96.5)	0.075[a]
Birthweight				
<1,500g	67 (0.8)	6 (1.0)	61 (0.7)	
1500g-2500g	401 (4.5)	38 (6.6)	362 (4.4)	
2500g-4500g	8,178 (92.3)	509 (88.8)	7,670 (92.6)	
>4500g	212 (2.4)	20 (3.5)	192 (2.3)	0.0138[c]
Breastfed on discharge from hospital	4,131 (46.6)	300 (52.4)	4,427 (53.4)	0.648[a]
Smoker in household during pregnancy	2,801 (31.6)	225 (39.3)	2,576 (31.1)	<0.001[a]
Maternal				
Age, mean (sd), y	32.0 (5.2)	30.6 (5.7)	32.1 (5.1)	<0.001[b]
BMI, mean (sd), kg/m^2	25.5 (4.7)	25.9 (5.0)	25.5 (4.7)	0.072[b]
Self-rated health				
Excellent	2,798 (31.6)	159 (27.7)	2,639 (31.9)	
Very good	3,505 (39.6)	223 (38.9)	3,282 (39.6)	
Good	2,020 (22.8)	150 (26.2)	1,870 (22.6)	
Fair	475 (5.4)	35 (6.1)	440 (5.3)	
Poor	60 (0.6)	6 (1.0)	54 (0.7)	0.097[c]
Current smoker - yes	2,028 (22.9)	174 (30.4)	1,854 (22.4)	<0.001[a]
Alcohol intake				
Never	1,457 (16.4)	192 (33.5)	1,265 (15.3)	
<1/month	2,409 (27.2)	178 (31.1)	2,231 (26.9)	
1-2/month	2,398 (27.1)	119 (20.8)	2,279 (27.5)	
1-2/week	2,176 (24.6)	73 (12.7)	2,103 (25.4)	
3-4/week	353 (4.0)	10 (1.7)	343 (4.1)	
5-6/week	49 (0.6)	<5	48 (0.6)	

Daily	16 (0.2)	<5	16 (0.2)	<0.001[c]
Maternal depression, median (IQR)	1 (3)	1 (3)	1 (3)	0.0384[d]
Chronic illness	1,019 (11.5)	55 (9.6)	964 (11.6)	0.159[a]
Highest level of education				
Primary or less	186 (2.1)	43 (7.5)	143 (1.7)	
Lower secondary	739 (8.3)	77 (13.4)	662 (8.0)	
Upper secondary	1,608 (18.2)	131 (22.9)	1,477 (17.8)	
Technical/vocational	1,194 (13.5)	114 (19.9)	1,080 (13.0)	
Non Degree	1,787 (20.2)	74 (12.9)	1,713 (20.7)	
Degree or equivalent	1,617 (18.3)	68 (11.9)	1,549 (18.7)	
Post degree cert/diploma	1,047 (11.8)	36 (6.3)	1,011 (12.2)	
Post degree masters or PhD	680 (7.7)	30 (5.2)	650 (7.8)	<0.001[a]
Infant				
Male sex	4,479 (50.6)	347 (60.6)	4,132 (49.9)	<0.001[a]
Socioeconomic				
Household income, median (IQR)	48,000 (33,600)	31,200 (24,963)	48,000 (33,330)	<0.001[d]
Household social class				
Professional workers	1,711 (19.3)	52 (9.1)	1,659 (20.0)	
Managerial and technical	2,863 (32.3)	109 (19.0)	2,754 (33.2)	
Non-manual	1,472 (16.6)	85 (14.8)	1,387 (16.7)	
Skilled manual	1,168 (13.2)	122 (21.3)	1,046 (12.6)	
Semi-skilled	638 (7.2)	72 (12.6)	566 (6.8)	
Unskilled	119 (1.3)	15 (2.6)	104 (1.3)	
All others gainfully occupied	37 (0.4)	8 (1.4)	29 (0.4)	
Never worked	850 (9.6)	110 (19.2)	740 (8.9)	<0.001[a]
Accommodation occupancy				
Owner occupied	6,315 (71.3)	223 (38.9)	6,091 (73.5)	
Local authority housing	623 (7.0)	85 (14.8)	538 (6.5)	
Private rental	1,622 (18.3)	249 (43.5)	1,374 (16.6)	
Living with parents	146 (1.6)	8 (1.4)	138 (1.7)	
Occupied rent free	152 (1.7)	8 (1.4)	144 (1.7)	<0.001[a]
Number of bedrooms, mean (sd)	3.5 (0.9)	3.1 (0.9)	3.5 (0.9)	<0.001[b]
Connectedness in community				
Strongly agree	3,777 (42.6)	186 (32.5)	3,591 (43.3)	
Agree	4,292 (48.5)	320 (55.8)	3,972 (47.9)	
Disagree	618 (7.0)	54 (9.4)	564 (6.8)	
Strongly disagree	171 (1.9)	13 (2.3)	158 (1.9)	<0.001[a]

Early environment				
Has siblings	5,410 (61.1)	355 (62.0)	5,055 (61.0)	0.687[a]
Books in home				
None	29 (0.3)	6 (1.0)	23 (0.3)	
<10	585 (6.6)	111 (19.4)	474 (5.7)	
10-20	1,626 (18.4)	169 (29.5)	1,457 (17.6)	
21-30	1,562 (17.6)	105 (18.3)	1,457 (17.6)	
>30	5,056 (57.1)	182 (31.8)	4,874 (58.8)	<0.001[a]
Days/week learning activities done at home, mean (sd)	4.9 (1.2)	4.5 (1.4)	4.9 (1.2)	<0.001[b]
Level of support				
Get enough	6,446 (72.8)	344 (60.0)	6,102 (73.7)	
Don't get enough	856 (9.7)	49 (8.6)	807 (9.7)	
Don't get any	482 (5.4)	36 (6.3)	446 (5.4)	
Don't need any	458 (5.2)	51 (8.9)	407 (4.9)	
Family not in country	616 (7.0)	93 (16.2)	523 (6.3)	<0.001[a]
English maternal native language – yes	7,512 (84.8)	309 (53.9)	7,203 (86.9)	<0.001[a]
Partner in household	7,929 (89.5)	491 (85.7)	7,438 (89.8)	0.003[a]

Abbreviations: NICU, neonatal intensive care unit; SCBU, special care baby unit; sd, standard deviation

[a]Pearson's Chi-squared test, [b]Welch two sample t-test, [c]Wilcoxon rank sum test, [d]Fisher's Exact Test (p-value simulated where cell number too small).

Table 7.4 Performance of final models on unseen test set

	Model 1 (8 features)[a]	Model 2 (15 features)[b]	Model 3 (23 features)[c]
Accuracy	0.86 (0.85-0.88)	0.87 (0.85-0.88)	0.87 (0.85-0.88)
Sensitivity	0.25	0.40	0.43
Specificity	0.91	0.90	0.89
Positive predictive value	0.16	0.22	0.22
Negative predictive value	0.95	0.96	0.96
AUROC (95% confidence interval)	0.69 (0.65-0.73)	0.77 (0.73-0.80)	0.78 (0.74-0.81)

All performance metrics are calculated at a decision threshold of 0.5

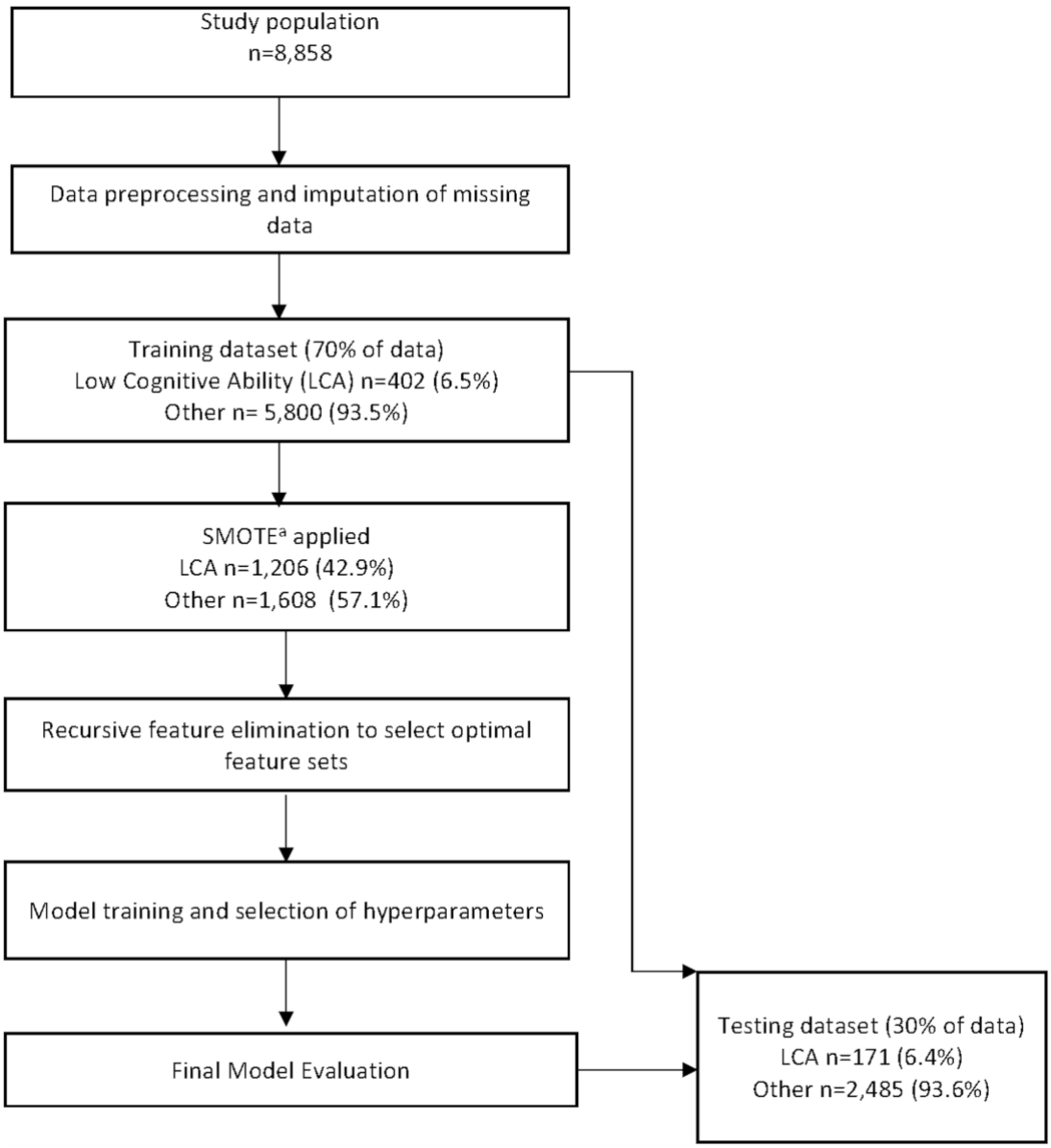
Abbreviations: AUROC, area under receiver operating characteristic curve

[a]Model 1 features – Household social class, household equivalised income, PCG age, gestational age, PCG BMI, PCG alcohol intake, maternal depression score, PCG highest education.

[b]Model 2 features – Household social class, household equivalised income, PCG age, gestational age, PCG BMI, PCG alcohol intake, maternal depression score, PCG highest education, number bedrooms in home, English native language of PCG, birthweight category, accommodation ownership, number in household smoking during pregnancy, community connectedness, PCG self-rated health

[c]Model 3 features - Household social class, household equivalised income, PCG age, gestational age, PCG BMI, PCG alcohol intake, maternal depression score, PCG highest education, number bedrooms in home, English native language of PCG, birthweight category, accommodation ownership, number in household smoking during pregnancy, community connectedness, PCG stress score, hours spent on learning activities, books in the home, hours of sleep for PCG, quality of attachment score, PCG employment status, verbal interaction between PCG and baby, level of support received by PCG, gestational diabetes diet.

Figure 7.4 Overview of modelling process



Footnote Figure 7.1

[a] Synthetic Minority Oversampling Technique

Figure 7.5 Ten most important features in random forest Model 2

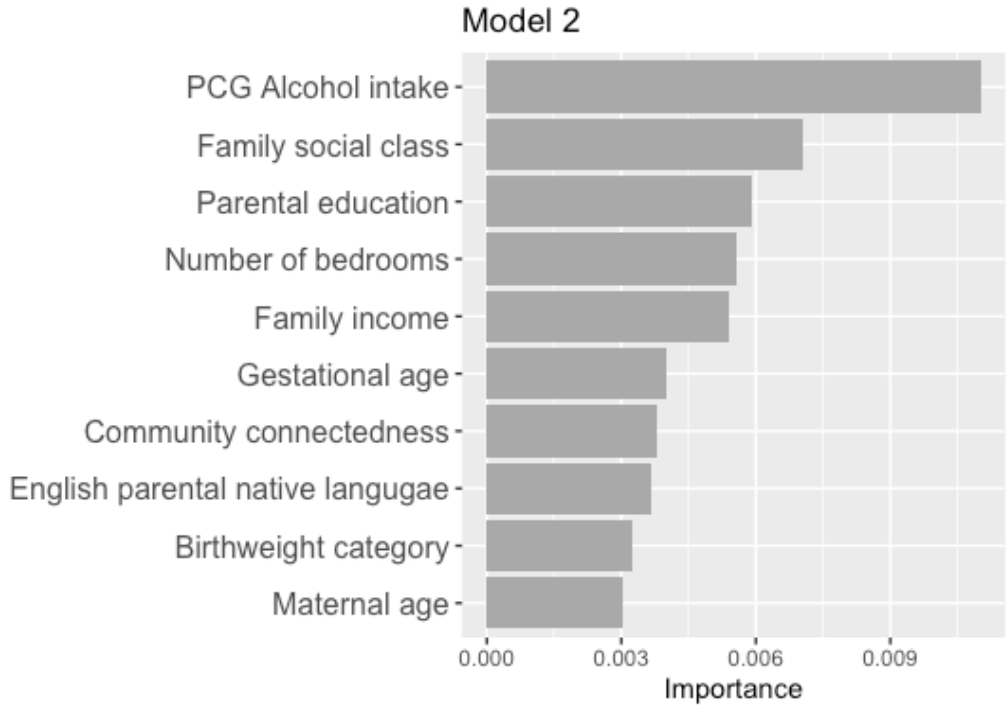


Figure 7.2 Legend
This figure plots the ten most important features used in random forest model 2 for predicting low cognitive ability at age 5. Importance was measured using the permutation method and importance values were averaged across ten simulations.

Figure 7.6 Partial dependence plots and histograms examining the relationship between the six most important features in random forest Model 2 and low cognitive ability

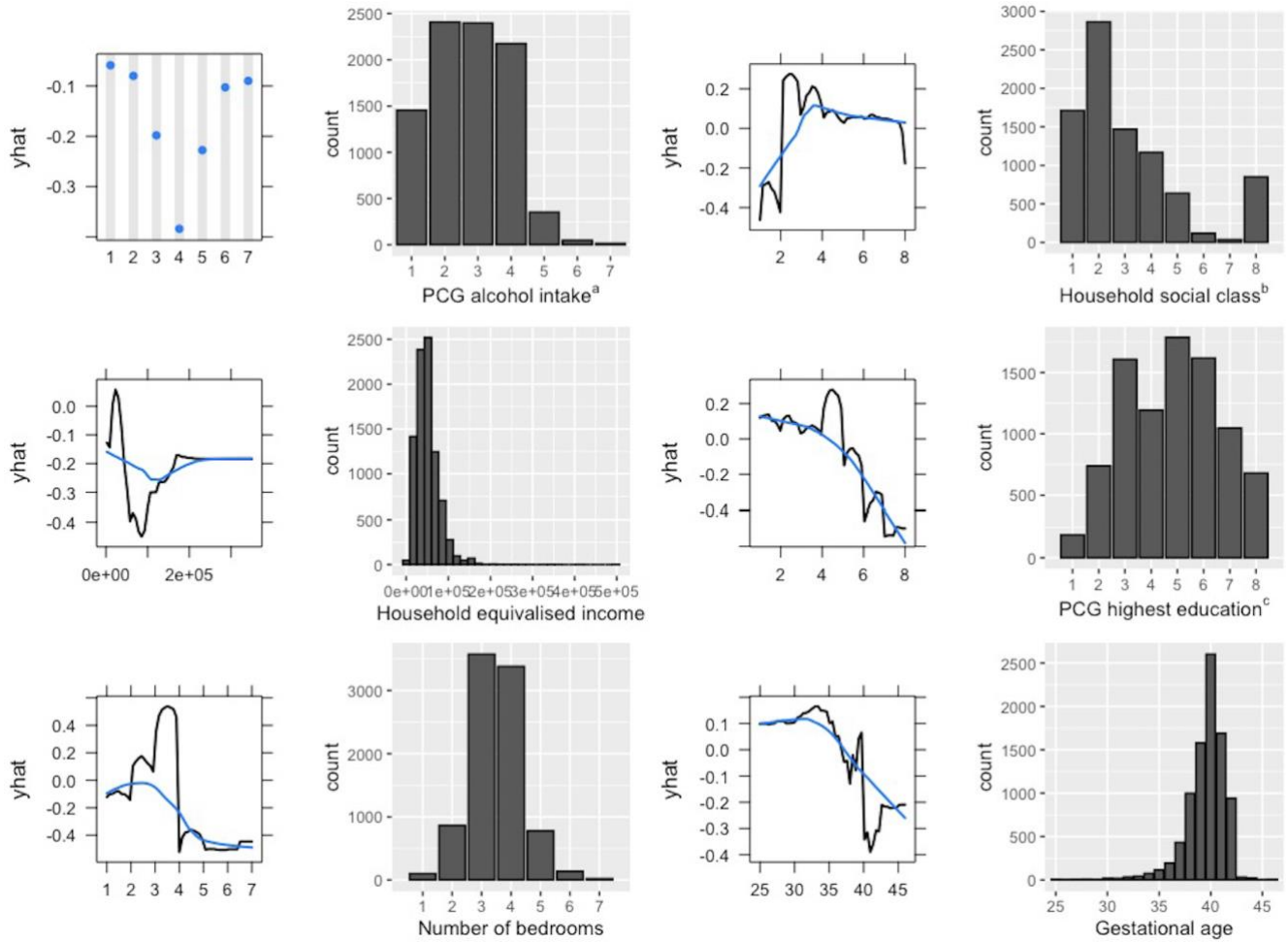


Figure 7.3 Legend

Shown in this figure are partial dependence plots (PDPs) for the six most important features in Model 2. Each PDP is accompanied by a histogram showing the distribution of the feature in the original unbalanced dataset. The PDP shows the cumulative effect (y-axis), according to our model, of each predictor's individual values (x-axis) on classification of the outcome. The absolute values on the y-axis are affected by class imbalance. The overall shape and directionality are important for interpretation. The effect of features of class factor are shown in dots (e.g. PCG alcohol intake) and the effect of features of class numeric are shown with a black line. The blue line represents a locally estimated scatterplot smoothing line (LOESS) which attempts to capture the general relationship while reducing noise. For example, in the PDP plot for household social class we can see that the highest social class categories predict low cognitive ability less strongly than the lower social class categories.

Footnote Figure 7.3

[a]Primary caregiver alcohol intake categories: 1, Never; 2, Less than once per month; 3, One to two times per month; 4, One to two times per week; 5, Three to four times per week; 6, Five to six times per week; 7, Everyday.

[b]Household social class categories: 1, Professional workers; 2, Managerial and technical; 3, Non-manual; 4, Skilled manual; 5, Semi-skilled; 6, Unskilled; 7, All others gainfully employed; 8, Never worked at all, no social class.

[c]Primary caregiver highest education categories: 1, Primary or less; 2, Lower secondary; 3, Upper secondary; 4, Technical or vocational qualification; 5, Non-Degree; 6, Degree or equivalent; 7, Post degree professional qualification; 8, Postgraduate masters or PhD.

7.7. References

1. Cioni G, Inguaggiato E, Sgandurra G. Early intervention in neurodevelopmental disorders: underlying neural mechanisms. Dev Med Child Neurol. 2016 Mar;58 Suppl 4:61-6. doi: 10.1111/dmcn.13050.
2. Ramey CT, Ramey SL. Prevention of intellectual disabilities: early interventions to improve cognitive development. Prev Med. 1998 Mar-Apr;27(2):224-32. doi: 10.1006/pmed.1998.0279. PMID: 9579000.
3. Pungello EP, Kainz K, Burchinal M, Wasik BH, Sparling JJ, Ramey CT, et al. Early educational intervention, early cumulative risk, and the early home environment as predictors of young adult outcomes within a high-risk sample. Child Dev. 2010 Jan-Feb;81(1):410-26. doi: 10.1111/j.1467-8624.2009.01403.x.
4. Spencer N, Raman S, O'Hare B, Tamburlini G. Addressing inequities in child health and development: towards social justice. BMJ Paediatr Open. 2019 Aug 1;3(1):e000503. doi: 10.1136/bmjpo-2019-000503.
5. Lager A, Bremberg S, Vågerö D. The association of early IQ and education with mortality: 65 year longitudinal study in Malmö, Sweden. BMJ. 2009 Dec 10;339:b5282. doi: 10.1136/bmj.b5282.
6. Alesi M, Rappo G, Pepi A. Emotional Profile and Intellectual Functioning: A Comparison Among Children With Borderline Intellectual Functioning, Average Intellectual Functioning, and Gifted Intellectual Functioning. SAGE Open. 2015; *5*(3). https://doi.org/10.1177/2158244015589995
7. Forrest LF, Hodgson S, Parker L, Pearce MS. The influence of childhood IQ and education on social mobility in the Newcastle Thousand Families birth

cohort. BMC Public Health. 2011 Nov 25;11:895. doi: 10.1186/1471-2458-11-895. PMID: 22117779; PMCID: PMC3248886.

8. Furnham A, Cheng H. Childhood Cognitive Ability Predicts Adult Financial Well-Being. J Intell. 2016 Dec 27;5(1):3. doi: 10.3390/jintelligence5010003.

9. Whalley LJ, Deary IJ. Longitudinal cohort study of childhood IQ and survival up to age 76. BMJ. 2001 Apr 7;322(7290):819. doi: 10.1136/bmj.322.7290.819.

10. Hirai AH, Kogan MD, Kandasamy V, Reuland C, Bethell C. Prevalence and Variation of Developmental Screening and Surveillance in Early Childhood. JAMA Pediatr. 2018 Sep 1;172(9):857-866. doi: 10.1001/jamapediatrics.2018.1524.

11. Campbell FA, Ramey CT, Pungello E, Sparling J, Miller-Johnson S. Early Childhood Education: Young Adult Outcomes From the Abecedarian Project. Appl Dev Sci. 2002; 6(1): 42–57. doi: 10.1207/S1532480XADS0601_05

12. Spittle AJ, Orton J, Doyle LW, Boyd R. Early developmental intervention programs post hospital discharge to prevent motor and cognitive impairments in preterm infants. Cochrane Database Syst Rev. 2007 Apr 18;(2):CD005495. doi: 10.1002/14651858.CD005495.pub2.

13. Whitehouse AJO, Varcin KJ, Pillar S, Billingham W, Alvares GA, Barbaro J, et al. Effect of Preemptive Intervention on Developmental Outcomes Among Infants Showing Early Signs of Autism: A Randomized Clinical Trial of Outcomes to Diagnosis. JAMA Pediatr. 2021 Nov 1;175(11):e213298. doi: 10.1001/jamapediatrics.2021.3298. Epub 2021 Nov 1.

14. Office of the Assistant Secretary for Planning and Evaluation. Poverty Guidelines [Internet]. 2023 [cited 2023 Sep 01]. Available at https://aspe.hhs.gov/topics/poverty-economic-mobility/poverty-guidelines

15. Eriksen HL, Kesmodel US, Underbjerg M, Kilburn TR, Bertrand J, Mortensen EL. Predictors of intelligence at the age of 5: family, pregnancy and birth characteristics, postnatal influences, and postnatal growth. PLoS One. 2013 Nov 13;8(11):e79200. doi: 10.1371/journal.pone.0079200.

16. Camargo-Figuera FA, Barros AJ, Santos IS, Matijasevich A, Barros FC. Early life determinants of low IQ at age 6 in children from the 2004 Pelotas Birth Cohort: a predictive approach. BMC Pediatr. 2014 Dec 16;14:308. doi: 10.1186/s12887-014-0308-1.

17. Azzolina D, Baldi I, Barbati G, Berchialla P, Bottigliengo D, Bucci A, et al. Machine learning in clinical and epidemiological research: Isn't it time for biostatisticians to work on it? Epidemiol Biostat Public Health.2019;16(4).doi: 10.2427/13245.

18. Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. N Engl J Med. 2016 Sep 29;375(13):1216-9. doi: 10.1056/NEJMp1606181. PMID: 27682033; PMCID: PMC5070532.

19. Bowe AK, Lightbody G, Staines A, Murray DM. Big data, machine learning, and population health: predicting cognitive outcomes in childhood. Pediatr Res. 2023 Jan;93(2):300-307. doi: 10.1038/s41390-022-02137-1. Epub 2022 Jun 9. PMID: 35681091; PMCID: PMC7614199.

20. Quail A, Williams J, McCrory C, Murray A, Thornton M. A Summary Guide to Wave 1 of the Infant Cohort of Growing Up in Ireland [Internet]. Dublin;

2011 [cited 2023 Mar 01] Available at: https://www.growingup.gov.ie/pubs/Summary-Guide_Infant-Cohort_Wave-1.pdf

21. Elliott CD, Smith P, McCulloch K. British Ability Scales Second Edition (BAS II). Administration and Scoring Manual.1996; London: Nelson.

22. Williams J, Thornton M, Murray A, Quail A. Growing Up in Ireland Design, Instrumentation and Procedures for Cohort '08 at Wave 3 (5 Years) [Internet].Dublin; 2019 [cited 2023 Sep 25]. Available at https://www.growingup.gov.ie/pubs/20190404-Cohort-08-at-5years-design-instrumentation-and-procedures.pdf

23. Jones EM, Schoon I. Child cognition and behaviour. Millennium Cohort Study Third Survey: A user's guide to initial findings [Internet]. 2008 Oct:118-44. [cited 2023 Sep 26]. Available at https://core.ac.uk/download/pdf/111052317.pdf#page=126

24. World Health Organisation. ICD-11 for Mortality and Morbidity Statistics: 6A00 Disorders of intellectual development.[Internet] 2023 [cited 2023 Sep 05]. Available at: https://icd.who.int/browse11/l-m/en#/http%253A%252F%252Fid.who.int%252Ficd%252Fentity%252F605267007

25. Karrasch M, Tiitta P, Hermann B, Joutsa J, Shinnar S, Rinne J, et al. Cognitive Outcome in Childhood-Onset Epilepsy: A Five-Decade Prospective Cohort Study. J Int Neuropsychol Soc. 2017 Apr;23(4):332-340. doi: 10.1017/S1355617716001077. Epub 2017 Jan 10.

26. Stekhoven DJ, Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012 Jan 1;28(1):112-8. doi: 10.1093/bioinformatics/btr597. Epub 2011 Oct 28. PMID: 22039212.

27. McCormick BJJ, Caulfield LE, Richard SA, Pendergast L, Seidman JC, Maphula A, et al; MAL-ED NETWORK INVESTIGATORS. Early Life Experiences and Trajectories of Cognitive Development. Pediatrics. 2020 Sep;146(3):e20193660. doi: 10.1542/peds.2019-3660. Epub 2020 Aug 17.

28. Bowe AK, Lightbody G, Staines A, Kiely ME, McCarthy FP, Murray DM. Predicting Low Cognitive Ability at Age 5-Feature Selection Using Machine Learning Methods and Birth Cohort Data. Int J Public Health. 2022 Nov 10;67:1605047. doi: 10.3389/ijph.2022.1605047.

29. Blagus R, Lusa L. SMOTE for high-dimensional class-imbalanced data. BMC Bioinformatics. 2013 Mar 22;14:106. doi: 10.1186/1471-2105-14-106.

30. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res. 2002; 16: 321–357. https://doi.org/10.1613/jair.953

31. Greenwell B, Boehmke B. Variable Importance Plots - An Introduction to the vip Package [Internet]. R J. 2020 [cited 2023 Sep 25]; 12: 343–366. Available from: https://journal.r-project.org/archive/2020/RJ-2020-013/RJ-2020-013.pdf

32. Greenwell B. pdp: An R Package for Constructing Partial Dependence Plots [Internet]. R J. 2017 [cited 2023 Sep 25]; 9: 421–436. Available from: https://journal.r-project.org/archive/2017/RJ-2017-016/RJ-2017-016.pdf

33. Bowe AK, Hourihane J, Staines A, Murray DM. The predictive value of the ages and stages questionnaire in late infancy for low average cognitive ability at age 5. Acta Paediatr. 2022 Jun;111(6):1194-1200. doi: 10.1111/apa.16309. Epub 2022 Mar 3. PMID: 35202483; PMCID: PMC9314849.

34. Charkaluk ML, Rousseau J, Calderon J, Bernard JY, Forhan A, Heude B, et al; EDEN Mother–Child Cohort Study Group. Ages and Stages Questionnaire at 3 Years for Predicting IQ at 5-6 Years. Pediatrics. 2017 Apr;139(4):e20162798. doi: 10.1542/peds.2016-2798. PMID: 28360034.

35. Schonhaut BL, Pérez RM, Castilla FAM, Castro MS, Salinas AP, Armijo RI. Validez del Ages & Stages questionnaires para predecir el desempeño cognitivo en los primeros años de educación escolar [Predictive value of Ages & Stages Questionnaires for cognitive performance at early years of schooling]. Rev Chil Pediatr. 2017 Feb;88(1):28-34. Spanish. doi: 10.1016/j.rchipe.2016.08.008. PMID: 28288224.

36. Daly MC, Duncan GJ, McDonough P, Williams DR. Optimal indicators of socioeconomic status for health research. Am J Public Health. 2002 Jul;92(7):1151-7. doi: 10.2105/ajph.92.7.1151. Erratum in: Am J Public Health 2002 Aug;92(8):1212. PMID: 12084700; PMCID: PMC1447206.

37. van Roode T, Sharples K, Dickson N, Paul C. Life-Course Relationship between Socioeconomic Circumstances and Timing of First Birth in a Birth Cohort. PLoS One. 2017 Jan 13;12(1):e0170170. doi: 10.1371/journal.pone.0170170.

38. Bradley RH, Corwyn RF. Socioeconomic status and child development. Annu Rev Psychol. 2002;53:371-99. doi: 10.1146/annurev.psych.53.100901.135233. PMID: 11752490.

39. Larson K, Russ SA, Nelson BB, Olson LM, Halfon N. Cognitive ability at kindergarten entry and socioeconomic status. Pediatrics. 2015 Feb;135(2):e440-8. doi: 10.1542/peds.2014-0434. Epub 2015 Jan 19. PMID: 25601983.

40. Tong S, Baghurst P, Vimpani G, McMichael A. Socioeconomic position, maternal IQ, home environment, and cognitive development. J Pediatr. 2007 Sep;151(3):284-8, 288.e1. doi: 10.1016/j.jpeds.2007.03.020. Epub 2007 Jun 22. PMID: 17719939.

41. von Stumm S, Plomin R. Socioeconomic status and the growth of intelligence from infancy through adolescence. Intelligence. 2015;48:30–6. doi:10.1016/J.INTELL.2014.10.002

42. Troller-Renfree SV, Costanzo MA, Duncan GJ, Magnuson K, Gennetian LA, Yoshikawa H, et al. The impact of a poverty reduction intervention on infant brain activity. Proc Natl Acad Sci U S A. 2022 Feb 1;119(5):e2115649119. doi: 10.1073/pnas.2115649119.

43. Evans GW. A multimethodological analysis of cumulative risk and allostatic load among rural children. Dev Psychol. 2003 Sep;39(5):924-33. doi: 10.1037/0012-1649.39.5.924.

44. Evans GW, Li D, Whipple SS. Cumulative risk and child development. Psychol Bull. 2013 Nov;139(6):1342-96. doi: 10.1037/a0031808.

45. Benavente-Fernández I, Synnes A, Grunau RE, Chau V, Ramraj C, Glass T, et al. Association of Socioeconomic Status and Brain Injury With Neurodevelopmental Outcomes of Very Preterm Children. JAMA Netw Open. 2019 May 3;2(5):e192914. doi: 10.1001/jamanetworkopen.2019.2914. PMID: 31050776; PMCID: PMC6503490.

46. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders [Internet]. Fifth Edition. Arlington, VA, 2013 [cited 2023 Sep 5]. Available from: https://doi/book/10.1176/appi.books.9780890425596

47. Deary IJ, Strand S, Smith P, Fernandes, C. Intelligence and educational achievement. Intelligence. 2007;35(1):13–21. https://doi.org/10.1016/j.intell.2006.02.001

48. Bowe AK, Staines A, Murray DM. Below Average Cognitive Ability-An under Researched Risk Factor for Emotional-Behavioural Difficulties in Childhood. Int J Environ Res Public Health. 2021 Dec 8;18(24):12923. doi: 10.3390/ijerph182412923.

49. Hoff E, Core C. What Clinicians Need to Know about Bilingual Development. Semin Speech Lang. 2015 May;36(2):89-99. doi: 10.1055/s-0035-1549104. Epub 2015 Apr 29. PMID: 25922994; PMCID: PMC5021218.

50. Twomey DM, Wrigley C, Ahearne C, Murphy R, De Haan M, Marlow N, et al. Feasibility of using touch screen technology for early cognitive assessment in children. Arch Dis Child. 2018 Sep;103(9):853-858. doi: 10.1136/archdischild-2017-314010. Epub 2018 Mar 13. PMID: 29535111.

51. The British Psychological Society. Test Review British Ability Scales Second Edition (BAS II) Early Years.[Internet].2014 [cited 2023 Sep 25]. Available at https://explore.bps.org.uk/binary/bpsworks/8c6cc30167a91a68/49da8e25db69400d9493abefd966956872cb302b488b2e54a8b2acdf2c769cb6/1854335138.pdf

52. Pansieri C, Pandolfini C, Clavenna A, Choonara I, Bonati M. An Inventory of European Birth Cohorts. Int J Environ Res Public Health. 2020 Apr 28;17(9):3071. doi: 10.3390/ijerph17093071. PMID: 32354101; PMCID: PMC7246583.

53. Jaddoe VWV, Felix JF, Andersen AN, Charles MA, Chatzi L, Corpeleijn E, et al; LifeCycle Project Group. The LifeCycle Project-EU Child Cohort Network: a federated analysis infrastructure and harmonized data of more than 250,000 children and parents. Eur J Epidemiol. 2020 Jul;35(7):709-724. doi: 10.1007/s10654-020-00662-z.

54. Murray A, Williams J, Quail A, Neary M, Thornton M. Growing Up in Ireland A Summary Guide to Wave 3 of the Infant Cohort (at 5 Years) [Internet]. Dublin; 2019 [cited 2023 Sep 27]. Available at: https://www.growingup.gov.ie/pubs/Summary-Guide-_Infant-Cohort_Wave-3.pdf

55. Gallegos D, Parkinson J, Duane S, Domegan C, Jansen E, Russell-Bennett R. Understanding breastfeeding behaviours: a cross-sectional analysis of associated factors in Ireland, the United Kingdom and Australia. Int Breastfeed J. 2020 Dec 2;15(1):103. doi: 10.1186/s13006-020-00344-2. PMID: 33267900; PMCID: PMC7709394.

56. San José B, van de Mheen H, van Oers JA, Mackenbach JP, Garretsen HF. The U-shaped curve: various health measures and alcohol drinking patterns. J Stud Alcohol. 1999 Nov;60(6):725-31. doi: 10.15288/jsa.1999.60.725.

57. Jussim L, Harber KD. Teacher expectations and self-fulfilling prophecies: knowns and unknowns, resolved and unresolved controversies. Pers Soc Psychol Rev. 2005; 9: 131–155. https://doi.org/10.1207/s15327957pspr0902_3

58. Shifrer D. Stigma of a label: educational expectations for high school students labeled with learning disabilities. J Health Soc Behav. 2013;54(4):462-80. doi: 10.1177/0022146513503346. PMID: 24311756.

59. Dean B, Ginnell L, Ledsham V, Tsanas A, Telford E, Sparrow S, et al. Eye-tracking for longitudinal assessment of social cognition in children born preterm. J Child Psychol Psychiatry. 2021 Apr;62(4):470-480. doi: 10.1111/jcpp.13304. Epub 2020 Jul 29. PMID: 32729133.

60. Novak I, Morgan C, Adde L, Blackman J, Boyd RN, Brunstrom-Hernandez J, et al. Early, Accurate Diagnosis and Early Intervention in Cerebral Palsy: Advances in Diagnosis and Treatment. JAMA Pediatr. 2017 Sep 1;171(9):897-907. doi: 10.1001/jamapediatrics.2017.1689. Erratum in: JAMA Pediatr. 2017 Sep 1;171(9):919. PMID: 28715518; PMCID: PMC9641643.

61. Maccow G. Overview of the Differential Ability Scales – Second Edition. Pearson Clinical [Internet].[cited 2023 Sep 25]. Available from https://pearsonclinical.in/wp-content/uploads/2018/03/DAS_II_WebinarHandout1192011.pdf

Chapter 8 Predicting 2-year cognitive outcomes in very preterm infants using machine learning methods

Andrea K. Bowe[1], Gordon Lightbody[1,2], Anthony Staines[3], Deirdre M. Murray[1,4], Mikael Norman[5]

[1] INFANT Research Centre, University College Cork, Cork, Ireland

[2] Department of Electrical and Electronic Engineering, University College Cork, Cork, Ireland

[3] School of Nursing, Psychotherapy, and Community Health, Dublin City University, Ireland

[4] Department of Paediatrics, Cork University Hospital, Cork, Ireland

[5] Department of Clinical Science, Intervention, and Technology, Karolinska Institutet, Stockholm, Sweden

[6] Department of Neonatal Medicine, Karolinska University Hospital, Stockholm, Sweden

Published in JAMA Network Open

December 2023

8.1.Abstract

Importance

Early intervention can improve cognitive outcomes for very preterm infants but is resource intensive. Identifying those who need it most is important.

Objective

To develop a model for use in very preterm infants to predict cognitive delay at 2 years of age using routinely available clinical and sociodemographic data.

Design

This prognostic study was based on the Swedish Neonatal Quality Register. Nationwide coverage of neonatal data was reached in 2011, and registration of follow-up data opened in January 2015 with inclusion herein ending in September 2022. A variety of machine learning models were trained and tested to predict cognitive delay.

Setting

All neonatal units in Sweden.

Participants

Surviving infants with a gestational age <32 weeks and complete data for the Bayley Scales of Infant Development 3rd Edition cognitive index or cognitive scale scores at 2 years of corrected age. Infants with major congenital anomalies were excluded.

Exposures

A total of 90 variables (containing sociodemographic and clinical information on conditions, investigations, and treatments initiated during pregnancy, delivery, and neonatal unit admission) were examined for predictability.

Main outcome

Cognitive function at 2 years, categorized as screening positive for cognitive delay (cognitive index score <90) or exhibiting typical cognitive development (score ≥90).

Results

The median (IQR) birthweight was 880 (380) g, 53% were boys, extremely preterm infants (gestational age <28 weeks) constituted 72% of children with complete follow-up data, and the prevalence of cognitive delay was 21.8% (n=231/1062). A logistic regression model containing 26 predictive features achieved an area under the receiver operating curve of 0.77 (95% confidence interval 0.71-0.83). The five most important features for cognitive delay were non-Scandinavian family language, prolonged duration of hospitalisation, low birthweight, discharge to other destination than home, and the infant not receiving breastmilk on discharge. At discharge from the neonatal unit, the full model identified 93% of infants with cognitive delay at 2-year follow up, with a specificity of 46%.

Conclusions and Relevance

Predictive modelling in neonatal care could enable early and targeted intervention for very preterm infants most at risk for developing cognitive impairment.

8.2.Introduction

Each year, 2.2 million infants are born very preterm (VPT), that is, at a gestational age (GA) younger than 32 weeks.[1] Accompanying improved survival, more children are at risk for long-term impairments associated with prematurity, such as cognitive difficulties.[2-5] The mean IQ, across childhood, for VPT infants has been reported to be almost 1 standard deviation (SD) lower than their full-term counterparts, equivalent to a mean difference of approximately 13 IQ points.[6] This finding has important implications for academic achievement, socioeconomic status, and physical and mental health.[7-12] Since 1990, the preterm infant mortality rate in high-income countries has decreased by 4% per year, yet no significant improvements in cognitive outcomes have been achieved.[2,6,13] The causal pathways underpinning the association between VPT birth and cognitive impairment remain unclear, making it difficult to implement primary prevention strategies.[6] Secondary prevention, in the form of early intervention, has so far been the only effective way forward.[14] However, many of the interventions studied were highly resource intensive—commencing in the hospital, followed by frequent visits to the families in their homes, for durations of up to 3 years.[14] Early and accurate identification of infants most in need of these interventions has therefore become increasingly important. Biological and environmental factors, as well as clinical events in the perinatal period, contain prognostic information about cognitive function in preterm children younger than 5 years.[15] Various prognostic models have been reported, but few have been able to fully account for nonlinear relationships and interactions between predictors or have validated their models on unseen data.[16] The aim of this study was to develop an explainable algorithm for use in VPT infants to predict

cognitive delay (CD) at 2 years of age. For this purpose, perinatal characteristics of those who did and did not experience CD at 2 years of age were described. A data-driven approach to identify the relevant predictors of cognitive outcome was then used, and a variety of machine learning models were trained and subsequently tested on an unseen test set. Finally, the most important predictive features were identified.

8.3. Methods

The Swedish Ethical Review Authority approved this study on March 28, 2022 (application number 022-01351-01). This authority also waived informed consent from parents or caregivers, who had been informed that perinatal and follow-up data were registered, with a possibility to opt out at any time (used by <5 families in the study period). The Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) guidelines were followed in the reporting of this study.[17]

Data

Data were from the Swedish Neonatal Quality Register (SNQ). The SNQ is a nationwide, population-based register that captures all infants admitted for neonatal care. The register provides detailed clinical information on conditions, investigations, and treatments initiated during pregnancy, delivery, and neonatal unit admission. Nationwide coverage of neonatal data was reached in 2011, and registration of follow-up data opened on January 1, 2015, with inclusion ending on September 31, 2022. Details on data reporting, completeness and validity of the register have been previously published.[18] As part of a national follow-up program

launched in 2015, high-risk survivors of neonatal intensive care were recommended to receive neurodevelopmental assessment, including the Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III) scheduled at corrected age of 2 years. The primary aim of the follow-up was to screen for different impairments, particularly in those with a GA younger than 28 weeks.[19] The results of the assessments should be reported to SNQ.

Study population

Eligible for inclusion in this study were VPT 2-year-old children with data for either BSID-III cognitive index or scale score. Children with a cognitive assessment but having a major congenital anomaly as defined by the European Registration of *Congenital* Anomalies and Twins[20] (EUROCAT) (n=142) were excluded. There were 59 VPT children who attended for BSID-III assessment but could not complete the test and reasons are contained in the footnote to Table 8.1.

Outcome

The outcome was cognitive ability at 2-year follow-up, measured using the BSID-III. The BSID-III has not been standardised on a Swedish population. However, in healthy Swedish controls born at term (n=366) and assessed at 30 months of age, the mean (SD) cognitive index score was 104 (10.6).[21] On the basis of that study, the SNQ follow-up programme applies a threshold score of less than 90 (approximately a –1 SD cut-off), to identify children who may benefit from further investigation or intervention.[19] In this study, children with a cognitive index score less than 90 were categorised as CD and those with a score of 90 or greater as typical cognitive development (TCD).

Descriptive statistics

The characteristics of participants were compared using Pearson χ2 and Fisher exact tests (categorical data), Welch's 2 sample t test (normally distributed continuous data), and Wilcoxon rank sum test (data with nonnormal distribution). Hypothesis tests were 2-tailed, and a priori significance was set at $P < .05$.

Data preparation

Data were handled with the R statistical software package, version 4.1.1 (R Project for Statistical Computing).[22] A study data set containing 97 potential predictive features, chosen based on SNQ contents, previous literature, plausible hypotheses, and senior author input, was created. Any feature with more than 25% missing values was removed from the data set (Table A13.1, Appendix 13), leaving 90 features categorized as pregnancy; birth; neonatal resuscitation; neonatal respiratory, neurological, infection, or other illnesses; and sociodemographic factors. Remaining missing values (Table A13.2, Appendix 13) were imputed using the missforest package, a random forest imputation method.[23,24]

Feature Selection

An "all relevant" approach, aiming to identify all features relevant for the classification, was adopted and the Boruta algorithm was used.[25,26] For each feature, a corresponding "shadow" feature was created by randomly shuffling the values of the original feature, thereby destroying any possible association with the outcome. A random forest classifier was trained on the original and shadow features. The importance value of each original feature was compared with a threshold—defined as the maximum importance value recorded among the shadow features. If a feature's importance was higher than this threshold, it was recorded in a vector as a hit. This process was iterated. Features that significantly

outperformed the threshold were confirmed important, and those that significantly underperformed were rejected and removed from further iterations. The algorithm stopped when a predefined maximum number of 600 iterations was reached and the remaining features were labelled as tentative.[25,26]

Correlated Features

Pearson correlation coefficients were calculated and plotted in a pairwise manner for all features identified as important or tentative. Features with coefficients greater than 0.70 were examined for redundancy. In choosing which features to retain, robustness, dimensionality, correlation with the outcome, expert opinion, and effect on prediction were considered.

Modelling

The dataset was stratified by the outcome and randomly split into a training set containing 70% of the data, and a testing set containing 30%. The outcome classes in the dataset were imbalanced, which risked producing a predictive model with a bias toward the majority class. To address this, the synthetic minority oversampling technique (SMOTE) was applied to the training dataset.[27,28]

Using all relevant features identified by Boruta, a random forest, logistic regression, support vector machine, and gradient boosting machine algorithm were trained and optimal hyperparameters selected in cross validated grid search using the 'caret' package (Models A-D).[29] These four algorithms were chosen as they have been widely used in healthcare research, are more likely to be recognised by the clinicians intended to use them, are efficient, and have potential for explainability.[30,31]

Evaluation

Internal validation was performed by examining accuracy in 10-fold cross validation. This involved splitting the training dataset into 10 parts containing equal numbers of observations. At each iteration the model was trained using 90% (9 parts) of the data and validated on the remaining 10%. The procedure was repeated ten times using a different 10% for validation each time and results were averaged across the folds. The decision threshold was set at a probability of 0.5.

The final models with optimally tuned hyperparameters were then externally validated by testing their performance on the unseen test dataset. Accuracy, balanced accuracy, sensitivity, and specificity were compared. To examine performance across all possible decision thresholds the area under the receiver operating curve (AUROC) was used to examine performance across all possible decision thresholds.[32] The modelling process is summarised in Figure 8.1.

Explainability

The explainability of the best performing models on external validation was explored by examining feature importance and feature effect. Feature importance plots were created using the test dataset and the permutation method in the 'vip' package.[33] Baseline model performance was measured using AUROC. The feature of interest was then randomly shuffled, and model performance was measured again. The difference between the 2 measures was used as a measure of feature importance. For each feature, shuffling was simulated ten times and importance was averaged across the simulations. Feature effect and more detailed methods are described in online-only material (Appendix 13).

8.4. Results

Characteristics of the study population

The perinatal, neonatal, and family characteristics of the 1062 children (median [IQR] birth weight, 880 [720-1100] g; 566 [53.3%] male and 496 [46.7%] female) included in the modelling process are shown in Table 8.1. A description of the cohort across all 90 features included in the selection process is contained in Table A13.3, Appendix 13. A total of 231 children (21.8%) screened positive for CD. The proportion of male children was higher in the CD group than the TCD group (152 [65.8%] vs 414 [49.8%]; P < .001). Infants in the CD group had a lower mean (SD) GA (26.0 [2.2] vs 26.7 [2.2] weeks; P < .001) and a lower median (IQR) birth weight (800 [660-980] vs 900 [760-1125] g; P < .001) than the TCD group. A higher proportion of children with CD than TCD were intubated during neonatal resuscitation (106 [46.7%] vs 265 [32.1%]; P < .001). There were also significant differences between the groups in terms of the respiratory complications and oxygenation requirements. The proportion with bronchopulmonary dysplasia was significantly higher in the CD than in the TCD group (162 [70.1%] vs 475 [57.2%]; P < .001), as was the proportion requiring postnatal steroids (103 [44.6%] vs 238 [28.6%]; P < .001). The median durations of continuous positive airway pressure (CPAP) and conventional or high-frequency oscillatory ventilation were longer for the CD group. Overall, the incidence of intraventricular haemorrhage (IVH) of any grade was higher (83 [36.7%] vs 217 [26.5%]), and higher grades were more frequently seen, in the CD than in the TCD group. The incidence of non-IVH intracranial haemorrhage was 1.1% (12 of 1062), and the incidence of cystic periventricular leukomalacia was 1.6% (17 of 1062), with a higher incidence in the CD group. Infants in the CD group had longer mean (SD) neonatal hospital stay (83 [42.5]

vs 72 [40.0] days; P < .001), and a lower proportion of these infants received any breastmilk at discharge (99 [53.8%] vs 494 [67.8%]). Lower proportions of parents of the CD than the TCD group reported more than 12 years of education (81 [44.3%] vs 395 [58.4%]; P < .001) and Scandinavian language as being the family language (100 [52.4%] vs 541 [74.8%]; P < .001).

Training and testing datasets

The original training data set contained 744 infants, of whom 162 (21.8%) screened positive for CD. After SMOTE was applied, the training set consisted of 1134 participants, of whom 468 (41.3%) had CD. The test set was composed of 318 infants, of whom 69 (21.7%) had CD. SMOTE was not applied to the test set.

Feature Selection

The Boruta algorithm confirmed 27 of 90 features as important, with a further 4 labelled tentative (Figure A13.1, Appendix 13). Among these 31 features, there were 12 features with correlation coefficients greater than 0.70 (Figure A13.2, Appendix 13). Five features were removed to reduce collinearity (Apgar score at 5 minutes, total duration of any mechanical ventilation, duration of supplemental oxygen, duration of hospital and homecare combined, and z-score for birthweight) leaving 26 features detailed in Table A13.4, Appendix 13. Duration of hospitalisation and duration of CPAP, as well as birthweight, GA, and head circumference, were highly correlated but predictive performance was better with all features included. The z-score for birthweight was removed as this feature was derived from both GA and birthweight, and predictive performance was better when the source features were included.

Model Training and Evaluation

The internal validation results are shown in Table A13.5, Appendix 13. As shown in Table 8.2 and Figure A13.3 Appendix 13, on the unseen test data, all models achieved an AUROC of greater than 0.70. Model D (gradient boosting machine) had an accuracy of 0.77 for identifying infants as having either CD or TCD at 2-year follow-up, with a sensitivity of 0.55 and a specificity of 0.83 at a decision threshold of 0.5. Model B (logistic regression) had an accuracy of 0.76, with a sensitivity of 0.48 and a specificity of 0.84. Model A (random forest) had an accuracy of 0.74, and model C (support vector machine) had an accuracy of 0.70. Although Model D achieved the highest accuracy of 0.77 at the 0.5 threshold, the AUROC curve suggested that model B could achieve the highest sensitivity with alteration of the threshold. Model B, a logistic regression model containing 26 features, achieved an AUROC of 0.77 (95% CI, 0.71-0.83).

Application of Model B

The decision threshold of Model B was lowered to 0.20 to improve sensitivity. A contingency table showing a hypothetical application is shown in Table 8.3. Among 3,000 VPT infants without major malformations, it would be expected that approximately 650 would have CD at 24 months. At discharge from the neonatal unit, model B could correctly identify 605 of 650 infants who would have CD at 24 months (sensitivity 0.93), and 1081 of 2350 who would not (specificity 0.46). Among 1874 infants predicted to have CD, 605 (32.3%) would go on to have it. Among 1126 infants predicted not to have CD, 1081 (96.0%) would not have it.

Feature Importance

Feature importance plots for Models A-D are shown in Figure 8.2. The 10 most important features for prediction in Model B were (listed in order of importance):

family language, duration of hospitalisation, birthweight, whether infant was discharged to home, whether infant was receiving breastmilk on discharge, grade of IVH, sex, head circumference, use of an umbilical artery catheter, and non-IVH intracranial haemorrhage.

Family language, birthweight, sex, duration of hospitalisation, whether infant was receiving breastmilk at discharge, and whether the infant was discharged directly home were included among the 10 most important features in all models. Feature effect plots are contained in Figure A13.4 and Figure A13.5, Appendix 13.

8.5. Discussion

Using clinical data that are easily available, we have shown that a predictive model containing 26 features could already at discharge from the neonatal unit identify 93% of surviving VPT infants who will screen positive for CD at 24 months, with a corresponding specificity of 46%. Such a model could be incorporated into the electronic health record of the unit to allow alerts for the requirement of early targeted intervention.

Early intervention can improve the cognitive outcome of a preterm infant by approximately half of a SD on a standardised cognitive test at preschool age (equivalent to 7.5 IQ points).[14] For 25% of infants with CD in this study, a half a SD improvement on the BSID-III would bring their cognitive function into the normal range, comparable with that of their typically developing peers. Alternatively, if these children are not identified at birth, they may pass early developmental checks[34] and miss the window of opportunity for early intervention. Their cognitive

difficulties may not come to light until they present with academic failure or behavioural difficulties in early school years.[35]

At the proposed decision threshold, a specificity of 46% would result in a substantial proportion of false positive screenings. However, the acceptability of false-positive results among families is likely to be high given that all screening-positive infants are offered effective intervention, and that the estimated number to treat to avoid 1 individual with cognition below the normal range is low (n = 12). Evidence-based programs enhancing parent-infant relationship and enriching the early environment extend their benefits to domains other than cognition, such as improved infant motor and social skills, and reduced parental anxiety and depression.[14,36-38] Accordingly, any overtreatment could still be considered as cost-effective.

The statistical approach used in this study was designed to optimise prediction, not investigate causal relationships. Important predictors such as intubation at birth, use of umbilical catheters, insulin treatment for hyperglycaemia, durations of ventilatory support and hospitalisation, and whether the infant was receiving breastmilk, are likely to reflect severity of illness after birth. However, these factors are modifiable and it cannot be excluded that the duration of time spent in the neonatal setting, where invasive procedures, artificial ventilation, and other noxious sensory exposures regularly occur, may have an effect on later cognitive outcomes.[39] Developmental care, designed to adjust the neonatal environment to reduce stress and promote neural growth may be beneficial[40], as may efforts to reduce duration of invasive ventilation and facilitate

discharge to home. Early discharge for stable preterm infants has demonstrated both safety and improved parental well-being.[41,42]

Limitations

Our study has important limitations. The study population consisted of survivors with a registered BSID-III cognitive score at 2 years of age. This population does not represent all VPT infants in Sweden. Although the follow-up program was launched in 2015, the completeness of BSID-III assessments has been limited by organisational and resource constraints, particularly by lack of testing capacity in remote areas. Most of the included children were born extremely preterm (<28 weeks), and resided in the 3 more densely populated urban areas.[43] Selection biases may therefore have introduced an overestimation of CD rates in very preterm infants, and the results may not be valid for all VPT children. However, extremely preterm infants assessed with BSID-III in 2016 to 2019 did not differ significantly from those non-assessed with regard to sex, GA, birth weight, small for gestational age status, or incidence of IVH stagers 3to 4 in the neonatal period (M.N., unpublished data, November 1, 2023).

The outcome used in this study was cognitive function but VPT infants often have deficits in multiple domains.[44] Although individual cognitive trajectories will vary by familial and socioenvironmental factors, most extremely preterm survivors with low BSID-III scores at 2 years will have below-average IQ scores in later childhood.[45,46] However, a significant proportion categorized as typically developed at 2 years of age will experience later worsening of cognitive function.[47]

Parental education and family language were important features in this study, and more detailed information on the socioeconomic and home

environment may have improved model performance. [48] Foreign family language may be a surrogate marker of low socioeconomic or immigration status. The administration of the BSID-III is heavily language dependent and children taking the test in a non-native language or through an interpreter may be disadvantaged, with apparent poor performance.[49]

8.6. Conclusion

This prognostic study found that it is possible to identify CD before discharge from the neonatal unit in VPT infants.. Future work should include further model validation on VPT infants. Other research groups using neurophysiological, microbiome, or imaging data should consider inclusion of the 26 clinical features identified herein in coming predictive models.

Author contributions

AB – Conceptualization, methodology, formal analysis, writing – original draft preparation, writing – review and editing

GL – Methodology, formal analysis, writing – review and editing

AS - Methodology, writing – review and editing

DM – Conceptualization, methodology, writing – review and editing

MN – Data acquisition, writing – review and editing

Table 8.4 Features assessed as important or tentative for future cognitive function in very preterm infants (GA <32 weeks).

Variable	Valid no.	Total (n = 1,062)	Typical cognitive development (n = 831)	Cognitive delay (n = 231)	P value
Pregnancy					
Gestational diabetes, No. (%)	1,062	14 (1.3)	8 (1.0)	6 (2.6)	0.094[a]
Birth					
GA, mean (SD), wk	1,062	26.5 (2.2)	26.7 (2.2)	26.0 (2.2)	<0.001[b]
Birthweight, median (IQR), g	1,055	880 (720 to 1,100)	900 (760 to 1,125)	800 (660 to 980)	<0.001[c]
z Score for birthweight, median (IQR)	1,056	-1.1 (-2.3 to -0.3)	-1.0 (-2.2 to -0.2)	-1.2 (-2.5 to -0.4)	0.157[c]
Head circumference, mean (SD), cm	883	24.7 (2.5)	24.9 (2.5)	24.2 (2.5)	<0.001[b]
Sex, No. (%)					
Male	1,062	566 (53.3)	414 (49.8)	152 (65.8)	
Female	1,062	496 (46.7)	417 (50.2)	79 (34.2)	<0.001[d]
Apgar score, median (IQR) 5 min	1,036				
5 min	1,036	8 (6 to 9)	8 (6 to 9)	7 (5 to 9)	0.001[c]
10 min	1,033	9 (8 to 10)	9 (8 to 10)	9 (7 to 10)	0.010[c]
Neonatal Resuscitation					
Intubation in delivery room, No. (%)	1,053	371 (35.2)	265 (32.1)	106 (46.7)	<0.001[d]
Respiratory					
Bronchopulmonary dysplasia, No. (%)[e]	1,062	637 (60.0)	475 (57.2)	162 (70.1)	<0.001[d]
Extra oxygen, median (IQR), d (days)	1,062	49 (15 to 80)	45 (12 to 74)	67 (34.5 to 98.5)	<0.001[c]
Duration CPAP, median (IQR), d (days)	1,038	26 (9 to 39)	25 (8 to 38)	30 (14 to 44)	<0.001[c]
Duration of mechanical ventilation, median (IQR), d	1,062	4 (0 to 17)	3 (0 to 14)	12 (0 to 26)	<0.001[c]
Duration conventional mechanical ventilation, median IIQR), d	1,062	1 (0 to 6)	1 (0 to 5)	3 (0 to 9)	<0.001[c]
Duration HFOV, median (IQR), d	1,062	0 (0 to 8)	0 (0 to 5)	3 (0 to 15)	<0.001[c]

Postnatal steroids (systemic), No. (%)	1,062	341 (32.1)	238 (28.6)	103 (44.6)	<0.001[d]
Infection					
Duration of antibiotics, median (IQR), d	1,062	14 (7 to 27)	12 (6 to 25)	20 (8 to 33)	<0.001[c]
Neurology					
IVH grade, No. (%)					
No confirmed IVH	1,045	745 (71.3)	602 (73.5)	143 (63.3)	
1	1,045	129 (12.3)	100 (12.2)	29 (12.8)	
2	1,045	97 (9.3)	73 (8.9)	24 (10.6)	
3	1,045	36 (3.4)	24 (2.9)	12 (5.3)	
4	1,045	38 (3.6)	20 (2.4)	18 (8.0)	<0.001[d]
Non-IVH intracranial haemorrhage, No. (%)	1,062	12 (1.1)	6 (0.7)	6 (2.6)	0.028[a]
cPVL, No. (%)	1,062	17 (1.6)	8 (1.0)	9 (3.9)	0.004[d]
Other treatments					
Umbilical artery catheter, No. (%)	1,062	861 (81.1)	679 (81.7)	182 (78.8)	0.364[d]
Insulin treatment for hyperglycaemia, No. (%)	1,062	131 (12.3)	88 (10.6)	43 (18.6)	0.002[d]
ROP treated, No. (%)	1,000	124 (12.4)	84 (10.8)	40 (17.9)	0.007[d]
Duration of inotrope administration, median (IQR), d	1,062	0 (0 to 0)	0 (0 to 0)	0 (0 to 1.5)	<0.001[c]
No. of plasma transfusions, median (IQR)	1,050	2 (1 to 12)	1 (1 to 12)	2 (1 to 20)	<0.001[c]
Hospital stay and discharge					
Duration of hospital and home care, median (IQR), d	1,062	75 (56 to 96)	73 (54 to 94)	85 (65 to 108)	<0.001[c]
Duration of hospitalisation, median (IQR), d	1,062	75 (56 to 96)	72 (53 to 93)	83 (65.5 to 108)	<0.001[c]
Discharged to home, No. (%)	1,062	927 (87.3)	746 (89.8)	181 (78.4)	<0.001[d]
Receiving breastmilk on discharge, No. (%)					
None	913	320 (35.0)	235 (32.2)	85 (46.2)	
Partially	913	355 (38.9)	284 (39.0)	71 (38.6)	
Completely	913	238 (26.1)	210 (28.8)	28 (15.2)	<0.001[d]
Sociodemographic					
Swedish or other Scandinavian	914	641 (70.1)	541 (74.8)	100 (52.4)	<0.001[d]

language as the family language, No. (%)					
Parental education, No. (%)					
0-3 y	859	8 (0.9)	5 (0.7)	3 (1.6)	
3-6 y	859	3 (0.3)	2 (0.3)	1 (0.5)	
7-9 y	859	50 (5.8)	31 (4.6)	19 (10.4)	
10-12 y	859	322 (37.5)	243 (35.9)	79 (43.2)	
>12 y	859	476 (55.4)	395 (58.4)	81 (44.3)	<0.001[a]
BSID-III[f]					
Corrected age at test, mean (SD), y	1,024	2.2 (0.2)	2.2 (0.2)	2.2 (0.2)	0.593[b]
Cognition index score, mean (SD)	1,062	96.8 (16.1)	103.0 (10.5)	73.5 (10.0)	<0.001[b]

Abbreviations: BSID-III, Bayley Scales of Infant and Toddler Development, Third Edition; CPAP, continuous positive airway pressure; cPVL, cystic periventricular leukomalacia; GA, gestational age; HFOV, highfrequency oscillatory ventilation; IVH, intraventricular haemorrhage; ROP, retinopathy of prematurity.

[a]Fishers exact test, [b]Welch 2-sample t test, [c]Wilcoxon rank sum test, [d]Pearson χ2 test, [e]Defined as a registered International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) code for bronchopulmonary dysplasia (P27.1) or a registration of supplemental oxygen use at 36 weeks of postmenstrual age, [f]Reasons for not completing BSID-III cognitive scale (n = 59) were as follows: child declined or would not participate (n = 23); inattention, hyperactivity, or fatigue (n = 6); language barrier (n = 1); parent declined (n = 4); resources or administrative reason (n = 6); child unable to complete (n = 13); or reason unclear or not recorded (n = 6).

Table 8.5 External validation on unseen test data for Models A-D

Model	Algorithm	No. of features	Hyperparameter final value	Sensitivity	Specificity	AUROC (95% CI)	Balanced accuracy	Accuracy
A	Random forest	26	Mtry (6)	0.49	0.81	0.74 (0.68-0.81)	0.65	0.74
B	Logistic regression	26		0.48	0.84	0.77 (0.71-0.83)	0.66	0.76
C	Support vector machine	26	C(2), Sigma (0.05)	0.64	0.72	0.75 (0.69-0.81)	0.68	0.70
D	Gradient boosting machine	26	Ntrees (800) Interaction depth (15) Shrinkage (0.1) n.minobsinnode (10)	0.55	0.83	0.76 (0.70-0.82)	0.69	0.77

Abbreviations: AUROC, area under the receiver operating characteristic curve; NA, not applicable; CI, confidence interval.

Table 8.6 Contingency table for Model B at a decision threshold of 0.20

		Cognitive development at 24 months		
		Cognitive delay	No cognitive delay	Total
Prediction from Model B	Cognitive delay	605	1,269	1,874
	No cognitive delay	45	1,081	1,126
		650	2,350	3,000

Numbers are based on a hypothetical population of 3,000 very preterm infants without major malformations.

Figure 8.3 Overview of modelling process

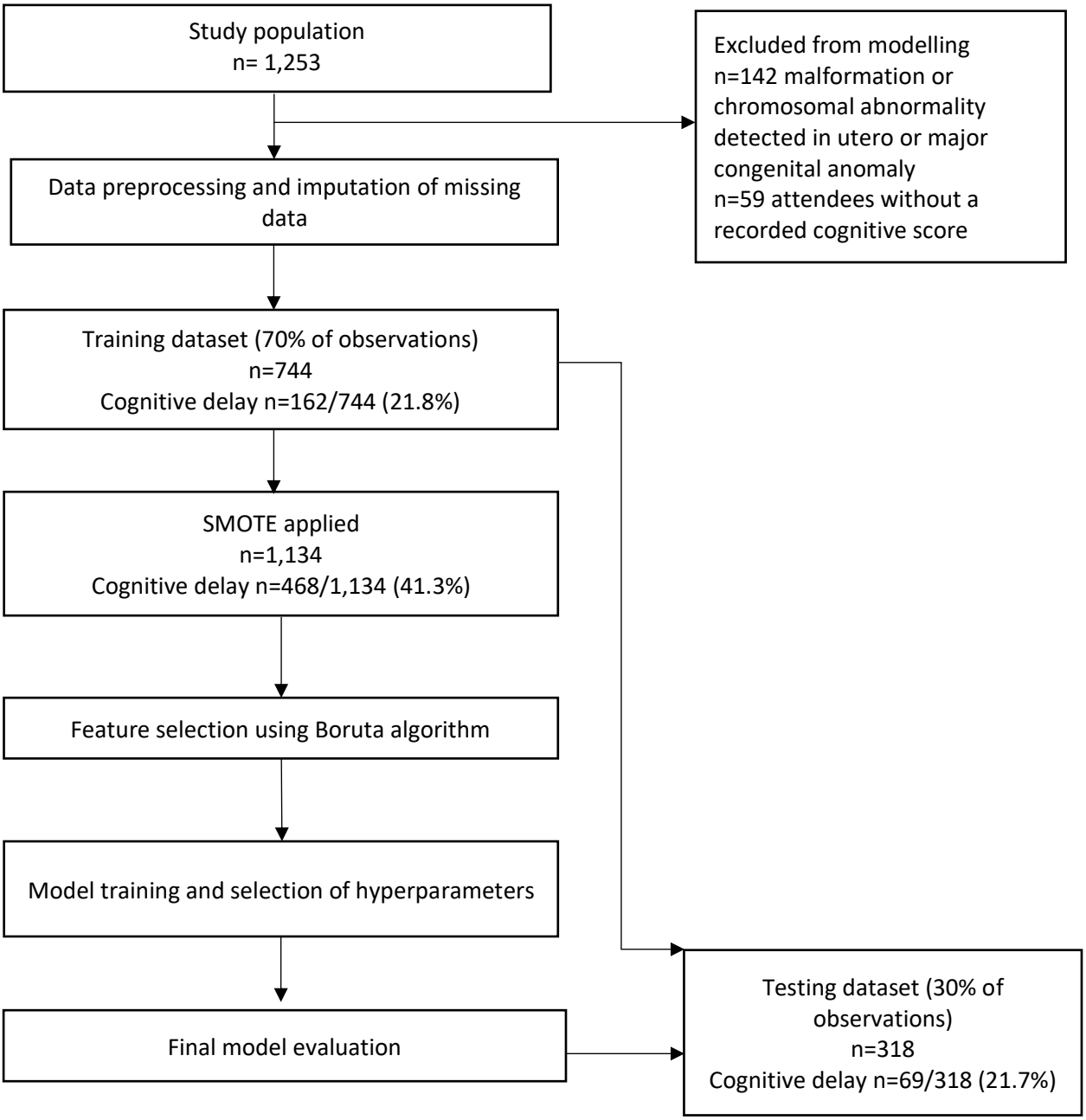


Figure 8.1 Legend
[a]Synthetic Minority Oversampling Technique

Figure 8.4 Feature importance plots for Model A-D



Figure 8.2 Legend

Importance was measured using the test dataset and permutation method with importance values averaged across ten simulations. Infants not discharged home were discharged to another hospital or unit or to home care. Abbreviations: HFOV, High frequency oscillatory ventilation; CPAP, Continuous positive airway pressure; Umb Art Cath, Umbilical artery catheter; Non IVH ICH, Non intraventricular haemorrhage intracranial haemorrhage; IVH, Intraventricular haemorrhage; NR_Inubation, Intubation during neonatal resuscitation.

8.7. References

1. Walani SR. Global burden of preterm birth. Int J Gynaecol Obstet. 2020 Jul;150(1):31-33. doi: 10.1002/ijgo.13195. PMID: 32524596.
2. Bhutta AT, Cleves MA, Casey PH, Cradock MM, Anand KJ. Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. JAMA. 2002 Aug 14;288(6):728-37. doi: 10.1001/jama.288.6.728. PMID: 12169077.
3. Larroque B, Ancel PY, Marret S, Marchand L, André M, Arnaud C, Pierrat V, Rozé JC, Messer J, Thiriez G, Burguet A, Picaud JC, Bréart G, Kaminski M; EPIPAGE Study group. Neurodevelopmental disabilities and special care of 5-year-old children born before 33 weeks of gestation (the EPIPAGE study): a longitudinal cohort study. Lancet. 2008 Mar 8;371(9615):813-20. doi: 10.1016/S0140-6736(08)60380-3. PMID: 18328928.
4. Wolke D, Meyer R. Cognitive status, language attainment, and prereading skills of 6-year-old very preterm children and their peers: the Bavarian Longitudinal Study. Dev Med Child Neurol. 1999 Feb;41(2):94-109. doi: 10.1017/s0012162299000201. PMID: 10075095.
5. van Noort-van der Spek IL, Franken MC, Weisglas-Kuperus N. Language functions in preterm-born children: a systematic review and meta-analysis. Pediatrics. 2012 Apr;129(4):745-54. doi: 10.1542/peds.2011-1728. Epub 2012 Mar 19. PMID: 22430458.
6. Twilhaar ES, Wade RM, de Kieviet JF, van Goudoever JB, van Elburg RM, Oosterlaan J. Cognitive Outcomes of Children Born Extremely or Very Preterm Since the 1990s and Associated Risk Factors: A Meta-analysis and

Meta-regression. JAMA Pediatr. 2018 Apr 1;172(4):361-367. doi: 10.1001/jamapediatrics.2017.5323. PMID: 29459939; PMCID: PMC5875339.

7. Lager A, Bremberg S, Vågerö D. The association of early IQ and education with mortality: 65 year longitudinal study in Malmö, Sweden. BMJ. 2009 Dec 10;339:b5282. doi: 10.1136/bmj.b5282. PMID: 20008007; PMCID: PMC2792333.

8. Alesi M, Rappo G, Pepi A. Emotional Profile and Intellectual Functioning: A Comparison Among Children With Borderline Intellectual Functioning, Average Intellectual Functioning, and Gifted Intellectual Functioning. Sage Open 2015; 5. https://doi.org/10.1177/2158244015589995

9. Forrest LF, Hodgson S, Parker L, Pearce MS. The influence of childhood IQ and education on social mobility in the Newcastle Thousand Families birth cohort. BMC Public Health. 2011 Nov 25;11:895. doi: 10.1186/1471-2458-11-895. PMID: 22117779; PMCID: PMC3248886.

10. Furnham A, Cheng H. Childhood Cognitive Ability Predicts Adult Financial Well-Being. Journal of Intelligence . 2017; 5. doi:10.3390/jintelligence5010003.

11. Whalley LJ, Deary IJ. Longitudinal cohort study of childhood IQ and survival up to age 76. BMJ. 2001 Apr 7;322(7290):819. doi: 10.1136/bmj.322.7290.819. PMID: 11290633; PMCID: PMC30556.

12. Twilhaar ES, de Kieviet JF, van Elburg RM, Oosterlaan J. Academic trajectories of very preterm born children at school age. Arch Dis Child Fetal Neonatal Ed. 2019 Jul;104(4):F419-F423. doi: 10.1136/archdischild-2018-315028. Epub 2018 Sep 27. PMID: 30262464.

13. Sefidkar R, Zayeri F, Kazemi E, Salehi M, Dehnad A, Hafizi M. A Trend Study of Preterm Infant Mortality Rate in Developed and Developing Countries over 1990 to 2017. Iran J Public Health. 2021 Feb;50(2):369-375. doi: 10.18502/ijph.v50i2.5353. PMID: 33748001; PMCID: PMC7956079.

14. Spittle A, Orton J, Anderson PJ, Boyd R, Doyle LW. Early developmental intervention programmes provided post hospital discharge to prevent motor and cognitive impairment in preterm infants. Cochrane Database Syst Rev. 2015 Nov 24;2015(11):CD005495. doi: 10.1002/14651858.CD005495.pub4.

15. Linsell L, Malouf R, Morris J, Kurinczuk JJ, Marlow N. Prognostic Factors for Poor Cognitive Development in Children Born Very Preterm or With Very Low Birth Weight: A Systematic Review. JAMA Pediatr. 2015 Dec;169(12):1162-72. doi: 10.1001/jamapediatrics.2015.2175. PMID: 26457641; PMCID: PMC5122448.

16. Crilly CJ, Haneuse S, Litt JS. Predicting the outcomes of preterm neonates beyond the neonatal intensive care unit: What are we missing? Pediatr Res. 2021 Feb;89(3):426-445. doi: 10.1038/s41390-020-0968-5. Epub 2020 May 19. PMID: 32428926; PMCID: PMC7276948.

17. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015 Jan 6;162(1):W1-73. doi: 10.7326/M14-0698.

18. Norman M, Källén K, Wahlström E, Håkansson S; SNQ Collaboration. The Swedish Neonatal Quality Register - contents, completeness and validity.

Acta Paediatr. 2019 Aug;108(8):1411-1418. doi: 10.1111/apa.14823. Epub 2019 May 20. PMID: 31006126.

19. Swedish Neonatal Society. National guidelines for follow-up after neonatal intensive care. Swedish Neonatal Society [Internet]. 2015. [cited July 12 2023]. Available from https://neo.barnlakarforeningen.se/riktlinjer-vardprogram/uppfoljning-av-neonatala-riskbarn-2/.

20. European Commission. EUROCAT Guide 1.5 Chapter 3.3 EUROCAT Subgroups of Congenital Anomalies [Internet]. 2022. [cited November 2, 2023]. Available from https://eu-rd-platform.jrc.ec.europa.eu/system/files/public/eurocat/Guide_1.5_Chapter_3.3.pdf.

21. Serenius F, Källén K, Blennow M, Ewald U, Fellman V, Holmström G, et al; EXPRESS Group. Neurodevelopmental outcome in extremely preterm infants at 2.5 years after active perinatal care in Sweden. JAMA. 2013 May 1;309(17):1810-20. doi: 10.1001/jama.2013.3786. PMID: 23632725.

22. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. 2021 [cited 2023 Nov 2]. R Foundation for Statistical Computing, Vienna. Available from https://www.R-project.org.

23. Stekhoven DJ, Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012 Jan 1;28(1):112-8. doi: 10.1093/bioinformatics/btr597. Epub 2011 Oct 28. PMID: 22039212.

24. Hong S, Lynn HS. Accuracy of random-forest-based imputation of missing data in the presence of non-normality, non-linearity, and interaction. BMC

Med Res Methodol. 2020; 20: 199. https://doi.org/10.1186/s12874-020-01080-1

25. Kursa MB, Rudnicki WR. Feature Selection with the Boruta Package. J Stat Softw. 2010; 36: 1–13. doi: 10.18637/jss.v036.i11

26. Degenhardt F, Seifert S, Szymczak S. Evaluation of variable selection methods for random forests and omics data sets. Brief Bioinform. 2019 Mar 22;20(2):492-503. doi: 10.1093/bib/bbx124. PMID: 29045534; PMCID: PMC6433899.

27. Chawla N, Bowyer K, Hall L, Kegelmeyer W. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res. 2002; 16: 341–348. https://doi.org/10.1613/jair.953

28. Blagus R, Lusa L. SMOTE for high-dimensional class-imbalanced data. BMC Bioinformatics. 2013 Mar 22;14:106. doi: 10.1186/1471-2105-14-106. PMID: 23522326; PMCID: PMC3648438.

29. Kuhn M. Variable Selection Using The caret Package [Internet]. 2009.[cited Jun 6 2023] Available from https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=aa4e6bb8bc727b7aa373620445764cb4f097c767

30. Bohr A, Memarzadeh K. The rise of artificial intelligence in healthcare applications. Artificial Intelligence in Healthcare. 2020:25–60. doi: 10.1016/B978-0-12-818438-7.00002-2. Epub 2020 Jun 26. PMCID: PMC7325854.

31. Deo RC. Machine Learning in Medicine. Circulation. 2015 Nov 17;132(20):1920-30. doi: 10.1161/CIRCULATIONAHA.115.001593. PMID: 26572668; PMCID: PMC5831252.

32. Hajian-Tilaki K. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. Caspian J Intern Med. 2013 Spring;4(2):627-35. PMID: 24009950; PMCID: PMC3755824.

33. Greenwell B, Boehmke B. Variable Importance Plots - An Introduction to the vip Package [Internet]. R J. 2020 [cited 2023 Nov 10]; 12: 343–366. Available from: https://journal.r-project.org/archive/2020/RJ-2020-013/RJ-2020-013.pdf

34. Bowe, AK, Hourihane, J, Staines, A, Murray, DM. The predictive value of the ages and stages questionnaire in late infancy for low average cognitive ability at age 5. Acta Paediatr. 2022; 111: 1194– 1200. doi:10.1111/apa.16309

35. Bowe AK, Staines A, Murray DM. Below Average Cognitive Ability-An under Researched Risk Factor for Emotional-Behavioural Difficulties in Childhood. Int J Environ Res Public Health. 2021 Dec 8;18(24):12923. doi: 10.3390/ijerph182412923. PMID: 34948532; PMCID: PMC8702024.

36. Benzies KM, Magill-Evans JE, Hayden KA, Ballantyne M. Key components of early intervention programs for preterm infants and their parents: a systematic review and meta-analysis. BMC Pregnancy Childbirth. 2013;13 (Suppl 1):S10. doi: 10.1186/1471-2393-13-S1-S10. Epub 2013 Jan 31. PMID: 23445560; PMCID: PMC3561170.

37. Herd M, Whittingham K, Sanders M, Colditz P, Boyd RN. Efficacy of preventative parenting interventions for parents of preterm infants on later child behavior: a systematic review and meta-analysis. Infant Ment Health J. 2014 Nov-Dec;35(6):630-41. doi: 10.1002/imhj.21480. Epub 2014 Oct 20. PMID: 25798512.

38. Ritchie K, Bora S, Woodward LJ. Social development of children born very preterm: a systematic review. Dev Med Child Neurol. 2015 Oct;57(10):899-918. doi: 10.1111/dmcn.12783. Epub 2015 Apr 27. PMID: 25914112.

39. Soleimani F, Azari N, Ghiasvand H, Shahrokhi A, Rahmani N, Fatollahierad S. Do NICU developmental care improve cognitive and motor outcomes for preterm infants? A systematic review and meta-analysis. BMC Pediatr. 2020 Feb 13;20(1):67. doi: 10.1186/s12887-020-1953-1. PMID: 32054469; PMCID: PMC7017495.

40. Khurana S, Kane AE, Brown SE, Tarver T, Dusing SC. Effect of neonatal therapy on the motor, cognitive, and behavioral development of infants born preterm: a systematic review. Dev Med Child Neurol. 2020;62: 684-692. https://doi.org/10.1111/dmcn.14485

41. Brooten D, Kumar S, Brown LP, Butts P, Finkler SA, Bakewell-Sachs S, Gibbons A, Delivoria-Papadopoulos M. A randomized clinical trial of early hospital discharge and home follow-up of very-low-birth-weight infants. N Engl J Med. 1986 Oct 9;315(15):934-9. doi: 10.1056/NEJM198610093151505. PMID: 3531852.

42. Sáenz P, Cerdá M, Díaz JL, Yi P, Gorba M, Boronat N, et al. Psychological stress of parents of preterm infants enrolled in an early discharge

programme from the neonatal intensive care unit: a prospective randomised trial. Arch Dis Child Fetal Neonatal Ed. 2009 Mar;94(2):F98-F104. doi: 10.1136/adc.2007.135921. Epub 2008 Jul 17. PMID: 18635681.

43. Swedish Neonatal Quality Register. Annual report from the Swedish Neonatal Quality Register 2021 (in Swedish) [Internet]. Stockholm: June 2021 [cited 2023 October 30]. Available from: https://www.medscinet.com/PNQ/uploads/website/Neonatalv%C3%A5rdsregistrets%20%C3%85rsrapport%202021%20(final%20version).pdf

44. van Baar AL, van Wassenaer AG, Briët JM, Dekker FW, Kok JH. Very preterm birth is associated with disabilities in multiple developmental domains. J Pediatr Psychol. 2005 Apr-May;30(3):247-55. doi: 10.1093/jpepsy/jsi035. Epub 2005 Feb 23. PMID: 15784921.

45. Lowe J, Bann CM, Dempsey AG, Fuller J, Taylor HG, Gustafson KE, et al; Eunice Kennedy Shriver National Institute of Child Health and Human Development Neonatal Research Network. Do Bayley-III Composite Scores at 18-22 Months Corrected Age Predict Full-Scale IQ at 6-7 Years in Children Born Extremely Preterm? J Pediatr. 2023 Aug 26;263:113700. doi: 10.1016/j.jpeds.2023.113700. Epub ahead of print. PMID: 37640232.

46. Serenius F, Ewald U, Farooqi A, Fellman V, Hafström M, Hellgren K, et al; Extremely Preterm Infants in Sweden Study Group. Neurodevelopmental Outcomes Among Extremely Preterm Infants 6.5 Years After Active Perinatal Care in Sweden. JAMA Pediatr. 2016 Oct 1;170(10):954-963. doi: 10.1001/jamapediatrics.2016.1210. PMID: 27479919.

47. Koller H, Lawson K, Rose SA, Wallace I, McCarton C. Patterns of cognitive development in very low birth weight children during the first six years of life. Pediatrics. 1997 Mar;99(3):383-9. doi: 10.1542/peds.99.3.383. PMID: 9041293.

48. Benavente-Fernández I, Synnes A, Grunau RE, Chau V, Ramraj C, Glass T, Cayam-Rand D, Siddiqi A, Miller SP. Association of Socioeconomic Status and Brain Injury With Neurodevelopmental Outcomes of Very Preterm Children. JAMA Netw Open. 2019 May 3;2(5):e192914. doi: 10.1001/jamanetworkopen.2019.2914. PMID: 31050776; PMCID: PMC6503490.

49. Lowe JR, Nolen TL, Vohr B, Adams-Chapman I, Duncan AF, Watterberg K. Effect of primary language on developmental testing in children born extremely preterm. Acta Paediatr. 2013 Sep;102(9):896-900. doi: 10.1111/apa.12310. Epub 2013 Jun 19. PMID: 23735043; PMCID: PMC4108617.

Chapter 9 Discussion

9.1. Introduction

This thesis began by examining the relationship between below average cognitive ability and emotional-behavioural function in childhood. Specifically, this first study aimed to elucidate the change in emotional-behavioural function as children transition to formal education. The next study aimed to further explore the impact of below average cognitive ability on the child's experience of education, their self-concept, and their engagement and behaviours in the classroom. Then, the predictive value of a routinely used infant developmental screening tool for identifying those who go on to have below average cognitive ability in childhood was investigated. Finally, three studies were performed which focussed on the application of machine learning to population-based perinatal and early life data for the prediction of below average cognitive ability in infancy and childhood. This had not previously been fully investigated in the published literature. This chapter reviews the findings of these studies, considers the strengths and limitations of this thesis, and explores the public health implications of the findings and directions for future research.

9.2.Summary of data sources used

There were three data sources used in the original studies in this thesis and these are summarised in Table 9.1.

Table 9.1 Data sources used in thesis

Name	Type	Location	Year of recruitment	Participants recruited	Chapter used
Growing up in Ireland	National prospective infant cohort	Ireland	2008	11,134	3,4 & 7
BASELINE	Regional prospective birth cohort	Cork, Ireland	2008	2,184	5 & 6
Swedish Neonatal Quality register	National population-based register	Sweden	2011	207,000	8

9.3.Summary of main findings

The overall aim of this thesis was to explore the impact of below average cognitive ability on outcomes in childhood and to examine current and alternative methods, namely the use of machine learning, for early identification of children at risk of below average cognitive ability in childhood. Specifically, this work includes six original investigations which address specific gaps identified in the literature in this field. A narrative review exploring the potential role of machine learning for predicting cognitive outcomes in childhood was also performed and published (Appendix 14). A summary of the original studies in this thesis is contained in Table 9.2.

Table 9.2 Summary of original studies in thesis

Chapter	Study title	Data	Predictor	Outcome	Cognitive ability cut off applied in main analysis	Cognitive ability cut off applied in supplementary analysis
3	Below Average Cognitive Ability – An Under Researched Risk Factor For Emotional Behavioural Difficulties In Childhood	GUI	Cognitive ability Age 3	Emotional-behavioural function Age 3,5, and 9	1 SD below mean (excludes scores >2 SD below mean)	1-2 SD below mean 2 SD below mean
4	The Relationship Between Below Average Cognitive Ability at Age 5 years and the Child's Experience of School at Age 9	GUI	Cognitive ability Age 5	Child and teacher-reported outcomes Age 9	1 SD below mean	1-2 SD below mean 2 SD below mean
5	The Predictive Value of The Ages and Stages Questionnaire in Late Infancy for Low Average Cognitive Ability at Age 5	BASE	ASQ 24/27 months	Cognitive ability Age 5	1 SD below mean	NA
6	Predicting Low Cognitive Ability at Age 5 – Feature Selection Using Machine Learning Methods and Birth Cohort Data	BASE	Perinatal and early life data	Cognitive ability Age 5	1.5 SDs below mean	NA
7	Predicting Low Cognitive Ability at Age 5 Years using Perinatal data and Machine Learning	GUI	Perinatal and early life data	Cognitive ability Age 5	1.5 SDs below mean	1 SD below mean
8	Predicting 2-year cognitive outcomes in very preterm infants using machine learning methods	SNQ	Perinatal and early life data	Cognitive ability Age 2	1 SD below mean	NA

Abbreviations: ASQ, Ages and Stages Questionnaire; GUI, Growing up in Ireland; BASE, BASELINE; SNQ, Swedish neonatal quality register; SD, Standard deviation; NA, not applicable.

9.3.1. Summary of findings: Below Average Cognitive Ability – An Under Researched Risk Factor For Emotional Behavioural Difficulties In Childhood (Chapter 3)

Perhaps the most important milestone in early childhood, for both children and parents, is the transition into primary school which, in Ireland, typically occurs around five years of age.[1] In this population-based cohort study of 7,134 infants, cognitive ability was measured at age 3 years using the Picture Similarities subtest of the British Ability Scales, and emotional-behavioural function was measured at ages 3, 5, and 9 years using the parent-reported Strengths and Difficulties Questionnaire. Children scoring more than 1 SD below the mean were categorised as below average cognitive ability. Those scoring more than 2 SD below the mean were excluded from the main analysis and examined separately in the supplementary analysis.

We hypothesised that the change in emotional-behavioural function between age 3 and 9 years would differ between those with and without below average cognitive ability. We postulated that for children with below average cognitive ability, starting school would result in a worsening of emotional-behavioural function as they struggle to cope with increasingly complex academic demand and may experience school failure. The results of the study supported this hypothesis.

The predicted change in mean SDQ score between age 3 and 9 years differed according to cognitive ability group, and the period between age 5 and 9 years accounted for most of this difference. Between age 5 and 9 years, the time coinciding with starting school, there was on average a non-significant decrease in total SDQ score for those in the average cognitive ability group (estimated mean difference -

0.06, p=0.488). For those in the below average group, there was on average a statistically significant increase, or worsening, in total SDQ score (estimated mean difference 0.48, p=0.016). Compared to children with average or above cognitive ability, the odds of experiencing a clinically significant emotional-behavioural problem were 1.4 times higher for those with below average ability (AOR 1.39, 95% CI 1.17–1.66, $p < 0.001$), after adjusting for gender, gestational age, the presence of a partner in the household, maternal education, maternal age, and maternal depression.

Compared to those with average or above cognitive ability, children with below average cognitive ability at age 3 in the GUI cohort were significantly more likely to be from disadvantaged backgrounds. They were significantly more likely to be in single parent households, to be from lower income and lower social class families, and to have a mother with second level or lower education.

A challenge in this study was establishing whether a causal relationship exists between below average cognitive ability and emotional-behavioural function. Firstly, there is some overlap between the constructs making it difficult to sharply distinguish between certain elements of both abilities. Secondly, there are many confounding variables in the relationship between cognitive ability and emotional-behavioral development and adjusting for all possible confounders was challenging.

Nonetheless, across a variety of statistical models there remained a significant association between early cognitive ability and emotional-behavioural function in childhood. There was a significant difference in the change in emotional-behavioural function over time between the cognitive ability groups, with the difference largely occurring between age 5 and 9 years, the time which coincides with starting school.

9.3.2. Summary of findings: The Relationship Between Below Average Cognitive Ability at Age 5 years and the Child's Experience of School at Age 9 (Chapter 4)

The second population-based cohort study focussed on examining the causal relationship between below average cognitive ability at age 5 years and child- and teacher-reported outcomes at age 9 years. Having established in Chapter 3, that children with below average cognitive ability experience, on average, a worsening in emotional-behavioural function during the period which coincides with their transition to school and early school years, this study aimed to further explore the early school experience for children with and without below average cognitive ability. Our hypothesis in Chapter 3 was that children with below average cognitive ability experienced a worsening in emotional-behavioural function during this time due to a negative early school experience, where due to their lower cognitive ability they struggle to keep up with peers, and may experience difficulty or failure with academic tasks in the classroom. This study further explored this hypothesis.

Acknowledging and learning from the challenges with confounding in Chapter 3, in this study directed acyclic graphs (DAGS) were used to transparently document our assumptions about the causal relationships between the variables and to identify the confounders requiring adjustment to estimate the causal association between below average cognitive ability and the outcomes of interest.

The study included 7,392 children from the Growing Up in Ireland infant cohort, who were categorised based on their cognitive ability at age 5 into below average cognitive ability (those with a score more than 1 SD below the mean), and typical cognitive development (those with a score above this). The causal

associations between below average cognitive ability at age 5 and child-reported experience of school, child self-concept, teacher-reported class engagement, and teacher-reported emotional-behavioural function at age 9 were examined using logistic regression analyses.

After adjusting for confounding, compared to children with typical cognitive development, those with below average cognitive ability had significantly higher odds of never liking school (AOR 1.82, 95% CI 1.37-2.43, p<0.001), of rating their school performance as poorly (AOR 3.04, 95% CI 1.85-4.98, p<0.001), of being picked on (AOR 1.27, 95% CI 1.09-1.48) or picking on others (AOR 1.53, 95% CI 1.27-1.84). They also had significantly higher odds of experiencing low self-concept at age 9 (AOR 1.20, 95% CI 1.02-1.42).

With regard to their teacher-reported outcomes, children with below average cognitive ability had higher odds of poor class engagement. After adjustment for confounding, they had significantly higher odds of: hardly ever or never being interested, motivated and excited to learn (AOR 2.29, 95% CI 1.70-3.10); hardly ever or never being confident to try new activities, initiate ideas, and speak in a familiar group (AOR 2.73, 95% CI 2.12-3.50); of hardly ever or never maintaining attention and concentration (AOR 2.42, 95% CI 1.85-3.16); and of hardly ever or never sustaining involvement and persevering particularly when trying to solve a problem (AOR 2.75, 95% CI 2.17-3.50). Compared to their peers with typical cognitive development, they also had significantly higher odds of teacher-reported emotional-behavioural problems (AOR 1.34, 95% CI 1.10-1.63).

This study provides evidence that, compared to their peers with typical cognitive development, children with below average cognitive ability are more likely

to have a negative experience of early education, to have low self-concept, poor classroom engagement, and problematic emotional-behavioural function. The use of DAGs and careful adjustment for confounding adds strength to the argument that below average cognitive ability is a causative factor in these outcomes. The findings of this study, combined with those of the previous chapter, suggest that below average cognitive ability has a negative impact on important outcomes, and this impact is seen early in childhood. This emphasises the importance of early intervention and support for these children and their families, prior to school entry and prior to overt signs of difficulty. It highlights the potential value of early identification of children whose cognitive development may benefit from early pre-emptive intervention.

9.3.3. Summary of findings: The Predictive Value of The Ages and Stages Questionnaire in Late Infancy for Low Average Cognitive Ability at Age 5 (Chapter 5)

The first step in exploring methods of early identification of children at risk of below average cognitive ability was to establish whether currently used developmental screening tools are valid for this purpose. In Ireland, one of the most widely used developmental screening tools is the parent-reported Ages and Stages Questionnaire, which is designed to discriminate between infants and children with developmental delay and those with typical development.

Using data from 755 participants in the Cork BASELINE birth cohort we examined the predictive value of the Ages and Stages Questionnaire (ASQ) in late infancy for identifying children who progressed to have below average cognitive

ability, defined as a score more than 1 SD below the mean on the Kauffman Brief Intelligence Test, at 5 years of age. In practice, the criterion used in the ASQ for onward referral for further professional assessment is an age-specific score, in any domain, that falls at least 2 SD below the mean in the reference population. Using this criterion, when the ASQ was performed at 24 or 27 months, the overall sensitivity for identifying below average cognitive ability at 5 years was 20.8% and the specificity was 91.1%. Using a total score cut-off point increased the sensitivity to 46.6% and 71.4% at 24 and 27 months, but specificity fell to 74.1% and 67.2%, respectively. This study confirmed that the ASQ, as it is currently used, would not detect the majority of children with below average cognitive ability at age 5.

Using a total score cut-off point could improve sensitivity but at the cost of reduced specificity. These results were based on the test being performed at 24 and 27-months, but mounting evidence on intervention would suggest that early intervention should ideally be started within the first year of life. In addition, developmental screening tools such as the ASQ rely on the presence of a delay in a developmental domain, which precludes the possibility of pre-emptive intervention. The findings of this study provided a rationale for proceeding with the investigation of alternative population-based methods of early identification of children at risk of below average cognitive ability in childhood.

9.3.4. Summary of findings: Predicting Low Cognitive Ability at Age 5 – Feature Selection Using Machine Learning Methods and Birth Cohort Data (Chapter 6)

The application of a variety machine learning methods to population-based cohort data for the prediction of cognitive ability in childhood had not previously been

investigated in the published literature. In this study, we applied the random forest algorithm to data from 1,070 participants in the BASELINE birth cohort to train a model to predict below average cognitive ability at 5 years of age and to identify the important predictive features. A random forest model using 11 perinatal features correctly classified the cognitive ability of 95% of participants at age 5. However, the relatively small sample size in the BASELINE cohort precluded the performance of external validation on unseen data, increasing the likelihood of overfitting. Therefore, the model performance must be interpreted with caution. There were 13 features consistently identified as important for prediction across the range of importance measures. These were total years of maternal schooling, socioeconomic index, family income, maternal age, maternal relationship status, maternal BMI, weekly units of alcohol in first trimester, maternal depression score, Apgar score at 1 minute, Apgar score at 5 minutes, infant birthweight, infant gestational age, and infant head circumference.

Similar to the characteristics of those with below average cognitive ability in the GUI cohort, in the BASELINE cohort a higher proportion of those in the below average cognitive ability group were from backgrounds with a significantly lower socioeconomic index. Compared to those with average or above cognitive ability, a higher proportion of the below average cognitive ability group were from the lowest income category (12.7% v 4.6%), were from single parent households (9.1% v 3.1%), lived in rental or government accommodation (27.3% v 18.4%), and were formula fed on discharge after birth (39.4% v 24.7%).

This study provided an opportunity to introduce important machine learning methodology and terminology. The findings provided a foundation for further

exploration of the role of machine learning in predicting cognitive ability in childhood, and the features identified helped guide feature selection in the subsequent studies. Having obtained excellent predictive performance on internal validation in this study, the question which remained was could this be replicated in external validation on unseen data.

9.3.5. Summary of findings: Predicting Low Cognitive Ability at Age 5 Years using Perinatal data and Machine Learning (Chapter 7)

Having established that a machine learning model could achieve excellent prediction on internal validation in one regional cohort and having identified a number of important features for prediction, we then used a larger population-based dataset to train and test a variety of machine learning models to predict low cognitive ability at age 5.

The study included 8,858 participants from the GUI cohort, whose cognitive ability was measured at age 5 using the Naming Vocabulary subtest and the Picture Similarities subtest of the British Ability Scales. Principal component analysis was used to produce a general cognitive ability score. Those who scored more than 1.5 SD below the mean were defined as low cognitive ability. A separate analysis was conducted using a cut off of 1 SD below the mean. In total 50 features were considered in the modelling process.

A random forest model containing 15 participant-reported features in the first year of infant life, achieved an AUROC of 0.77 for predicting, on unseen data, low cognitive ability, defined as a general cognitive ability score more than 1.5 SD below the mean, at age 5. At a decision threshold of 0.5, the model could correctly

predict in the first year of life, on previously unseen data, the cognitive outcome, at age 5 for 87% of infants. However, the accuracy was largely driven by the high specificity of 90% with a relatively low sensitivity of 40%. Altering the decision threshold to an optimal threshold of 0.28, calculated using Youden's index, improved the sensitivity to 72% but specificity fell to 66%.

Using the alternative cut off of a score more than 1 SD below the mean to define low cognitive ability, a random forest model containing 18 features that could be collected in the perinatal period achieved an AUROC of 0.68. At a decision threshold of 0.5 it correctly predicted the cognitive outcome of 82% of children at age 5, but again this was driven by a high specificity of 92% while the sensitivity was low at 29%.

Unlike the BASELINE cohort used in Chapter 6, the GUI cohort is not a birth cohort and the earliest data collection began when the infant was aged 9months. While there were differences in the predictive features available in the two cohorts, there was significant overlap in the most important predictive features identified in the two studies. As shown in Table 9.3, seven of the ten most important features in the GUI study were also identified among the most important features in the BASELINE study. Other important features in the BASELINE study such as Apgar scores at 1 and 5 minutes and infant head circumference at birth were not measured in the GUI study. Similarly, community connectedness, number of bedrooms in the home, and native language of parents were measured in the GUI cohort but were not measured in BASELINE.

Table 9.3 Ten most important features in random forest models predicting low cognitive ability at age 5 years using the Growing Up in Ireland and BASELINE cohorts.

Growing Up in Ireland model[a]	**BASELINE model[b]**
Maternal alcohol intake at 9 months	Maternal alcohol intake first trimester
Social class	Socioeconomic index
Maternal education	Total years maternal schooling
Family income	Family income
Gestational age	Gestational age
Birthweight	Birthweight
Maternal age	Maternal age
Number bedrooms in the home	Infant Apgar score at 1 minute
Community connectedness	Infant head circumference
Maternal native language	Infant Apgar score at 5 minutes

[a]Ten most important features using permutation feature importance across ten simulations, [b]Ten most important features using the mean decrease in Gini.

Depending on the threshold used, at a population level the models developed in the study would either miss a large proportion of infants who would go on to have low cognitive ability at age 5 (low sensitivity at a decision threshold of 0.5), or at a lower decision threshold would create a large number of false positive cases (low specificity at lower decision thresholds). In Ireland, in 2023 there were approximately 54,000 births, of whom approximately 6% (n=3,240) would have cognitive ability more than 1.5 SD below the mean at age 5 years.[2] At a lower decision threshold of 0.28, our model would identify, in the first year of life, 72% (2,332/3,240) of these infants who would go on to have low cognitive ability at age 5 years. The model would correctly

identify, as not at risk, 66% (33,502/50,760) of those with cognitive ability above a -1.5 SD cut off at age 5. However, 17,258/50,760 would be incorrectly identified as being at risk of later low cognitive ability. The resource implications of this false positive rate would be unsustainable and unacceptable at a population level. However, among a smaller high risk population of infants a prediction tool with this false positive rate would not have the same resource implications. However, a smaller high risk cohort may have different risk factors or predictive variables than those identified among the general paediatric population, and this would require training of a different predictive model.

9.3.6. Summary of findings: Predicting 2-year cognitive outcomes in very preterm infants using machine learning methods (Chapter 8)

Building on the results of Chapter 6, in this study we identified a high-risk population of very preterm infants who, compared to the general population, would be at increased risk for poor cognitive outcomes. This study was based on data from the Swedish Neonatal Quality Register, a nationwide, population-based register, that captures data on all infants admitted for neonatal intensive care in Sweden. Included in the study were 1,062 infants born at a gestational age <32 weeks and who completed the Bayley Scales of Infant and Toddler Development Third Edition (BSID-III) at a corrected age of 2 years. Cognitive delay, defined as a score <90 (equivalent to approximately a 1 SD cut-off), was present in 21.8%.

A total of 90 features containing sociodemographic and clinical information on conditions, investigations, and treatments initiated during pregnancy, delivery, and neonatal unit admission were examined for predictability. A range of machine

learning models were trained and tested. A logistic regression model containing 26 features achieved an AUROC of 0.77. At discharge from the neonatal unit, the full model could identify 93% of infants with cognitive delay at 2-year follow up, with a specificity of 46%. The five most important features for cognitive delay in this study were non-Scandinavian family language, prolonged duration of hospitalisation, low birthweight, discharge to destination other than home, and the infant not receiving breastmilk on discharge.

9.3.7. Application of conceptual framework in consideration of early life factors affecting cognitive ability in childhood

The conceptual framework outlined in Figure 1.1, Chapter 1 was considered in the design of each original study. It provided a framework for considering confounding variables in Chapters 3, 4 & 5, and provided a framework for feature selection in Chapters 6-8. While not all factors identified in Figure 1.1 were available in the data sources used, it was often the case that there was an alternative variable which could be considered a correlate or surrogate marker. For example, in the GUI cohort detailed information on obstetric complications was not provided. Unlike in the SNQ where this information was recorded by a medical professional using diagnostic codes, in the GUI study the limited information on birth complications was parent-reported and therefore subject to recall bias. However, admission to the neonatal intensive care unit, which is likely a more reliable parent-reported variable than a specific parent-reported diagnosis, was recorded and was considered a surrogate marker for perinatal complications affecting the infant.

9.4. Interpretation of findings

The eight key findings of this thesis are:

1. Children with below average cognitive ability at age 3 have, on average, a different trajectory of emotional-behavioural function between age 3 and 9 years compared to those with typical cognitive development. Between age 5 and 9 years, the period coinciding with school entry, children with below average cognitive ability, on average, experience a worsening in emotional-behavioural function scores, whereas those without experience no significant change. (Chapter 3)
2. Children with below average cognitive ability at age 3 have higher odds of experiencing a clinically significant emotional-behavioural problem between the age of 3 and 9 years compared to those with typical cognitive development. (Chapter 3)
3. Having below average cognitive ability at age 5, the time most children in Ireland start school, is associated with increased odds of low self-concept, of poor classroom engagement, and of teacher-reported emotional-behavioural problems at age 9 years. (Chapter 4)
4. The routinely used Ages and Stages Questionnaire, when performed at 24 or 27 months, will not detect the majority of children who will go on to have below average cognitive ability at age 5. (Chapter 5)
5. Children with below average cognitive ability are more likely to be characterised by socioeconomic disadvantage. (Chapters 3, 4, 5, 6, 7).
6. At a population level, a random forest model containing 15 features, which could be collected in the perinatal period, can correctly identify the cognitive

outcome at age 5, defined as above or below a -1.5 SD cut off, of 87% of infants. However, it is better at predicting those who will not have cognitive difficulties. (Chapter 7)

7. The most important perinatal features for predicting childhood cognitive outcomes at a population-level are maternal alcohol intake, social class, maternal education, maternal age, family income, number of bedrooms in the home, native language of the mother, community connectedness, gestational age, birthweight, Apgar scores at 1 and 5 minutes, and infant head circumference. (Chapters 6 & 7)
8. Among very preterm infants admitted to neonatal units in Sweden, it is possible to identify upon discharge from the neonatal unit 93% of infants who will have cognitive delay at 24 months with a specificity of 46%, using a logistic regression model containing 26 features. (Chapter 8)

These key findings must be interpreted in the context of possible bias and confounding.

9.4.1. Bias

9.4.1.1 Selection bias

There were three data sources used in this thesis – one national population-based infant cohort, one regional birth cohort, and one population register and each is subject to potential selection bias.

BASELINE cohort

The BASELINE cohort recruited pregnant women in Cork, Ireland's largest county and second largest city. This is therefore a regional cohort and may not be representative

of the population of pregnant women in Ireland as a whole. There were two streams of recruitment in the study, with the majority of participants (73%) recruited through stream 1. Primiparity was one of the eligibility criteria in this recruitment stream, therefore the majority of women in the BASELINE cohort are primiparous mothers. It is possible that primiparity would affect the cognitive outcome of infants in the BASELINE study. The resource dilution theory posits that firstborn children have sole access to a family's resources, including parental attention, books and toys, and finances. When siblings are present, these resources are distributed among them.[3,4] A previous study has shown that firstborn children performed better in cognitive assessments at age 4, compared to non-firstborn children.[5] This is one possible explanation for the relatively high mean IQ of children in the BASELINE cohort which was 104 with a standard deviation of 9. One would expect the mean IQ of a population to be approximately 100 with a standard deviation of 15.[6]

Another possible explanation is the over-representation of mothers with high educational backgrounds in the BASELINE cohort. Of mothers recruited to the BASELINE study, 55% had university level education, whereas census data from 2011 suggests this figure was closer to 28% among women age 15-44 nationally.[7]

Alternatively, the higher IQ may be due to the widely accepted Flynn effect, which refers to an observed rise in IQ scores over time of about 3 IQ points per decade. The Flynn effect would suggest that a standardised IQ test will overestimate an individual's IQ score by approximately 0.3 points per year between the year in which the test was normed and the year which it was administered.[8] The KBIT-2 IQ test used in BASELINE was published in 2004 and was administered between 2013-

2016, which according to the Flynn effect would account for approximately a 3-4 point increase in IQ.[9]

Growing Up in Ireland Cohort

The GUI cohort was a national cohort and a careful sampling strategy was used to select a representative sample of participants. However, there was an under-representation of single parents and mothers from low education backgrounds recruited to the original cohort.[10] At the 9 year follow-up, attrition was higher among younger mothers, and those with lower levels of education, lower income, lower social class, single parent families, those in social housing, those in the Dublin region, primary caregivers who smoked daily, and infants who were never breastfed.[11] The GUI study team calculated reweighting variables which can be applied to analyses to adjust for non-response and attrition bias. Studies in this thesis using the GUI cohort (Chapter 3 & 4) reweighted the data accordingly thereby minimising the effect of attrition and non-response. In Chapter 7, for the machine learning analyses the data were not reweighted. This would not affect the predictive relationships seen between features and the outcome. However, it may mean that relationships and patterns between features with very few observations and the outcome may not be uncovered by the algorithm.

There are certain members of the population, who are unlikely to participate and remain engaged in a cohort study and are therefore unlikely to be represented in either the BASELINE or GUI cohorts, but who are at high risk of having a child with low cognitive ability.[12-14] These include mothers experiencing homelessness, those with drug and alcohol addiction, and those in temporary accommodation such as

direct provision. These are likely to be important features in predicting this outcome but some are not represented in the models developed.

Swedish Neonatal Quality Register

The SNQ is an opt-out, nationwide, population-based register, covering all infants admitted for neonatal care in Sweden.[15] However, due to organisational and resource constraints formal neurodevelopmental assessment is not performed for all infants in the SNQ. The majority of infants who receive neurodevelopmental follow-up are those born extremely preterm (<28 weeks). Therefore, selection bias may have introduced an overestimation of the prevalence of cognitive delay in our study of very preterm infants born at <32 weeks. The model should ideally be validated in the entire population of infants born at <32 weeks in Sweden, and in international populations, to ensure predictive performance remains similar.

9.4.1.2 Measurement bias

Cognitive ability was measured at age 5 in the BASELINE cohort using the Kauffman Brief Intelligence Test 2nd Edition, and in the GUI cohort using the British Ability Scales 2nd edition. Both tests included a measure of verbal ability. In the KBIT-2 verbal ability was measured using the verbal knowledge and riddles subtests. The verbal knowledge subtest requires the child to point to a picture, from a selection of six, that best represents a vocabulary word spoken by the examiner in the English language. For the riddles subtest the child is asked to point to a picture or provide a word that solves a riddle spoken by the examiner in the English language. Together, these two subtests primarily measure crystalised verbal ability which largely comes from prior learning and experience.[9,16] In the GUI study the naming vocabulary

subtest from the BAS II was used to measure verbal ability and consists of the child naming everyday items in the English language displayed from a picture book.[17]

In the BASELINE cohort, the majority of participants reported living in Ireland for more than two generations. Among the below average cognitive ability group 4.6% reported either one or both parents of the child immigrated compared to 3% of the typical cognitive development group. This difference was not statistically significant. In the GUI study, 84.8% of all mothers reported English as their native language, but the proportion was significantly lower among mothers of infants with low cognitive ability compared to those with typical cognitive development (53.9% v 86.9%, $p<0.001$).

Children taking either the KBIT-2 or the BAS II, for whom English is not the native language of their parents or for whom English is not spoken in the home, may be disadvantaged in these tests with apparent poor performance. This means the predictive model may identify children with typical cognitive development, but who perform poorly on a measure of English vocabulary. A more appropriate measure of cognitive ability in children from multilingual backgrounds may include total vocabulary across their languages, which is unlikely to be feasible in a large population based study, or the use of novel language-independent cognitive assessments.[18,19]

Similarly, neurodevelopmental follow-up of high risk infants from the SNQ register is performed using the Bayley Scales of Infant and Toddler Development 2nd Edition, which is also heavily language dependent. Children taking the test in non-native language or through an interpreter may be disadvantaged. Among those with cognitive delay a significantly lower proportion reported Scandinavian as being the

family language compared to those with typical cognitive development (52.4% vs 74.8%, p<0.001).[20] In all three cohorts it is possible that the family language adversely affected the performance of the child on the language component of the test. However, family language may also represent the immigration background of the family and therefore may be a surrogate marker of socioeconomic status. It is plausible that children from non-English speaking families are at a truly increased risk of adverse cognitive outcomes due to socioeconomic disadvantage.

9.4.1.3 Recall bias

In the BASELINE cohort neonatal data such as gestational age, birthweight, head circumference, admission to NICU and Apgar scores were recorded by trained research nurses two days following delivery of the baby. Similarly, in the SNQ data related to the mother, pregnancy, and delivery were automatically extracted from medical records and transferred via the Swedish Pregnancy Register. Data on the neonatal unit admission were entered manually into the register by neonatal units, using an electronic platform which warns users of unreasonable values, blocks impossible values, and provides pop-ups with definitions and explanations.[15] In the GUI study, information related to the pregnancy, delivery and postnatal period, such as gestational age, birthweight, and admission to NICU, were reported by the mother at the 9-month survey and are therefore subject to recall-bias. Previous research has shown that 94% of mothers recall the birthweight within plus-or-minus 0.5 pounds correctly, and 86% recall gestational age within plus-or-minus 1 week correctly.[21]

9.4.1.4 Social desirability bias

Social desirability bias refers to the tendency of an individual to underreport socially undesirable behaviours or opinions, and to overstate more desirable ones.[22]

Self-reported surveys, such as those used in BASELINE and GUI, are vulnerable to this type of bias. It is particularly relevant for self-reported information on individual behaviours such as smoking and alcohol intake, and for sensitive information such as household income and mental or physical health.[23] It is possible that this type of bias may distort the relationship between certain predictive features and the outcome. In the GUI study, efforts were made to minimise social desirability bias through the use of self-complete questionnaires for both children and parents when information deemed to be sensitive was collected.[10]

9.4.2. Confounding

In Chapters 3 & 4, the studies examined the association between a predictor, namely cognitive ability, and an outcome. These studies are therefore subject to confounding. In Chapter 3, perhaps the most important finding was that children with below average cognitive ability have, on average, a worsening in their emotional-behavioural function during the time which coincides with starting school. There are several possible explanations for this. The hypothesis put forward in this study was that the increased complexity of cognitive tasks experienced during this time may uncover difficulties for the child that had previously gone unnoticed. The child may have trouble completing the tasks put before them, may struggle to keep up with peers, and may experience failure, all of which could have detrimental effects on their emotional-behavioural function.

An alternative explanation is that children with below average cognitive ability, who are also those more likely to be exposed to socioeconomic disadvantage, experience the early years of school differently compared to peers from more

affluent backgrounds. Children from disadvantaged backgrounds may not receive the same parental support and engagement, and may lack resources which others have access to, for example access to sports and after-school activities. This could also contribute to worsening emotional-behavioural function during this time. To account for this confounding, the generalised linear mixed models and logistic regression models in this study were adjusted for maternal education and presence of a partner in the household. They were also adjusted for maternal depression which may affect parenting behaviour. However, residual confounding is still possible.

In Chapter 4, directed acyclic graphs were used to carefully document our assumptions about relationships between the predictor, the outcomes, and the confounding variables. The use of DAGs allowed us to systematically tease out which variables were considered confounders, and which were considered mediators i.e. variables on the causal pathway between the predictor and the outcome.[24] This approach minimises the chance of residual confounding provided all confounders in the minimal adjustment set are adjusted for.[24]

The socioeconomic environment of the child is one of the most important confounders in the relationship between cognitive ability and school outcomes. Literature on the most appropriate method of adjusting for the socioeconomic environment was reviewed prior to this study.[25,26] In keeping with recommendations in the literature, collinearity between socioeconomic variables was examined prior to inclusion of a range of variable representing different domains of the socioeconomic environment – parental education, family income, presence of a partner in the household, and the cultural background of the family.[25] This approach

adds weight to the argument for causality between the exposure and outcome in this study.

In Chapters 6-8, the studies focussed on prediction, which is not subject to confounding in the same way as studies trying to establish a causal association. However, it is crucially important to consider confounding when trying to interpret and explain the models predictions. For example, in the random forest model developed for predicting low cognitive ability among the general population in Chapter 7, the primary caregivers alcohol intake at infant age 9-months was one of the ten most important predictors. The partial dependence plot, which shows the cumulative effect, according to the model, of each predictors individual values on the classification of the outcome, shows that the risk of low cognitive ability was higher among infants whose mother reported never or very rarely drinking alcohol, compared to infants of mothers who reported moderate alcohol intake. Without consideration of confounding, one could interpret this plot as meaning a moderate amount of maternal alcohol consumption will protect against low cognitive ability in offspring. However, we know that this is biologically implausible. This U-shaped relationship, which has been reported in many studies examining the relationship between alcohol and health outcomes, is likely confounded by religious, immigration, and socioeconomic background. Those who reported no alcohol intake may also be those more likely affected by immigration and potentially adverse socioeconomic conditions, which are likely the underlying causal factors in the relationship. The GUI data supported this hypothesis.

9.4.3. Chance findings

It is unlikely that the key findings in this study are due to random error. In Chapter 3, the study examining the association between cognitive ability and emotional-behavioural function, the study population included 7,134 infants. The confidence intervals around point estimates in both the generalised linear mixed models and the logistic regression models were typically narrow. In Chapter 4, which examined the association between cognitive ability and school outcomes, all analyses contained at least 6,000 participants and again the confidence intervals around point estimates were narrow. In Chapter 5, the study population consisted of infants from the BASELINE cohort who had completed either the 24-month (n=468) or 27-month (n=287) ASQ and the KBIT-2 at age 5. The smaller study population did result in wider confidence intervals around estimates of the AUROC and performance characteristics. In Chapters 6-8 the use of cross-validation and external validation on unseen data make the findings reported in these studies unlikely to be due to chance.

9.5. Strengths and Limitations

9.5.1. Strengths

This thesis includes six original cohort studies, all of which address a clear gap in the existing literature. Each of these six studies have important individual findings, but each study also provided a strong rationale for proceeding with the subsequent studies in the thesis. The findings from Chapters 3 & 4 provide a strong rationale for early intervention, prior to school entry, for children at risk of below average cognitive ability. Chapter 5 confirms that current developmental screening tools will

not identify the majority of children at risk and highlights the importance of investigation of alternative methods. In Chapters 6-8, the application of a variety of machine learning methods, using readily-available population data, for the prediction of later low cognitive ability was explored. This has not been previously investigated and reported in the literature. In addition, the ML studies were designed with clinical application in mind and focussed on using readily-available early life data so that models would have the potential to be used in clinical practice and in very early life.

The thesis used three population-based data sources from two European countries.[27-29] This adds considerable weight to findings that were consistent across cohorts. For example, in both the GUI and BASELINE cohort children with below average cognitive ability were characterised by socioeconomic disadvantage, and there was considerable overlap in the most important predictive features for below average cognitive ability in each cohort. Additionally, each cohort used an objective measure of cognitive ability, with the KBIT-2 used in the BASELINE cohort, the BAS-II in the Growing up in Ireland cohort, and the BSID-III in the Swedish Neonatal Quality register. Each test has demonstrated reliability and validity as a measure of cognitive ability.[30-32]

9.5.2. Limitations

9.5.2.1 Measuring cognitive ability

While the measures of cognitive ability used in this thesis were all reliable and valid measures of cognitive ability, there are many important limitations to their use which must be considered. Firstly, in the BASELINE and GUI cohorts the tests were

administered in the English language and as previously outlined this risks misclassifying children, for whom English is not the native language in their home environment, as being at risk of cognitive difficulty. Similarly, in the SNQ the BSID-III was administered in Swedish and the same risk applies to infants whose family native language is not Swedish. It is crucially important that this limitation is taken into account in the design of future cohort studies. Prior to the Covid-19 pandemic, global migration was at its highest level recorded in history and this has resumed post pandemic, with the increasing threat of climate change likely to be a major contributor.[33] Future cohort studies measuring cognitive ability in childhood must take the increased prevalence of children in multilingual households into account when designing studies and choosing instruments to measure cognitive ability.

Secondly, in the GUI study, due to restricted contact time in the home of participants, only two core scales of the Early Years Battery of the British Ability Scales were used to derive measures of verbal and non-verbal ability in children.[28] The full battery consists of 11 scales and includes scales designed primarily to measure numerical ability, spatial ability, perceptual speed, and memory.[30] Performing the full battery of BAS subtests would provide a more robust assessment of cognitive ability. In addition, performing the full battery would enable the scores to be summed to produce a composite score as recommended by the scale authors.[30] However, as only two core scales were administered principal components analysis was used to produce a composite score for use in the studies in this thesis. This approach has limitations – it results in information loss as it reduces the dimensionality of the data, the principal components can be difficult to interpret, and

the use of PCA would not be an acceptable method for diagnosing an intellectual developmental disorder.

Thirdly, it is generally assumed that the measurement of cognitive ability among a population will form a normal distribution, with approximately 68% of the population having a general cognitive ability score within 1 standard deviation of the mean.[34] However, if a cohort study is not representative of the whole population, then the distribution of cognitive ability scores may not form a normal distribution or the curve may be shifted to the right or left. This risks misclassifying children in that cohort. In the BASELINE cohort the mean IQ at age 5 years was 104 IQ points with a standard deviation of 9 points. This may be due to selection bias or it could be due to the Flynn effect. Without standardising each test on a representative sample of the Irish or Swedish population, one cannot know for sure. This lack of standardisation also makes it very difficult to make meaningful comparisons across different studies and cohorts. An individual scoring below a 1 SD cut-off in one cohort may score within the range of reported norms or within the normal range in another cohort.

Finally, this thesis focussed primarily on one form of ability in childhood – cognitive ability. We have outlined the existing evidence in Chapter 1 Section 1.5, and provided additional new evidence in Chapters 3 & 4, as to its importance as a predictor and cause of outcomes throughout the life course. However, focussing on cognitive ability alone risks placing undue emphasis on this one ability and ignoring multiple other abilities in childhood which require nurturing and attention. Success is not determined solely by one's cognitive ability, nor should it be used as a marker of success in life. In fact, many argue that emotional intelligence, one's ability to

understand and use emotions positively to communicate, empathise, problem solve and manage conflict, is equally as important as general intelligence.[35] However, it would be remiss not to say that a minimal level of cognitive ability is required for almost all tasks in daily life. The work in this thesis is centred on ensuring that every child is given the opportunity to reach their individual cognitive ability, irrespective of the biological and environmental risks with which they are born.

9.5.2.2 Categorising cognitive ability

In this thesis we applied cut-off points of 1, 1.5, or 2 standard deviations below the mean to categorise and define different groups of cognitive ability. From a statistical perspective these cut-offs should identify those with cognitive ability below approximately 84%, 94%, and 98% of the population respectively. From a historical perspective, a cut off of 1 SD below the mean corresponds to what was previously defined as an intellectual disability in the first version of the Diagnostic and Statistical Manual of Mental Disorders.[36] In later versions of both the DSM and the International Statistical Classification of Diseases and Related Health Problems (ICD), those scoring between 1 – 2 SD below the mean were defined as having borderline intellectual function.[37] The ICD-11 currently applies a cut off of two or more SDs below the mean as part of their definition of a disorder of intellectual development.[38] From a research perspective, previous studies examining low cognitive ability in childhood have used cut-offs of both 1 and 1.5 SDs below the mean to define low cognitive ability.[39,40]

There is an increasingly large body of research and expert opinion that criticises the use of these arbitrary cut offs in defining disorders of intellect.[36] There has been a move away from solely relying on IQ to define an intellectual developmental disorder and both the ICD and DSM manuals now include deficits in

adaptive behaviour, which includes social, conceptual, and practical skills, as part of their definition.[38] In both research, and in the real world, measuring adaptive behaviour is challenging. Assessment of adaptive behaviour should be an extensive examination of both observed behaviour and third-party reported behaviour, usually from a parent or teacher, and should be performed across multiple settings.[41] It is subjective and expectations about adaptive behaviour can vary according to sociocultural group, age, and setting.[42] In practice it would appear that IQ has more weight in decisions made in relation to diagnostic labels, and adaptive behaviour is seldom used to correct for "false negative" cases, where a person may be above the IQ cut-off but should be considered to have an IDD due to poor adaptive functioning.[36]

The use of IQ cut-offs and "black and white" categorisation is bureaucratically more appealing as it makes decisions around entitlements and eligibility for services clear cut. However, as explained eloquently by Greenspan et al it ignores important realities - "bureaucracies have a need both for clear and rigid ''yes'' or ''no'' decision rules that do not take into account the true nature of scientific inquiry, which is always about ambiguity and evolution, rather than certainty and stasis".[36] An important limitation in this thesis is the use of arbitrary cut-off points to define levels of cognitive ability in the absence of measures of adaptive behaviour, which were not included in the data sources used in this thesis. Furthermore, different terminology was used in different studies to label those scoring below these arbitrary cut-off points. This was largely due to the absence of a clearly defined nomenclature in the literature, the use of different terms among different disciplines, and due to different preferences of the academic journals in which the studies were published.

9.5.2.3 Use of existing data

This thesis was based on existing data sources. A limitation of using existing data sources is that they have not been designed specifically to answer the research question proposed, and may therefore not include information deemed important to answer the research question. For example, cognitive ability has a strong genetic component, and genetics are estimated to account for 20-50% of the variability in childhood intelligence.[43] A surrogate measure of genetic influence, which is feasible to collect in a population cohort study, is parental IQ which previous studies have shown to explain a large proportion of variation in childhood IQ.[44] Inclusion of this measure may have improved predictive performance of models.

9.5.2.4 Limitations of machine learning

The machine learning analyses applied in Chapters 6-8 were designed to optimise predictive performance. The analyses were not designed to investigate causal relationships. The features used by the model may themselves be causal factors or may be surrogate markers for other underlying causative factors.[45] A limitation of this is that intervention programmes cannot be based solely on addressing the predictive features in a machine learning model. Designing an intervention programme should not be based on predictive factors but should be focussed on addressing the underlying causal factors. For some features identified in our ML models, such as socioeconomic status, there already exists an evidence base demonstrating the causal relationship with low cognitive ability in childhood.[46,47]

However, for other predictors this evidence does not exist. For example, in Chapter 8, among the most important predictors of cognitive delay at age 2 in very preterm infants was the duration of hospitalisation. This may be a surrogate marker

of severity of illness after birth or it is possible that prolonged hospitalisation has a direct causal association with later cognitive delay. It is possible that the duration of time spent in the neonatal setting, where invasive procedures, artificial ventilation, and noxious sensory exposure to noise and bright lights regularly occurs, has a direct causal effect on later cognitive function. It is possible that developmental care designed to adjust the neonatal environment, minimise stress and duration of hospitalisation may improve cognition but this would require further research focussed specifically on causality. The implementation of ML models into clinical practice and the acceptability of the models to both clinicians and patients can be affected by this 'black box' nature.[48]

9.6. Public health and policy implications of findings

9.6.1. Below average cognitive ability as a risk factor for adverse outcomes

The results of this thesis show that children with below average cognitive ability may be at increased risk of adverse outcomes in childhood compared to their peers with typical cognitive development. On average, they will experience a worsening in emotional-behavioural function in the early years of school, in contrast to children with typical cognitive development who on average experience no significant change. The odds of poor engagement at school, of having low self-concept, and of experiencing a clinically significant emotional-behavioural problem are higher for these children compared to their peers with typical cognitive development.

For families and teachers of children experiencing disengagement from school, low self-concept and problematic behaviours or emotions, consideration should be given to their cognitive functioning. Families and teachers should work together to

determine the root causes of these problems, and below average cognitive ability should be considered as one potential cause. If further investigation of a child's cognitive ability is deemed necessary, then teachers should have access without undue delay or administrative barriers to prompt assessment. In addition, the results of this thesis clearly demonstrate, that children with below average cognitive ability are more likely to require additional psychological support in the early years of school. Prompt access to these services is essential.

9.6.2. Population-screening in early life for below average cognitive ability in childhood

At a population level, the random forest model developed in Chapter 7, using only data that could be collected in the perinatal period achieved an AUROC of 0.77 on an unseen dataset for identifying children with cognitive ability more than 1.5 SDs below the population mean, and an AUROC of 0.68 for identifying those with cognitive ability more than 1 SD below the mean. The resource implications of a low specificity model have been previously outlined. If the alternative threshold of 0.50 were used, the model would identify 40% of those, in the first year of life, who would go on to have cognitive ability more than 1.5 SDs below the mean, and would correctly identify as not at risk 96% of those with cognitive ability above this at age 5. Among 54,000 births, the model would identify 12% (n=6,480) of babies as being at risk of later low cognitive ability.[2] Of the babies identified as being at risk, 22% (1,426/6,480) would be true positives and would receive needed intervention. The remaining 5,054/6,480 would be false positives. The model would identify 88% of babies (n=47,520) as not being at high risk, and among them 96% would be true negatives

but 4% (1,901/47,520) would be false negatives and would not receive needed intervention.

There are a number of strengths to the model which must be considered. Firstly, the model uses data which could be easily collected in the perinatal period and therefore intervention, to those correctly identified, could occur in the first year of life during the period of optimal neuroplasticity. Acknowledging that the cognitive cut-offs used differed by 0.5 SD, the sensitivity of this model was higher than the reported sensitivity, in Chapter 5, of the routinely used Ages and Stages Questionnaire when performed at 24 or 27 months (sensitivity 21%, specificity 91%). Thirdly, the model uses routinely available perinatal data that could be easily incorporated into an electronic clinical decision support system if electronic patient records were in use.

However, at this decision threshold the model would still miss the majority of children with low cognitive ability. The number of false positives would have implications on already scarce early intervention resources and would come at an opportunity cost to other services. There are also important ethical issues around labelling, particularly when that label is incorrect.

Nonetheless, the model is a step in the direction of early, individualised, risk prediction which could enable early intervention and improve cognitive outcomes for many children. Improvement in predictive performance is clearly required. A population-based risk tool has important distinctions from a population-based screening test. A screening test generally involves performing a test on an asymptomatic population or subgroup to determine their likelihood of having a condition.[49] A prediction tool, involves performing a test on an asymptomatic

population to determine their likelihood of acquiring a condition. However, the performance characteristics of currently used population-based screening tools should be considered as a guideline in the development of a prediction tool. In Ireland, population-based screening programmes are in place for bowel cancer, cervical cancer, and breast cancer.[50] Screening for cervical cancer is performed now using a test for human papillomavirus, which according to a Cochrane review has a sensitivity of approximately 90% and a corresponding specificity of 90%.[51] The faecal immunohistochemical test for bowel cancer has a sensitivity of approximately 79% and a specificity of 94% for detecting colorectal cancer.[52]

The model developed in Chapter 8 for identifying very preterm infants at risk of cognitive delay at age 2 years could be considered for implementation as a clinical decision support tool, after wider external validation. At a lower decision threshold of 0.20 the model could identify, prior to discharge from the neonatal unit, 93% of very preterm infants who would go on to have cognitive delay at age 2 years. The low specificity of 46% would result in a high proportion of false positives, however, given that the proportion of infants born at <32 weeks is relatively small, in Ireland approximately 1.3% of births[53,54], this would not have the same resource implications as a screening tool for the entire population. In addition, parents of very preterm infants are often aware their child is at higher risk of developmental delay and the acceptability of applying this label may be higher in this population.[55]

9.6.3. Machine learning in public health

Almost all public health challenges we face result from a complex interplay between individual, community, and wider socioeconomic, cultural, and environmental

factors.[56] Machine learning methods can be harnessed in public health to capture the complex relationships between these determinants. The findings of this thesis show that machine learning methods can be a useful tool in addressing public health challenges, particularly those involving prediction problems. ML techniques, particularly data rebalancing, feature selection, and pattern recognition, will become increasingly important as the quantity and diversity of data available for prediction problems increases. For some prediction problems, such as that in Chapter 8, more traditional algorithms such as logistic regression may prove to have equal or superior predictive performance than more complex ML algorithms. However, it is important to note that in Chapter 8, a number of ML techniques to rebalance data and identify the important predictive features were first applied prior to training the logistic regression model.

The findings of this thesis show that ML algorithms, should be considered, in addition to more traditional statistical models, when using prediction to try address a particular public health challenge. To overcome the inherent bias in cohort data that is now largely used to train and validate ML models, there is an urgent need in Ireland to develop electronic health records, to introduce unique population health identification numbers, and to follow in the footsteps of Scandinavian countries with the establishment of linked population health registries.

There are however significant challenges with implementing ML algorithms into clinical practice. Firstly, they must be acceptable to both clinicians and patients. Research in this area consistently points to a lack of trust between clinicians and clinical AI, with the majority of clinicians believing they themselves are more accurate than a ML algorithm.[57] The interpretability of the model is a key factor in a clinicians

willingness to use it.[57] There are methods, such as the partial dependence plots used in this thesis, which can provide some explanation on how a model is making its prediction, however their interpretation is not always straightforward. There is also an argument that rather than trying to explain black box models in a post hoc manner, interpretable models should be created in the first instance.[58] Cross-discipline training and collaboration between clinicians and computer scientists may help improve clinicians understanding and knowledge of AI, but due to competing demands on the time of clinicians this may not be realistic depending on the health setting.

Another barrier to implementation is the lack of clarity around medico-legal responsibility for errors made by AI tools. A survey of clinicians and members of the public in the United States reported the majority of both believed the clinician should be held responsible for an AI error.[59] If healthcare systems wish to incorporate AI into clinical practice, these issues will need to be carefully addressed with full clarity around medico-legal responsibilities.

9.6.4. Policy implications of findings

Cognitive ability in childhood is not fixed and is not predetermined solely by genetics.[60] It is malleable and can change in response to the socioenvironmental exposures of the child.[61] Across two Irish population-based cohort studies, children with below average cognitive ability were characterised by markers of socioeconomic disadvantage. In the BASELINE cohort study, children with lower socioeconomic index scores, those from single parent families, those from lower income households, children living in government housing, children whose mothers consumed higher

levels of alcohol and cigarettes in the first trimester of pregnancy, and children who were predominantly formula fed were more likely to have below average cognitive ability at age 5.

In the Growing Up in Ireland cohort, children with low cognitive ability at age 5 had a similar socioeconomic profile. Among those with low cognitive ability, a higher proportion had mothers who smoked, a higher proportion had mothers with only primary or secondary education, and a higher proportion had mothers for whom English was not the native language. Children with low cognitive ability had a lower median family income, tended to be from lower social classes, and a significantly lower proportion reported living in owner occupied accommodation.

Efforts to address inequity in childhood cognitive outcomes must be multi-faceted, cross-sectoral, and must include primary, secondary, and tertiary prevention. While the main focus of this thesis was on early identification and intervention for at risk children, a form of secondary prevention, the findings of this thesis and a large body of international literature have consistently shown that socioeconomic inequity is a driving force in the disparities seen in childhood cognitive outcomes.[61,62] Primary prevention must focus on addressing these underlying drivers.

Self-actualisation, the fulfilment of one's potential, was described in 1943 by Maslow as a higher need, which cannot be met before basic lower level needs are met.[63] To fulfil one's cognitive potential, first basic physiological needs must be met. In January 2025, there were 2,164 families experiencing homelessness in Ireland, including 4,603 children.[64] The number of children experiencing homelessness in Ireland has increased steadily since 2021 and this now represents the highest number

of children ever recorded as experiencing homeless in Ireland.[65] It is important to note that these figures do not include families living in direct provision centres, families in domestic refuges, or families sleeping rough.[64,65]

The emergency accommodation for families experiencing homelessness in Ireland may be a bed and breakfast, hotel, hostel, or family hub. It is temporary and unstable. Families may share a single room, may or may not have access to workspaces for homework, and may or may not have access to a kitchen.[66] The Irish Governments 'Housing for All' policy aims to eradicate homelessness by 2030 but with the current rental and housing situation in Ireland this looks increasingly unlikely.[67] All children require stable and suitable accommodation to reach their full cognitive potential and this must be a policy priority in Ireland.

No child can be expected to reach their cognitive potential, or to participate to the best of their ability in school, if the basic need for food is not met. In the 2018 Health Behaviour in School Age Children Study, a representative cross-sectional survey of 15,557 children age 8-18 years in Ireland, 19% reported either going to school or bed hungry because there was not enough food at home.[68] Data from the Irish Central Statistics Office in 2022 indicated that 15.2% of children were at risk of poverty, an increase of 1.6% from 2021, and 19.9% experienced deprivation, an increase of 2.9% from 2021.[69] Poverty is intimately intertwined with substance abuse, depression and mental health problems, and domestic violence – exposures which all have detrimental effects on the cognitive development of a child.[70-73]

The Government in Ireland have made steps toward addressing child poverty and in Spring 2023 established 'The Child Poverty and Well-Being Programme Office' to coordinate the government response to child poverty. The initial programme plan

set out six priority areas in which to focus efforts to reduce child poverty – income support and joblessness, early learning and childcare, education costs, family homelessness, service integration, participation in arts, culture, and sport.[74] Actions in these areas, particularly that of income support, joblessness, and homelessness, are fundamental to achieving equity in early childhood development and cognition.

The findings of this thesis, and of the literature reviewed as part of this thesis, all point to the importance of early intervention, prior to starting formal education, for children at risk of poor cognitive outcomes. In Ireland, there have been positive steps taken to provide universal access to early education. The Early Childhood Care and Education Scheme, first introduced in 2010, provides state-funded care and education, for three hours a day five days per week for children between age 2 years 8 months until they transfer to primary school.[75] However, moving toward a system of progressive universalism, whereby everyone is entitled to a baseline universal service, with additional services and resources provided according to need, could help narrow the gap in early cognitive functioning for children additional biological or socioeconomic risk factors.[76] The predictive models developed in this study are a step towards enabling this approach. While, at a population level, the predictive performance of the model developed in Chapter 7 is not yet good enough to be implemented, it is a step towards enabling an individualised, risk-based approach which could be part of a broader system of progressive universalism, whereby those with most needs are identified and provided with additional care and resources.

9.7. Ethical issues arising from this thesis

9.7.1. Ethical issues related to measuring cognitive ability

Cognitive ability is undoubtedly an important determinant of a broad range of outcomes throughout the life course. However, it alone does not determine achievement, success, or happiness in life.[77] Focussing a public health intervention solely on this one domain of ability risks minimising the importance of other equally important domains of ability and perpetuating a system of education which places higher value on cognitive ability and achievement than on other non-cognitive abilities. Gardner's theory of multiple intelligences describes at least eight forms of intelligence - linguistic-verbal, logical-mathematical, visuo-spatial, body-kinaesthetic, musical, interpersonal, intrapersonal, and naturalistic intelligences.[78] By adopting a narrow focus on predominantly verbal and logical-mathematical cognitive abilities, we risk endorsing current educational and societal structures which disproportionately value cognitive achievement. To truly encourage, nurture, and value other non-cognitive abilities and all forms of intelligence would require a systematic shift in both the structure of our education system, including teaching and assessment, as well as a shift in public attitudes toward achievements and abilities. Establishing, developing, and encouraging each child's individual strengths and abilities could help children with cognitive difficulties engage with the education system, and improve self-esteem and confidence within the school setting.

9.7.2. Ethical issues related to early labelling

One of the overall aims of this thesis was to devise a method for the early identification of children at risk of below average cognitive ability in childhood using

perinatal and early life information. Achieving this would mean applying some form of label to an infant in the first year of life. The associated risks must be carefully considered. The results of this thesis clearly demonstrate that these children are disproportionately characterised by socioeconomic disadvantage, and the application of a label so early in life risks further stigmatisation and systematically pathologising children from certain socioeconomic backgrounds.[79] Identifying a child as being at risk of poor cognitive outcomes in childhood can detrimentally affect parental expectations which have been shown to be an important determinant of a child's actual educational achievement.[80] Children whose parents hold high expectations have been shown to achieve higher scores on standardised tests, remain in school longer, and have higher motivation than children whose parents hold lower expectations.[80,81] Equally, school feedback on academic performance is also an important determinant of parental expectation about their child's intellectual ability and future.[82] It is not only parental expectations that could be altered by early labelling, but teachers expectations too, which can also be affected by race, gender, and other sociodemographic factors. It is well documented that these biases can result in self-fulfilling prophecies, whereby a teacher's inaccurate expectation can hinder a child's learning resulting in lower educational gains and thus a self-fulfilling prophecy.[83] These risks must be carefully considered, and they are especially pertinent in the case of children who would be identified falsely as being at risk of later cognitive difficulties with a prediction tool.

The risks must be balanced against the risk of not identifying at risk children and not providing early intervention, the status quo in Ireland, which continues to result in educational inequality, itself a cause of socioeconomic inequality and the

intergenerational transmission of inequality. Identifying children and families who would benefit from early intervention must always come from a place of support and be cognisant of risks of further stigmatisation. If a model with acceptable predictive performance were developed, public patient involvement in how this model could be implemented in real life would be a critical step. To be of benefit, early identification needs to be linked with early supports and effective intervention appropriate for each child and family.

9.7.3. Ethical issues related to the use of machine learning in public health

The black box nature of machine learning models predisposes them to inherent algorithmic bias which can be invisible without careful scrutiny of the model and the processes used to develop it. There are examples in the literature which demonstrate how the use of ML algorithms can actually exacerbate health inequalities. Obermeyer et al rigorously examined a widely-used commercial machine learning algorithm in the United States which aimed to identify and support those with complex health needs.[84] Their analysis demonstrated inherent racial bias in the algorithm which, if left unaddressed, would result in widening racial health inequalities. The racial bias in the algorithm resulted from the use of an inappropriate surrogate marker. The model used health care cost as a surrogate marker intended to represent the underlying health care needs of the individual. However, it is widely established that health care utilisation, and therefore cost, does not equal health care need. This is largely due to barriers in accessing healthcare.[85] Unequal access to healthcare in the United States meant more money was spent caring for a white patient, who in general was more likely to be able to access healthcare, than for a black patient with

the same underlying conditions. This resulted in an algorithmic bias in favour of white patients.[84]

Algorithmic bias can be introduced to a ML model in its design, its training, and its application. One important source, relevant to this thesis, is the bias introduced by using data that is not representative of the whole population in which the model is intended to be used. Cohort data tends to exclude the most vulnerable groups within our population, such as those living with homelessness, addiction, or those living in temporary accommodation such as direct provision. A ML algorithm is trained on the data it sees, and if these individuals are not represented in that data then it is possible that algorithmic bias will be introduced and these at risk individuals may not be identified by the model. Using linked population-based registers can help overcome this challenge, but these are not currently in use in Ireland.

9.8. Future Research

Cognitive ability is an important construct to measure in population cohorts studies, as it represents both a marker of, and a predictor of biological, social, and psychological processes throughout the life course. However, as acknowledged throughout this thesis, its measurement in cohort studies is often flawed. Perhaps the most obvious bias in its measurement is the bias in favour of those tested in their native language. Most tests of cognitive ability either directly or indirectly test verbal ability. International migration has increased steadily over the past five decades and in Ireland, 16% of our population are citizens of other countries.[86] Future cohort studies must acknowledge the increasing proportion of children being raised in multilingual homes, and tests of cognitive ability in Irish cohorts must not

disadvantage those who native or home language is not English. Language-free, gaming based, touchscreen technologies have recently been developed and studied as an alternative measure of cognitive ability but require further validation among large, diverse populations. Future research should focus on developing these novel language-independent methods of measuring cognitive ability.[87]

A limitation in this thesis was the use of cognitive ability as an outcome in the absence of a measure of adaptive behaviour. Adaptive behaviour was not measured in the three data sources used. The measurement of adaptive behaviour is arguably more complex than cognitive testing as ideally it should measure a child's behaviour and skills in a variety of settings based on feedback from a variety of people including parents, the child, and teachers.[88,89] Scales do exist, such as the Vineland Adaptive Behaviour Scales and the Adaptive Behaviour Assessment System but do not appear to be as widely used in cohort studies as measures of cognitive ability.[88,89] Measures of adaptive behaviour would help us to further understand the lived experience for those with below average cognitive ability. Outside of cohort studies, qualitative research involving individuals with below average cognitive ability and their families, particularly individuals whose cognitive ability lies between 1 and 2 standard deviations below the mean, would be particularly valuable in shaping directions for future research. This qualitative work should also focus on what individuals and families with lived experience of having below average cognitive ability deem the important outcomes to be and early intervention programmes should be tailored to their needs.

One of the main questions arising from this thesis is how can we improve our ability to identify, at the earliest possible stage, infants and families who will benefit

from early intervention programmes. The models developed in this study, using readily available data which could be collected in the perinatal period, demonstrate that these features alone provide a large amount of predictive information about cognitive outcomes in childhood. However, improvement in predictive validity would be required for introduction at a population level. There are a number of possible avenues for future research, although feasibility and practicality at a population level must be considered. Firstly, future research should consider what additional features could be added to the predictive model developed in Chapter 7 to improve performance at a general population level.

Approximately 25%-50% of intellectual developmental disorders have a genetic cause. When an IDD is accompanied by additional clinical symptoms or physical anomalies, it may lead to identification of a known syndrome such as Trisomy 21 or DiGeorge syndrome (22q11.2 deletion). There are many genetic disorders which cause non-syndromic IDDs.[90] Similarly, a substantial proportion of below average cognitive ability can be attributed to specific genetic or chromosomal disorders.[91] Ongoing advances in genomic sequencing technologies coupled with declining costs, means access to genetic testing, such as whole genome sequencing (WGS), is becoming increasingly realistic at a population level. WGS, which refers to the sequencing of the entire genome of an individual, can be used to diagnose suspected genetic syndromes and also to search for etiological causes for unexplained intellectual developmental disorders.[92] The use of WGS at birth, at a broader population level, is currently under investigation by Genomics England in the Generation Study. The study aims to sequence the genomes of 100,000 newborn babies and to screen for over 200 genetic conditions, many of which are linked to low

cognitive ability.[93] Combining this type of genomic data, with the perinatal and socioeconomic predictors identified in this thesis could improve the accuracy of a screening algorithm for below average cognitive ability at birth.

Intelligence is a heritable trait. Unlike the genetic disorders discussed, there is no single gene responsible for the heritability of intelligence but instead it is the result of many genes of small effect.[94] Including parental IQ in a predictive model would likely improve model performance. There are cohorts, for example the Danish National Birth Cohort, where this information has been collected.[44] ML models incorporating this information in combination with the variables identified in our study in Chapters 6 & 7 should be trained and tested.

Eye-tracking is receiving attention in the developmental literature, largely due to improvements in eye-tracking technologies and analytic approaches.[95] Gaze location can be objectively measured in children from a few days old, and can reveal information about a child's attentional focus.[96] The pathways which control reflexive orienting to a stimulus are mature at birth and can be used to study cognition in very early life, while those controlling voluntary eye movements develop during the first six months.[97] This information can provide valuable insights into the developing brain. For example, a study which tracked gaze location in infants age 4-12 months as they watched a female speaking in their language, found that between age 4-8 months the infants focus of attention was on the mouth of the female, while at 12-months the focus was on the eyes. The authors proposed that at age 4-8 months the infant focuses on the mouth in an attempt to learn to speak and produce sound, while at later ages they focus on the eyes to search for social cues.[96] Previous research has already demonstrated significant correlations between visual reaction

time and anticipation at age 3.5 months and childhood IQ at age 4 years.[97] The advances in hardware and software for measuring eye tracking mean it is increasingly feasible to perform these tests at a broader population-level. Incorporating this information may improve prediction of a ML model.

The role of the infant microbiome in neurodevelopment is also receiving attention in the literature, and there is now a body of research examining the microbiome-gut-brain axis.[98] Much of the evidence to date comes from animal experiments but increasingly human studies are being published. The first prospective human study examining the relationship between the infant gut microbiome and later cognition was published in 2018, and found significant associations between distinct gut microbial patters at age 1 and cognitive outcomes at age 2.[99] This relationship is clearly confounded by breastfeeding, which in Ireland acts as a surrogate marker for socioeconomic background. However, after adjustment for feeding type and other sociodemographic variables the association in this study remained significant. The microbiome could provide additional predictive value for future cognition but the feasibility of sampling microbiome samples at a population level would need to be established.

In addition to investigating how predictive performance could be improved with addition of novel predictors, future research must ensure that high quality data is being used in ML algorithms. This data must be representative of the population in which the algorithm is intended for use. Population registers, which can be linked to health registers using unique identification numbers, such as those in use in Sweden, can ensure data is representative of the population as a whole.[100] One disadvantage is that they often lack the granular detail which may be contained in smaller cohort

studies. Cohort studies can tailor their data collection toward research questions of interest, can include serum biobanking, genomics, imaging, and more novel assessments which will not be contained in health registers. However, to train and test unbiased ML models data must be representative and there should ideally be large volumes. Researchers, and those funding cohort studies, must make representativeness a key objective in the design and implementation of their study, and consider how minority populations can be encouraged to participate and remain engaged in these studies. To truly harness the value of cohort studies, methods of linking and harmonising multiple international cohorts are required. Through amalgamating data sources, large data repositories could be created which would be particularly valuable for all fields of research, but especially for AI. Work in this area is ongoing for pregnancy and childhood cohort studies.[101]

9.9.Conclusion

The overall aim of this thesis was to explore the impact of below average cognitive ability on outcomes in childhood, and to examine current and alternative methods, namely the use of machine learning, for the early identification of children at risk of below average cognitive ability in childhood.

There is a large body of published literature examining the relationship between early cognitive ability and later life outcomes such as educational attainment, social mobility, physical and mental health, and occupational status. However, far less has been published on how early cognitive ability impacts important outcomes in childhood. Much of what has been published focuses on those with diagnosed intellectual developmental disorders and excludes the larger

proportion of often overlooked children with cognitive ability between 1 and 2 standard deviations below the mean. In this thesis we examined important childhood outcomes related to emotional-behavioural functioning and the early experience of education for children whose cognitive ability falls between 1 and 2 standard deviations below the mean, as well as for those with cognitive ability more than 2 standard deviations below the mean.

The findings of this thesis suggest that children whose cognitive ability lies between 1 and 2 standard deviations below the mean represent a risk group, who compared to their peers with typical cognitive development, on average experience a worsening in their emotional-behavioural function during the period coinciding with their transition into school. This longitudinal relationship between early cognitive ability and emotional-behavioural function during childhood, particularly around the school transition, has not previously been described in the published literature for children with below average cognitive ability.

The relationship uncovered may be explained, at least in part, by a negative early experience of school where children with unrecognised below average cognitive ability struggle to cope with new and increasingly complex academic demands. This may be reinforced by repeated failure at school, and constant comparison to typically developing peers. The findings of the thesis support this hypothesis and show that children who begin school with below average cognitive ability at age 5 have higher odds of poor classroom engagement, of low self-concept, particularly intellectual self-concept, and of teacher-reported emotional-behavioural problems at age 9 years. These findings have important implications for public health and early educational policy. They show that adverse outcomes associated with early

low cognitive ability are present not only in adulthood, but also early in childhood. These findings add further impetus for the provision of early, pre-emptive intervention, prior to starting school for at risk children. Properly resourced, early intervention programmes have been shown in both experimental and real world settings to improve cognitive outcomes in at risk children. The challenge which remains is identifying at a population-level the children who will benefit most from this early intervention.

We have shown in this thesis that the current infant developmental screening tool routinely used in Ireland, the Ages and Stages Questionnaire, will not detect the majority of children who will have below average cognitive ability at school-age. It is well established that many of the most important factors for predicting cognitive outcomes in childhood occur, or are present in the perinatal period. Multivariable models, using epidemiological data, for predicting low cognitive ability in childhood have been published in the literature but to date, there had been no published studies exploring a range of machine learning models for this purpose. The hypothesised benefit of using machine learning algorithms for this purpose was their ability to model potentially important interactions between variables, without the need for pre-specification, and their ability to model non-linear relationships which we know exist between certain predictors and the outcome. At a general population level the random forest algorithm had superior predictive ability for predicting cognitive ability more than 1.5 SD below the mean at age 5 using perinatal data, compared to the more traditional logistic regression model. However, among a very preterm population logistic regression outperformed a range of machine learning algorithms for predicting cognitive delay at age 2. It should be noted that a variety of

machine learning techniques were used in the feature selection process prior to performing logistic regression in this study.

The findings of this thesis demonstrate the value of applying machine learning techniques for prediction problems in public health. It is likely that the value of machine learning in public health will continue to increase as the quantity and complexity of data we collect increases. While at a population level, the model developed in this thesis did not achieve a sufficient level of predictive performance that would demand its introduction into clinical practice, it is a step toward early, personalised, risk stratification that could enable delivery of an early intervention system based on the principal of proportionate universalism. This, however, is not without its consequences and throughout this thesis important ethical considerations were raised. The unintended negative consequences of early labelling must be carefully considered, as must the potential of overemphasising the importance of this one form of ability.

Children with below average cognitive ability, including those whose cognitive ability lies between 1 and 2 standard deviations below the mean, represent a risk group with increased odds of experiencing adverse emotional-behavioural and school outcomes in childhood. There is a window of opportunity in the first 1,000 days of life, when the infant brain has an amazing ability to adapt and respond to its environmental stimuli, to provide additional support and environmental enrichment to the infant and their family. The key outcome determinant, like many things in life, is early recognition of a problem. Early identification and intervention within this 1,000 day window could alter the cognitive trajectory for that child for their entire adult life. Paradoxically, the largest obstacle to the reliable identification of many of

these children with below average cognitive ability is that their difficulties are much more subtle than children with diagnosed intellectual developmental disorders. Further research is needed to determine the optimal method for identifying, at a population level, the children who could benefit most from early intervention. One promising avenue for further research is the combination of the epidemiological predictive data identified in this thesis with genetic data, as this becomes increasingly feasible to obtain at a population level.

Intervention at this individual level alone will not achieve equity in childhood cognitive outcomes. It must occur alongside primary prevention efforts to address the upstream drivers of adverse cognitive outcomes – namely homelessness, unemployment, income inequality, and addiction. Finally, while every child should have the opportunity to reach their individual cognitive potential, this potential will differ throughout our population. Equally musical, artistic, physical, social, and emotional abilities will differ. That is the nature and beauty of human diversity. Our education system should strive to nurture, praise, and develop each individual child's array of abilities. To return to the opening line of this thesis - "The question is not whether we can afford to invest in every child; it is whether we can afford not to".

9.10. References

1. McIntyre LL, Blacher J, Baker BL. The transition to school: adaptation in young children with and without intellectual disability. J Intellect Disabil Res. 2006 May;50(Pt 5):349-61. doi: 10.1111/j.1365-2788.2006.00783.x. PMID: 16629928.
2. Central Statistics Office Ireland. Vital Statistics Yearly Summary 2023 [Internet]. May 2024 [cited 2025 March 03]. Available at: https://www.cso.ie/en/releasesandpublications/ep/p-vsys/vitalstatisticsyearlysummary2023/
3. Luo R, Song L, Chiu IM. A Closer Look at the Birth Order Effect on Early Cognitive and School Readiness Development in Diverse Contexts. Front Psychol. 2022 Jun 2;13:871837. doi: 10.3389/fpsyg.2022.871837.
4. Blake J. Family size and the quality of children. Demography. 1981 Nov;18(4):421-42. PMID: 7308532.
5. Barreto FB, Sánchez de Miguel M, Ibarluzea J, Andiarena A, Arranz E. Family context and cognitive development in early childhood: A longitudinal study. Intelligence.2017; 65:11-22. https://doi.org/10.1016/j.intell.2017.09.006
6. Hegelund ER, Teasdale TW, Okholm GT, Osler M, Sørensen TIA, Christensen K, et al. The secular trend of intelligence test scores: The Danish experience for young men born between 1940 and 2000. PLoS One. 2021 Dec 9;16(12):e0261117. doi: 10.1371/journal.pone.0261117.

7. Central Statistics Office. Central Statistics Office Census 2011. 2011 [cited 2023 Sep 07]. Available at: https://www.cso.ie/en/census/census2011reports/
8. Trahan LH, Stuebing KK, Fletcher JM, Hiscock M. The Flynn effect: a meta-analysis. Psychol Bull. 2014 Sep;140(5):1332-60. doi: 10.1037/a0037173.
9. Bain SK, Jaspers KE. Review of Kaufman Brief Intelligence Test, Second Edition [Review of the software Kaufman brief intelligence test, second edition, by A. S. Kaufman & N. L. Kaufman]. Journal of Psychoeducational Assessment. 2010; 28(2), 167–174. https://doi.org/10.1177/0734282909348217
10. Thornton M, Williams J, McCrory C, Murray A, Quail A. Growing Up in Ireland. Design, Instrumentation and Procedures for the Infant Cohort at Wave One [Internet]. Dublin; 2013 [cited 2023 Oct 05]. Available at: https://www.growingup.gov.ie/pubs/BKMNEXT252.pdf
11. Quail A, O'Reilly C, Watson D, McNamara E, O'Mahony D, Murray A. Growing Up in Ireland A Summary Guide to Cohort '08 at 9 Years [Internet]. Dublin; 2019 [cited 2023 Oct 05]. Available at: https://www.growingup.gov.ie/pubs/Summary-Guide-Infant-Cohort-Wave-5.pdf
12. Parks RW, Stevens RJ, Spence SA. A systematic review of cognition in homeless children and adolescents. J R Soc Med. 2007 Jan;100(1):46-50. doi: 10.1177/014107680710000116. PMID: 17197688; PMCID: PMC1761678.

13. Kodituwakku PW. Neurocognitive profile in children with fetal alcohol spectrum disorders. Dev Disabil Res Rev. 2009;15(3):218-24. doi: 10.1002/ddrr.73.
14. Nygaard E, Moe V, Slinning K, Walhovd KB. Longitudinal cognitive development of children born to mothers with opioid and polysubstance use. Pediatr Res. 2015 Sep;78(3):330-5. doi: 10.1038/pr.2015.95.
15. Norman M, Källén K, Wahlström E, Håkansson S; SNQ Collaboration. The Swedish Neonatal Quality Register - contents, completeness and validity. Acta Paediatr. 2019 Aug;108(8):1411-1418. doi: 10.1111/apa.14823. Epub 2019 May 20. PMID: 31006126.
16. Meyer EM, Reynolds MR. Multidimensional Scaling of Cognitive Ability and Academic Achievement Scores. J Intell. 2022 Dec 1;10(4):117. doi: 10.3390/jintelligence10040117. PMID: 36547504; PMCID: PMC9785841.
17. Elliott CD, Smith P, McCulloch K. British Ability Scales Second Edition (BAS II). Administration and Scoring Manual. London: Nelson;1996
18. Hoff E, Core C. What Clinicians Need to Know about Bilingual Development. Semin Speech Lang. 2015 May;36(2):89-99. doi: 10.1055/s-0035-1549104.
19. Twomey DM, Wrigley C, Ahearne C, Murphy R, De Haan M, Marlow N et al. Feasibility of using touch screen technology for early cognitive assessment in children. Arch Dis Child. 2018 Sep;103(9):853-858. doi: 10.1136/archdischild-2017-314010.
20. Lowe JR, Nolen TL, Vohr B, Adams-Chapman I, Duncan AF, Watterberg K. Effect of primary language on developmental testing in children born extremely preterm. Acta Paediatr. 2013 Sep;102(9):896-900. doi:

10.1111/apa.12310. Epub 2013 Jun 19. PMID: 23735043; PMCID: PMC4108617.

21. Chin HB, Baird DD, McConnaughey DR, Weinberg CR, Wilcox AJ, Jukic AM. Long-term Recall of Pregnancy-related Events. Epidemiology. 2017 Jul;28(4):575-579. doi: 10.1097/EDE.0000000000000660.
22. Latkin CA, Edwards C, Davey-Rothwell MA, Tobin KE. The relationship between social desirability bias and self-reports of health, substance use, and social network factors among urban substance users in Baltimore, Maryland. Addict Behav. 2017 Oct;73:133-136. doi: 10.1016/j.addbeh.2017.05.005.
23. Teh WL, Abdin E, Siva Kumar FD, Roystonn K, Wang P, Shafie S, et al. Measuring social desirability bias in a multi-ethnic cohort sample: its relationship with self-reported physical activity, dietary habits, and factor structure. BMC Public Health. 2023 Mar 1;23(1):415. doi: 10.1186/s12889-023-15309-3.
24. Digitale JC, Martin JN, Glymour MM. Tutorial on directed acyclic graphs. J Clin Epidemiol. 2022 Feb;142:264-267. doi: 10.1016/j.jclinepi.2021.08.001. Epub 2021 Aug 8.
25. Braveman PA, Cubbin C, Egerter S, Chideya S, Marchi KS, Metzler M, et al. Socioeconomic status in health research: one size does not fit all. JAMA. 2005 Dec 14;294(22):2879-88. doi: 10.1001/jama.294.22.2879. PMID: 16352796.
26. Green MJ, Popham F. Interpreting mutual adjustment for multiple indicators of socioeconomic position without committing mutual adjustment fallacies.

BMC Public Health. 2019 Jan 3;19(1):10. doi: 10.1186/s12889-018-6364-y.

27. O'Donovan SM, Murray DM, Hourihane JO, Kenny LC, Irvine AD, Kiely M. Cohort profile: The Cork BASELINE Birth Cohort Study: Babies after SCOPE: Evaluating the Longitudinal Impact on Neurological and Nutritional Endpoints. Int J Epidemiol. 2015 Jun;44(3):764-75. doi: 10.1093/ije/dyu157. Epub 2014 Aug 7. PMID: 25102856.
28. Williams J, Thornton M, Murray A, Quail A. Growing Up in Ireland Design, Instrumentation and Procedures for Cohort '08 at Wave 3 (5 Years).[Internet].Dublin; 2019 [cited 2023 Sep 25]. Available at https://www.growingup.gov.ie/pubs/20190404-Cohort-08-at-5years-design-instrumentation-and-procedures.pdf
29. Norman M, Hakansson S, Swartling Schlinzig L. Swedish Neonatal Quality Register [Internet]. [cited 2023 Sep 12]. Available from: https://www.medscinet.com/PNQ/
30. Elliott CD, Smith P, McCulloch K. British Ability Scales (2nd ed.). NFER-Nelson; 1996.
31. Kaufman AS, Kaufman NL. Kaufman Brief Intelligence Test, Second Edition. Bloomington, MN: Pearson, Inc;2004
32. Bayley N. Bayley scales of infant and toddler development, third edition: Administration manual. San Antonio, TX: Harcourt;2006
33. Squires A, Thompson R, Sadarangani T, Amburg P, Sliwinski K, Curtis C, et al. International migration and its influence on health. Res Nurs Health. 2022 Oct;45(5):503-511. doi: 10.1002/nur.22262.

34. Neisser U, Boodoo G, Bouchard Jr. TJ, Boykin AW, Brody N, Ceci SJ, et al. Intelligence: Knowns and unknowns. Am Psychol. 1996;51:77–101. https://doi.org/10.1037/0003-066X.51.2.77

35. Drigas AS, Papoutsi C. A New Layered Model on Emotional Intelligence. Behav Sci (Basel). 2018 May 2;8(5):45. doi: 10.3390/bs8050045.

36. Greenspan S, Harris J. Intellectual disability is “a condition, not a number”: Ethics of IQ cut-offs in psychiatry, human services and law. Ethics, Medicine and Public Health. 2015 August;1(3):312-324. doi: 10.1016/j.jemep.2015.07.004.

37. Wieland J, Zitman FG. It is time to bring borderline intellectual functioning back into the main fold of classification systems. BJPsych Bull. 2016 Aug;40(4):204-6. doi: 10.1192/pb.bp.115.051490. PMID: 27512590; PMCID: PMC4967780.

38. World Health Organisation. ICD-11 for Mortality and Morbidity Statistics: 6A00 Disorders of intellectual development.[Internet] 2023 [cited 2023 Sep 05]. Available at: https://icd.who.int/browse11/l-m/en#/http%253A%252F%252Fid.who.int%252Ficd%252Fentity%252F605267007

39. Camargo-Figuera FA, Barros AJ, Santos IS, Matijasevich A, Barros FC. Early life determinants of low IQ at age 6 in children from the 2004 Pelotas Birth Cohort: a predictive approach. BMC Pediatr. 2014 Dec 16;14:308. doi: 10.1186/s12887-014-0308-1. PMID: 25510879; PMCID: PMC4272809.

40. Karrasch M, Tiitta P, Hermann B, Joutsa J, Shinnar S, Rinne J, et al. Cognitive Outcome in Childhood-Onset Epilepsy: A Five-Decade Prospective Cohort

Study. J Int Neuropsychol Soc. 2017 Apr;23(4):332-340. doi: 10.1017/S1355617716001077. Epub 2017 Jan 10.

41. National Research Council (US) Committee on Disability Determination for Mental Retardation. Mental Retardation: Determining Eligibility for Social Security Benefits. Reschly DJ, Myers TG, Hartel CR, editors. Washington (DC): National Academies Press (US); 2002. PMID: 25032290.
42. Tassé MJ, Kim M. Examining the Relationship between Adaptive Behavior and Intelligence. Behav Sci (Basel). 2023 Mar 13;13(3):252. doi: 10.3390/bs13030252.
43. Deary IJ, Cox SR, Hill WD. Genetic variation, brain, and intelligence differences. Mol Psychiatry. 2022; 27:335–353 https://doi.org/10.1038/s41380-021-01027-y
44. Eriksen HL, Kesmodel US, Underbjerg M, Kilburn TR, Bertrand J, Mortensen EL. Predictors of intelligence at the age of 5: family, pregnancy and birth characteristics, postnatal influences, and postnatal growth. PLoS One. 2013 Nov 13;8(11):e79200. doi: 10.1371/journal.pone.0079200.
45. Mooney SJ, Keil AP, Westreich DJ. Thirteen Questions About Using Machine Learning in Causal Research (You Won't Believe the Answer to Number 10!). Am J Epidemiol. 2021 Aug 1;190(8):1476-1482. doi: 10.1093/aje/kwab047..
46. Hackman DA, Farah MJ, Meaney MJ. Socioeconomic status and the brain: mechanistic insights from human and animal research. Nat Rev Neurosci. 2010 Sep;11(9):651-9. doi: 10.1038/nrn2897. PMID: 20725096; PMCID: PMC2950073.

47. McLoyd VC. Socioeconomic disadvantage and child development. Am Psychol. 1998 Feb;53(2):185-204. doi: 10.1037//0003-066x.53.2.185.
48. Poon AIF, Sung JJY. Opening the black box of AI-Medicine. J Gastroenterol Hepatol. 2021 Mar;36(3):581-584. doi: 10.1111/jgh.15384. PMID: 33709609.
49. Maxim LD, Niebo R, Utell MJ. Screening tests: a review with examples. Inhal Toxicol. 2014 Nov;26(13):811-28. doi: 10.3109/08958378.2014.955932.
50. Health Service Executive. National Screening Service [Internet]. [cited 2023 Oct 13] Available at: https://www2.healthservice.hse.ie/organisation/nss/
51. Koliopoulos G, Nyaga VN, Santesso N, Bryant A, Martin-Hirsch PP, Mustafa RA, et al. Cytology versus HPV testing for cervical cancer screening in the general population. Cochrane Database Syst Rev. 2017 Aug 10;8(8):CD008587. doi: 10.1002/14651858.CD008587.pub2.
52. Lee JK, Liles EG, Bent S, Levin TR, Corley DA. Accuracy of fecal immunochemical tests for colorectal cancer: systematic review and meta-analysis. Ann Intern Med. 2014 Feb 4;160(3):171. doi: 10.7326/M13-1484. PMID: 24658694; PMCID: PMC4189821.
53. European Foundation for the Care of Newborn Infants. Country snapshot Ireland [Internet]. [cited 2023 Oct 13]. Available at: https://www.efcni.org/wp-content/uploads/2018/03/EFCNI_ireland_light.pdf_copyright.pdf
54. Field D, Draper ES, Fenton A, Papiernik E, Zeitlin J, Blondel B, et al; MOSAIC research group. Rates of very preterm birth in Europe and neonatal mortality rates. Arch Dis Child Fetal Neonatal Ed. 2009 Jul;94(4):F253-6. doi: 10.1136/adc.2008.150433.

55. Matos J, Amorim M, Silva S, Nogueira C, Alves E. Prematurity-related knowledge among mothers and fathers of very preterm infants. J Clin Nurs. 2020 Aug;29(15-16):2886-2896. doi: 10.1111/jocn.15361.

56. Dyar OJ, Haglund BJA, Melder C, Skillington T, Kristenson M, Sarkadi A. Rainbows over the world's public health: determinants of health models in the past, present, and future. Scand J Public Health. 2022 Nov;50(7):1047-1058. doi: 10.1177/14034948221113147.

57. Chen M, Zhang B, Cai Z, Seery S, Gonzalez MJ, Ali NM, et al Ren R. Acceptance of clinical artificial intelligence among physicians and medical students: A systematic review with cross-sectional survey. Front Med (Lausanne). 2022 Aug 31;9:990604. doi: 10.3389/fmed.2022.990604.

58. Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell. 2019; 1: 206–215. https://doi.org/10.1038/s42256-019-0048-x

59. Khullar D, Casalino LP, Qian Y, Lu Y, Chang E, Aneja S. Public vs physician views of liability for artificial intelligence in health care. J Am Med Inform Assoc. 2021 Jul 14;28(7):1574-1577. doi: 10.1093/jamia/ocab055.

60. Petrill SA, Lipton PA, Hewitt JK, Plomin R, Cherny SS, Corley R, DeFries JC. Genetic and environmental contributions to general cognitive ability through the first 16 years of life. Dev Psychol. 2004 Sep;40(5):805-12. doi: 10.1037/0012-1649.40.5.805. PMID: 15355167; PMCID: PMC3710702.

61. Nilsen FM, Ruiz JDC, Tulve NS. A Meta-Analysis of Stressors from the Total Environment Associated with Children's General Cognitive Ability. Int J

Environ Res Public Health. 2020 Jul 29;17(15):5451. doi: 10.3390/ijerph17155451. PMID: 32751096; PMCID: PMC7432904.

62. von Stumm S, Plomin R. Socioeconomic status and the growth of intelligence from infancy through adolescence. Intelligence 2015; 48: 30–36. doi:10.1016/j.intell.2014.10.002
63. Kenrick DT, Griskevicius V, Neuberg SL, Schaller M. Renovating the Pyramid of Needs: Contemporary Extensions Built Upon Ancient Foundations. Perspect Psychol Sci. 2010 May;5(3):292-314. doi: 10.1177/1745691610369469. PMID: 21874133; PMCID: PMC3161123.
64. Department of Housing, Local Government and Heritage. Monthly Homelessness Report January 2025 [Internet]. Department of Housing, Local Government and Heritage; 2025 February [cited 2025 March 4th]. Available from: https://data.gov.ie/dataset/homelessness-report-january-2025
65. Holland K. Homeless figures reach record high with 14,429 people in emergency accommodation. Irish Times [Internet]. 2024 Aug 30 [cited 2025 Mar 05] Available from: https://www.irishtimes.com/ireland/social-affairs/2024/08/30/homeless-figures-reach-record-high-with-14429-people-in-emergency-accommodation/
66. Citizens Information. Housing and other supports for homeless people [Internet]. Citizens Information [updated 2019 May 2nd, cited 2023 December 4th]. Available from https://www.citizensinformation.ie/en/housing/losing-your-home-and-homelessness/supports-for-homeless-people/

67. Department of Housing, Local Government and Heritage. Housing for all – A New Housing Plan for Ireland.[Internet]. Department of Housing, Local Government and Heritage; 2021 September 2nd [cited 2023 December 4th]. Available from: https://www.gov.ie/pdf/?file=https://assets.gov.ie/197237/29edec3e-6664-4e62-86b2-af2e77f2f609.pdf#page=null
68. Költő A, Gavin A, Molcho M, Kelly C, Walker L, Nic Gabhainn S. The Irish Health Behaviour in School-aged Children (HBSC) Study 2018.[Internet]. Department of Health and National University of Ireland, Galway; 2020 January [cited 2023 December 5th]. Available from: https://www.universityofgalway.ie/media/healthpromotionresearchcentre/hbscdocs/nationalreports/2018-report---online-version-interactive---updated.pdf
69. Central Statistics Office. Survey on Income and Living Conditions (SILC) 2022. [Inetnet]. Central Statistics Office; 2023 February 22nd [cited 2023 December 5th]. Available from: https://www.cso.ie/en/releasesandpublications/ep/p-silc/surveyonincomeandlivingconditionssilc2022/poverty/
70. Mamluk L, Edwards HB, Savović J, Leach V, Jones T, Moore THM, et al. Prenatal alcohol exposure and pregnancy and childhood outcomes: a systematic review of alternative analytical approaches. Lancet. 2016;388:S73. https://doi.org/10.1016/S0140-6736(16)32309-1
71. Yeoh SL, Eastwood J, Wright IM, Morton R, Melhuish E, Ward M, Oei JL. Cognitive and Motor Outcomes of Children With Prenatal Opioid Exposure: A Systematic Review and Meta-analysis. JAMA Netw Open. 2019 Jul

3;2(7):e197025. doi: 10.1001/jamanetworkopen.2019.7025. PMID: 31298718; PMCID: PMC6628595.

72. Rogers A, Obst S, Teague SJ, Rossen L, Spry EA, Macdonald JA, et al. Association Between Maternal Perinatal Depression and Anxiety and Child and Adolescent Development: A Meta-analysis. JAMA Pediatr. 2020 Nov 1;174(11):1082-1092. doi: 10.1001/jamapediatrics.2020.2910.
73. Abel KM, Heuvelman H, Rai D, Timpson NJ, Sarginson J, Shallcross R, Mitchell H, Hope H, Emsley R. Intelligence in offspring born to women exposed to intimate partner violence: a population-based cohort study. Wellcome Open Res. 2019 Jul 10;4:107. doi: 10.12688/wellcomeopenres.15270.1. PMID: 31681855; PMCID: PMC6820818.
74. Department of the Taoiseach. Child Poverty and Well-Being Programme Office [Internet]. Department of the Taoiseach; 2023 September 1st [updated 2023 November 16th; cited 2023 December 5th]. Available from https://www.gov.ie/en/campaigns/7c189-child-poverty-and-well-being-programme-office/#
75. Citizens Information. Early Childhood Care and Education Scheme (ECCE) [Internet]. Citizens Information [updated 2023 October 5th; cited 2023 December 5th]. Available from: https://www.citizensinformation.ie/en/education/pre-school-education-and-childcare/early-childhood-care-and-education-scheme/
76. Moore T, McDonald M, McHugh H. Evidence review: Early childhood development and the social determinants of health inequities. [Internet]. Victorian Health Promotion Foundation; 2015 September [cited 2023

December 5th]. Available from: https://www.rch.org.au/uploadedFiles/Main/Content/ccch/151014_Evidence-review-early-childhood-development-and-the-social-determinants-of-health-inequities_Sept2015.pdf

77. Duckworth AL, Quirk A, Gallop R, Hoyle RH, Kelly DR, Matthews MD. Cognitive and noncognitive predictors of success. Proc Natl Acad Sci U S A. 2019 Nov 19;116(47):23499-23504. doi: 10.1073/pnas.1910510116.
78. Gardner H. Taking a multiple intelligences (MI) perspective. Behav Brain Sci. 2017 Jan;40:e203. doi: 10.1017/S0140525X16001631.
79. Werkhoven S, Anderson JH, Robeyns IAM. Who benefits from diagnostic labels for developmental disorders? Dev Med Child Neurol. 2022 Aug;64(8):944-949. doi: 10.1111/dmcn.15177.
80. Yamamoto Y, Holloway SD. Parental Expectations and Children's Academic Performance in Sociocultural Context. Educ Psychol Rev. 2010; 22: 189–214 https://doi.org/10.1007/s10648-010-9121-z
81. Davis-Kean PE. The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment. J Fam Psychol. 2005 Jun;19(2):294-304. doi: 10.1037/0893-3200.19.2.294.
82. Seginer R. Parents' Educational Expectations and Children's Academic Achievements: A Literature Review [Internet]. Merrill-Palmer Quarterly. 1983 [cited 2023 Dec 5]; 29(1): 1–23. Available from: https://www.jstor.org/stable/23086188

83. Jussim L, Harber KD. Teacher expectations and self-fulfilling prophecies: knowns and unknowns, resolved and unresolved controversies. Personal Soc Psychol Rev an Off J Soc Personal Soc Psychol Inc. 2005;9(2):131–55. https://doi.org/10.1207/s15327957pspr0902_3

84. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019 Oct 25;366(6464):447-453. doi: 10.1126/science.aax2342. PMID: 31649194.

85. National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Board on Health Care Services; Committee on Health Care Utilization and Adults with Disabilities. Health-Care Utilization as a Proxy in Disability Determination. Washington (DC): National Academies Press (US); 2018 Mar 1. PMID: 29782136.

86. Government of Ireland. Migration – the facts [Internet]. Government of Ireland; 2024 [cited 2025 Feb 26]. Available from: https://www.gov.ie/ga/bailiuchan/aeea0-migration-the-facts/#migration-in-numbers

87. Casey T, Thachuthara AJ, Fogarty L, Livingstone V, De Haan M, Marlow N, Kiely ME, Murray DM. Validation of a touchscreen assessment tool to screen for cognitive delay at 24 months. Dev Med Child Neurol. 2023 Sep;65(9):1206-1214. doi: 10.1111/dmcn.15555. Epub 2023 Feb 19. PMID: 36808732.

88. Tassé MJ, Schalock RL, Balboni G, Bersani H Jr, Borthwick-Duffy SA, Spreat S, Thissen D, Widaman KF, Zhang D. The construct of adaptive behavior: its conceptualization, measurement, and use in the field of intellectual

disability. Am J Intellect Dev Disabil. 2012 Jul;117(4):291-303. doi: 10.1352/1944-7558-117.4.291. PMID: 22809075.

89. Tamm L, Day HA, Duncan A. Comparison of Adaptive Functioning Measures in Adolescents with Autism Spectrum Disorder Without Intellectual Disability. J Autism Dev Disord. 2022 Mar;52(3):1247-1256. doi: 10.1007/s10803-021-05013-9. Epub 2021 Apr 26. PMID: 33900538; PMCID: PMC9255431.
90. Kaufman L, Ayub M, Vincent JB. The genetic basis of non-syndromic intellectual disability: a review. J Neurodev Disord. 2010 Dec;2(4):182-209. doi: 10.1007/s11689-010-9055-2. Epub 2010 Jul 29. PMID: 21124998; PMCID: PMC2974911
91. Kim HW. Etiology of Borderline Intellectual Functioning. Soa Chongsonyon Chongsin Uihak. 2024 Jul 1;35(3):188-191. doi: 10.5765/jkacap.240013. PMID: 38966196; PMCID: PMC11220479
92. Lu Y, Li M, Gao Z, Ma H, Chong Y, Hong J, Wu J, Wu D, Xi D, Deng W. Advances in Whole Genome Sequencing: Methods, Tools, and Applications in Population Genomics. Int J Mol Sci. 2025 Jan 4;26(1):372. doi: 10.3390/ijms26010372. PMID: 39796227; PMCID: PMC11719799
93. Genomics Education Programme, National Health Service England. The Generation Study [Internet]. National Health Service England; March 2024 [cited 2025 Jan 28]. Available from: https://www.genomicseducation.hee.nhs.uk/genotes/knowledge-hub/the-generation-

study/#:~:text=The%20Generation%20Study%20will%20evaluate,access%20to%20care%20and%20treatment.

94. Plomin R, Deary IJ. Genetics and intelligence differences: five special findings. Mol Psychiatry. 2015 Feb;20(1):98-108. doi: 10.1038/mp.2014.105. Epub 2014 Sep 16. PMID: 25224258; PMCID: PMC4270739.
95. Eckstein MK, Guerra-Carrillo B, Miller Singley AT, Bunge SA. Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? Dev Cogn Neurosci. 2017 Jun;25:69-91. doi: 10.1016/j.dcn.2016.11.001.
96. Lewkowicz DJ, Hansen-Tift AM. Infants deploy selective attention to the mouth of a talking face when learning speech. Proc Natl Acad Sci USA. 2012 Jan 31;109(5):1431-6. doi: 10.1073/pnas.1114783109.
97. Dougherty TM, Haith MM. Infant expectations and reaction time as predictors of childhood speed of processing and IQ. Dev Psychol. 1997 Jan;33(1):146-55. doi: 10.1037//0012-1649.33.1.146.
98. Warner BB. The contribution of the gut microbiome to neurodevelopment and neuropsychiatric disorders. Pediatr Res. 2019 Jan;85(2):216-224. doi: 10.1038/s41390-018-0191-9.
99. Carlson AL, Xia K, Azcarate-Peril MA, Goldman BD, Ahn M, Styner MA, et al. Infant Gut Microbiome Associated With Cognitive Development. Biol Psychiatry. 2018 Jan 15;83(2):148-159. doi: 10.1016/j.biopsych.2017.06.021.
100. Ludvigsson JF, Almqvist C, Bonamy AK, Ljung R, Michaëlsson K, Neovius M, et al. Registers of the Swedish total population and their use in

medical research. Eur J Epidemiol. 2016 Feb;31(2):125-36. doi: 10.1007/s10654-016-0117-y.

101. Jaddoe VWV, Felix JF, Andersen AN, Charles MA, Chatzi L, Corpeleijn E, et al; LifeCycle Project Group. The LifeCycle Project-EU Child Cohort Network: a federated analysis infrastructure and harmonized data of more than 250,000 children and parents. Eur J Epidemiol. 2020 Jul;35(7):709-724. doi: 10.1007/s10654-020-00662-z.

Appendix 1 Data access Growing Up in Ireland Infant Cohort (Chapter 3)

DATA REQUEST FORM FOR RESEARCH PURPOSES

Section 1: contact details

[Please type]

First name *	Andrea	Last name *	Bowe
Address	46 Cappaghmore, Clondalkin, Dublin 22		
Country *	Ireland		
Institution/Company *	INFANT Research Centre, University College Cork		
Email *	Andrea.bowe@ucdconnect.ie		
Phone	0876153190		

section 2: dataset requested* [Please type]

Please refer to the full list of datasets on the ISSDA website: www.ucd.ie/issda/data/

Please include Study Number (SN) where appropriate

Infant Cohort '08 of Growing Up in Ireland Study
SN: 0019-01
SN: 0019-02
SN 0019-03
SN: 0019-04
SN:0019-05

Section 3: INTENDED USE OF DATA [Please type]

Subject area *	☐ Economics	☐Education	☐Engineering	☐Environment/ Agriculture
	✓Health/ Well-being	☐Law	☐Policy	☐Social issues
	☐ Sport	Technology	☐Other [please specify] _____	

Please provide a short description (approx. 100 words) of your intended use of the dataset/s, including title and details of your intended research project*:

This data will be used as part of two mini projects which will contribute to an overall PhD on the topic of cognitive development in children.
The first project will examine the utility of the Ages and Stages Questionnaire in detecting borderline cognitive functioning as measured by a score of 1-2 standard deviations below the mean on the British Ability Scales at Age 5.
The second project will examine the association between cognitive function on the British Ability Scales and emotional-behavioural problems, measured by the Strengths and Difficulties Questionnaire.
Each project will adjust for the relevant confounding variables which will include among other variables markers of socio-economic status and gender.

Type of user *	☐Academic Staff	☐ Post Doc	☑PhD	☐Masters
	☐Undergraduate	☐Independent Researcher	☐Government/ Policy researcher	
	☐Other [please specify] _____			

Estimated end date for this use (maximum 5 years from date of	21/08/2022
Funder (where applicable)	I am a public health medicine specialist registrar completing a PhD as part of the Irish Clinical Academic Training (ICAT) Programme which is jointly funded by the HSE and HRB. No other external funding applies.

SECTION 4: PERSONS PERMITTED TO ACCESS THE DATA UNDER ARTICLE 3 OF TERMS OF USE*

The data may only be used for the purpose as outlined in Section 3. Data must be re-applied for in the case of new projects being undertaken.

Please indicate the number of users who will have access to and use of the data.

I – Andrea Bowe will have access to the data.

- All users are bound by the same terms and conditions as the End User
- Users may not be given a copy of the data for use outside of the scope of the research project outlined in Section 3
- Users wishing to use the data outside of the scope of the research project outlined in Section 3 can apply directly to ISSDA for a copy of the data, subject to the usual terms and conditions
- The data must be deleted from all computers once the research project outlined in Section 3 has been completed

Section 5: Register of use*

The Irish Longitudinal Study on Ageing (TILDA) and Growing Up in Ireland (GUI) only

ISSDA would like to facilitate researchers using other datasets to collaborate, where appropriate. If you consent to allowing your details to being published on our website, please tick one of the following options:

☐ Description of project

☑ No details

End User Licence

This Agreement is made between [Andrea Bowe] (the "End User") and the ISSDA in order to provide the "End User" with the right to use the data collections provided via the ISSDA according to the terms below:

terms of use

These terms of use cover the use of data files and all related material (hereafter "the data collections") supplied by ISSDA in accordance with the Data Request Form.

The End User undertakes and agrees*:*

1. To use and to make personal copies of any part of the data collections only for the purposes specified in the part on intended use of data. If it is wished to use the data for any other purpose enquiries should be made to the ISSDA.
2. That this Licence does not operate to transfer any interest in intellectual property from the data collection funders, the ISSDA, the original data creators or depositors, copyright or other right holders to the End User.
3. To ensure that strictly only the End User, and those persons referred to in Section 4, have access to the data provided under this agreement. The End User shall be responsible for controlling access to the data.
4. To ensure that the means of access to the data collections (such as passwords) are kept secure and not disclosed to a third party (excluding those listed in Section 5) except by special written permission or licence obtained from the ISSDA.
5. Not to use the data collections to attempt to obtain or derive information relating specifically to an identifiable individual or household.
6. To be aware at all times of the risk of inadvertently disclosing information, which might result in the identification of an individual. All use of the data and production of all analysis and output should be sensitive to this risk.
7. To undertake to abide by the conditions laid out in the Statistics Act, 1993, and, in particular, Section 34 thereof. Please see http://www.irishstatutebook.ie/1993/en/act/pub/0021/index.html
8. To acknowledge, in any work based in whole or part on resources provided by the ISSDA, the original data creators, depositors or copyright holders and the ISSDA, and to declare, in any such work, that those who carried out the original analysis and collection of the data bear no responsibility for the further analysis or interpretation of it.
9. To ensure that all such works acknowledge that copyright and all other intellectual property rights in the data and associated documentation are vested in the original data creators or depositors. Please see the ISSDA website (www.ucd.ie/issda) for appropriate wording.
10. To ensure that all such works acknowledge the ISSDA in the following way: "Accessed via the Irish Social Science Data Archive - www.ucd.ie/issda".

11. To reference the recommended bibliographic citation in any publication that employs resources provided by the ISSDA.
12. To send to the ISSDA citations of any publication based in whole or part on resources provided by the ISSDA for inclusion in a database of related publications.
13. That ISSDA may submit details of the End User's research, as outlined in Section 3, for statistical purposes to the original data creator or depositor.
14. To notify the ISSDA of any errors discovered in the data or accompanying documentation.
15. Any breach of any of the provisions of this Agreement will lead to immediate termination of the End User's access to all services provided by the ISSDA either permanently or temporarily, and may result in legal action being taken against the End User. The End User acknowledges that the data depositor will be notified in the event of a breach coming to the notice of the End User within 72 hours. Permission to use the data for the specified purpose may be withdrawn by the ISSDA at any time, without notice and without cause assigned, by written notice to the End User, signed by or on behalf of the Director of the ISSDA.

INDEMNITY

End User agrees to indemnify and shall keep indemnified each member of the ISSDA against any costs, actions, claims, demands, liabilities, expenses, damages or losses (including without limitation consequential losses and loss of profit, and all interest, penalties and legal and other professional costs and expenses) arising from or in connection with any third party claim made against any member of the ISSDA relating to End Users use of the data collections or any other activities in relation to the data collection where such use is in breach of this licence.

DISCLAIMERS

To the extent that applicable law permits:

a. The ISSDA bears no legal responsibility for the accuracy or comprehensiveness of the data supplied.

b. The ISSDA accepts no liability for, and the End User will not be entitled to claim against them in respect of, any direct, indirect, consequential or incidental damages or losses arising from use of

   the data collections, or from the unavailability of, or break in access to, the service, for whatever reason.

c. Whilst steps have been taken to ensure all licences, authorisation and permissions required for the granting of this Licence have been obtained, this may not have been possible in all cases, and no warranties or assurance are given in this regard. To the extent that additional licences, authorisations and permissions are required to use the data collections in accordance with this Licence, it is the End User's responsibility to obtain them.

I acknowledge that by signing my name below, I confirm that I have read, understood, and agree to abide by all of the terms and conditions in the ISSDA End User Licence (including the obligations imposed on the End User).

Signed _____	**Signed** 
Administrator	
Irish Social Science Data Archive	**END USER**
Date _____	**Date 21/08/20**
For office use only	

Please note that any reference to signature and date in this document can be read as meaning the typed name and date where such an application is forwarded electronically.

Definitions

"data collection funder" means the persons or organisations that funded the collection and/or creation of the data collections. The data collection funder for a particular data collection is identified in the metadata applicable to that data collection;

"original data creator or depositor" means the persons or organisations that originally collected, created or deposited the materials making up the data collections and/or who own the intellectual property rights in the data collections. The original data creator or depositor for a particular data collection is identified in the metadata applicable to that data collection;

"metadata" means any additional or bibliographic information about one or more of the data collections, as notified to the End User from time to time. Metadata may be supplied by electronic means

Appendix 2 Data access Growing Up in Ireland Infant Cohort (Chapter 4)

DATA REQUEST FORM FOR RESEARCH PURPOSES

Section 1: contact details [Please type]

First name *	Andrea	Last name *	Bowe
Address	INFANT Research Centre, University College Cork, Cork		
Country *	Ireland		
Institution/Company *	INFANT Research Centre UCC		
Email *	abowe@ucc.ie		
Phone	0876153190		

section 2: dataset requested* [Please type]

Please refer to the full list of datasets on the ISSDA website: www.ucd.ie/issda/data/

Please include Study Number (SN) where appropriate
1.11.1. GUI Infant Cohort Wave 1 Study number (SN): 0019-01 (Wave 1)

GUI Infant Cohort Wave 1 Study number (SN): 0019-02 (Wave 2)
GUI Infant Cohort Wave 1 Study number (SN): 0019-03 (Wave 3)
GUI Infant Cohort Wave 1 Study number (SN): 0019-05 (Wave 5)

Do you wish to be contacted if the dataset is updated or in the case of new releases of data? Please tick one only.
X Yes ☐No

Section 3: INTENDED USE OF DATA [Please type]

Subject area *	☐ Economics	☐Education	☐Engineering	☐Environment/ Agriculture
	X Health/ Well-being	☐Law	☐Policy	☐Social issues
	☐ Sport	Technology	☐Other [please specify]	

Please provide a short description (approx. 100 words) of your intended use of the dataset/s, including title and details of your intended research project*:

Title: The impact of below average cognitive ability at school entry on child and teacher-reported outcomes at age 9
Details: This study will examine the effect of having below average cognitive ability at age 5 on school outcomes at age 9 years. It will use the naming vocabulary and picture similarities subtests at age 5 as the primary predictor. Gender, school class, and primary caregiver education level will be the main covariates. The outcomes will be child-reported school outcomes at age 9, child self concept at age 9, teacher reported class engagement and emotional behavioural function at age 9.

Type of user *	☐Academic Staff	☐ Post Doc	☑PhD	☐Masters
	☐Undergraduate	☐Independent Researcher	☐Government/ Policy researcher	
	☐Other [please specify]			

Estimated end date for this use (maximum 5 years from date of	August 2024
Funder (where applicable)	Funded by the Irish Clinical Academic Training Programme

SECTION 4: PERSONS PERMITTED TO ACCESS THE DATA UNDER ARTICLE 3 OF TERMS OF USE*

The data may only be used for the purpose as outlined in Section 3. Data must be re-applied for in the case of new projects being undertaken.

Please indicate the number of users who will have access to and use of the data.

Andrea Bowe

- All users are bound by the same terms and conditions as the End User
- Users may not be given a copy of the data for use outside of the scope of the research project outlined in Section 3
- Users wishing to use the data outside of the scope of the research project outlined in Section 3 can apply directly to ISSDA for a copy of the data, subject to the usual terms and conditions
- The data must be deleted from all computers once the research project outlined in Section 3 has been completed

Section 5: Register of use*

The Irish Longitudinal Study on Ageing (TILDA) and Growing Up in Ireland (GUI) only

ISSDA would like to facilitate researchers using other datasets to collaborate, where appropriate. If you consent to allowing your details to being published on our website, please tick one of the following options:

☑ Description of project

☐ No details

End User Licence

This Agreement is made between [Andrea Bowe] (the "End User") and the ISSDA in order to provide the "End User" with the right to use the data collections provided via the ISSDA according to the terms below:

terms of use

These terms of use cover the use of data files and all related material (hereafter "the data collections") supplied by ISSDA in accordance with the Data Request Form.

The End User undertakes and agrees:

16. To use and to make personal copies of any part of the data collections only for the purposes specified in the part on intended use of data. If it is wished to use the data for any other purpose enquiries should be made to the ISSDA.
17. That this Licence does not operate to transfer any interest in intellectual property from the data collection funders, the ISSDA, the original data creators or depositors, copyright or other right holders to the End User.
18. To ensure that strictly only the End User, and those persons referred to in Section 4, have access to the data provided under this agreement. The End User shall be responsible for controlling access to the data.
19. To ensure that the means of access to the data collections (such as passwords) are kept secure and not disclosed to a third party (excluding those listed in Section 5) except by special written permission or licence obtained from the ISSDA.
20. Not to use the data collections to attempt to obtain or derive information relating specifically to an identifiable individual or household.
21. To be aware at all times of the risk of inadvertently disclosing information, which might result in the identification of an individual. All use of the data and production of all analysis and output should be sensitive to this risk.
22. To undertake to abide by the conditions laid out in the Statistics Act, 1993, and, in particular, Section 34 thereof. Please see http://www.irishstatutebook.ie/1993/en/act/pub/0021/index.html
23. To acknowledge, in any work based in whole or part on resources provided by the ISSDA, the original data creators, depositors or copyright holders and the ISSDA, and to declare, in any such work, that those who carried out the original analysis and collection of the data bear no responsibility for the further analysis or interpretation of it.
24. To ensure that all such works acknowledge that copyright and all other intellectual property rights in the data and associated documentation are vested in the original data creators or depositors. Please see the ISSDA website (www.ucd.ie/issda) for appropriate wording.
25. To ensure that all such works acknowledge the ISSDA in the following way: "Accessed via the Irish Social Science Data Archive - www.ucd.ie/issda".

26. To reference the recommended bibliographic citation in any publication that employs resources provided by the ISSDA.
27. To send to the ISSDA citations of any publication based in whole or part on resources provided by the ISSDA for inclusion in a database of related publications.
28. That ISSDA may submit details of the End User's research, as outlined in Section 3, for statistical purposes to the original data creator or depositor.
29. To notify the ISSDA of any errors discovered in the data or accompanying documentation.
30. Any breach of any of the provisions of this Agreement will lead to immediate termination of the End User's access to all services provided by the ISSDA either permanently or temporarily, and may result in legal action being taken against the End User. The End User acknowledges that the data depositor will be notified in the event of a breach coming to the notice of the End User within 72 hours. Permission to use the data for the specified purpose may be withdrawn by the ISSDA at any time, without notice and without cause assigned, by written notice to the End User, signed by or on behalf of the Director of the ISSDA.

INDEMNITY

End User agrees to indemnify and shall keep indemnified each member of the ISSDA against any costs, actions, claims, demands, liabilities, expenses, damages or losses (including without limitation consequential losses and loss of profit, and all interest, penalties and legal and other professional costs and expenses) arising from or in connection with any third party claim made against any member of the ISSDA relating to End Users use of the data collections or any other activities in relation to the data collection where such use is in breach of this licence.

DISCLAIMERS

To the extent that applicable law permits:

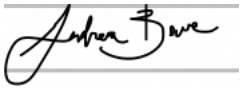
d. The ISSDA bears no legal responsibility for the accuracy or comprehensiveness of the data supplied.

e. The ISSDA accepts no liability for, and the End User will not be entitled to claim against them in respect of, any direct, indirect, consequential or incidental damages or losses arising from use of

the data collections, or from the unavailability of, or break in access to, the service, for whatever reason.

f. Whilst steps have been taken to ensure all licences, authorisation and permissions required for the granting of this Licence have been obtained, this may not have been possible in all cases, and no warranties or assurance are given in this regard. To the extent that additional licences, authorisations and permissions are required to use the data collections in accordance with this Licence, it is the End User's responsibility to obtain them.

I acknowledge that by signing my name below, I confirm that I have read, understood, and agree to abide by all of the terms and conditions in the ISSDA End User Licence (including the obligations imposed on the End User).

Signed _____ _____	**Signed** 
Administrator	
Irish Social Science Data Archive	**END USER**
Date _____	**Date 06/08/2023**

For office use only

Please note that any reference to signature and date in this document can be read as meaning the typed name and date where such an application is forwarded electronically.

Definitions

"data collection funder" means the persons or organisations that funded the collection and/or creation of the data collections. The data collection funder for a particular data collection is identified in the metadata applicable to that data collection;

"original data creator or depositor" means the persons or organisations that originally collected, created or deposited the materials making up the data collections and/or who own the intellectual property rights in the data collections. The original data creator or depositor for a particular data collection is identified in the metadata applicable to that data collection;

"metadata" means any additional or bibliographic information about one or more of the data collections, as notified to the End User from time to time. Metadata may be supplied by electronic means

Appendix 3 Data access Growing Up in Ireland Infant Cohort (Chapter 7)

DATA REQUEST FORM FOR RESEARCH PURPOSES

Section 1: contact details

[Please type]

First name *	Andrea	Last name *	Bowe
Address	INFANT Research Centre, Cork University Hospital, Cork, Ireland		
Country *	Ireland		
Institution/Company *	University College Cork		
Email *	Andrea.bowe@ucdconnect.ie		
Phone	0876153190		

Section 2: dataset requested* [Please type]

Please refer to the full list of datasets on the ISSDA website: www.ucd.ie/issda/data/

Please include Study Number (SN) where appropriate

Growing Up in Ireland Infant Cohort Waves 1-5
Study number (SN): 0019-01 (Wave 1)

Study number (SN): 0019-02 (Wave 2)

Study number (SN): 0019-03 (Wave 3)

Study number (SN): 0019-04 (Wave 4)

For office use only:
Request ID
Date Received
Date Sent

Section 3: INTENDED USE OF DATA [Please type]

Subject area *	☐ Economics	☐Education	☐Engineering	☐Environment/ Agriculture
	☑ Health/ Well-being	☐Law	☐Policy	☐Social issues
	☐ Sport	Technology	☐Other [please specify] _____	

Please provide a short description (approx. 100 words) of your intended use of the dataset/s, including title and details of your intended research project*:

Title: Perinatal Predictors of Cognitive Outcomes in Childhood
Summary: This project will examine whether machine learning can be applied to perinatal data to predict cognitive outcomes in childhood. Many children will not achieve their cognitive potential due to adverse socioeconomic, environmental, or perinatal conditions. The early identification of children who may benefit from early intervention is a public health priority. This research will examine whether novel statistical methods, such as machine learning, can be applied to large epidemiological datasets for predicting children at risk of poor cognitive outcomes in childhood. The project will use data from the infant cohort of the growing up in Ireland study.

Type of user *	☐Academic Staff	☐ Post Doc	☑PhD	☐Masters
	☐Undergraduate	☐Independent Researcher	☐Government/ Policy researcher	
	☐Other [please specify] _____			

Estimated end date for this use (maximum 5 years from date of	July 2024
Funder (where applicable)	Irish Clinical Academic Training Programme

SECTION 4: PERSONS PERMITTED TO ACCESS THE DATA UNDER ARTICLE 3 OF TERMS OF USE*

The data may only be used for the purpose as outlined in Section 3. Data must be re-applied for in the case of new projects being undertaken.

Please indicate the number of users who will have access to and use of the data.

3 – Professor Gordon Lightbody UCC, Professor Anthony Staines DCU, Dr Andrea Bowe

- All users are bound by the same terms and conditions as the End User
- Users may not be given a copy of the data for use outside of the scope of the research project outlined in Section 3
- Users wishing to use the data outside of the scope of the research project outlined in Section 3 can apply directly to ISSDA for a copy of the data, subject to the usual terms and conditions
- The data must be deleted from all computers once the research project outlined in Section 3 has been completed

Section 5: Register of use*

The Irish Longitudinal Study on Ageing (TILDA) and Growing Up in Ireland (GUI) only

ISSDA would like to facilitate researchers using other datasets to collaborate, where appropriate. If you consent to allowing your details to being published on our website, please tick one of the following options:

☐ Description of project

☑ No details

End User Licence

This Agreement is made between [Andrea Bowe] (the "End User") and the ISSDA in order to provide the "End User" with the right to use the data collections provided via the ISSDA according to the terms below:

terms of use

These terms of use cover the use of data files and all related material (hereafter "the data collections") supplied by ISSDA in accordance with the Data Request Form.

The End User undertakes and agrees:

31. To use and to make personal copies of any part of the data collections only for the purposes specified in the part on intended use of data. If it is wished to use the data for any other purpose enquiries should be made to the ISSDA.
32. That this Licence does not operate to transfer any interest in intellectual property from the data collection funders, the ISSDA, the original data creators or depositors, copyright or other right holders to the End User.
33. To ensure that strictly only the End User, and those persons referred to in Section 4, have access to the data provided under this agreement. The End User shall be responsible for controlling access to the data.
34. To ensure that the means of access to the data collections (such as passwords) are kept secure and not disclosed to a third party (excluding those listed in Section 5) except by special written permission or licence obtained from the ISSDA.
35. Not to use the data collections to attempt to obtain or derive information relating specifically to an identifiable individual or household.
36. To be aware at all times of the risk of inadvertently disclosing information, which might result in the identification of an individual. All use of the data and production of all analysis and output should be sensitive to this risk.
37. To undertake to abide by the conditions laid out in the Statistics Act, 1993, and, in particular, Section 34 thereof. Please see http://www.irishstatutebook.ie/1993/en/act/pub/0021/index.html
38. To acknowledge, in any work based in whole or part on resources provided by the ISSDA, the original data creators, depositors or copyright holders and the ISSDA, and to declare, in any such work, that those who carried out the original analysis and collection of the data bear no responsibility for the further analysis or interpretation of it.
39. To ensure that all such works acknowledge that copyright and all other intellectual property rights in the data and associated documentation are vested in the original data creators or depositors. Please see the ISSDA website (www.ucd.ie/issda) for appropriate wording.
40. To ensure that all such works acknowledge the ISSDA in the following way: "Accessed via the Irish Social Science Data Archive - www.ucd.ie/issda".

41. To reference the recommended bibliographic citation in any publication that employs resources provided by the ISSDA.
42. To send to the ISSDA citations of any publication based in whole or part on resources provided by the ISSDA for inclusion in a database of related publications.
43. That ISSDA may submit details of the End User's research, as outlined in Section 3, for statistical purposes to the original data creator or depositor.
44. To notify the ISSDA of any errors discovered in the data or accompanying documentation.
45. Any breach of any of the provisions of this Agreement will lead to immediate termination of the End User's access to all services provided by the ISSDA either permanently or temporarily, and may result in legal action being taken against the End User. The End User acknowledges that the data depositor will be notified in the event of a breach coming to the notice of the End User within 72 hours. Permission to use the data for the specified purpose may be withdrawn by the ISSDA at any time, without notice and without cause assigned, by written notice to the End User, signed by or on behalf of the Director of the ISSDA.

INDEMNITY

End User agrees to indemnify and shall keep indemnified each member of the ISSDA against any costs, actions, claims, demands, liabilities, expenses, damages or losses (including without limitation consequential losses and loss of profit, and all interest, penalties and legal and other professional costs and expenses) arising from or in connection with any third party claim made against any member of the ISSDA relating to End Users use of the data collections or any other activities in relation to the data collection where such use is in breach of this licence.

DISCLAIMERS

To the extent that applicable law permits:

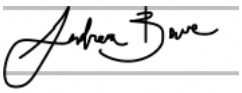
g. The ISSDA bears no legal responsibility for the accuracy or comprehensiveness of the data supplied.

h. The ISSDA accepts no liability for, and the End User will not be entitled to claim against them in respect of, any direct, indirect, consequential or incidental damages or losses arising from use of

the data collections, or from the unavailability of, or break in access to, the service, for whatever reason.

i. Whilst steps have been taken to ensure all licences, authorisation and permissions required for the granting of this Licence have been obtained, this may not have been possible in all cases, and no warranties or assurance are given in this regard. To the extent that additional licences, authorisations and permissions are required to use the data collections in accordance with this Licence, it is the End User's responsibility to obtain them.

I acknowledge that by signing my name below, I confirm that I have read, understood, and agree to abide by all of the terms and conditions in the ISSDA End User Licence (including the obligations imposed on the End User).

Signed _____	**Signed** 
Administrator	
Irish Social Science Data Archive	**END USER**
Date _____	**Date 11/01/2022**

For office use only

Please note that any reference to signature and date in this document can be read as meaning the typed name and date where such an application is forwarded electronically.

Definitions

"data collection funder" means the persons or organisations that funded the collection and/or creation of the data collections. The data collection funder for a particular data collection is identified in the metadata applicable to that data collection;

"original data creator or depositor" means the persons or organisations that originally collected, created or deposited the materials making up the data collections and/or who own the intellectual property rights in the data collections. The original data creator or depositor for a particular data collection is identified in the metadata applicable to that data collection;

"metadata" means any additional or bibliographic information about one or more of the data collections, as notified to the End User from time to time. Metadata may be supplied by electronic means

Appendix 4 Data access BASELINE Birth Cohort

RESEARCH APPLICATION FORM
BASELINE RESEARCH PROJECT

Project Title: Predictors of Borderline Intellectual Functioning and Cognitive Outcomes in Childhood

Project PI: Dr. Andrea Bowe supervised by Prof. Deirdre Murray Principal Supervisor

Co-authors (in order): Prof. Anthony Staines, Dr. Fergus McCarthy, Dr. Gordon Lightbody

Date of Submission: 2nd December 2020

Funding required to do project ✓ YES NO (tick as appropriate)

If yes, please complete Table below

Funding Status	**Yes**	**No**	**N.A.**	**Funder/Amount funded/Date**
Applied for funding				
Planning to apply for funding				
Funder/Potential Funder				
Date expect to hear outcome of funding application				
Date heard outcome of application				
Funded	✓			ICAT Programme funding
Amount funded				

Ethics
Ethical Approval for the BASELINE study was granted by the Cork Research Ethics Committee. Projects examining existing data and biobanked samples should be covered by this consent, but there maybe occasion when additional ethical approval is required.
Ethics Approval required: no yes ✓

Ethical approval was sought and granted from the Clinical Research Ethics committee of the Cork Teaching Hospitals for Dr Andrea Bowe to be added as co-investigator to the BASELINE study.

Ethics Approval completed: no ☐ yes ✓

Collaboration with Commercial Institution or Academic Inst. outside UCC and BASELINE research committee

NO ☐ YES ✓

If yes, please complete this table:

Commercial Details	**Complete details**
Name of Company/ Academic Institution	Dublin City University
Type of Collaboration	Co-supervisor Prof Staines is Professor of Health Systems at DCU

IP ownership solved	Yes, study is part of PhD which will be submitted through UCC
If IP ownership solved, please summarise solution	
If IP ownership not solved, summarise issues	

Note: If more than 1 institution involved please copy the rows in table and complete for each institution

Research Proposal (max 1 page)

Aims/Hypotheses:
Aim: The overall aim of the study is to examine the antenatal and perinatal factors associated with cognitive outcomes in childhood and to determine if a risk prediction algorithm can be created to accurately predict those at greatest risk of cognitive difficult in childhood.

Objectives:
1. To examine the utility of the routinely administered Ages and Stages Questionnaire for detecting mild cognitive difficulty in childhood
2. To examine the association between IQ and emotional and behavioural problems in childhood
3. To examine the association between antenatal and perinatal factors associated with IQ and in particular with borderline intellectual function.

Background: (1 paragraph)
Borderline intellectual function (BIF) is a heterogenous clinical entity, previously defined by an IQ of 70-85, which is both under recognised and under studied (1). Without diagnosis and recognition, individuals with BIF may struggle to meet expected standards with respect to education and social function (1). BIF has been shown to be associated with higher rates of poor social functioning, mental health problems and poor occupational and educational outcomes (2, 3). Early intervention and early exposure to school related materials has been shown to improve IQ in childhood (4). In order to intervene early, we must be able to recognise those who are most at risk. The overall aim of this study is to investigate the potential for an algorithm based on ante- and peri-natal factors to predict those at risk of cognitive difficulty in childhood. The study will also examine the utility of the routinely used Ages and Stages Questionnaire for predicting cognitive outcome in childhood and for detecting those with BIF. The study will examine the association between cognitive function and emotional and behavioural problems in childhood.

1. Karande S, Kanchan S, Kulkarni M. Clinical and psychoeducational profile of children with borderline intellectual functioning. The Indian Journal of Pediatrics. 2008;75(8):795.
2. Baglio F, Cabinio M, Ricci C, Baglio G, Lipari S, Griffanti L, et al. Abnormal development of sensory-motor, visual temporal and parahippocampal cortex in children with learning disabilities and borderline intellectual functioning. Frontiers in human neuroscience. 2014;8:806-.

3. Chen C-Y, Lawlor JP, Duggan AK, Hardy JB, Eaton WW. Mild cognitive impairment in early life and mental health problems in adulthood. American journal of public health. 2006;96(10):1772-8.
4. Gorey KM. Early childhood education: A meta-analytic affirmation of the short- and long-term benefits of educational opportunity. School Psychology Quarterly. 2001;16(1):9-30.

Proposed methods

Type(s) and Numbers of Participants:
Overall Study: All participants with IQ measured by the Kauffman Brief Intelligence Test at age 5

Proposed methods: This study will be a secondary analysis of data contained in the Baseline dataset. R statistical software and R studio will be used throughout project.

Data required from BASELINE database:

1. Paediatric data
List summary of information and timepoints (eg feeding method at 2 months and 6 months), body composition data at birth.

Description	Form	Data code
Infant feeding at 6 months	17	f17_6mo_Current_infant_feeding_type
Infant feeding at 12 months	30	f30_12mo_Current_infant_feeding_type
Age at 5 year assessment	59	f59_yr5_Age_assessment_mo
Preschool attendance	67	f67_yr5_Mont_preschool_attendance
Age preschool	67	f67_yr5_Age_preschool_years
Age preschool	67	f67_yr5_Age_preschool_months
School	67	f67_yr5_School_attendance
Age school	67	f67_yr5_Age_school_years
Learning disabilities	68	f68_yr5_Learning_disabilities
Speech problems	68	f68_yr5_Speech_problems
Kaufmann in age months	70	f70_yr5_Kaufmann_age_months
Verbal in knowledge value	70	f70_yr5_Verbal_knowledge_value
Riddles value	70	f70_yr5_Riddles_value
Matrices value	70	f70_yr5_Matrices_value
ASQ3 complete	70	f70_yr5_ASQ3_complete
ASQ3 date	70	f70_yr5_ASQ3_date
ASQ3 location	70	f70_yr5_ASQ3_location
ASQ3 in age months	70	f70_yr5_ASQ3_age_months
ASQ3 in age days	70	f70_yr5_ASQ3_age_days
ASQ3 in age admin	70	f70_yr5_ASQ3_age_admin
ASQ3 in communication score	70	f70_yr5_ASQ3_communication_score
ASQ3 in grossmotor score	70	f70_yr5_ASQ3_grossmotor_score
ASQ3 in finemotor score	70	f70_yr5_ASQ3_finemotor_score

ASQ3 in problemsolving score	70	f70_yr5_ASQ3_problemsolving_score
ASQ3 in socialskills score	70	f70_yr5_ASQ3_socialskills_score
CBCL complete	70	f70_yr5_CBCL_complete
CBCL date	70	f70_yr5_CBCL_date
CBCL location	70	f70_yr5_CBCL_location
CBCL in age months	70	f70_yr5_CBCL_age_months
CBCL in age days	70	f70_yr5_CBCL_age_days
CBCL in informant relationship	70	f70_yr5_CBCL_informant_relationship
CBCL in emotionally reactive	70	f70_yr5_CBCL_emotionally_reactive
CBCL in anx depressed	70	f70_yr5_CBCL_anx_depressed
CBCL in somatic complaints	70	f70_yr5_CBCL_somatic_complaints
CBCL withdrawn	70	f70_yr5_CBCL_withdrawn
CBCL in sleep problems	70	f70_yr5_CBCL_sleep_problems
CBCL in attention problems	70	f70_yr5_CBCL_attention_problems
CBCL in aggressive behaviour	70	f70_yr5_CBCL_aggressive_behaviour
CBCL internal	70	f70_yr5_CBCL_internal
CBCL external	70	f70_yr5_CBCL_external
CBCL total	70	f70_yr5_CBCL_total
CBCL in affective problems	70	f70_yr5_CBCL_affective_problems
CBCL in anxiety problems	70	f70_yr5_CBCL_anxiety_problems
CBCL in pervasive problems	70	f70_yr5_CBCL_pervasive_problems
CBCL in attention deficit	70	f70_yr5_CBCL_attention_deficit
CBCL in defiant problems	70	f70_yr5_CBCL_defiant_problems
CBCL questionnaire	70	f70_yr5_CBCL_questionnaire
Developmental in concerns teachers	70	f70_yr5_Developmental_concerns_teachers
Developmental in concerns specify	70	f70_yr5_Developmental_concerns_specify
Early in intervention assessment	70	f70_yr5_Early_intervention_assessment
Speech in lang therapist	70	f70_yr5_Speech_lang_therapist
Physio assessment	70	f70_yr5_Physio_assessment
Occ therapist	70	f70_yr5_Occ_therapist
Language delay	70	f70_yr5_Language_delay
Hyperactivity ADD	70	f70_yr5_Hyperactivity_ADD
Hearing difficulties	70	f70_yr5_Hearing_difficulties
Visual difficulties	70	f70_yr5_Visual_difficulties
Austistic spectrum	70	f70_yr5_Austistic_spectrum
Other in difficulties disorder	70	f70_yr5_Other_difficulties_disorder
Other in difficulties specify	70	f70_yr5_Other_difficulties_specify

2. Maternal data
List summary of information and timepoints

Mother DOB	1	DOB
Placental weight	2	f2_Birth_Placental_weight_g
Cord in arterial pH	2	f2_Birth_Cord_arterial_pH
Base deficit cord	2	f2_Birth_Base_deficit
Base excess cord	2	f2_Birth_Base_excess
Evidence infection	2	f2_Birth_Evidence_infection
Evidence infection 48h labour	2	f2_Birth_Evidence_infection_48h_afterbirth
Confirmed infection	2	f2_Birth_Confirmed_infection
Confimed infection 48h labour	2	f2_Birth_Confimed_infection_48h_afterbirth
Alcohol pregnancy any	9	f9_mo2_Alcohol_pregnancy
Alcohol in pregnancy comparison	9	f9_mo2_Alcohol_pregnancy_comparison
Maternal smoking	11	f11_mo2_Current_smoker
Maternal smoking	11	f11_mo2_Years_smoking
Maternal smoking	11	f11_mo2_No_cigarettes_perday
Maternal smoking	11	f11_mo2_Ex_smoker_no_cigarettes_day
Maternal smoking	11	f11_mo2_Stop_smoking_pregnancy
Maternal smoking	11	f11_mo2_No_cigarettes_prepregnancy
Maternal smoking	11	f11_mo2_No_cigarettes_during_pregnancy
No adults house labour	11	f11_mo2_No_adults_house_pregnancy
No children house labour	11	f11_mo2_No_children_house_pregnancy
No adults house at labour	12	f12_mo2_No_adults_house_at_present
No children house at labour	12	f12_mo2_No_children_house_at_present
Ever history of depression	69	f69_yr5_Mat_depression
Ever history of anxiety	69	f69_yr5_Mat_anxiety_disorder
Maternal depression	69	f69_yr5_Mat_depression
Maternal anxiety	69	f69_yr5_Mat_anxiety_disorder
Paternal depression	69	f69_yr5_Pat_depression
Paternal anxiety	69	f69_yr5_Pat_anxiety_disorder

3. Biosamples: Type(s) and volume of specimens required:

Specimen Type (e.g., serum/EDTA plasma/DNA)	Maternal (gestation) or umbilical cord blood sample	Number of Aliquots per maternal/fetal dyad*	Total of Aliquots in Expt

*All blood specimens 250ul aliquots, urine 0.9ml aliquot

Other Comments:

Data sought from SCOPE study

f1_partcpt_dob f1_mw_contact_date f6_partcpt_gest_at_del, f1_age, f7_hx_depression f10c_any_folate_prepreg f10c_any_folate_1st_trim f10c_any_folate_1st_vst f10c_cig_prepreg_3m_gp f10_cig_prepreg_3m f10_cig_1st_trim f10c_cig_1st_trim_gpf10_cig_wks_pre_1st_vst f10c_cig_wks_pre_1st_vst_gp f10_cig_1st_vst f10c_cig_1st_vst_gp f10c_smoking_status_1st_vst_gps f10c_any_cig_1st_vst f10c_cig_1st_trim f10c_cig_tot_number_pre_1st_vst f10c_cig_1st_trim f10_alcohol_prepreg_3m f10_alcohol_1st_trim f10c_alcohol_1st_trim f10_othdrugs_prepreg_3m f10_othdrugs_1st_trim f11_wgt f11_bmi f11c_bmi_gp f11_hgt f19c_any_folate_2nd_vst f13c_depress f13c_depress_cat f19_cig_2nd_vst f19c_cig_2nd_vst_gp f19c_cig_2nd_vst f19_cig_2nd_vst f19_alcohol_2nd_vst f19c_alcohol_2nd_vst f19_othdrugs_2nd_vst f19c_any_folate_2nd_vst f22c_depress f22c_depress_cat f5_tertiary_edu f5_mat_occupation, f5_major_grp f5c_main_ethnicity_reduced f5c_main_ethnicity_reduced_gp f5_ethnicity_custcentiles f5_partnr_migration f5c_edu_le_12yr f5_oth_tertiary_edu f5c_any_tertiary_edu f5c_edu_gp f13c_stai f13c_stai_gt_90c f24c_gdm_hapo f24_last_Wgt f24_gest_at_last_wgt f34_pet f34_sptb f34c_gest_diag_pet f34c_gest_dev_gh

Planned commencement date: Study objectives are in progress. The variables requested above are additional data.

Expected finish date: The overall project is part of a PhD which will be completed over a 4year period as part of the ICAT programme. Expected finish date July 2024.

Comments from Research committee: discussion date:

Any changes suggested:

Is SCOPE RAF required:	**YES**		**NO**	
Is statistical input required:	**YES**	✓	**NO**	☐
Level of statistical input:	**Planning phase**			☐
	Throughout project			✓
	Final review prior to submission			☐

If input required throughout project, statistician as co-author ✓

Clinical Research Ethics Committee Of The Cork Teaching Hospitals

AMENDMENT APPLICATION FORM (D39)

Please note that submission of an incomplete amendment application pack will result in the application being returned to the study Chief Investigator without review.

The study title on this amendment application form must be the registered title of the study as per CREC approval letter.

Study documents that have changed due to the amendment must clearly show the changes (highlighted or tracked) and must contain a new version number and date. If the revised documents do not contain this information the application will be returned.

Applications cannot be handwritten.

Chief Investigator Details

Name of Chief Investigator:

Professor Deirdre Murray

Hospital and Department:

Department of Paediatrics, Cork University Hospital
Department of Paediatrics and Child Health, University College Cork

Study Title (title must be the exactly the same as the original registered study title):

Parent Study: Cork Baseline Birth Cohort

Current Study: The current study is based on secondary data analysis of data contained within the Baseline Birth Cohort and aligns with original objectives of the Baseline Birth Cohort study.

Title: Examining the antenatal and perinatal factors associated with cognitive outcomes at Age 5
Aim: To examine the antenatal and perinatal factors associated with cognitive outcomes at age 5
Objectives: 1. To examine the prevalence of intellectual disability, borderline intellectual functioning, and normal intellectual functioning
2. To examine association between the results of developmental assessment (Ages and Stages Questionnaire) at age 24 months and IQ at age 5

3. To assess the association between cognitive outcome at age 5 as measured by IQ and emotional/behavioural problems at age 5 as measured by the Child Behaviour Checklist
4. To examine the antenatal and perinatal factors associated with cognitive outcome and determine whether a risk prediction algorithm using these factors can be created to accurately predict those at risk of adverse cognitive outcomes

---

The following changes are proposed for the study:

Answer Yes or No to the following:

Chief Investigator: No

Co-investigator(s): Yes. Dr Andrea Bowe to be added as co-investigator. Dr Bowe is a Public Health Clinical Fellow who has been awarded an ICAT fellowship to focus on this area of research. She is based in UCD, but will be registered with UCC for her PhD.

Dosage: No

Treatment Procedures: No

Drug/Device: No

Study Population: No

Number of Subjects: No

Risks: No

Is a revised study protocol necessary as a result of this amendment?
No

Is a revised Participant Information Leaflet and Consent Form necessary as a result of this amendment?
No

Please list the specific changes from the previously approved protocol and provide sufficient rationale for each change to allow the committee to make a decision.

Chief Investigator Signature: 

Date: _____31/07/2020_

(This form must bear the original signature of the Chief Investigator)

Post Application to: Lancaster Hall, 6 Little Hanover Street, Cork

Appendix 5 Ethical approval from the Cork Research Ethics Committee for Cork Teaching Hospitals

COISTE EITICE UM THAIGHDE CLINICIÚIL
Clinical Research Ethics Committee of the Cork Teaching Hospitals

Tel: +353-21-4901901
Email: crec@ucc.ie

University College Cork
Lancaster Hall
6 Little Hanover Street
Cork
Ireland

Our ref: ECM 3 (ggg) 11/08/2020

5th August 2020

Professor Deirdre Murray
Department of Paediatrics
Cork University Hospital
Wilton
Cork

Re: Baseline study. Babies after SCOPE; Evaluating the longitudinal impact using neurological and nutritional endpoints

Dear Professor Murray

- Cover Letter dated 4th August 2020
- Addition of Dr Andrea Bowe, Public Health Clinical Fellow as a co-investigator in the study.

Yours sincerely



Professor David Kerins
Chairman
Clinical Research Ethics Committee
of the Cork Teaching Hospitals

Appendix 6 Ethics application to the Swedish Ethical Review Board

Prediktion av senare kognitivt utfall hos nyfödda barn som intensivvårdats för svåra komplikationer genom att tillämpa maskininlärning på populationsbaserade registerdata.

2022-01351-01
Grundansökan
Grundansökan
Avvakta avgift - inkommen ansökan

Mikael Norman

1.2 Ansvarig huvudman för forskningen (forskningshuvudman)

Region Stockholm (232100-0016)

1.3 Behörig företrädare för forskningshuvudman

Lars Navér, docent, överläkare, verksamhetschef, Karolinska universitetssjukhuset

1.3.1 Behörig företrädare – titel som innebär ett verksamhetsansvar

Verksamhetschef

1.4 Har projektet fler forskningshuvudmän?

Nej

1.5 Hemvist för forskningen

Medicinsk Enhet Neonatologi, Karolinska universitetssjukhuset och Institutionen för klinisk vetenskap, intervention och teknik (CLINTEC), enheten för pediatrik, Karolinska Institutet.

1.6 Huvudansvarig forskare för projektet (kontaktperson)

Mikael Norman

1.6.1 Institution/hemvist som huvudansvarig forskare är verksam vid

Medicinsk Enhet Neonatologi, Karolinska universitetssjukhuset och Institutionen för klinisk vetenskap, intervention och teknik (CLINTEC), enheten för pediatrik, Karolinska Institutet.

1.7 Är den huvudansvariga forskaren disputerad?

Ja

Frågor för avgiftskategori

1.9 Hur många forskningshuvudmän kommer att ingå i forskningsprojektet?

En

1.10 Avser forskningen klinisk läkemedelsprövning?

Nej

1.11 Ska endast befintliga personuppgifter behandlas i projektet?

Ja

2.1 Avser ansökan forskning som inbegriper äggdonation?

Nej

2.2 Avser ansökan forskning med läkemedel för genterapi eller somatisk cellterapi eller läkemedel som innehåller genetiskt modifierade organismer?

Nej

2.3 Avser ansökan forskning med xenogen cellterapi?

Nej

2.4 Kommer joniserande strålning ingå i forskningsprojektet?

Nej

2.5 Kommer biologiskt material från människor att nyinsamlas för projektet?

Nej

2.6 Planerar projektet att använda biologiskt material från människor från en eller flera befintliga provsamlingar?

Nej

2.7 Avser forskningen klinisk prövning av medicinteknisk produkt?

Nej

2.8 Gör en egen bedömning och ange på vilka punkter nedan som forskningen omfattas av 3-4 §§ etikprövningslagen. Observera att myndigheten kan komma att göra en annan bedömning.

✓ 3 § 1 Forskningen kommer att samla in känsliga personuppgifter.

2.8.1 [Om 3 § 1] Gör en egen bedömning och ange vilken typ av känsliga personuppgifter som kommer att behandlas i projektet. Observera att myndigheten kan komma att göra en annan bedömning.

✓ hälsa

2.9 Önskas ett rådgivande yttrande?

Ja

2.10 Söker projektet förtur med motivering att projektet potential för nytta vad gäller behandling och förebyggande av COVID-19?

Nej

3.1 Skriv en populärvetenskaplig sammanfattning av forskningsprojektet.

Tidigt i livet utvecklas hjärnan snabbt. Denna tidiga utveckling kännetecknas av både sårbarhet och möjligheter. Idag saknas fortfarande detaljerad kunskap om hur olika faktorer/exponeringar och genetik samverkar när det gäller påverkan på hjärnans tidiga utveckling och uppkomst av kognitiva funktionshinder. Exempelvis uppvisar tidig födda barn nedsatt kognition (logisk tankeförmåga/IQ) senare livet även utan uppenbara tecken på hjärnskada i nyföddhetsperioden (såsom syrebrist eller blödning i hjärnan). Orsakerna är än så länge i stora delar dåligt kända. Däremot vet vi att nedsatt kognition redan vid 2 års ålder sannolikt blir bestående och prognostiserar på gruppnivå skolproblem och försämrad socioekonomi och hälsa i vuxenlivet.

Konsensus råder om att åtgärder insatta tidigt - särskilt under barnets första levnadsmånader och år - har positiva effekter på hjärnas utveckling och senare kognitiv förmåga hos barn i farozonen. Problemet har varit att identifiera de barn som är mest sårbara och att angripa rätt orsaker för att förebygga nedsatt kognition. Lösningen på dessa problemställningar är komplex och innefattar analyser av hur biologiska, miljörelaterade, sociala och genetiska faktorer interagerar. Detta har tidigare varit mycket svårt att utföra med traditionella statistiska metoder.

Detta projekt avser att bearbeta redan insamlade registerdata med nya avancerade statistiska metoder, så kallad maskininlärning eller artificiell intelligens. Målsättningen är att skapa ny kunskap och bättre förståelse för olika riskfaktorers betydelse tidigt i livet och tillhandahålla nya kliniskt tillgängliga verktyg för prognostisering av neurokognitiva långtidsutfall hos barn som intensivvårdats i nyföddhetsperioden.

3.2 Vad är det vetenskapliga syftet med projektet?

Det övergripande syftet är att förbättra neonatalvårdade barns kognition och hälsa på kort och lång sikt. Genom avancerad databearbetning av ett större antal riskfaktorer än vad som tidigare varit möjligt att göra, och tillskapande av nya prognosverktyg kan vården av dessa barn förbättras.

Det övergripande forskningsproblem som projektet avser bearbeta gäller vilka prediktorer finns för väsentligt nedsatt kognition hos barn som intensivvårdats i nyföddhetsperioden? Projektet avgränsas av de variabler och antal barn som finns insamlade i befintliga register.

3.3 Vilka är de vetenskapliga frågeställningarna?

Vilka riskfaktorer har störst betydelse och vilka kan användas för prediktion av framtida kognition bland barn som intensivvårdats i nyföddhetsperioden?
Vår övergripande hypotes är att insamlade registerdata kan användas för att med avancerad statistisk metodik (artificiell intelligens/maskininlärning) förbättra prognostisering av bestående kognitionsnedsättning hos sjuka nyfödda barn.

4.1 Redogör för metod inkl. proceduren, tekniken eller behandlingen.

Studiedesign: nationell, populationsbaserad kohortstudie baserad på prospektivt definierade och insamlade registerdata.

Datakälla: Neonatalregistret SNQ (Svenskt Neonatal Kvalitetsregister). SNQ är nationellt heltäckande sedan 2011 och uppgifter i registret förs in prospektivt.

Inklusions- och exklusionskriterier: Sedan 2016 har barn som exponerats för svåra komplikationer i nyföddhetsperioden följts upp med avseende på utveckling och kognition vid 2 års ålder, och resultaten av dessa uppföljningar har rapporterats till SNQ. Studiepopulationen i den här studien kommer att utgöras av alla barn som fullgjort formell testning vid 2 års ålder med testverktyget "Bayley Scales of Infant Development (BSID), version III".
En subgruppsanalys kommer att innefatta tidigt födda barn (födda före 37:e graviditetsveckan).

Riskfaktorer: totalt kommer ett 60-tal variabler inkluderas i maskininlärningsanalyserna. Gemensamt för dessa är att de ska vara tillgängliga vid födelsen för att prediktionsverktyget ska kunna användas så tidigt som möjligt under vårdförloppet och bidra med underlag för tidig intervention.

Utfallsmått: primärt utfall är kognitionspoäng (Barn-IQ) <85 (mer än en standarddeviation under medel) vid 2 års ålder. Sekundära utfall innefattar samma avgränsning för språk/kommunikation, motorik och ordförståelse vid 2 års ålder.

För att besvara forskningsfrågan om vilka riskfaktorer som har störst betydelse och som kan användas för prediktion av långsiktigt nedsatt kognition hos barn som intensivvårdats i nyföddhetsperioden kommer data att analyseras i olika steg: först i en s.k. tränings- eller modellutvecklingskohort, därefter sker en intern validering i en testkohort, för att slutligen mynna ut i en algoritm för ett prognosverktyg som ska kunna användas för undervisning och rådgivning, och som underlag för vårdåtgärder på gruppnivå.

Bortfall: Barn som ingår i riskgrupp efter neonatalvård men som inte fullgjort så kallad "riskbarnsuppföljning efter komplicerad neonatalvård", eller som inte genomgått formell testning vid 2 års ålder med testverktyget "Bayley Scales of Infant Development (BSID), version III", eller vars

karakteristika under graviditet och förlossning, givna behandlingar av barnet och utfall under neonatalvård, samt uppföljningsdata vid 2 års ålder.

Datauttag för denna studie kommer begäras enligt de bestämmelser som Neonatalregistret upprättat. Ansökan om datauttag till Neonatalregistret ska ske skriftligt och med obligatoriska bilagor:
•Ansökan om etikprövning inkl bilagor
•Beslut från etikprövningsmyndighet
•Projektplan
•Variabelförteckning

6.2 Redogör för det statistiska underlaget för studiepopulationen/ undersökningsmaterialets storlek.

Sample size: Beräkningar av sample size för maskininlärning är svåra att göra då storleken på studien/antal deltagande forskningspersoner beror på komplexiteten i algoritmerna vilken inte är känd på förhand. En tumregel är att det bör finnas åtminstone 15-20 observationer per riskfaktor. Baserat på denna regel skulle det behövas 100-200 barn med nedsatt kognition i en multivariabel modell med 10 riskfaktorer. Preliminära uppgifter från SNQ visar att det finns ca 350 barn registrerade med IQ <85 vilket vi uppskattar ska räcka till en analys innefattande ett 50-tal riskfaktorer.

6.3 Hur kommer undersökningsprocedurerna att dokumenteras?

Ej relevant för denna studie.

6.4 Hur kommer insamlad data att hanteras och förvaras?

Neonatalregistret omfattas av samma sekretesskrav som journaluppgifter. Datasäkerheten har utformats i enlighet med patientdatalagen och Datainspektionens föreskrifter. Neonatalregistret stöds av Sveriges kommuner och regioner (SKL), v.g. se:
https://kvalitetsregister.se/hittaregister/registerarkiv/svensktneonataltkvalitetsregistersnq.2338.html

Forskaruttaget och samtliga analyser görs på pseudoanonymiserade data. Inga personnummer eller födelsedatum kommer att begäras ut. Databasen kommer förvaras på lösenordsskyddad server som endast ansvariga forskare kommer att ha tillgång till. Forskningsdatabasen kommer att sparas i tio år, därefter kommer en utvärdering göras huruvida forskningsdatabasen kan användas för fler/annat ändamål eller om den ska förstöras.

7.1 Vilka risker kan ett deltagande medföra för de forskningspersoner som ingår i forskningsprojektet?

INTEGRITETSSKYDD: Vid analyser och presentation av resultat kommer inga enskilda personer kunna identifieras. I manuskript/publikation redovisas resultat på gruppnivå. Databasen innehåller inte

uppgift om identitet och förvaras säkert utan åtkomst för obehöriga.
FRAMTIDSFRÄMJANDE: Det övergripande målet är förbättra hälsa för kommande generationer barn. Varje barn som ådrar sig en skada i nyföddhetsperioden kan innebära ett (livs)långt lidande för barnet/individen och familjen, och stora kostnader för vården och samhället. Studien syftar till att identifiera undvikbara risker för att i framtiden kunna förebygga skador hos intensivvårdade nyfödda barn och till att påvisa hälsobefrämjande faktorer.
ENHETLIGHETSSTRÄVAN: Studien syftar mot enhetlighet och utgör underlag för framtagande av nya/reviderade nationella riktlinjer som ska underlätta för berörda professioner att säkra processen för beslut gällande evidensbaserad diagnostik och behandling. Detta förväntas öka sjuka nyfödda barns möjligheter att överleva utan bestående men.Vår förhoppning är att resultaten även kommer att kunna användas i undervisning och komma till nytta i ett internationellt perspektiv.

7.2 Vilken nytta kan ett deltagande medföra för de forskningspersoner som ingår i forskningsprojektet?

Vi ser ingen nytta för personerna som ingår i projektet då studien bygger på historiska data.

7.3 Gör en värdering av förhållandet mellan riskerna och nyttan av projektet.

Någon risk för skada av forskningspersoner i samband med projektet finns inte. Den gravida kvinnans eller barnets karakteristika, undersökningar och behandlingar påverkas inte av studien, som är en ren observationsstudie av kvalitetssäkringskaraktär.

7.4 Beskriv hur projektet har utformats för att minimera riskerna för forskningspersonerna.

Det kommer inte vara möjligt att identifiera enskilda personer i samband analys, författande av rapport eller vid offentliggörande/ publicering - endast avidentifierade och aggregerade data kommer att användas och publiceras.

7.5 Identifiera och precisera om eventuella etiska problem (nackdelar/fördelar) kan uppstå i ett vidare perspektiv genom forskningsprojektet.

Beroende på vad som framkommer i studien kan vissa resultat eventuellt uppfattas som kontroversiella av berörda familjer, profession och av allmänhet. Å andra sidan förväntas den nya kunskapen kunna användas för att förbättra vården och patientsäkerheten till nytta för patient, profession, vårdgivare och samhälle. Detta uppväger eventuella kortsiktiga nackdelar som resultatet av projektet skulle kunna medföra.

8.1 Hur görs urvalet av forskningspersoner?

Studiepopulationen i den här studien kommer att utgöras av alla barn som fullgjort så kallad "riskbarnsuppföljning efter komplicerad neonatalvård", som genomgått formell testning vid 2 års ålder med testverktyget "Bayley Scales of Infant Development (BSID), version III", och vars

testresultat inrapporterats till det nationella kvalitetsregistret (drygt 2000 barn enligt preliminära uppgifter från Neonatalregistret SNQ).

8.2 Hur många forskningspersoner kommer att inkluderas i forskningsprojektet?

Uppgifter från drygt 2000 barn som finns registrerade som uppföljda med psykologtest vid 2 års ålder.

8.3 Vilka urvalskriterier kommer att användas för inklusion?

Alla barn som fullgjort så kallad "riskbarnsuppföljning efter komplicerad neonatalvård", som genomgått formell testning vid 2 års ålder med testverktyget "Bayley Scales of Infant Development (BSID), version III", och vars testresultat inrapporterats till det nationella kvalitetsregistret.

8.4 Vilka urvalskriterier kommer att användas för exklusion?

Exklusion: Barn födda med fosterskador (missbildningar eller kromosomfel), barn med kongenitala infektioner, barn födda utanför Sverige och som efter födelsen överförts till svenska sjukhus.

8.5 Ange relationen mellan forskare och forskningspersonerna.

Forskare-forskningsperson.

8.6 Vilket försäkringsskydd finns för de forskningspersoner som deltar i forskningsprojektet?

Ej relevant, forskningen sker på redan insamlade data.

8.7 Redogör för den beredskap som finns för att hantera oväntade bifynd eller händelser under forskningsprocessen som kan äventyra forskningspersonernas säkerhet.

Ej relevant, forskningen sker på redan insamlade data.

8.8 Kommer ekonomisk ersättning eller andra förmåner betalas ut till forskningspersonerna?

Nej

9.1 Kommer forskningspersonerna att informeras om forskningsprojektet och tillfrågas om de vill vara med eller inte?

Nej

9.1.2 [Om Nej 9.1] Motivera varför forskningspersonerna inte ska informeras och tillfrågas.

Studien bygger på data som är på förhand insamlade med kvalitetssäkring som ändamål. Enligt Socialstyrelsens författning SOSFS 2011:9 "Ledningssystem för systematiskt kvalitetsarbete" har vårdgivaren ett ansvar att kvalitetssäkra ("planera, leda, kontrollera, följa upp, utvärdera och förbättra") vården. Egenkontrollen kan enligt §2 innefatta "granskning av journaler, akter och annan information ". Vi ser detta projekt som ett viktigt led i vårdgivarens arbete att kvalitetssäkra vården av svårt sjuka nyfödda på nationell nivå.
Berörda patienter är utspridda över landet och antalet är förhållandevis stort (dryga 2000). Föräldrar har under vårdtiden på neonatalavdelning och vid insamlandet av data till Neonatalregistret SNQ informerats om registrets syften, vg se bilaga 5 och:
https://www.medscinet.com/pnq/foraldrainformation.aspx.

Vi avser inte att informera föräldrar en gång till och avser inte att inhämta samtycke från berörda vårdnadshavare för studiens genomförande. Detta förfarande har godkänts av EPM i flera tidigare ansökningar med liknande projekt.

9.2 Kommer barn under 18 år att ingå i forskningsprojektet?

Ja

9.2.1 [Om Ja 9.2] Ange barnens ålder.

0-2 år.

9.3 Kommer forskningspersoner, vars mening på grund av sjukdom, psykisk störning, försvagat hälsotillstånd eller något annat liknande förhållande inte kan inhämtas, att ingå i forskningsprojektet?

Nej

10.1 Kommer projektet att begära ut uppgifter från ett befintligt register?

Ja

10.1.1 [Om Ja 10.1] Ur vilket eller vilka register kommer uppgifterna att begäras?

Neonatalregistret SNQ (www.snq.se). Huvudman/CPUA: Karolinska Universitetssjukhuset i Stockholm.

10.1.2 [Om Ja 10.1] Vilka uppgifter kommer att begäras ut och varför?

Registerdata för barn neonatalvårdade under åren 2012-2020 är aktuella.

Vi kommer att begära ut följande uppgifter om modern och graviditeten:
moderns ålder, bostadsregion, och under graviditeten förekomst av epilepsi, diabetes, kronisk lungsjukdom/astma, högt blodtryck/preeklampsi, blödningar, infektion (chorioamnionit), flerbörd, intrauterin tillväxthämning, immunisering, tidig vattenavgång, användning av antenatala steroider, förlossningsenhet (kategori liten, medel, stor) och förlossningssätt.

För barnet kommer följande uppgifter begäras ut:
Födelseår, graviditetslängd, födelsevikt, födelselängd, huvudomfång, kön, Apgarpoämg vid 1, 5 och 10 minuter, navelsträngsblodgaser, HLR-åtgärder vid födelsen, avled före intagning, ankomsttemperatur vid inläggning på neonatalavdelning, född utanför regionklinik, CRIB (Clinical risk index for babies)-score),
lungsjukdomar, behandling av lungsjukdomar (extra O2 dgr, inhalerad NO antal dagar, surfaktantbehandling, steroidbeh pga lungsjd, CPAP-dygn, respiratordygn, högflödesgrimma), sepsis, behandling med antibiotika (dygn), farmakologisk eller kirurgisk behandling av öppetstående ductus arteriosus, prematuritetsretinopati (ROP), nekrotiserande enterokolit (NEC), parenteral nutrition, kärlkateterar, antal erytrocyt och plasmatransfusioner, inotroptstöd (antal dagar), vikt vid olika åldrar, periventrikulär leukomoalaci, intraventrikulär hjärnblödning, annan hjärnblödning eller infarkt, hypoxisk-ischemisk encefalopati, kramper, kylbehandling, antiepileptikabehandling, antal P-glukos värden <2,6 mmol/L, insulinbehandling för hyperglykemi, hyperbilirubinemi, amning vid utskrivning, avliden och dödsålder. Vidare ålder för 2-årskontroll, resultat av psykologtest (Bayley-III), motorik, syn, hörsel, ev CP och beteendeproblem.

Uppgifterna behövs för att kunna utföra de planerade statistiska modelleringarna.

11.1 Finns det relevanta resultat från djurförsök?

Ej aktuellt

12.1 Hur garanteras tillgång till data för forskningshuvudmannen och medverkande forskare?

Data är redan insamlade och finns att tillgängliggöras (för forskningshuvudmannen och medverkande forskare) i Neonatalregistret SNQ.

12.2 Vem eller vilka ansvarar för databearbetning och skriftlig redovisning av resultaten?

Statistisk databearbetning kommer ske i samarbete med Andrea Bowe, Anthony Staines och Gordon Lightbody. Samtliga medverkande forskare kommer att delta vid resultatanalys och rapportskrivning. Den huvudansvariga forskaren kommer att ha full tillgång till alla data.

12.3 Hur och när planeras resultaten att offentliggöras?

Projektet förväntas generera 1-2 vetenskapliga rapporter publicerade i internationellt framstående tidskrifter som använder sig av peer review. Den första publikationen förväntas under 2023-24.

12.4 På vilket sätt garanteras forskningspersonernas rätt till integritet när materialet offentliggörs?

Det kommer inte vara möjligt att identifiera enskilda personer i samband med offentliggörande eller publicering av resultaten - endast helt avidentifierade och aggregerade data kommer att publiceras.

13.1 Redovisa eventuella ekonomiska överenskommelser med bidragsgivare eller andra finansiärer (namn och belopp).

Ansvarig forskare (MN) finansieras dels av ett treårigt s.k. ALF-anslag för åren 2021-2023 som beviljats av Region Stockholm (totalt 1 miljon kr årligen), dels av ramanslag från Stiftelsen Barnhuset av Svenska Frimurarorden.

Andrea Bowe finansieras av the Irish Clinical Academic Training Programme (Grant Number 203930/B/16/Z), the Health Service Executive, National Doctors Training and Planning and the Health and Social Care, Research and Development Division, Northern Ireland. Det finns inga övriga överenskommelser om finansiering att redovisa i samband med forskningen.

13.2 Redovisa forskningshuvudmannens, huvudansvarig forskares och medverkande forskares egna ekonomiska intressen.

Det finns inga egna ekonomiska intressen att redovisa i samband med forskningen.

Forskningsplan

Den sammanfattande beskrivningen av forskningsprojektet ska förstås av fackmän. Den kan lämpligen utformas enligt följande:
Vetenskaplig frågeställning: En redogörelse för det övergripande syftet med det föreslagna forskningsprojektet samt specifika mål (primära och sekundära frågeställningar).
Områdesöversikt: Ge ett sammandrag av egna och andras forskning och tidigare resultat inom forskningsområdet. Översikten ska tydliggöra det aktuella projektets relevans. Nyckelreferenser ska anges.
Projektbeskrivning: Gör en sammanfattning av projektets/motsvarande uppläggning. Urval av forskningspersoner, procedurer, metoder med mera ska tydligt redovisas. Det ska framgå hur metoder, urval och procedurer kan ge svar på de specifika frågeställningarna. Om flera delprojekt avses anges sekvens för genomförande och på vilket sätt ett efterföljande delprojekts uppläggning kan bero av resultaten av ett föregående.
Betydelse: Ge en kortfattad redogörelse för projektets betydelse för forskningsområdet.
Preliminära resultat: Kan i förekommande fall anges.
SKA VARA PÅ SVENSKA ELLER ENGELSKA.

Project_plan.docx.pdf
98.42KB

Information till forskningspersoner och samtyckesformulär

Etikprövningsmyndigheten behöver alltid ta del av all information som kommer att ges till forskningspersonen i samband med tillfrågan om deltagande. Både den information som ska ges muntligt och den som ska ges skriftligt. Om vårdnadshavare ska samtycka till deltagande ska även den information som ges till vårdnadshavarna bifogas. Om anhörig ska ges möjlighet att motsätta sig deltagande ska även den information som ges till anhörig bifogas.
Etikprövningsmyndigheten rekommenderar att vår stödmall för forskningspersonsinformation används, den hittar du på www.etikprovning.se
SKA VARA PÅ SVENSKA.

SNQ_Frldrainformation_4.0.pdf
216.62KB

Variabellista

Variabellista bör bifogas ansökan om data ska begäras ut från befintliga register.
SKA VARA PÅ SVENSKA.

Variabellista.xlsx.pdf
122.74KB

Variabellista_2.xlsx.pdf
12KB

CV för ansvarig forskare

Bifoga CV för ansvarig forskare.
I undantagsfall kan icke disputerad forskare godtas om annan medverkande disputerad forskare uttalat att forskningen sker under aktivt överinseende av denne. Uttalandet intygas vid signering av disputerad forskare. CV för den disputerade ska även bifogas.
SKA VARA PÅ SVENSKA ELLER ENGELSKA.

Short_CV_Mikael_Norman_March_1_2022.pdf
850.49KB

I och med att ansökan undertecknas intygar du som är ansvarig forskare samt du som är behörig företrädare följande;

- Att den information som lämnas i ansökan om etikprövning och samtliga medföljande bilagor är riktig och fullständig.

- Att verksamhetsansvariga i samtliga medverkande verksamheter är informerade om forskningsprojektets innehåll och utförande och att de har samtyckt till att delta i studien.
- Att du säkerställt att det i samtliga medverkande verksamheter finns resurser som garanterar forskningspersonernas säkerhet och integritet vid genomförandet av den forskning som beskrivs i ansökan.
- Att ansvarig forskare ges rätt att företräda huvudmannen i alla framtida kontakter med Etikprövningsmyndigheten som rör detta forskningsprojekt samt ansöka om ändringar i forskningsprojektet.
- Att du tagit del av Etikprövningsmyndighetens information om hantering av personuppgifter på myndighetens webbplats.

Behörig företrädare för forskningshuvudmannen

Lars Navér

Signatur behörig företrädare

Signatur-behorig-foretradare.pdf
31.59KB

Är behörig företrädare ordinarie eller tillförordnad?

Ordinarie företrädare

Signatur huvudansvarig forskare

Signatur-huvudansvarig-forskare.pdf
30.39KB

Beslut och handlingar från Etikprövningsmyndigheten

Beslutsbrev och andra handlingar från Etikprövningsmyndigheten i relation till denna ansökan

2022-01351-01_Avgiftsavisering.pdf
35.39KB

Study Title
Prediction of Poor Cognitive Outcomes at 24-months in Infants Admitted to Neonatal Units: Applying Machine Learning to a Population Based Neonatal Registry

Collaborators

Dr Andrea Bowe MB, BCh, BAO, MPH, (PhD Candidate)
INFANT Research Centre, University College Cork, Cork, Ireland

Professor Anthony Staines MB, BCh, BAO, PhD
School of Nursing, Psychotherapy and Community Health
Dublin City University, Dublin, Ireland

Dr Gordon Lightbody Meng, PhD
INFANT Research Centre, University College Cork, Cork, Ireland

Professor Deirdre Murray MB, BCh, BAO, PhD
INFANT Research Centre, University College Cork, Cork, Ireland

Professor Mikael Norman MD, PhD
Department of Clinical Science, Intervention, and Technology, Karolinska Institutet

Funding: Dr Andrea Bowe is funded through the Irish Clinical Academic Training Programme, supported by the Wellcome Trust and the Health Research Board (Grant Number 203930/B/16/Z), the Health Service Executive, National Doctors Training and Planning and the Health and Social Care, Research and Development Division, Northern Ireland.

Etikprövningsmyndigheten
2022-01351-01-242720
2022-03-08

Background

Early life is the most crucial period for shaping the developing brain, and represents a window of both vulnerability and opportunity.[1] Without intervention, it is the period during which most health and social inequities will originate. A failure to achieve early foundational cognitive skills may result in a permanent loss of opportunity to achieve full cognitive potential.

Previous studies have found that the level of cognitive functioning achieved early in life is predictive of that at school-age.[2] This in turn is an important predictor of outcomes throughout the life course, including educational attainment, socioeconomic position, and health.[3-5] Optimising cognitive development in very early life is therefore an important public health strategy.

Across epidemiological, animal, and human research there is consensus agreement that early life is a period of fundamental importance in cognitive development, and strategies for intervention should ideally be targeting this unique period. [6] Early pre-emptive interventions, based on the presence of a risk factor or very early sign, have the potential to alter the trajectory of neurodevelopmental outcomes.[7]

Etikprövningsmyndigheten
2022-01351-01-242720
2022-03-08

Population-based screening programmes based on the principal of universalism apply structured developmental screening, independent of pre-existing risk factors, to all children. While this approach has many merits, there are significant limitations. Firstly, the opportunity for pre-emptive intervention is lost as these infants are already behind their typically developing counterparts at the time of detection, with valuable opportunities in a period of optimal neuroplasticity being lost. Secondly, universal screening programmes tend not to provide a systematic, robust assessment of the environmental, social, and demographic factors, to which the infant will be exposed in early life which are the strongest and most consistent predictors of cognitive outcomes in childhood.[8,9]

Adverse cognitive outcomes in childhood are complex, heterogenous, and result from highly interactive relationships between biological, environmental, and social factors, the mechanisms of which are often poorly understood. This real-world complexity, which may be difficult to model, particularly with more traditional statistical methods, may have to date acted as a deterrent for researches investigating risk prediction tools for this purpose. However, there is now increasing potential to statistically model such complex interactions with machine learning and large datasets, which are increasingly able to record granular data on individuals at a population level.

One subgroup who are at increased risk of poor cognitive outcomes are infants born preterm, at <37 weeks gestation. [10,11] Among preterm infants there are biological, environmental, and family factors which affect cognitive outcomes.[12] For example a high level of parental education may mitigate the effect of gestational age on cognitive outcomes, but a low level combined with prematurity may exacerbate poor cognitive outcomes.[13]

Aim

The aim of this study is to investigate whether biological, environmental ,and social factors at birth can be combined in a machine learning algorithm to identify infants admitted to a neonatal unit at birth who are at risk of poor cognitive outcomes at 24 months.

Objectives

1. Train a machine learning classification algorithm to identify infants at high risk of poor cognitive outcomes based on data available at birth.
2. To identify the important features for predicting cognitive outcomes at 24 months
3. To examine the performance characteristics of this algorithm on data held back in the training process.
4. To repeat this training and testing process and develop a specific algorithm for a subgroup confined to preterm infants born at <37 weeks gestation to determine

Methodology

Study population: The Swedish Neonatal Quality Register is the largest pediatric quality register in Sweden and includes all infants born alive in Sweden who were admitted for neonatal care within 27 days of birth.[16] Since 2016, children who experienced complicated neonatal periods have been followed up with neurodevelopmental assessments at 2 years of age. The study population for this study will consist of all infants in the SNQ who have completed the Bayley Scales of Infant Development (BSID) at 24 months. A separate subgroup consisting of infants with a gestational age <37 weeks will also be analysed as part of this study.

Etikprövningsmyndigheten
2022-01351-01-242720
2022-03-08

Predictors of Interest: Machine learning will be used to train a model and identify the features of importance. As such, a broad array of potential predictors based on biological, environmental and social factors will be included as potential predictors. The broad range of potential predictors which will be considered are listed in Table 1. The criteria for inclusion of a predictor is that it is readily measurable at birth. This will ensure a model that could be implemented at the earliest possible stage for intervention.

Outcome of Interest: The primary outcome of interest will be performance on the cognitive domain of the BSID at 24 months. A score greater than 1 standard deviation below the mean (score <85) will be used to indicate a poor cognitive outcome at 24-months. The outcome variables required for this study are listed in Table 2.

Statistical Analysis

Sample size: Sample size calculations are difficult to perform for machine learning problems as they are dependent on the complexity of the algorithm. A widely proposed rule of thumb for multivariable classification problems is that there should be at least 10-20 events per candidate predictor parameter (EPP). The predictor parameter refers to the degrees of freedom of each variable.[15] Other studies have proposed a higher EPP of 20 for predictive models which contain low-prevalence predictors.[16,17] Based on this rule, for a multivariable classification model based on 10 parameters, there would need to be approximately 100 -200 infants with a poor cognitive outcome. This rule provides a very approximate guide to sample

size calculation for machine learning problems which are highly dependent on the variability within the dataset.

Training of model: The data will be split into a training and a testing dataset. The machine learning model will be trained on the training dataset using a variety of classification techniques including random forests, gradient boosting, and logistic regression. The relative importance of each variable will be assessed by examining the mean decrease in Gini index. K-fold cross validation will to tune model hyperparameters.

Testing of the model: The final model will then be tested on the testing dataset, which is held back in the training process. The performance of the model will be measured using sensitivity, specificity, receiver operating characteristic curves, and precision recall curves.

Roles of Authors in the Study

Dr Andrea Bowe is a medical doctor training in public health medicine and a PhD candidate funded by the Irish Clinical Academic Training Programme. Under the supervision of co-authors she will perform the cleaning of the dataset, the statistical analysis, and will be responsible for writing the manuscript.

Professor Anthony Staines is a Professor of Health Systems and Public Health. He will provide supervision with regard to statistical methods and will provide public health input to the project.

Dr Gordon Lightbody is a principal investigator at the INFANT Research Centre. His background is in engineering and he is an expert in machine learning. He will provide input on the statistical analysis of the data.

Professor Deidre Murray is a Consultant Paediatrician and is Principal Investigator with the BASELINE birth cohort. Her research has focussed on early brain injury and methods to assess and predict neurocognitive outcomes.

Professor Mikael Norman is the Principal Investigator in Sweden for this project. His research has focused on 1) preterm birth: risk factors, interventions and short and long-term outcome; 2) perinatal epidemiology; and 3) patient safety and quality of care. He will be responsible for regulatory approvals and data safety issues.

Etikprövningsmyndigheten
2022-01351-01-242720
2022-03-08

References

1. Spencer N, Raman S, O'Hare B, Tamburlini G. Addressing inequities in child health and development: towards social justice. *BMJ paediatrics open* 2019; **3**: e000503–e000503.
2. Maxwell JR, Yellowhair TR, Oppong AY, Camacho JE, Lowe JR, Jantzie LL, Ohls RK. Cognitive development in preterm infants: multifaceted deficits reflect vulnerability of rigorous neurodevelopmental pathways. Minerva Pediatr. 2017 Aug;69(4):298-313. doi: 10.23736/S0026-4946.17.04905-2. Epub 2017 Feb 17. PMID: 28211648.
3. Lager A, Bremberg S, Vågerö D. The association of early IQ and education with mortality: 65 year longitudinal study in Malmö, Sweden. *BMJ (Clinical research ed)* 2009; **339**: b5282.
4. Forrest LF, Hodgson S, Parker L, Pearce MS. The influence of childhood IQ and education on social mobility in the Newcastle Thousand Families birth cohort. *BMC public health* 2011; **11**: 895.
5. Whalley LJ, Deary IJ. Longitudinal cohort study of childhood IQ and survival up to age 76. *BMJ (Clinical research ed)* 2001; **322**: 819.
6. Ramey CT, Ramey SL. Prevention of intellectual disabilities: early interventions to improve cognitive development. *Preventive medicine* 1998; **27**: 224–232.
7. Whitehouse AJO, Varcin KJ, Pillar S, Billingham W, Alvares GA, Barbaro J *et al*. Effect of Preemptive Intervention on Developmental Outcomes Among Infants Showing Early Signs of Autism: A Randomized Clinical Trial of Outcomes to Diagnosis. *JAMA Pediatrics* 2021; **175**: e213298–e213298
8. Flensborg-Madsen T, Falgreen Eriksen H-L, Mortensen EL. Early life predictors of intelligence in young adulthood and middle age. *PloS one* 2020; **15**: e0228144.
9. Eriksen H-LF, Kesmodel US, Underbjerg M, Kilburn TR, Bertrand J, Mortensen EL. Predictors of intelligence at the age of 5: family, pregnancy and birth characteristics, postnatal influences, and postnatal growth. *PloS one* 2013; **8**: e79200.
10. Maxwell JR, Yellowhair TR, Oppong AY, Camacho JE, Lowe JR, Jantzie LL, Ohls RK. Cognitive development in preterm infants: multifaceted deficits reflect vulnerability of rigorous neurodevelopmental pathways. Minerva Pediatr. 2017 Aug;69(4):298-313. doi: 10.23736/S0026-4946.17.04905-2. Epub 2017 Feb 17. PMID: 28211648.
11. Ådén U, Sävman K, Norman M. De flesta för tidigt födda barn får normal utveckling och god hälsa - Men viktigt med tidig behandling och rätt stöd för dem som riskerar funktionspåverkan och sämre hälsa på sikt [Children born too soon and outcomes]. Lakartidningen. 2019 Oct 7;116:FR6D. Swedish. PMID: 31593283.
12. Stålnacke SR, Tessma M, Böhm B, Herlenius E. Cognitive Development Trajectories in Preterm Children With Very Low Birth Weight Longitudinally Followed Until 11 Years of Age. Front Physiol. 2019 Apr 2;10:307. doi: 10.3389/fphys.2019.00307. PMID: 31001126; PMCID: PMC6454032.
13. Linsell L, Malouf R, Morris J, Kurinczuk JJ, Marlow N. Prognostic factors for poor cognitive development in children born very preterm or with very low birth weight: a systematic review. *JAMA Pediatr*. 2015;169(12):1162-1172. doi:10.1001/jamapediatrics.2015.2175
14. Bilsteen JF, Ekstrøm CT, Børch K, Nybo Andersen AM. The role of parental education on the relationship between gestational age and school outcomes. Paediatr Perinat

Etikprövningsmyndigheten
2022-01351-01-242720
2022-03-08

Epidemiol. 2021 Nov;35(6):726-735. doi: 10.1111/ppe.12766. Epub 2021 May 26. PMID: 34080707; PMCID: PMC8597013.
15. Benavente-Fernández I, Synnes A, Grunau RE, et al. Association of socioeconomic status and brain injury with neurodevelopmental outcomes of very preterm children. *JAMA Netw Open*. 2019;2(5):e192914. doi:10.1001/jamanetworkopen.2019.2914
16. Norman M, Källén K, Wahlström E, Håkansson S; SNQ Collaboration. The Swedish Neonatal Quality Register - contents, completeness and validity. Acta Paediatr. 2019 Aug;108(8):1411-1418. doi: 10.1111/apa.14823. Epub 2019 May 20. PMID: 31006126.
17. Riley RD, Ensor J, Snell KIE, Harrell FE Jr, Martin GP, Reitsma JB, Moons KGM, Collins G, van Smeden M. Calculating the sample size required for developing a clinical prediction model. BMJ. 2020 Mar 18;368:m441. doi: 10.1136/bmj.m441. PMID: 3218860
18. Pourhoseingholi MA, Vahedi M, Rahimzadeh M. Sample size calculation in medical studies. Gastroenterol Hepatol Bed Bench. 2013 Winter;6(1):14-7. PMID: 24834239; PMCID: PMC4017493.
19. Ogundimu EO, Altman DG, Collins GS. Adequate sample size for developing prediction models is not simply related to events per variable. J Clin Epidemiol. 2016 Aug 1;76:175–82.

Etikprövningsmyndigheten
2022-01351-01-242720
2022-03-08

Appendix 7 Ethical approval from the Swedish Ethical Review Board

Dnr 2022-01351-01

Göteborg avdelning 2 medicin

BESLUT

2022-03-28

Sökande forskningshuvudman
Region Stockholm

Forskare som genomför projektet
Mikael Norman

Projek tel
Predik on av senare kogni vt u all hos nyfödda barn som intensivvårdats för svåra komplika oner genom a llämpa maskininlärning på popula onsbaserade registerdata.

Uppgi er om ansökan
Ansökan inkom ll E kprövningsmyndigheten 2022-03-08 och blev valid 2022-03-11.

E kprövningsmyndigheten beslutar enligt nedan.

BESLUT

E kprövningsmyndigheten godkänner den forskning som anges i ansökan.

Etikprövningsmyndigheten
2022-01351-01-256458
2022-04-06

Appendix 8 Supplementary material for Chapter 2: Below average cognitive ability – An under researched risk factor for emotional-behavioural difficulties in childhood

Table A8.1 Descriptive table comparing children with below average cognitive ability and children with low cognitive ability at age 3 years

		Low Cognitive Ability ≥2 SD below mean (n=216)		Below Average Cognitive Ability 1-2 SD below mean (n=772)	
	Valid	n (%)	Valid	n (%)	p-value
Maternal Characteristics					
Age[a], mean (SD)	216	216, 29.6 (6.91)	772	772, 31.0 (5.7)	0.020[c]
No partner in household	216	72 (33.4)	772	165 (21.4)	<0.001[d]
Current Smoker	216	85 (39.2)	772	275 (35.6)	0.336[d]
Depression, n, mean (SD)	211	211, 3.7 (4.9)	754	754, 3.0 (4.2)	0.010[c]
Sociodemographic Characteristics					
Maternal Education	213		772		
Lower Secondary		88 (41.5)		239 (31.0)	
Upper Secondary		59 (27.8)		211 (27.3)	
Sub Degree		39 (18.1)		176 (22.8)	
Degree or higher		27 (12.5)		146 (18.9)	0.010[d]
Income Quintile	199		715		
Lowest		78 (39.0)		202 (28.3)	
2nd		55 (27.4)		164 (22.9)	
3rd		30 (15.0)		149 (20.8)	
4th		25 (12.5)		112 (15.7)	
Highest		12 (6.22)		88 (12.3)	0.003[d]
Infant Characteristics					
Male gender	216	149 (68.8)	772	477 (61.8)	0.057[d]
Birth weight	210		767		
<2,500g		12 (5.8)		54 (7.0)	
2,500-4,000g		179 (85.4)		606 (79.0)	
>4,000g		19 (8.8)		107 (14.0)	0.102[d]
Gestational age[b], mean (SD), wk	216	39.6 (2.12)	772	39.5 (2.3)	0.543[c]
Admitted to NICU at birth	216	34(15.6)	772	119 (15.4)	0.930[d]
Total SDQ Scores					
Age 3, mean (SD)		11.0 (5.05)		8.68 (4.62)	<0.001[c]
Age 5, mean (SD)		10.6 (5.30)		8.47 (5.47)	<0.001[c]
Age 9, mean (SD)		10.8 (6.54)		8.86 (5.97)	<0.001[c]

[a]Maternal age includes participants aged 40 years and over who were coded as 40 in original dataset, [b]Gestational age includes participants with a gestational age of 25 weeks or less which was recorded as 25. [c]Welch Two Sample t-test, [d]Pearson's Chi-squared test.

Table A8.2 Correlations between non-verbal cognitive ability t-scores at age 3 and SDQ total and subscale scores at age 3, age 5, and age 9

	Emotional	Peer-relationship	Conduct	Hyperactivity	Total SDQ
Age 3	-0.08	-0.10	-0.08	-0.14	-0.08
Age 5	-0.06	-0.10	-0.08	-0.12	-0.14
Age 9	-0.07	-0.11	-0.08	-0.12	-0.14

All correlation coefficients were significant at a p-value <0.05.

Table A8.3 Estimated marginal means and mean differences according to Model 5

	Average/above		Below average	
Age	Estimate	95% CI*	Estimate	95% CI*
3	7.32	7.19 – 7.45	8.19	7.78 – 8.59
5	6.86	6.73 – 7.00	7.78	7.36 – 8.21
9	6.80	6.65 – 6.95	8.26	7.79 – 8.74
3-9	-0.52	-0.67 - -0.37	0.08	-0.39 – 0.55
3-5	-0.46	-0.58 - -0.33	-0.40	-0.80 - -0.01
5-9	-0.06	-0.20 – 0.07	0.48	0.06 – 0.90

*Confidence intervals adjusted using Bonferroni method

Figure A8.1 Flowchart of study population

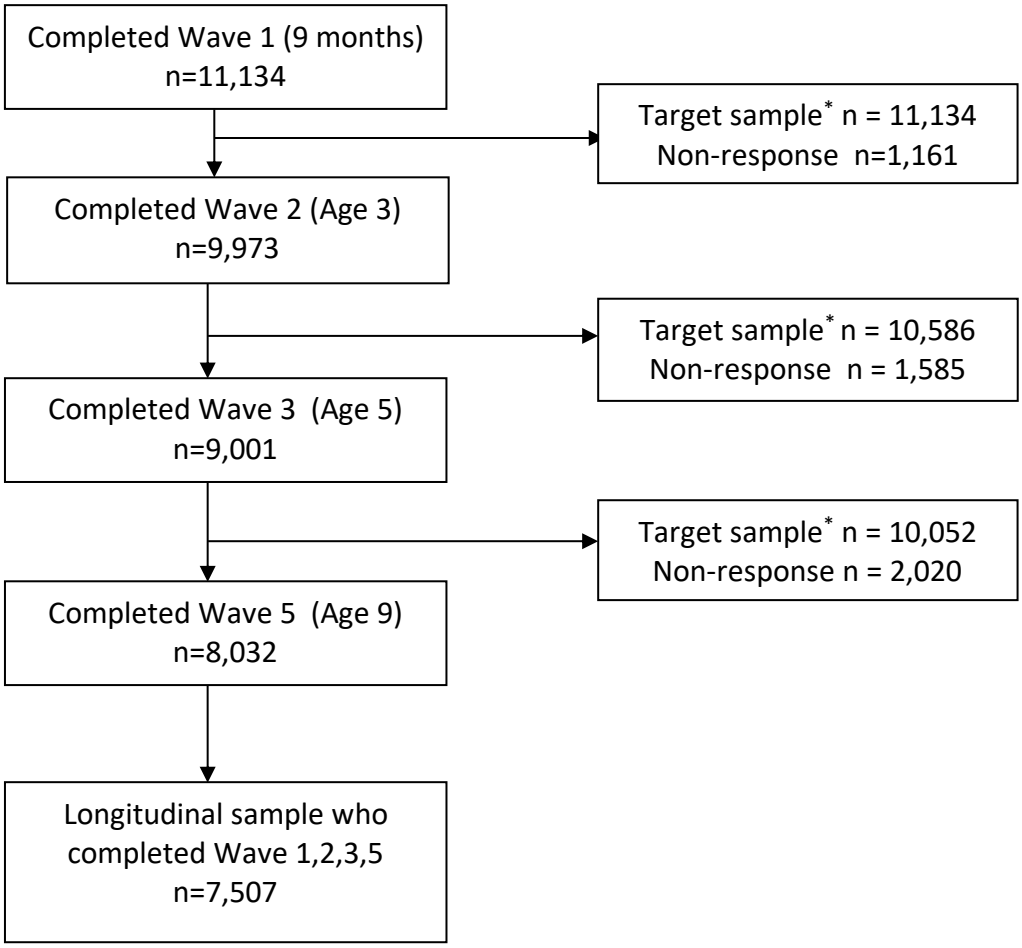


Figure A8.1 Legend
*Target samples included those who participated in the previous round and/or those who had participated at one of the earlier rounds of the study
**Wave 4 of the study, not shown in diagram above, consisted of a postal questionnaire only, which 5,344 participants completed.

Figure A8.2. Parameter estimates for adjusted model examining relationship between cognitive ability and longitudinal SDQ scores (includes those with low cognitive ability defined as a score 2 SD below mean or lower)



Figure A8.2 Legend

Categorical variables are represented by circles in the centre of a line. The circles in represent the point estimates and the lines represent 95% confidence intervals. Continuous variables a represented by a point only. Variables in green are predictors of lower longitudinal total SDQ scores. Variables in orange are predictors of higher longitudinal total SDQ scores. Reference group for cognitive ability, Cognitive ability average or above; Reference group for gender, male; Reference group for partner in household, no partner present; Reference group for maternal education, degree or higher.

Figure A8.3. Predicted change in total SDQ by cognitive ability

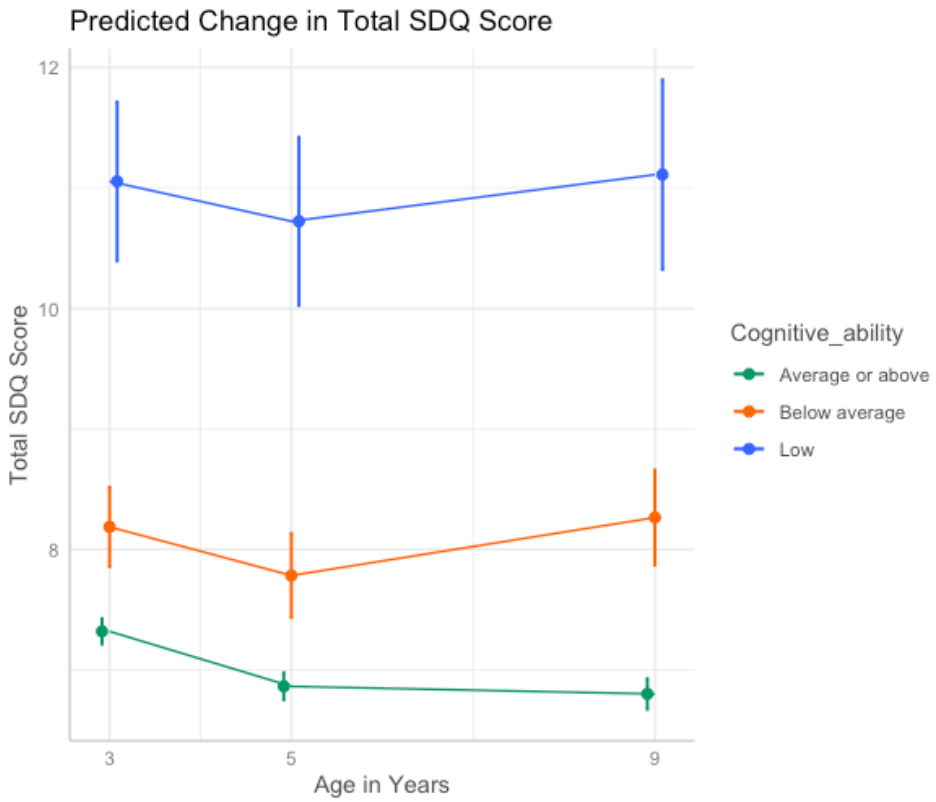
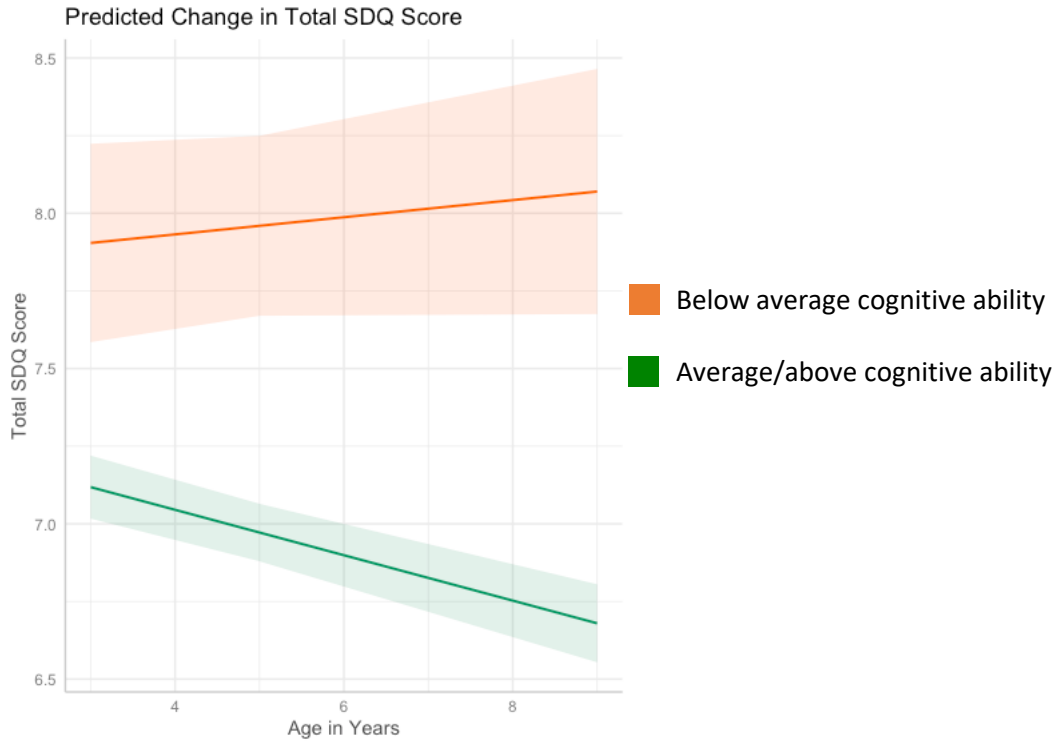


Figure A8.3 Legend

This figure shows the predicted change in total SDQ score between age 3-9 years by cognitive ability group. Average or above refers to participants with a cognitive ability score greater than 1 standard deviation below the mean. Below average refers to participants with a cognitive ability score between 1 and 2 standard deviations below the mean. Low refers to participants with a cognitive ability score more than 2 standard deviations below the mean.

Figure A8.4. Predicted change in total SDQ by cognitive ability (with age modelled as a continuous variable)



Appendix 9 Supplementary material for Chapter 4: The Relationship Between Below Average Cognitive Ability at Age 5 Years and the Child's Experience of School at Age 9

Table A9.1 Characteristics of study population and comparison of characteristics between those with typical cognitive development, below average cognitive ability, and low cognitive ability.

	Valid	Total n=7,392	Typical cognitive development (GCA ≥ 85) n=6,286	Below average cognitive ability (GCA 70-85) n=952	Low cognitive ability (GCA <70) n=154	p-value
Child characteristics						
Gender	7,392					
Male		3,783 (51.2)	3,131 (49.8)	562 (59.1)	89 (57.8)	<0.001[a]
Gestational age	7,375					
N, mean (sd)		39.5 (2.1)	6,277, 39.5 (2.1)	949, 39.6 (2.1)	149, 39.4 (2.6)	0.609[b]
Child temperament						
Sociability – n, median (IQR)	7,389	4.0 (3.0-5.0)	6,285, 4.0 (3.0-5.0)	950, 3.8 (3.0-4.8)	153, 4.0 (2.5-5.0)	0.050[c]
Persistence – n, median (IQR)	7,373	4.3 (3.5-5.0)	6,279, 4.3 (3.5-5.0)	947, 4.0 (3.0-4.8)	147, 3.8 (2.8-4.8)	<0.001[c]
Reactivity – n, median (IQR)	7,381	2.3 (1.8-3.0)	6,282, 2.3 (1.8-3.0)	950, 2.3 (1.8-3.0)	148, 2.8 (2.0-3.8)	<0.001[c]
General cognitive ability score	7,392					
N, mean (sd)		99.8 (14.9)	6,286, 103.8 (12.1)	951, 79.2 (4.2)	154, 64.6 (5.5)	<0.001[b]
Body mass index – kg/m^2	7,319	16.3 (1.7)	6,231, 16.3 (1.7)	936, 16.3 (1.7)	152, 16.8 (2.1)	0.002[b]
Sociodemographic characteristics						
Partner in household	7,392					
Yes		6,224 (84.2)	5,352 (85.1)	746 (78.4)	126 (81.8)	<0.001[a]
Siblings	7,391					
Yes		6,479 (87.7)	5,514 (87.7)	836 (87.9)	129 (83.8)	0.328[a]
Cultural background	7,375					
Irish		6,398 (86.8)	5,598 (89.1)	717 (75.9)	83 (56.5)	
Other white		642 (8.7)	457 (7.3)	149 (15.8)	37 (25.2)	
African or other black		156 (2.1)	93 (1.5)	55 (5.8)	8 (5.4)	

Chinese or other Asian		151 (2.0)	116 (1.8)	21 (2.2)	15 (10.2)	
Other		28 (0.4)	20 (0.3)	3 (0.3)	4 (2.7)	<0.001[a]
PCG Highest education	7,392					
Lower secondary or less		965 (13.1)	722 (11.5)	214 (22.5)	29 (18.8)	
Secondary		1,227 (16.6)	1,027 (16.3)	169 (17.8)	31 (20.1)	
Technical or vocational		1,430 (19.3)	1,169 (18.6)	217 (22.8)	45 (29.2)	
Certificate or diploma		1,605 (21.7)	1,421 (22.6)	163 (17.1)	21 (13.6)	
Primary degree		851 (11.5)	746 (11.9)	88 (9.2)	17 (11.0)	
Post degree qualification		1,313 (17.8)	1,201 (19.1)	101 (10.6)	11 (7.1)	<0.001[a]
Household social class	7,392					
Professional workers		934 (12.6)	858 (13.6)	63 (6.6)	13 (8.4)	
Managerial and technical		2,380 (32.2)	2,164 (34.4)	191 (20.1)	24 (15.6)	
Non-manual		1,334 (18.0)	1,126 (17.9)	186 (19.5)	22 (14.3)	
Skilled manual		1,115 (15.1)	868 (13.8)	213 (22.4)	34 (22.1)	
Semi-skilled		668 (9.0)	525 (8.4)	126 (13.2)	18 (11.7)	
Unskilled		110 (1.5)	95 (1.5)	15 (1.6)	0 (0.0)	
All others gainfully occupied		76 (1.0)	51 (0.8)	19 (2.0)	6 (3.9)	
Validly no social class[d]		775 (10.5)	599 (9.5)	139 (14.6)	37 (24.0)	<0.001[a]
Equivalised Income Quintile[e]	7,009					
Lowest		1,448 (20.7)	1,107 (18.5)	286 (32.1)	55 (39.9)	
2		1,472 (21.0)	1,194 (20.0)	238 (26.7)	40 (29.0)	
3		1,383 (19.7)	1,207 (20.2)	150 (16.8)	26 (18.8)	
4		1,360 (19.4)	1,231 (20.6)	118 (13.2)	11 (8.0)	
Highest		1,346 (19.2)	1,240 (20.7)	100 (11.2)	6 (4.3)	<0.001[a]
Relationship characteristics						
Parent-child relationship						
Closeness – n, median (IQR)	7,380	35.0 (33.0-35.0)	6,285, 35.0 (33.0-35.0)	947, 34.0 (33.0-35.0)	148, 34.0 (32.0-35.0)	<0.001[c]

Conflict – n, median (IQR)	7,384	14.0 (10.0-19.0)	6,284, 14.0 (10.0-18.0)	951, 15.0 (11.0-19.0)	149, 16.0 (11.0-21.0)	<0.001[c]

Abbreviations: GCA, General cognitive ability score; sd, standard deviation; IQR, interquartile range.
[a]Pearson's Chi-squared test, [b]Analysis of variance (ANOVA), [c]Independent samples Kruskal-Wallis test,
[d]Validly no social class consists of households where both caregivers are currently economically inactive and have not held any previous employment in the past, [e]Equivalised Income Quintile is based on the disposable household income (total gross income less statutory deductions), divided by the equivalised household size which takes account of differences in size and composition of households.

Table A9.2 Learning limitations and supports reported by child's teacher at age 5 and age 9

	Valid	Total n=7,392	Typical cognitive development (GCA ≥ 85) n=6,286	Below average cognitive ability (GCA 70-85) n=952	Low cognitive ability (GCA <70) n=154	p-value
Recognised limitation to learning at age 5	6,955					
No		5,420 (77.9)	4,854 (81.9)	525 (59.1)	41 (29.5)	
Yes		1,535 (22.1)	1,074 (18.1)	363 (40.9)	98 (70.5)	<0.001[a]
Special help or resources provided to child at age 5	6,954					
No		6,132 (88.2)	5,394 (91.0)	676 (76.1)	62 (44.6)	
Yes		822 (11.8)	533 (9.0)	212 (23.9)	77 (55.4)	<0.001[a]
Specific limitations to learning at age 9						
Physical/visual/hearing impairment	6,491	177 (2.7)	136 (2.5)	32 (3.9)	9 (7.0)	<0.001[a]
Speech impairment	6,487	155 (2.4)	103 (1.9)	37 (4.5)	15 (11.7)	<0.001[a]
Autism spectrum disorder	6,468	205 (3.2)	149 (2.7)	38 (4.6)	18 (14.2)	<0.001[a]
Specific learning difficulty e.g. dyslexia	6,439	428 (6.6)	331 (6.0)	76 (9.3)	21 (16.5)	<0.001[a]
General learning disability - mild	6,457	324 (5.0)	213 (3.9)	89 (10.9)	22 (17.5)	<0.001[a]
General learning disability – mod/sev/profound	6,444	80 (1.2)	53 (1.0)	17 (2.1)	10 (7.8)	<0.001[a]

Emotional or behavioural problem e.g. ADHD[b]	6,443	261 (4.1)	198 (3.6)	50 (6.1)	12 (10.2)	<0.001[a]
Home environment	6,137	326 (5.0)	239 (4.3)	79 (9.6)	8 (6.3)	<0.001[a]
Language	6,476	128 (2.0)	51 (0.9)	53 (6.4)	24 (18.9)	<0.001[a]
Discipline	6,468	192 (3.0)	144 (2.6)	43 (5.2)	5 (3.9)	<0.001[a]
Poor attendance	6,464	214 (3.3)	145 (2.6)	60 (7.3)	9 (7.0)	<0.001[a]
Other	5,633	297 (5.3)	253 (5.2)	38 (5.5)	6 (6.3)	0.872[a]

[a]Pearson's Chi-squared test

Table A9.3 Child-reported experience of school at age 9

	Valid	Total	Typical cognitive development (GCA ≥ 85) n=6,286	Below average cognitive ability (GCA 70-85) n=952	Low cognitive ability (GCA <70) n=154	P	BACA v TCD Odds ratio (95% CI)	BACA v TCD Adjusted odds ratio[b] (95% CI)	LCA v TCD Odds ratio (95% CI)	LCA v TCD Adjusted odds ratio[b] (95% CI)
Thoughts about school	7,305									
Always like it		2,409	2,088	275	46		Ref	Ref	Ref	Ref
		(33.3)	(33.6)	(29.1)	(31.5)					
Sometimes like it		4,494	3,837	581	76		1.15	1.06	0.90	0.67
		(61.5)	(61.7)	(61.5)	(52.1)		(0.99-1.34)	(0.89-1.25)	(0.62-1.31)	(0.43-1.05)
Never like		402 (5.5)	290	88	24		2.29	1.74	3.71	2.30
			(4.7)	(9.3)	(16.4)	<0.001[a]	(1.75-3.00)	(1.28-2.37)	(2.22-6.19)	(1.20-4.38)
School performance	7,298									
Well		4,013	3,488	448	77		Ref	Ref	Ref	Ref
		(55.0)	(56.2)	(47.5)	(53.5)					
Average		3,195	2,668	468	59		1.36	1.41	1.00	1.05
		(43.8)	(43.0)	(49.6)	(41.0)		(1.19-1.57)	(1.21-1.65)	(0.71-1.41)	(0.70-1.58)
Poorly		90 (1.2)	54	28	8		3.95	2.96	6.93	3.81
			(0.9)	(3.0)	(5.6)	<0.001[a]	(2.47-6.31)	(1.77-4.96)	(3.23-14.84)	(1.21-12.00)
Picked on in past year	6,978									
No		4,111	3,573	466	71		Ref	Ref	Ref	Ref
		(58.8)	(59.9)	(53.1)	(53.8)					
Yes		2,867	2,394	412	61		1.32	1.31	1.27	0.99
		(41.1)	(40.1)	(46.9)	(46.2)	<0.001[a]	(1.14-1.52)	(1.12-1.53)	(0.90-1.79)	(0.64-1.53)

Picked on other in past year	6,872								
No		5,819 (84.7)	5,045 (86.0)	681 (78.0)	92 (70.8)		Ref	Ref	Ref	Ref
Yes		1,053 (15.3)	822 (14.0)	192 (22.0)	38 (29.2)	<0.001[a]	1.73 (1.45-2.07)	1.48 (1.22-1.80)	2.55 (1.74-3.75)	2.04 (1.24-3.35)

Abbreviations: BACA, Below average cognitive ability; TCD, Typical cognitive development; LCA, Low cognitive ability; CI, Confidence interval
[a]Pearson's Chi-squared test, [b]Adjusted for child gender, gestational age, siblings, primary caregiver education, household income, parent-child relationship, child BMI, cultural background, child temperament, partner in household, maternal mental health

Table A9.4 Child's self-concept at age 9 years

	Valid	Total	Typical cognitive development (GCA ≥ 85) n=6,286	Below average cognitive ability (GCA 70-85) n=952	Low cognitive ability (GCA <70) n=154	p-value	BACA v TCD Odds ratio (95% CI)	BACA v TCD Adjusted odds ratio[b] (95% CI)	LCA v TCD Odds ratio (95% CI)	LCA v TCD Adjusted odds ratio[b] (95% CI)
Overall self-concept	6,702									
Average or high		4,860 (72.5)	4,236 (73.8)	568 (67.5)	56 (45.2)		Ref	Ref	Ref	Ref
Low		1,842 (27.5)	1,501 (26.2)	273 (32.5)	68 (54.8)	<0.001[a]	1.36 (1.16-1.59)	1.12 (0.94-1.34)	3.39 (2.37-4.85)	2.17 (1.38-3.41)
Intellectual and school status	6,671									
Average or high		4,960 (74.4)	4,335 (75.9)	563 (67.5)	62 (50.0)		Ref	Ref	Ref	Ref
Low		1,711 (25.6)	1,378 (24.1)	271 (32.5)	62 (50.0)	<0.001[a]	1.51 (1.29-1.77)	1.30 (1.09-1.55)	3.15 (2.21-4.50)	2.25 (1.44-3.52)
Behavioural adjustment	6,845									
Average[g]		4,388 (64.1)	3,841 (65.6)	496 (57.9)	51 (39.2)		Ref	Ref	Ref	Ref
Low		2,457 (35.9)	2,018 (34.4)	360 (42.0)	79 (60.8)	<0.001[a]	1.38 (1.19-1.60)	1.18 (1.00-1.38)	2.96 (2.08-4.23)	2.42 (1.55-3.79)
Physical appearance and attributes	6,799									

Average		5,155	4,449	628	78		Ref	Ref	Ref	Ref
		(75.8)	(76.6)	(72.8)	(61.4)					
Low		1,644	1,360	235	49		1.23	1.12	2.06	1.64
		(24.2)	(23.4)	(27.2)	(38.6)	<0.001[a]	(1.04-1.44)	(0.94-1.34)	(1.43-2.96)	(1.04-2.60)
Freedom from anxiety	6,878									
Average or high		5,317	4,584	654	79		Ref	Ref	Ref	Ref
		(77.3)	(77.9)	(75.4)	(60.8)					
Low		1,561	1,297	213	51		1.15	0.99	2.30	1.40
		(22.7)	(22.1)	(24.6)	(39.2)	<0.001[a]	(0.97-1.36)	(0.82-1.19)	(1.61-3.28)	(0.88-2.23)
Popularity	6,810									
Average or high		5,214	4,524	620	70		Ref	Ref	Ref	Ref
		(76.6)	(77.7)	(72.2)	(54.3)					
Low		1,596	1,298	239	59		1.34	1.12	2.96	2.20
		(23.4)	(22.3)	(27.8)	(45.7)	<0.001[a]	(1.14-1.58)	(0.94-1.34)	(2.08-4.20)	(1.41-3.42)
Happiness and satisfaction	6,877									
Average[c]		4,562	3,997	518	47		Ref	Ref	Ref	Ref
		(66.3)	(68.0)	(59.9)	(35.9)					
Low		2,315	1,885	347	84		1.42	1.17	3.80	2.96
		(33.7)	(32.0)	(40.1)	(64.1)	<0.001[a]	(1.23-1.64)	(1.00-1.38)	(2.65-5.46)	(1.90-4.62)

Abbreviations: BACA, Below average cognitive ability; TCD, Typical cognitive development; LCA, Low cognitive ability; CI, Confidence interval
[a]Pearson's Chi-squared test, [b]Adjusted for child gender, gestational age, siblings, primary caregiver education, household income, parent-child relationship, child BMI, cultural background, child temperament, partner in household, maternal mental health, [c]Where 'Average' is listed alone no child scored in the 'high average' or 'high' categories

Table A9.5 Teacher-reported classroom engagement and behaviour at age 9 years

	Valid	Total	TCD (GCA ≥ 85) n=6,286	BACA (GCA 70-85) n=952	LCA (GCA <70) n=154	p-value	BACA v TCD Odds ratio (95% CI)	BACA v TCD Adjusted odds ratio[b] (95% CI)	LCA v TCD Odds ratio (95% CI)	LCA v TCD Adjusted odds ratio[b] (95% CI)
Shows an interest in classroom activities	6,727									
Always/almost always		5,174 (76.9)	4,542 (79.1)	553 (65.0)	79 (58.5)		Ref			
Sometimes		1,406 (20.9)	1,094 (19.1)	263 (30.9)	49 (36.3)		1.97 (1.68-2.32)	1.49 (1.24-1.79)	2.58 (1.80-3.70)	1.51 (0.95-2.41)
Hardly ever/never		147 (2.2)	105 (1.8)	35 (4.1)	7 (5.2)	<0.001[a]	2.76 (1.87-4.08)	1.90 (1.20-3.00)	3.85 (1.74-8.51)	2.84 (1.13-7.11)
High level of involvement in self-chosen activities	6,732									
Always/almost always		5,071 (75.3)	4,468 (77.8)	535 (62.8)	68 (50.4)		Ref			
Sometimes		1,476 (21.9)	1,159 (20.2)	263 (30.9)	54 (40.0)		1.90 (1.61-2.23)	1.56 (1.30-1.87)	3.05 (2.13-4.39)	1.82 (1.14-2.90)
Hardly ever/never		185 (2.7)	118 (2.1)	54 (6.3)	13 (9.6)	<0.001[a]	3.81 (2.73-5.32)	2.62 (1.78-3.86)	6.92 (3.68-12.99)	5.69 (2.73-11.84)
Uses activities and resources independently	6,727									
Always/almost always		4,135 (61.5)	3,702 (64.5)	394 (46.1)	39 (28.7)		Ref			

Sometimes		2,254	1,797	379	78		1.98	1.65	4.14	2.75
		(33.5)	(31.3)	(44.4)	(57.4)		(1.70-2.31)	(1.39-1.96)	(2.81-6.11)	(1.71-4.40)
Hardly ever/never		338	238	81	19		3.20	2.23	7.51	5.79
		(5.0)	(4.1)	(9.5)	(14.0)	<0.001[a]	(2.44-4.21)	(1.63-3.07)	(4.26-13.23)	(2.91-11.52)
Interested, motivated, and excited to learn	6,724									
Always/almost always		4,379	3,918	406	55		Ref			
		(65.1)	(68.3)	(47.8)	(40.4)					
Sometimes		2,000	1,575	364	61		2.23	1.84	2.75	2.22
		(29.7)	(27.4)	(42.9)	(44.9)		(1.91-2.60)	(1.55-2.19)	(1.90-3.97)	(1.39-3.54)
Hardly ever/never		345	246	79	20		3.11	2.08	5.78	4.34
		(5.1)	(4.3)	(9.3)	(14.7)	<0.001[a]	(2.37-4.10)	(1.51-2.87)	(3.41-9.81)	(2.19-8.59)
Confident to try new things, initiate ideas and speak in familiar group	6,721									
Always/almost always		4,028	3,609	363	56		Ref			
		(59.9)	(62.9)	(42.6)	(42.1)					
Sometimes		2,193	1,785	363	45		2.02	1.72	1.64	1.61
		(32.6)	(31.1)	(42.6)	(33.8)		(1.73-2.36)	(1.45-2.04)	(1.11-2.44)	(0.99-2.61)
Hardly ever/never		500	342	126	32		3.66	2.47	6.13	4.92
		(7.4)	(6.0)	(14.8)	(24.1)	<0.001[a]	(2.91-4.61)	(1.89-3.23)	(3.92-9.58)	(2.78-8.69)
Maintains attention and concentrates	6,719									
Always/almost always		3,813	3,426	346	41		Ref			
		(56.7)	(59.8)	(40.6)	(30.4)					
Sometimes		2,420	1,950	397	73		2.02	1.75	3.10	1.75
		(36.0)	(34.0)	(46.6)	(54.1)		(1.73-2.36)	(1.47-2.09)	(2.12-4.55)	(1.09-2.80)

Hardly ever/never		487	356	109	21		3.05	2.30	4.91	3.35
		(7.2)	(6.2)	(12.8)	(15.6)	<0.001[a]	(2.39-3.88)	(1.73-3.05)	(2.87-8.39)	(1.76-6.37)
Perseveres when problem solving	6,714									
Always/almost always		3,676 (54.8)	3,324 (58.0)	319 (37.5)	33 (25.2)		Ref			
Sometimes		2,397	1,958	383	56		2.04	1.78	2.91	2.14
		(35.7)	(34.2)	(45.0)	(42.7)		(1.88-4.48)	(1.49-2.12)	(1.88-4.48)	(1.26-3.65)
Hardly ever/never		641	450	149	42		3.46	2.43	9.49	6.60
		(9.5)	(7.9)	(17.5)	(32.1)	<0.001[a]	(2.79-4.31)	(1.89-3.13)	(5.95-15.13)	(3.62-12.03)

Abbreviations: BACA, Below average cognitive ability; TCD, Typical cognitive development; LCA, Low cognitive ability; CI, Confidence interval
[a]Pearson's Chi-squared test, [b]Adjusted for child gender, gestational age, siblings, primary caregiver education, household income, parent-child relationship, child BMI, cultural background, child temperament, partner in household, maternal mental health

Table A9.6 Teacher-reported emotional behavioural function at age 9 years

	Valid	Total	TCD (GCA score ≥ 85) n=6,286	BACA (GCA score 70-85) n=952	LCA (GCA score <70) n=154	p-value	BACA v TCD Odds ratio (95% CI)	BACA v TCD Adjusted odds ratio[b] (95% CI)	LCA v TCD Odds ratio (95% CI)	LCA v TCD Adjusted odds ratio[b] (95% CI)
Total Difficulties	6,739									
Average		5,680 (84.3)	4,923 (85.6)	663 (77.6)	94 (69.1)		Ref	Ref	Ref	Ref
High		1,059 (15.7)	826 (14.4)	191 (22.4)	42 (30.9)	<0.001[a]	1.71 (1.44-2.05)	1.27 (1.03-1.56)	2.64 (1.82-3.83)	1.95 (1.20-3.17)
Emotional	6,740									
Average		5,632 (83.6)	4,853 (84.4)	675 (79.0)	104 (76.5)		Ref	Ref	Ref	Ref
High		1,108 (16.4)	897 (15.6)	179 (21.0)	32 (23.5)	<0.001[a]	1.38 (1.14-1.67)	1.16 (0.95-1.42)	2.23 (1.50-3.33)	1.89 (1.19-3.02)
Conduct	6,740									
Average		6,118 (90.8)	5,242 (91.2)	753 (88.2)	13 (90.4)		Ref	Ref	Ref	Ref
High		622 (9.2)	508 (8.8)	101 (11.8)	13 (9.6)	0.019[a]	1.38 (1.10-1.73)	0.97 (0.74-1.26)	1.08 (0.60-1.93)	0.58 (0.27-1.24)
Hyperactivity	6,739									
Average		5,641 (83.7)	4,895 (85.1)	662 (77.5)	84 (61.8)		Ref	Ref	Ref	Ref
High		1,098 (16.3)	854 (14.9)	192 (22.5)	52 (38.2)	<0.001[a]	1.66 (1.39-1.98)	1.29 (1.05-1.59)	3.52 (2.47-5.01)	2.58 (1.60-4.15)

Peer problems	6,739									
Average		5,799	5,000	700	99		Ref	Ref	Ref	Ref
		(86.1)	(87.0)	(82.0)	(73.3)					
High		940	750	154	36		1.46	1.06	2.42	2.29
		(13.9)	(13.0)	(18.0)	(26.7)	<0.001[a]	(1.21-1.77)	(0.85-1.32)	(1.64-3.57)	(1.44-3.64)
Prosocial	6,737									
Average		5,888	5,062	717	109					
		(87.4)	(88.1)	(84.2)	(80.1)					
High		849	687	135	27		1.39	1.00	1.83	0.92
		(12.6)	(11.9)	(15.8)	(19.9)	<0.001[a]	(1.14-1.70)	(0.79-1.26)	(1.19-2.82)	(0.52-1.63)

Abbreviations: BACA, Below average cognitive ability; TCD, Typical cognitive development; LCA, Low cognitive ability; CI, Confidence interval.
[a]Pearson's Chi-squared test, [b]Adjusted for child gender, gestational age, siblings, primary caregiver education, household income, parent-child relationship, child BMI, cultural background, child temperament, partner in household, maternal mental health

Table A9.7 Data dictionary

Respondent	Format of questionnaire	Age	Question	Potential Responses	Scoring or recoding if relevant
Covariates					
Parent	Self-complete[a]	9 months	**Center for Epidemiological Studies Depression Scale (CES-D)** a) I felt I could not shake off the blues even with help from my family or friends b) I felt depressed c) I thought my life had been a failure d) I felt fearful e) My sleep was restless f) I felt lonely g) I had crying spells h) I felt sad	Rarely or none of the time (less than 1 day) (0) Some or a little of the time (1-2 days) (1) Occasionally or a moderate amount of the time (3-4 days) (2) Most or all of the time (5-7 days) (3)	Score range 0-24
Parent	Interviewer administered[b]	9 months	After how many weeks of pregnancy was baby born?	25 – 25 weeks or less Otherwise numeric	
Parent	Interviewer administered	9 months	Can you tell me what is your ethnic or cultural background?	Irish Other white African/other black Chinese/other Asian Other	
Parent	Self-complete	5 years	**Adapted Short Temperament Scale** a) This child is shy with strange adults b) When this child starts a project such as a puzzle or model, he/she works on it	Almost never (1) Not often (2) Variable, usually does not (3)	Each of the three subscales of sociability, persistence and

			without stopping until it is completed, even if it takes a long time c) If this child wants a toy or sweet while shopping, he/she will easily accept something else instead d) This child is shy when first meeting new children e) This child likes to complete one task or activity before going onto the next f) When this child is angry about something, it is difficult to side-track him/her g) When in a park or visiting, this child will go up to strange children and join in their play h) This child stays with an activity(e.g. puzzle, construction kit, reading) for a long time i) When shopping together, if I do not buy what this child wants (e.g. sweets, clothing) he/she cries and yells j) When unknown adults visit our home, this child is immediately friendly and approaches them k) If this child is upset, it is hard to comfort him/her	Variable usually does (4) Frequently (5) Almost always (6)	reactivity consists of 4-items scored from 1-6. The average score from the four items is used as the score for that subscale.

			l) When a toy or game becomes difficult, this child quickly turns to another activity		
Parent	Interviewer administered	5 years	Do you have a spouse/partner who lives here with you in household?	Yes No	
Parent	Interviewer administered	5 years	Does child have any brothers or sisters?	Yes No	
Parent	Interviewer administered	5 years	What is the highest level of education (full-time or part-time) which you have completed to date?	No formal education Primary education Lower secondary Upper secondary Technical or vocational qualification National certificate, Diploma, Nursing Diploma (Non-Degree) Primary Degree Professional Qualification of at least degree status Both a degree and a professional qualification Postgraduate certificate or diploma Postgraduate degree (Masters) Postgraduate doctorate (PhD)	Lower secondary or less (a,b,c) Secondary complete (d) Technical or vocational (e) Certificate or diploma (f) Primary degree (g,h,i) Post degree qualification (j,k,l)
Parent	Derived	5 years	**Household social class**	1. Professional workers	

			A social class classification was generated for both the primary and secondary caregiver based on details they provided in the questionnaire about their occupation from current or previous employment. The household's social class was taken as the highest social class category of both caregivers in the household. Where both caregivers were economically inactive and had not had any previous employment they were classified as 'validly no social class'.	2. Managerial and technical 3. Non-manual 4. Skilled manual 5. Semi-skilled 6. Unskilled 7. All others gainfully occupied and unknown 8. Validly no social class	
Parent	Derived	5 years	**Equivalised household income** An equivalence scale was used to assign a weight of 1 to the first adult in the household, and a weight of 0.66 to each subsequent adult (age 14+) living in the household, with a weight of 0.33 assigned to each child (age less than 14). The sum of the weights was used as the equivalised size. Disposable household income, recorded as total gross household income reported less statutory deductions of income tax and social insurance contributions, was divided by the equivalised size to provide a measure of equivalised household income.		Quintiles Range 1-5 (Lowest – Highest)
Measured	Derived	5 years	**Child BMI** The child's weight was measured by the interviewer using a medically approved digital scale and height was measured using a	Numeric	

			standardised measuring stick. BMI was derived from these variables.		
Parent	Self-complete	5 years	**Pianta Child Parent Relationship Scale** a) I share an affectionate, warm relationship with my child b) My child and I always seem to be struggling with each other c) If upset, my child will seek comfort from me d) My child is uncomfortable with physical affection or touch from me e) My child values his/her relationship with me f) When I praise my child he/she beams with pride g) My child spontaneously shares information about his/herself h) My child easily becomes angry at me i) It is easy to be in tune with what my child is feeling j) My child remains angry or is resistant after being disciplined k) Dealing with my child drains my energy l) When my child is in a bad mood I know we're in for a long and difficult day m) My child's feelings toward me can be unpredictable or can change suddenly n) My child is sneaky or manipulative with me	Definitely does not apply (1) Does not really apply (2) Neutral not sure (3) Applies somewhat (4) Definitely applies (5)	Positive subscale score range 8-40 Conflict subscale score range 8-40

			o) My child openly shares his/her feelings and experiences with me		
Teacher	Posted questionnaire	5 years	Do any of the following limit the kind or amount of activity the Study Child can do at school? a) Physical disability or visual or hearing impairment b) Speech impairment c) Autism spectrum disorder d) General learning disability: mild e) General learning disability: moderate/severe/profound f) Specific learning difficulties (e.g. dyslexia) g) Emotional or behavioural problem (e.g. Attention Deficit Hyperactivity Disorder) h) Home environment/problems at home i) Has limited knowledge of the main language of instruction j) Discipline problems k) Poor attendance l) Other	Yes No	
Teacher	Posted questionnaire	5 years	If answered 'yes' to any of the above – Does the study child receive special help or resources in the school because of this (these) limitations?	Yes No	
Teacher	Posted questionnaire	9 years	Do any of the following limit the kind or amount of activity the Study Child can do at school? a) Physical disability or visual or hearing impairment b) Speech impairment	Yes No	

			c) Autism spectrum disorder d) General learning disability: mild e) General learning disability: moderate/severe/profound f) Specific learning difficulties (e.g. dyslexia) g) Emotional or behavioural problem (e.g. Attention Deficit Hyperactivity Disorder) h) Home environment/problems at home i) Has limited knowledge of the main language of instruction j) Discipline problems k) Poor attendance l) Other		
Outcome variables					
Child	Interviewer administered[c]	9 years	What do you think about school?	Always like it Sometimes like it Never like it	
Child	Interviewer administered	9 years	How well do you think you are doing in your school work?	Well Average/Ok Poorly	
Child	Self-complete[d]	9 years	Thinking back over the last year would you say that anyone (either a child or an adult) picked on you?	Yes No	
Child	Self-complete	9 years	Thinking back over the last year would you say that you picked on someone (either a child or an adult)	Yes No	
Child	Self- complete	9 years	Piers-Harris Self-Concept Scale Items cannot be published due to copyright	Yes No	'Very low', 'low, 'low average' categorised as 'Low'.

					'Average', 'high average', 'high' categorised as 'Average or high'.
Teacher	Posted questionnaire	9 years	In general would you say the study child shows an interest in classroom activities through observation or participation	Always/almost always Sometimes Hardly ever/Never	
Teacher	Posted questionnaire	9 years	In general would you say the study child displays high levels of involvement in self-chosen activities?	Always/almost always Sometimes Hardly ever/Never	
Teacher	Posted questionnaire	9 years	In general would you say the study child selects and uses activities and resources independently?	Always/almost always Sometimes Hardly ever/Never	
Teacher	Posted questionnaire	9 years	In general would you say the study child continues to be interested, motivated, and excited to learn?	Always/almost always Sometimes Hardly ever/Never	
Teacher	Posted questionnaire	9 years	In general would you say the study child is confident to try new activities, initiate ideas, and to speak in a familiar group?	Always/almost always Sometimes Hardly ever/Never	
Teacher	Posted questionnaire	9 years	In general would you say the study child maintains attention and concentrates?	Always/almost always Sometimes Hardly ever/Never	
Teacher	Posted questionnaire	9 years	In general would you say the study child sustains involvement and perseveres, particularly when trying to solve a problem or reach a satisfactory conclusion?	Always/almost always Sometimes Hardly ever/Never	
Teacher	Posted questionnaire	9 years	Strengths and Difficulties Questionnaire Items. Listed below is a set of statements which could be	Not true Somewhat True	Total difficulties ≥12 – High

			used to describe the Study Child's behaviour. For each item please mark box…. a) Considerate of other people's feelings b) Restless, overactive, cannot stay still for long c) Often complains of headaches, stomach-aches or sickness d) Shares readily with other children (treats, toys, pencils, etc) e) Often has temper tantrums or hot tempers f) Rather solitary, tends to play alone g) Generally obedient, usually does what adults request h) Many worries, often seems worried i) Helpful if someone is hurt, upset or feeling ill j) Constantly fidgeting or squirming k) Has at least one good friend l) Often fights with other children or bullies them m) Often unhappy, downhearted or tearful n) Generally liked by other children o) Easily distracted, concentration wanders p) Nervous or clingy in new situations, easily loses confidence q) Kind to younger children r) Often lies or cheats s) Picked on or bullied by other children	Certainly True	Emotional ≥4 – High Conduct ≥3 – High Hyperactivity ≥6 – High Peer problems ≥3 – High Prosocial ≤5 - High

			t) Often volunteers to help others (parents, teachers, other children) u) Thinks things out before acting v) Steals from home, school or elsewhere w) Gets on better with adults than with other children x) Many fears, easily scared y) Sees tasks through to the end, good attention span		

[a]Computer assisted self-completion interview was used for self-complete parental questionnaires, [b]Computer assisted personal interviewing was used for interviewer administered parent questionnaires, [c]Computer assisted personal interviewing was used for interviewer administered child questionnaires, [d]Self-complete child questionnaires were completed using a paper-based questionnaire.

Table A9.8. Child-reported experience of school at age 9 years

	Valid	Total	Typical cognitive development	Below average cognitive ability	p-value	Odds ratio (95% CI[b])	Adjusted odds ratio[c] (95% CI)	p-value
Thoughts about school	7,304							
Always like it		2,409 (33.3)	2,088 (33.6)	321 (29.5)		Ref	Ref	
Sometimes like it		4,494 (61.5)	3,837 (61.7)	657 (60.3)		1.11 (0.97-1.29)	1.01 (0.86-1.18)	0.944
Never like		401 (5.5)	290 (4.7)	111 (10.2)	<0.001[a]	2.49 (1.95-3.20)	1.82 (1.37-2.43)	<0.001
Self-reported performance	7,299							
Well		4,014 (55.0)	3,488 (56.2)	526 (48.3)		Ref	Ref	
Average		3,195 (43.8)	2,668 (43.0)	527 (48.4)		1.31 (1.15-1.49)	1.37 (1.18-1.58)	<0.001
Poorly		90 (1.2)	54 (0.9)	36 (3.3)	<0.001[a]	4.39 (2.85-6.76)	3.04 (1.85-4.98)	<0.001
Reported being picked on	6,978							
No		4,111 (58.8)	3,573 (59.9)	538 (53.2)			Ref	
Yes		2,867 (41.1)	2,394 (40.1)	473 (46.8)	<0.001[a]	1.31 (1.15-1.50)	1.27 (1.09-1.48)	0.002
Reported picking on someone	6,872							
No		5,819 (84.7)	5,045 (86.0)	774 (77.0)				
Yes		1,053 (15.3)	822 (14.0)	231 (23.0)	<0.001[a]	1.83 (1.55-2.16)	1.53 (1.27-1.84)	<0.001

[a]Pearson's Chi-squared test, [b]CI - Confidence Interval, [c]Adjusted for child gender, gestational age, siblings, primary caregiver education, household income, parent-child relationship, child BMI, cultural background, child temperament, partner in household, maternal mental health

Figure A9.1 Directed acyclic graph depicting the causal relationship between cognitive ability at age 5 and child-reported school-experience at age 9

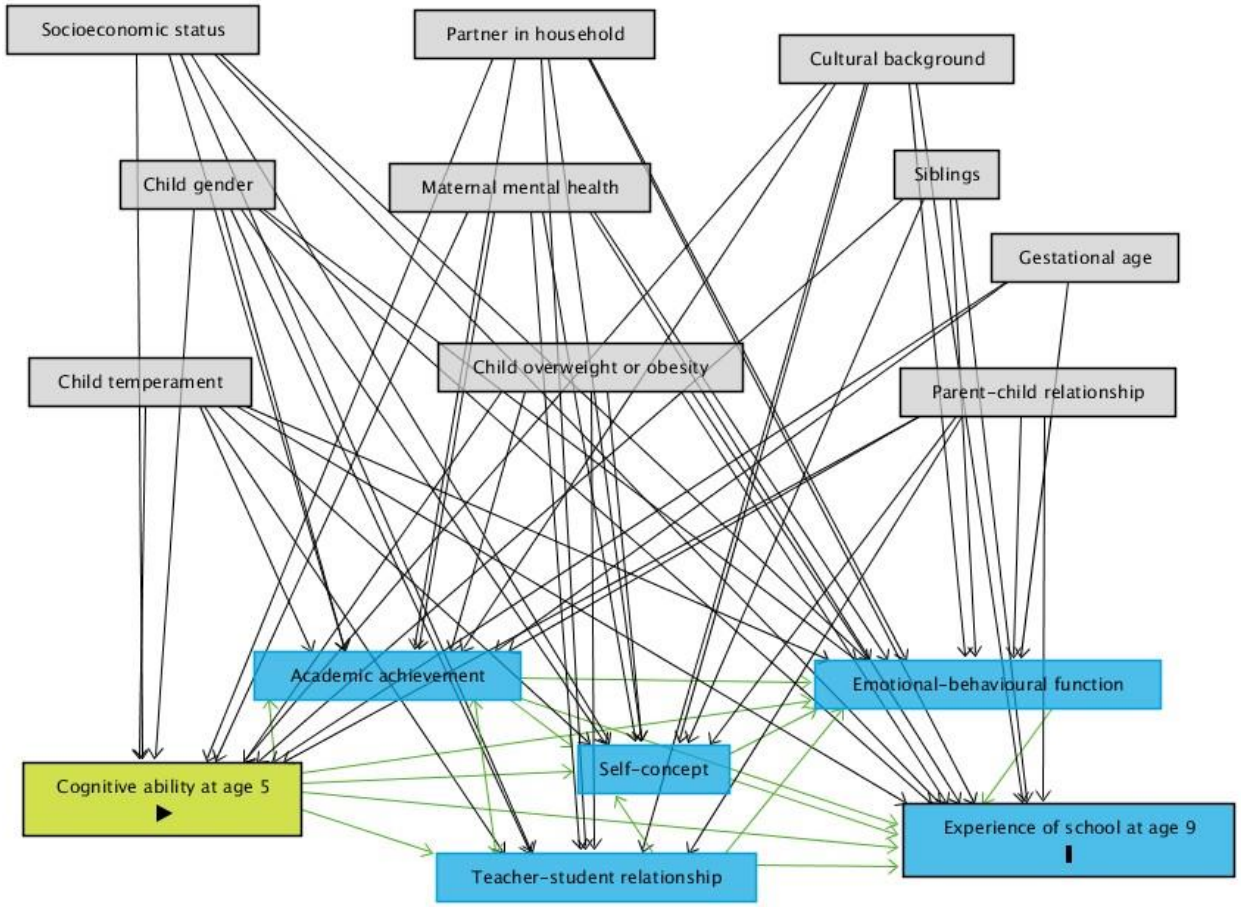


Figure S1 Legend

Exposure

Outcome

Ancestor of

Adjusted variable

Causal pathway

For clarity of presentation, lines representing relationships between confounding variables have been removed.

Estimand of interest: The total causal effect of cognitive ability at age 5 on the child-reported experience of school at age 9.

Minimally sufficient DAG implied adjustment set: Socioeconomic status, partner in household, cultural background, child gender, maternal mental health, siblings, child temperament, child overweight or obesity, parent-child relationship.

Figure A9.2 Directed acyclic graph depicting the causal relationship between cognitive ability at age 5 and child's self-concept at age 9

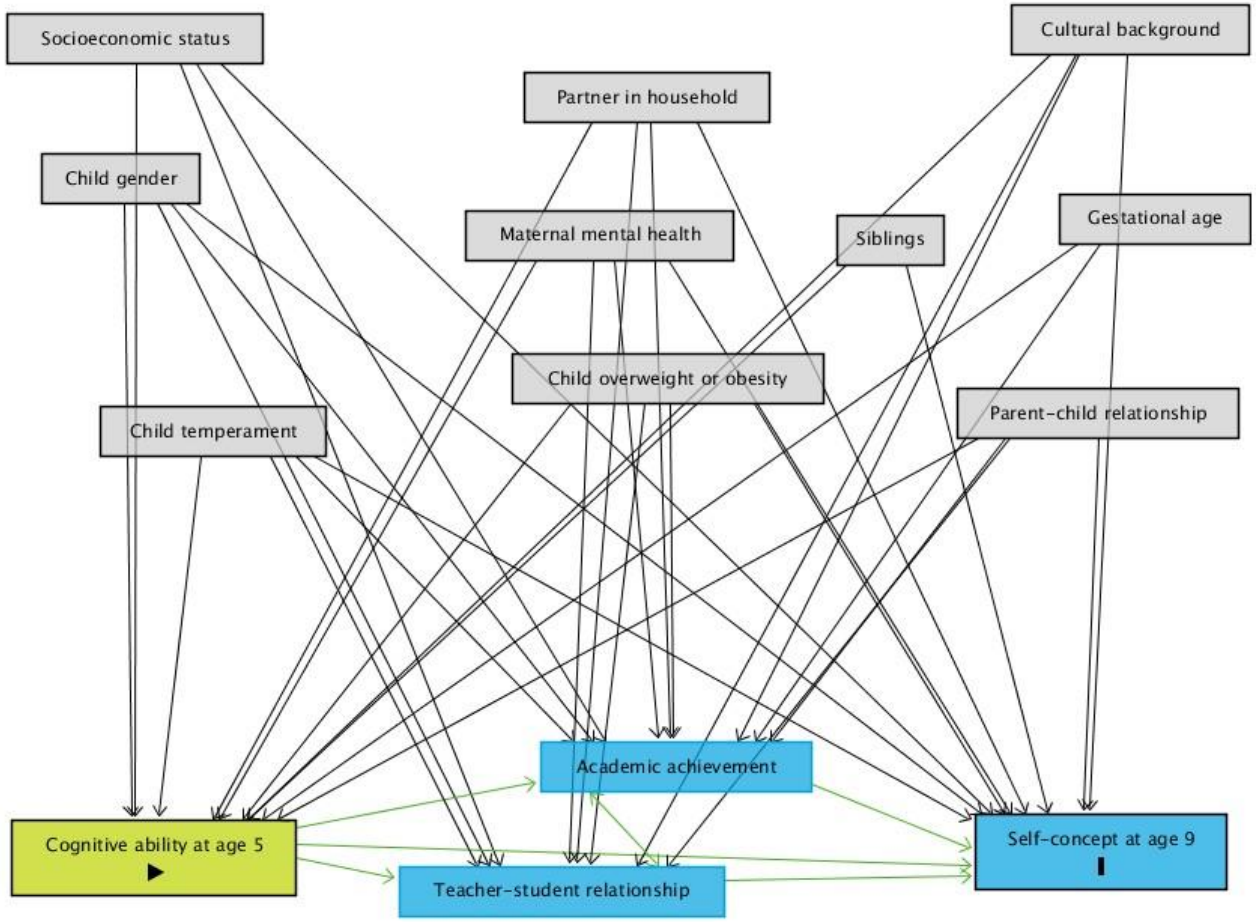


Figure S2 Legend

Exposure

Outcome

Ancestor of

Adjusted variable

Causal pathway

For clarity of presentation, lines representing relationships between confounding variables have been removed.

Estimand of interest: The total causal effect of cognitive ability at age 5 on the child self-concept at age 9

Minimally sufficient DAG implied adjustment set: Socioeconomic status, partner in household, cultural background, child gender, maternal mental health, siblings, child temperament, child overweight or obesity, parent-child relationship

Figure A9.3 Directed acyclic graph depicting the causal relationship between cognitive ability at age 5 and class engagement at age 9

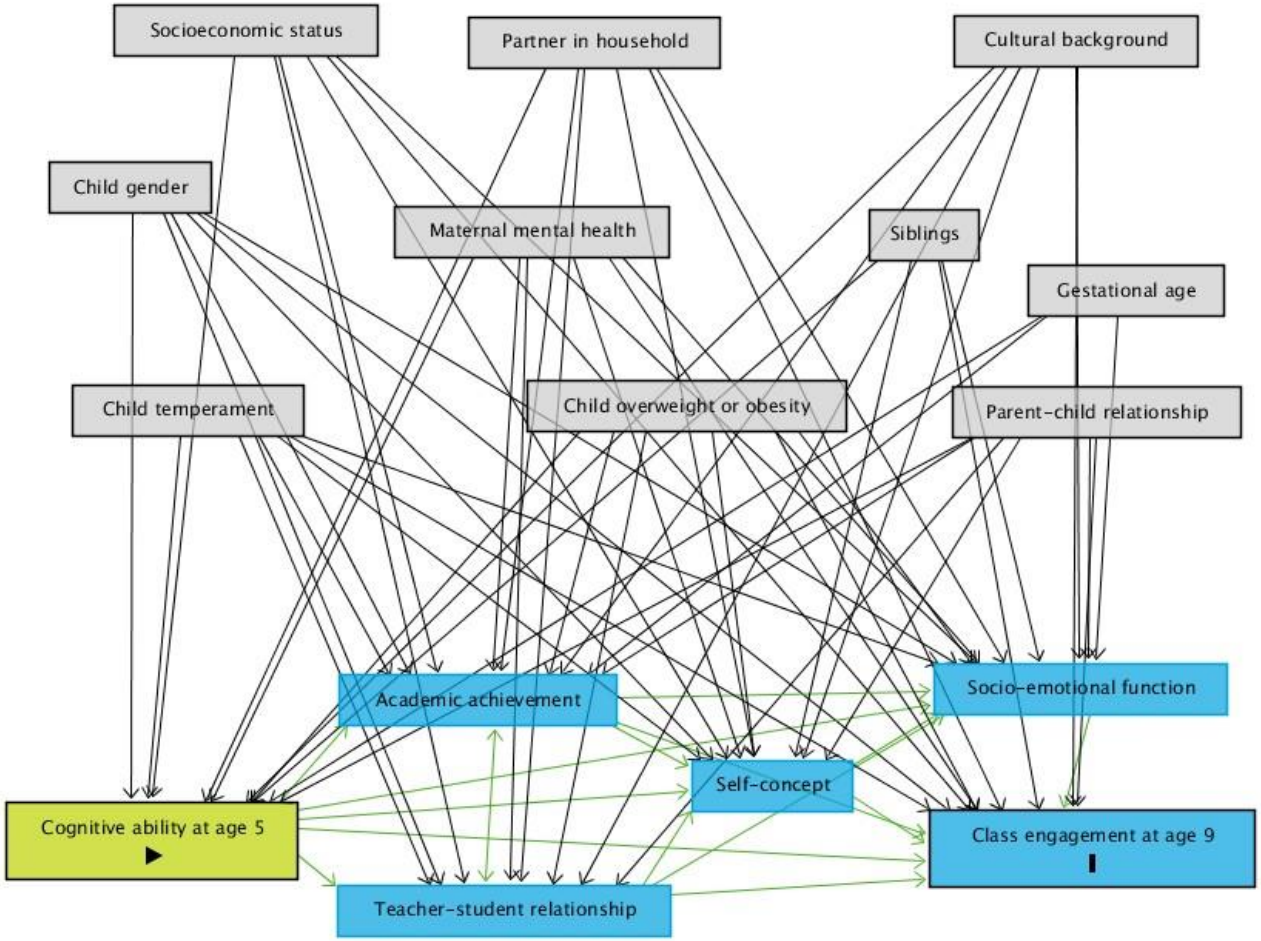


Figure S3 Legend

Exposure

Outcome

Ancestor of outcome

Adjusted variable

Causal pathway

For clarity of presentation, lines representing relationships between confounding variables have been removed.

Estimand of interest: The total causal effect of cognitive ability at age 5 on the teacher-reported class engagement at age 9

Minimally sufficient DAG implied adjustment set: Socioeconomic status, partner in household, cultural background, child gender, maternal mental health, siblings, child temperament, child overweight or obesity, parent-child relationship

Figure A9.4. Directed acyclic graph depicting the causal relationship between cognitive ability at age 5 and emotional-behavioural function at age 9

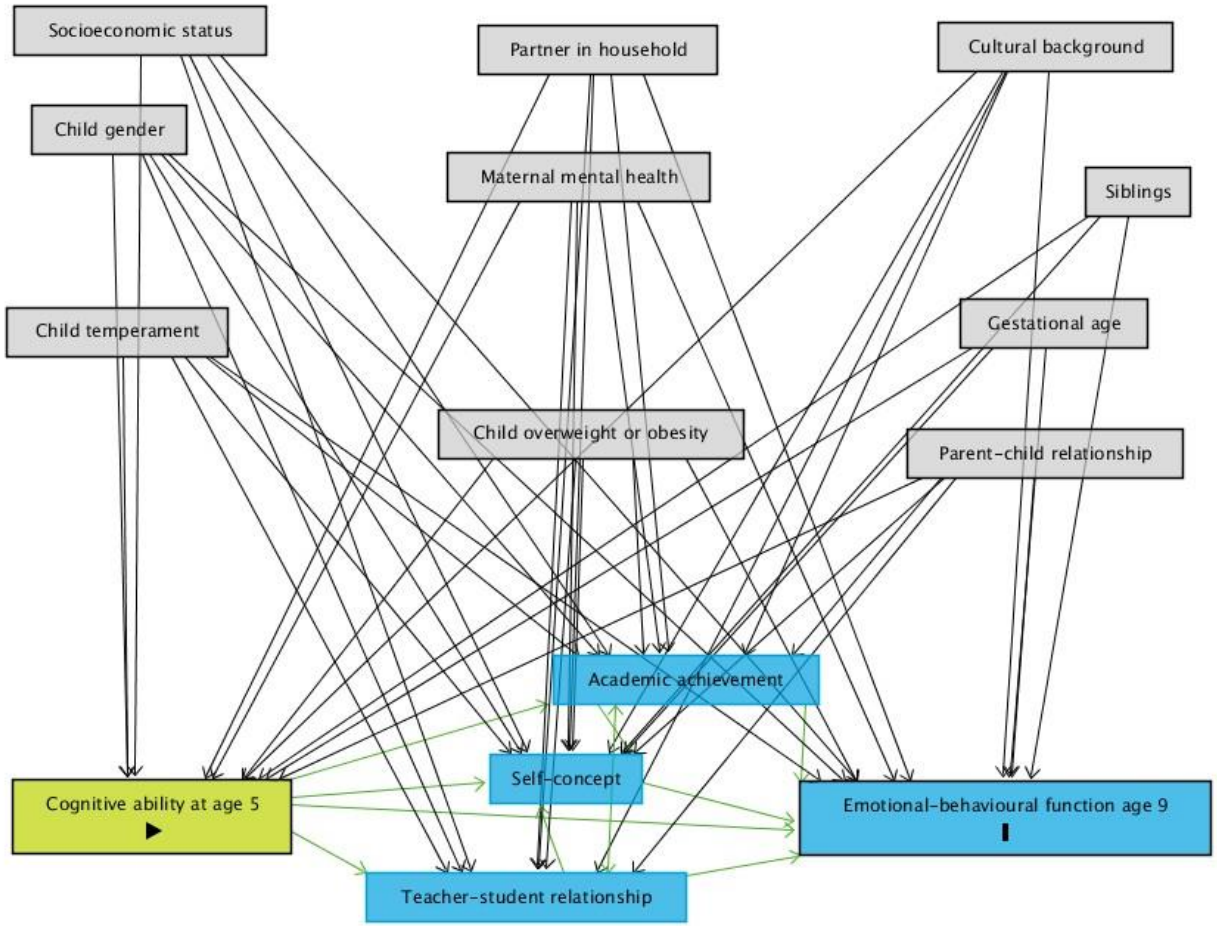


Figure S4 Legend

Exposure

Outcome

Ancestor of

Adjusted variable

Causal pathway

For clarity of presentation, lines representing relationships between confounding variables have been removed.

Estimand of interest: The total causal effect of cognitive ability at age 5 on the teacher-reported emotional-behavioural function at age 9

Minimally sufficient DAG implied adjustment set: Socioeconomic status, partner in household, cultural background, child gender, maternal mental health, siblings, child temperament, child overweight or obesity, parent-child relationship

Appendix 10 Supplementary material for Chapter 5: The Predictive Value of The Ages and Stages Questionnaire in Late Infancy for Low Average Cognitive Ability at Age 5

Table A10.1 Ages and Stages Questionnaire cut-offs for onward referral

	Communication	Gross Motor	Fine Motor	Problem Solving	Social skills
24 months	<25.17	<38.07	<35.16	<29.78	<31.54
27 months	<24.02	<28.01	<18.42	<27.62	<25.31

Table A10.2 Mean 24-month and 27-month ASQ scores by cognitive ability

	Low cognitive ability	Average/ above cognitive ability				Low cognitive ability	Average/ above cognitive ability		
	n=73	n=395				n=28	n=259		
24-month ASQ	Mean (sd)	Mean (sd)	Mean difference. (95% CI)^	p	27-month ASQ	Mean (sd)	Mean (sd)	Mean difference (95% CI)^	p
Total ASQ	244.6 (37.5)	261.2 (27.2)	16.6 (7.4-25.7)	0.001	Total ASQ	226.4 (42.2)	251.3 (29.9)	24.8 (8.1-41.5)	0.005
Communication	50.5 (12.6)	55.6 (9.0)	5.1 (2.0-8.2)	0.001	Communication	50.7 (10.9)	56.8 (6.1)	6.1 (1.8-10.4)	0.007
Gross Motor	53.4 (10.6)	55.7 (6.2)	2.3 (-0.2-4.9)	0.072	Gross Motor	48.9 (13.1)	50.5 (9.7)	1.6 (-3.6-6.8)	0.527
Fine Motor	49.4 (6.4)	51.8 (7.3)	2.4 (0.8-4.1)	0.005	Fine Motor	36.9 (11.3)	44.3 (12.3)	7.6 (2.8-11.9)	0.003
Problem Solving	43.9 (11.7)	47.4 (9.7)	3.5 (0.6-6.4)	0.017	Problem Solving	46.0 (11.8)	52.2 (8.0)	6.2 (1.5-10.9)	0.011
Social Skills	47.5 (11.1)	50.7 (8.5)	3.2 (0.5-5.9)	0.023	Social Skills	43.9 (10.5)	47.5 (9.1)	3.5 (-0.6-7.7)	0.095

95% CI, 95% Confidence Interval; p, p-value

Figure A10.1 Empirical receiver operating characteristic curve for 24-month ASQ for predicting low cognitive ability at age 5 years

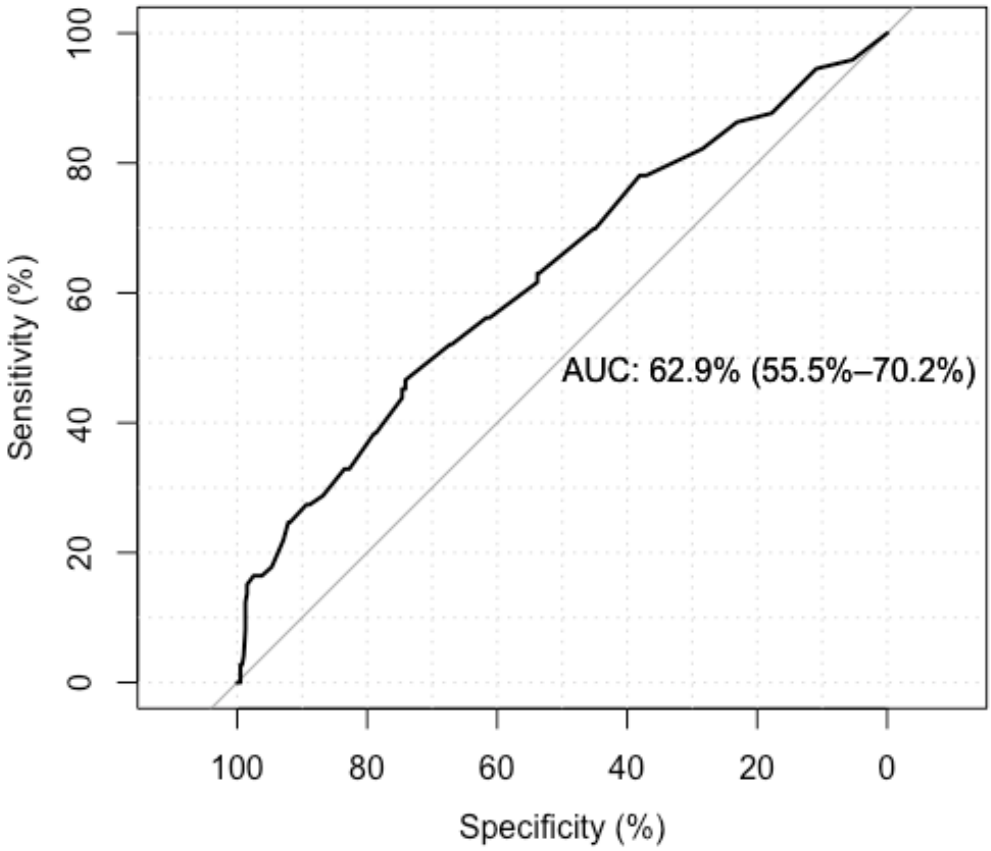


Figure A10.1 Legend
ROC curve performed using total ASQ score at 24-months for the prediction of a composite IQ score more than 1 standard deviation below the cohort mean at age 5 years. The ROC curve was generated using the "pROC" package in R. Analysis includes n=394 controls and n=73 cases. The Area Under the Curve was 0.629, 95% Confidence Interval 0.555-0.702.

Figure A10.2 Empirical receiver operating characteristic curve for 27-month ASQ for predicting low cognitive ability at age 5 years

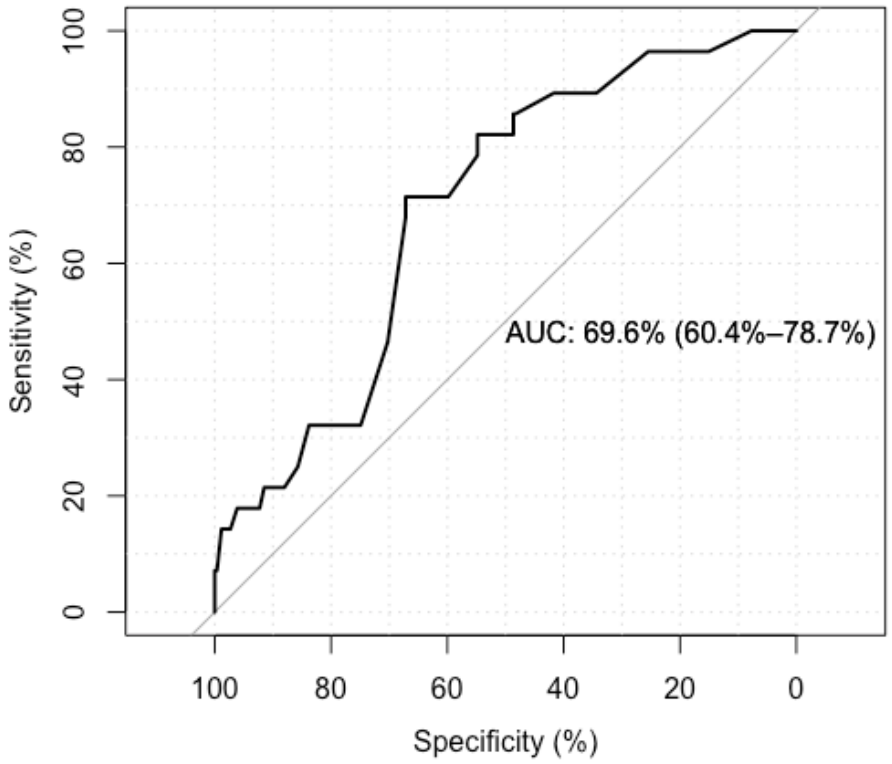


Figure A10.2 Legend
ROC curve performed using total ASQ score at 27-months for the prediction of a composite IQ score more than 1 standard deviation below the cohort mean at age 5 years. The ROC curve was generated using the "pROC" package in R. Analysis includes n=259 controls and n=28 cases. The Area Under the Curve was 0.696, 95% Confidence Interval 0.604-0.787.

Appendix 11 Supplementary material for Chapter 6: Predicting Low Cognitive Ability at Age 5 – Feature Selection Using Machine Learning Methods and Birth Cohort Data

Table A11.1 Features considered for random forest model predicting low cognitive ability at age 5

Category	**Feature**	**Time of measurement**	**Description**	**Type**
Sociodemographic	Maternal Age	15 weeks gestation or birth	Age in years	Numeric
	Maternal total years of schooling	Infant age 2 months	Total years of schooling for mother	Numeric
	Socioeconomic index	Infant age 2 months	New Zealand Socioeconomic Index Guide	Numeric
	Accommodation type	Infant age 2 months	own house/flat, private rental, council/government rental, other	Factor
	Family income	Infant age 2 months	<21k, 21-42k, 43-63k, 64-84k, 85-105k, 106-140k, >140k	Factor
	Maternal relationship status	Infant age 2 months	In a relationship, not in a relationship	Factor
	Maternal employment status	Infant age 2 months	Full time, Part time, Student, Homemaker, Unemployed/sickness	Factor
	Maternal migration history	Infant age 2 months	Mother immigrated, One or both parents immigrated, Lived in country >2gen	Factor
Behavioural/Lifestyle				
	Maternal BMI	15 weeks gestation	Measured weight in kg/measured height in m^2	Numeric
	Alcohol intake first trimester	15 weeks gestation	Units of alcohol per week in first trimester	Numeric
	Cigarettes smoked in first trimester	15 weeks gestation	Number of cigarettes per day in first trimester	Numeric

	Depression score	15 weeks gestation	Edinburgh Postnatal Depression Scale	Numeric
Birth/Delivery	Birthweight	Birth*	Birthweight in grams	Numeric
	Admission to Neonatal Intensive Care Unit	Birth	Yes, No	Factor
	High risk birth	Birth	Yes, No	Factor
	Gestational age	Birth	Gestational age in weeks	Numeric
	Apgar score 1 minute	Birth	Apgar score measured at 1 minute	Numeric
	Apgar score 5 minutes	Birth	Apgar score measured at 5 minutes	Numeric
	Head circumference at birth	Birth	Head circumference in cms	Numeric
	Infant feeding at discharge	Birth	Predominantly breastfed, Partially breastfed/mixed fed, Formula fed	Factor
	Infant Gender	Birth	Male, Female	Factor

*Day 2 after birth

Table A11.2 Sociodemographic, behavioural/lifestyle and birth/delivery characteristics of those who did and did not complete KBIT-2 at 60 months

	Valid	Completed IQ test at 60 months (n= 1,070)	Valid	Did not complete IQ test at 60 months n = (1,113)	
Sociodemographic					
Gender					
Male	1,070	545 (50.9)	1,113	555 (49.9)	
Female		525 (49.1)		558 (50.1)	0.648[a]
Relationship Status	1,070				
In a relationship		1,032 (96.4)		816 (92.5)	
Not in a relationship		38 (3.6)		66 (7.5)	<0.001[a]
Maternal migration hx	1,059		871		
Mother immigrated		102 (9.6)		183 (21.0)	
One parent immigrated		15 (1.4)		14 (1.6)	
Both parents immigrated		18 (1.7)		21 (2.4)	
Lived in country >2 generations		924 (87.3)		653 (75.0)	<0.001[a]
Total years of schooling, mean (sd)	1,070	13.41 (0.76)	882	13.16 (0.97)	<0.001[b]
University complete/attending	1,069		882		
Yes		612 (57.2)		447 (50.7)	
No		457 (42.8)		435 (49.3)	0.004[a]
Current Job Status	1,067		880		
Full time		869 (81.4)		635 (72.2)	
Part time		99 (9.3)		110 (12.5)	
Student		16 (1.5)		23 (2.6)	
Homemaker		37 (3.5)		53 (6.0)	
Unemployed/sickness benefit		46 (4.3)		59 (6.7)	<0.001[a]
Socioeconomic index, mean (sd)	1,070	44.13 (15.56)	882	40.82 (16.34)	<0.001[b]
Accommodation type	1,070		882		
Own house/flat		843 (78.8)		553 (62.8)	
Private rental		186 (17.4)		267 (30.3)	
Government/council rental		17 (1.6)		31 (3.5)	
Other		24 (2.2)		30 (3.4)	<0.001[a]
Family Income	1,042		833		
<21k		53 (5.1)		103 (12.4)	
21-42k		178 (17.1)		183 (22.0)	

43-63k		233 (22.4)		203 (24.4)	
64-84k		263 (25.2)		152 (18.2)	
85-105		164 (15.7)		110 (13.2)	
<106k		151 (14.5)		82 (9.8)	<0.001[a]
Maternal age, mean (sd), y	806	30.7 (3.9)	777	29.25 (4.82)	<0.001[b]
Behavioural/Lifestyle					
Maternal BMI, mean (sd), kg/m^2	806	24.97 (4.06)	777	24.80 (4.21)	0.415[b]
Depression score at 15 weeks gestation, mean (sd)	795	4.92 (4.26)	768	5.50 (4.35)	0.008[b]
Depression score at 15 weeks gestation, median (IQR)	795	4 (7)	768	5 (6)	0.004[c]
Units of alcohol/week 1st trimester, mean (sd)	806	4.49 (5.40)	777	5.07 (7.19)	0.070[b]
Units of alcohol/week 1st trimester, median (IQR)	806	3.0 (5.0)	777	3 (5.8)	0.926[c]
Smoking status pre pregnancy	806		777		
Non smoker		612 (75.9)		532 (68.5)	
Smoker		194 (24.1)		245 (31.5)	0.001[a]
Smoking status 1st trimester	806				
Non smoker		628 (77.9)		546 (70.3)	
Smoker		178 (22.1)		231 (29.7)	<0.001[a]
Birth/Delivery					
Apgar score at 1 minute, mean (sd)	1,061	8.5 (1.2)	1,099	8.5 (1.2)	0.892[b]
Apgar score at 1 minute, median (IQR)	1,061	9.0 (0.0)	1,099	9 (0.0)	0.873[c]
Birth head circumference, mean (sd), cm	1,055	34.90 (1.51)	1,099	34.75 (1.62)	0.031[b]
Admission to NICU	1,070				
Not Admitted		996 (93.1)		1,012 (90.9)	
Admitted		74 (6.9)		101 (9.1)	0.075[a]

[a]p-value calculated using Pearson's Chi-square test
[b]p-value calculated using Welch Two Sample t-test
[c]p-value calculated using Wilcoxon rank sum test

Figure A11.1 Correlation matrix of relevant variables

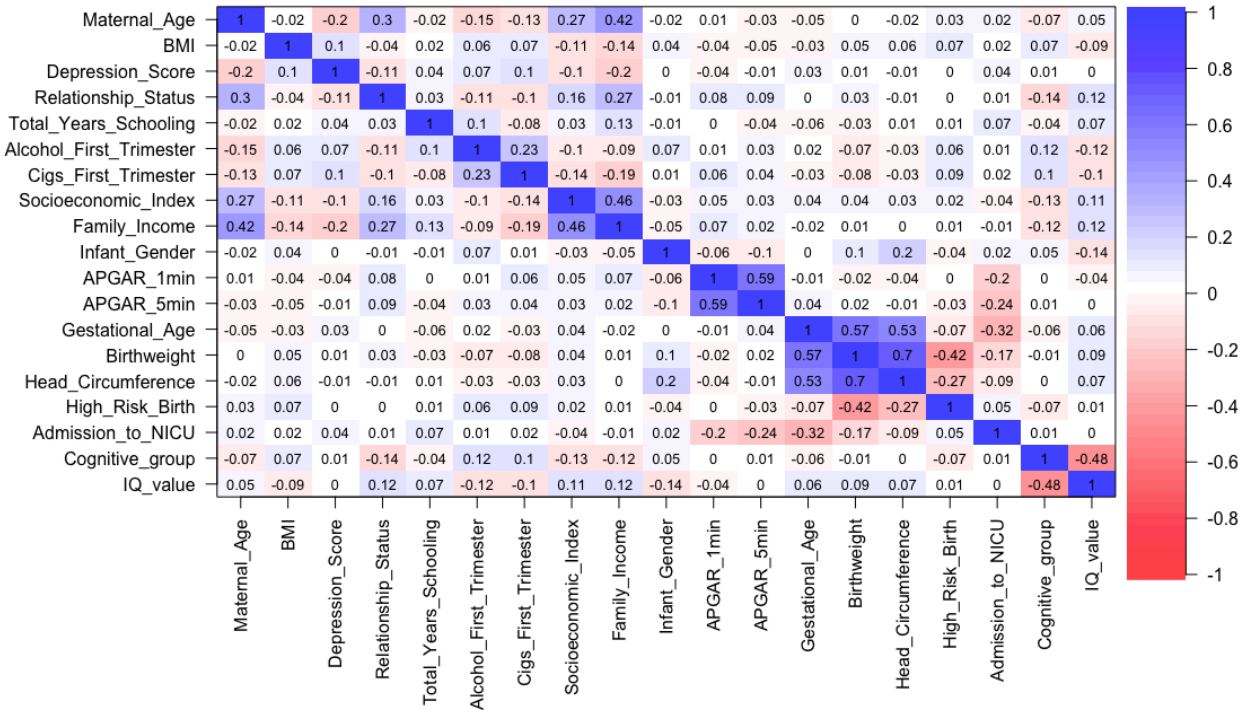


Figure A11.1 Legend

The correlation matrix shows the correlation between relevant variables on the original dataset before SMOTE was applied. Binary features coded as follows; Relationship_Status 0=not in relationship 1=in relationship; Family_Income 1=<€21,000 2=€21,000-€42,000 3=€43,000-€63,000 4=€64,000-€84,000 5=€85,000-€105,000 6=€106,000-€140,000 7=>€140,000; Infant_Gender 0=female 1=male; High_Risk_Birth 0=no 1=yes; Admission_to_NICU 0=no 1=yes; Cognitive group 0=average/above 1=low.

Low cognitive group was defined as those with an IQ ≤90 on the Kaufman Brief Intelligence Test. IQ value refers to the composite IQ value. The categorical features 'Infant_feeding_on_discharge', 'Maternal_migration_history', and 'Family_accomodation_type' are excluded.

Figure A11.2 Histogram of composite IQ scores at age 5 in the BASELINE cohort

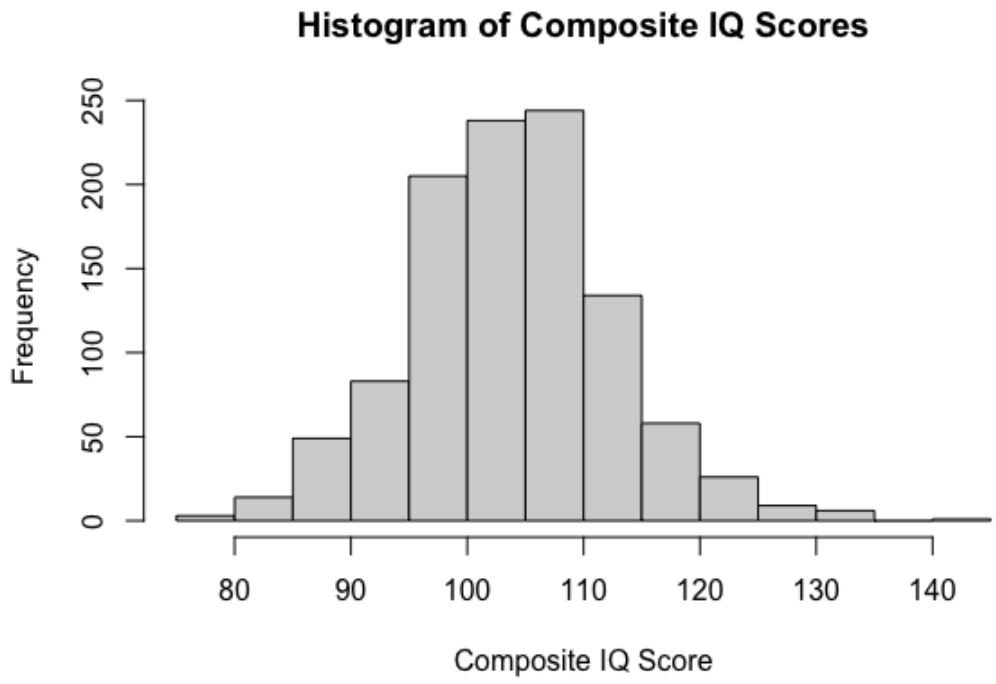


Figure A11.2 Legend
A histogram was generated based on the composite IQ scores of the 1,070 participants who completed the KBIT-2 at age 5. The mean IQ was 104.3 points (SD 8.8) and it ranged from 76-143 points.

Figure A11.3 Feature importance plot showing the mean minimal depth of the 15 most important features

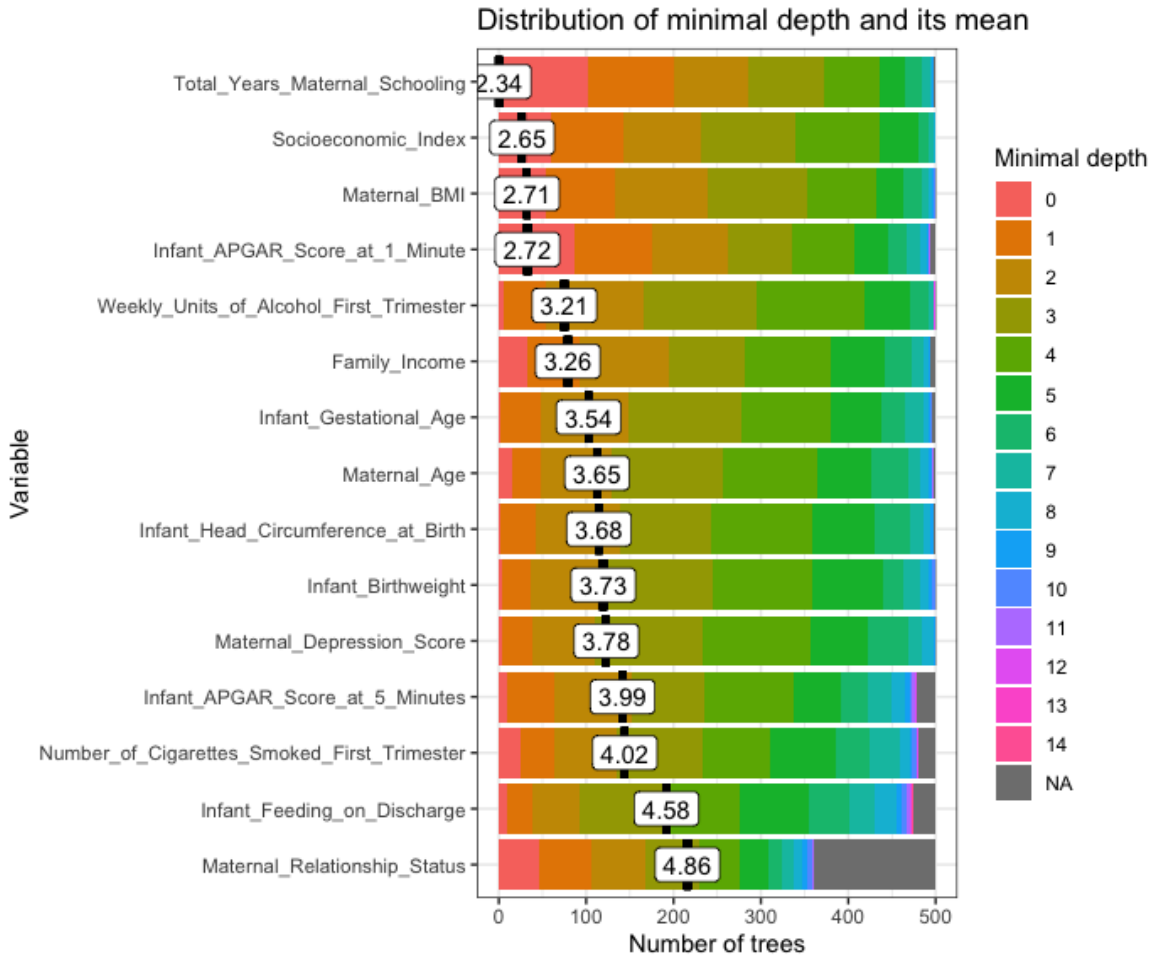


Figure A11.3 Legend

Features are ranked in descending order of importance. The mean minimal depth is the mean depth of the splitting node for that feature across all trees in the random forest. A feature that splits a large number of observations early in the decision making process is an important feature and has a low mean minimal depth.

Figure A11.4 Plot of correlations between importance rankings for 21 features according to different measures of importance

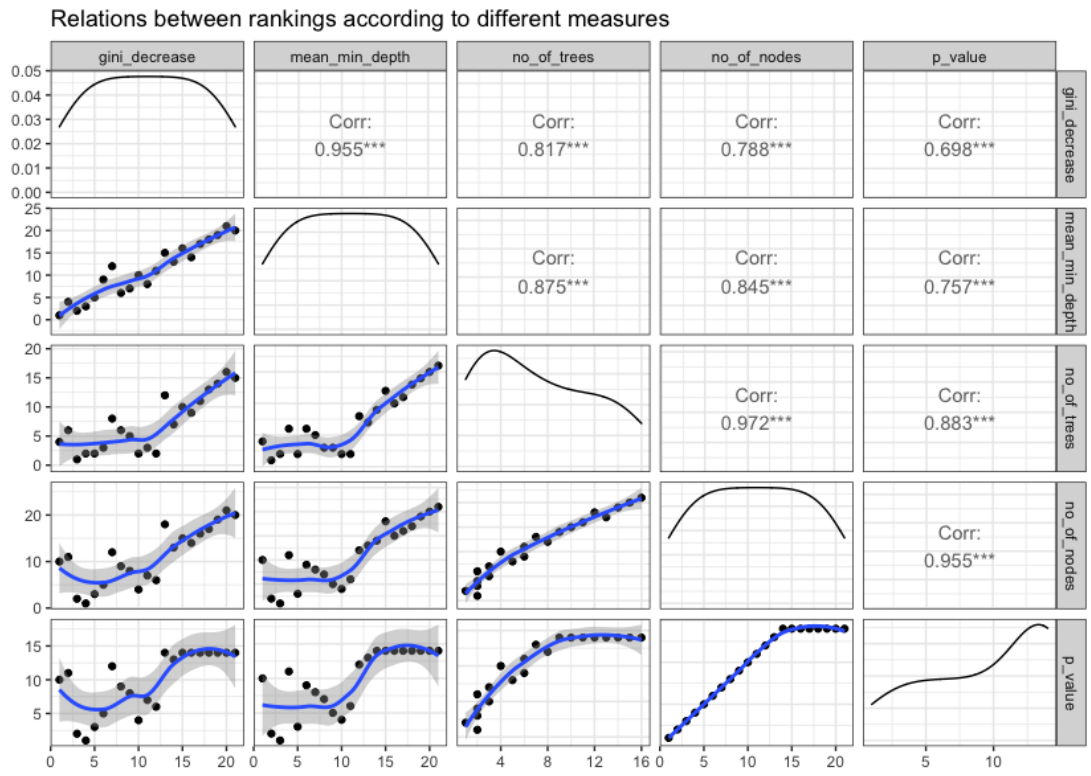


Figure A11.4 Legend: The correlations between the feature importance ranking generated for each feature with five different methods are shown both graphically and in text. The Gini decrease refers to the decrease in Gini impurity; the mean minimal depth refers to the mean depth of the splitting node for that feature across all trees in the forest; the number of trees refers to the total number of trees in which a split on that feature occurs; the number of nodes refers to the total number of nodes that use that feature for splitting and is often equal to the number of trees; the p-value indicates whether the number of nodes in which the feature was used for splitting exceeds the theoretical number if a binomial distribution was followed. The rankings according to the different measures are all strongly correlated.

Figure A11.5 Partial dependence plots showing the relationship between 13 important variables and the probability of low cognitive ability at age 5

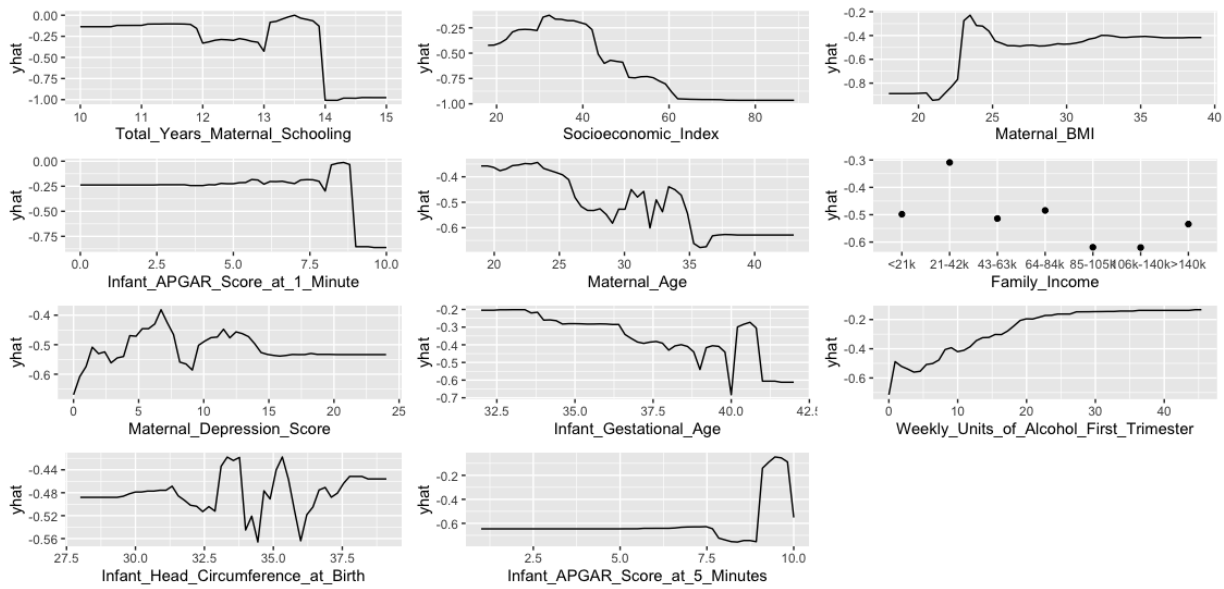
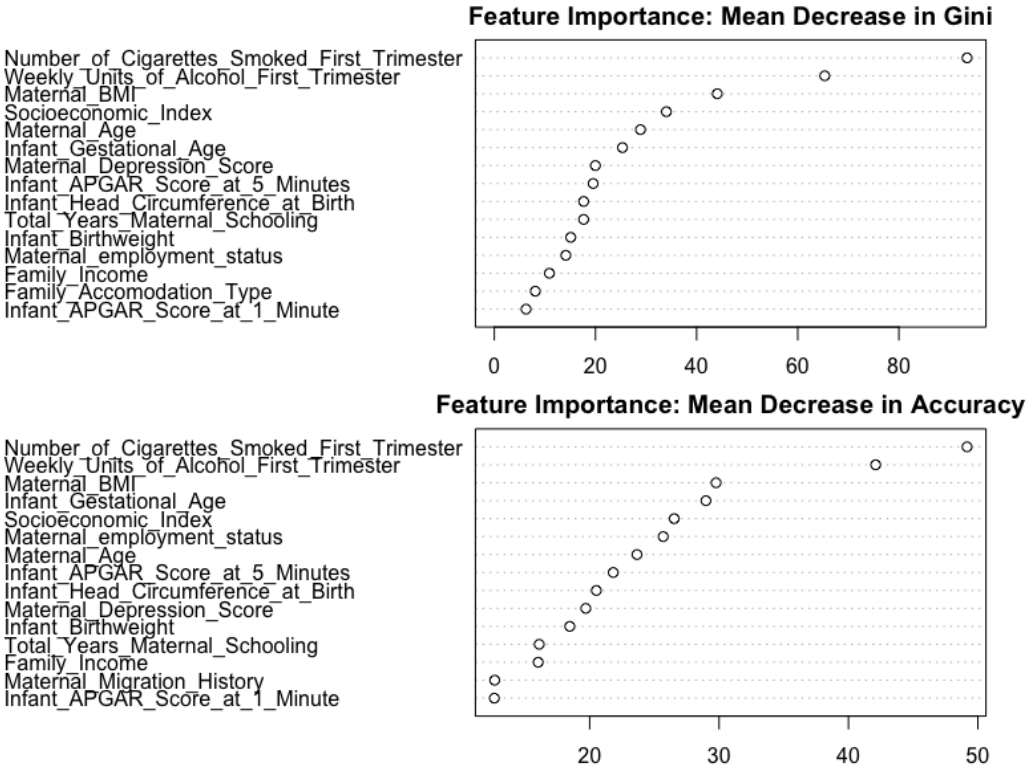


Figure A11.5 Legend

On the x-axis of the plots is the value for numeric features and the class for categorical features. On the y-axis is the value of the partial dependence function labelled yhat. The partial dependence function indicates the average marginal effect of a given feature on the predicted outcome of the random forest model. The plot demonstrates how the average prediction changes as each feature changes. The following limitations of PDP plots should be acknowledged in interpretation; PDP plots assume the feature of interest is not correlated with the other features in the model. If this assumption is violated spurious relationships may be seen. In regions with limited training data, for example Infant Apgar scores at 5 minutes <7, an accurate prediction of the marginal effect may not be possible in this range.

Figure A11.6 Feature importance plot showing the mean decrease in Gini and mean decrease in accuracy for the 15 most important features based on imputed dataset (supplemental data analysis)



Supplemental Data Analysis

Analysis of Imputed Dataset

Methods

There were 1,070 participants in the BASELINE study who completed the Kaufman composite IQ test at age 5. Imputation was performed with the 'missForest' package, which uses a random forest imputation algorithm. SMOTE was then applied to the imputed dataset to address the class imbalance.

Results

Shown below are the 15 most important features identified using mean decrease in Gini and accuracy based on this dataset. The random forest model was trained using the 'caret package' with 10-fold cross validation repeated 5 times. The optimal mtry was 4.

Based on 10-fold cross validation repeated 5 times, the accuracy of the model was 91.7%. The optimal model achieved a sensitivity of 86.1% and a specificity of 96.1%. There were some differences in the fifteen most important features identified using complete case analysis and imputation and these are shown in Figure A11.6. The 13 features identified consistently using complete case analysis, apart from maternal relationship status, were all among the 15 most important variables identified using the imputed dataset.

Supplemental Text

Hyperparameters

Hyperparameters refer to the parameters that can be "tuned" with the aim of improving model performance.[47] Default parameters are set by statistical packages but can be altered through methods such as grid-searching. The parameter 'mtry' refers to the number of features selected at random as candidate features at each split.[11] A low value can ensure heterogenous, less correlated trees, and may increase the selection of features with small or moderate effects that can contribute to predictive performance, but which otherwise may be masked by features with stronger effects.[11] A large mtry may ensure features with stronger predictive ability are selected.[11]

The number of trees in a forest, ntree, is not a tuning parameter as such but should be set sufficiently high so that each candidate feature has opportunity to be selected.[48] However it has been shown that there is likely an optimal threshold for the number of trees, beyond which there is no performance gain and only computational cost.

The parameter 'nodesize', specifies the minimum number of observations at a terminal node. A lower value increases the number of splits, and therefore the depth of the tree, whereas a higher value results in a smaller tree.[11] Sample size refers to the number of observations used for the training of each tree and replace refers to whether the observations are replaced after each sample is randomly selected.[49,50] With replacement, bootstrap samples are created whereas subsamples of the data are created if observations are not replaced.[11]

Feature Importance: Mean Minimal Depth

The minimal depth refers to the depth of the node which splits on that feature. Features with high importance are those that tend to divide large numbers of observations and these nodes tend to be present in the initial levels of a tree. The mean minimal depth averages the depth of the splitting node for that feature across all trees in the forest and a lower value is likely to indicate a more important feature.[49]

References for supplemental text

11 Couronné R, Probst P, Boulesteix AL. Random forest versus logistic regression: a large-scale benchmark experiment. BMC Bioinformatics. 2018 Jul 17;19(1):270.

47 Huang BFF, Boutros PC. The parameter sensitivity of random forests. *BMC Bioinformatics.* 2016; 17: 331. https://doi.org/10.1186/s12859-016-1228

48 Oshiro TM, Perez PS, Baranauskas JA. How Many Trees in a Random Forest? In: Perner, P. (eds) Machine Learning and Data Mining in Pattern Recognition. MLDM 2912. Lecture Notes in Computer Science (), vol 7376 Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-31537-4_13

49 Probst P, Wright M, Boulesteix A-L. Hyperparameters and tuning strategies for random forest. WIREs Data Mining Knowl Discov. 2019; 9:e1301. https://doi.org/10.1002/widm.1301

50 Aria M, Cuccurullo C, Gnasso A. A comparison among interpretative proposals for Random Forests. Machine Learning with Applications 2021; 6: 100094. https://doi.org/10.1016/j.mlwa.2021.100094

Appendix 12 Supplemental material for Chapter 7: Predicting Low Cognitive Ability at Age 5 Years using Perinatal data and Machine Learning

Table A12.1 Subtests of the British Ability Scales Second Edition Early Years Battery [21,61]

Name of Scale	**Ability being tested**
Naming vocabulary	Expressive verbal knowledge
Picture similarities	Non-verbal reasoning
Verbal comprehension	Receptive verbal knowledge
Early number concepts	Quantitative knowledge reasoning
Block building	Perceptual motor ability
Pattern construction	Spatial problem solving
Copying	Perceptual motor ability
Matching letter-like forms	Visual perceptual matching
Recall of digits forward	Short term auditory memory
Recall of objects	Short and intermediate term visual verbal memory
Recognition of pictures	Short term visual memory

[a]The total number of subtests administered is dependent on the age of the child being tested and ranges between 8-11 for children age 3 years 6 months – 5 years 11 months. The core subtests produce a measure of verbal ability (naming vocabulary and verbal comprehension), non-verbal ability (picture similarities and block building), and spatial ability (pattern construction and copying).

Table A12.2 Description of features considered in feature selection process

Feature number	Feature	Details	Type
	Pregnancy and birth		
1	Hypertension	Yes, No	Factor
2	Pre-eclampsia	Yes, No	Factor
3	Urinary tract infection	Yes, No	Factor
4	Persistent nausea and vomiting	Yes, No	Factor
5	Gestational diabetes (controlled by diet)	Yes, No	Factor
6	Gestational diabetes (controlled by insulin)	Yes, No	Factor
7	Bleeding late pregnancy	Yes, No	Factor
8	Vaginal infection	Yes, No	Factor
9	Intrauterine growth restriction	Yes, No	Factor
10	Mode delivery	Normal, forceps/suction assisted, planned c-section, emergency c-section	Factor
11	Gestational age	Gestational age in weeks	Numeric
12	Birthweight category	<1500g, 1500-2000g, 2500-4500g, >4500g	Ordinal. Treated as numeric in model.
13	Required neonatal intensive care unit (NICU) or special care baby unit (SCBU)	Yes, No	Factor
14	Breastfed on discharge from hospital	Yes, No	Factor
15	Number of people who smoked in household during pregnancy	Numeric (response of 5 = 5 or more)	Numeric
	Maternal		
16	Age	Age of PCG when infant age 9 months (response 40 = 40 or more years)	Numeric
17	Body mass index	Weight in kg/height m^2 derived from measured data	Numeric
18	Self-rated health	Excellent, very good, good, fair, poor	Ordinal. Treated as numeric in model

19	Depression score	Centre for Epidemiological Studies Depression Scale (short version)	Numeric
20	Chronic illness (includes physical, mental, disability)	Yes, No	Factor
21	Current smoker (daily or occasionally)	Yes, No	Factor
22	Alcohol intake	Never (1) < once a month (2) 1-2/month (3) 1-2 times/week (4) 3-4 times/week (5) 5-6 times/week (6) daily (7)	Ordinal. Treated as numeric in model.
23	Stress	Parental stress scale	Numeric
24	Born in Ireland	Yes, No	Factor
25	Ethnic background	Irish, any other white background, African or any other Black background, Chinese or any other Asian background, Other	Factor
	Infant		
26	Infant gender	Male, Female	Factor
	Socioeconomic		
27	Number of bedrooms	Number of separate bedrooms in accommodation	Numeric
28	Number in household who currently smoke	Number who currently smoke (response of 6 = 6 or more)	Numeric
29	Accommodation ownership	Owner occupied (with or without mortgage), being purchased from a local authority, rented from a local authority, rented from a voluntary body, rented from private landlord, living with and paying rent to parents, occupied free of rent	Factor
30	PCG employment status	Employee, self-employed, farmer, student full-time, state training scheme,	Factor

		unemployed actively looking for job, long-term sickness or disability, home duties, retired, other	
31	Household equivalised income	Calculated as disposable household income divided by equivalised household size.	Numeric
32	PCG highest education	Primary or less, lower secondary, upper secondary, technical/vocational, non-degree, degree or equivalent, post degree cert/diploma, postgrad masters or PhD	Ordinal. Treated as numeric in model.
33	Family's social class	Highest social class category of both parents in family (where relevant) based on their occupation. Professional workers (1), managerial and technical (2), non-manual (3), skilled manual (4), semi-skilled (5), unskilled (6), all others gainfully occupied (7), never worked at all (8)	Ordinal. Treated as numeric in model.
34	Connectedness in community	As a family are settled and part of community – Strongly agree, Agree, Disagree, Strongly disagree	Ordinal. Treated as numeric in model
	Early environment		
35	Siblings	Brothers or sisters living in household. Yes, No.	Factor
36	Parental sleep	Average hours of sleep nightly. Numeric (response of 3 = 3 or less)	Numeric
37	Verbal interaction with baby	Parent talks to baby when busy doing things	Ordinal. Treated as

		- Never, rarely, sometimes, often, always	numeric in model.
38	Native language of primary care giver (PCG) is English	Yes, No	Factor
39	Quality of parental attachment (PCG)	The Quality of Attachment subscale from the Maternal Postnatal Attachment Scale	Numeric
40	Level of PCG support	I get enough help, I don't get enough help, I don't get any help at all, I don't need any help, Family not living in the country	Factor
41	Books in the home	How many children's books does child have access to in the home now, including any library books – None, <10, 10-20, 21-30, >30	Factor
42	Reading	Days per week someone reads to child in the home	Numeric
43	Alphabet	Days per week someone does alphabet with child in the home	Numeric
44	Numeracy	Days per week someone does numbers or counting with child in the home	Numeric
45	Music	Days per week someone does songs, poems, or nursery rhymes with child in the home	Numeric
46	Games	Days per week someone play games (board games, jigsaws, card games etc) with child in the home	Numeric
47	Art	Days per week someone paints, draws,	Numeric

		colours, or plays with play-doh at home with child in the home	
48	Active play	Days per week someone plays active games with child (e.g. football)	Numeric
49	Average days per week learning activities performed at home	Mean of items 42-48	Numeric

All features (excluding features 41-49*) were measured at the 9-month questionnaire which was a self-complete questionnaire completed by the child's primary caregiver which in 99.7% of cases was the child's mother. *Features 41-49 were measured at the 3-year questionnaire.

Table A12.3 Feature sets used in Recursive Feature Elimination process

	Feature set 1	Feature set 2	Feature set 3
Pregnancy and Birth	Gestational age* Admission to NICU/SCBU	Hypertension Pre-eclampsia Urinary tract infection Persistent nausea/vomiting Gestational diabetes (controlled by diet) Gestational diabetes (controlled by insulin) Bleeding late pregnancy Vaginal infection Intrauterine growth restriction Mode of delivery Gestational age* Birthweight category* Admission to NICU[a]/SCBU[b] Number in household smoking during pregnancy* Breastfed on discharge from maternity hospital	Hypertension Pre-eclampsia Urinary tract infection Persistent nausea/vomiting Gestational diabetes(controlled by diet)* Gestational diabetes (controlled by insulin) Bleeding late pregnancy Vaginal infection Intrauterine growth restriction Mode of delivery Gestational age* Birthweight category* Admission to NICU/SCBU Number in household smoking during pregnancy* Breastfed on discharge from maternity hospital
Maternal	BMI* Highest level of education* Depression* Alcohol intake* Age*	BMI* Highest level of education* Depression* Alcohol intake* Age* Self-rated health* Chronic illness Smoking status	BMI* Highest level of education* Depression* Alcohol intake* Age* Self-rated health* Chronic illness Smoking status

			Stress*
Infant	Gender	Gender	Gender
Socioeconomic	Household social class* Household income* Partner in household	Household social class* Household income* Maternal employment status Number of bedrooms* Accommodation ownership* Connectedness in community* Partner in household	Household social class* Household income* Maternal employment status* Number of bedrooms* Accommodation ownership* Connectedness in community* Partner in household
Early environment		Siblings English is maternal native language*	Siblings English is maternal native language* Books in the home* Average days/week learning activities done at home* Verbal interaction* Parental sleep* Level of support* Quality of parental attachment*

*Selected in recursive feature elimination for final model; [a]Neonatal intensive care unit; [b]Special care baby unit

Table A12.4 Pregnancy, maternal, infant, socioeconomic and early environment characteristics of study population

Characteristics	Total	Low n=573 (6.5%)	Other n=8,285 (93.5%)	p-value
Pregnancy and Birth				
Mode of delivery				
Normal vaginal	5,239 (59.1)	367 (64.0)	4,875 (58.8)	
Forceps/suction assisted	1,299 (14.7)	60 (10.5)	1,238 (14.9)	
Planned c-section	1,151 (13.0)	61 (10.6)	1,090 (13.2)	
Emergency c-section	1,169 (13.2)	85 (14.8)	1,082 (13.1)	0.003[b]
Preeclampsia	620 (7.0)	35 (6.1)	585 (7.1)	0.436[b]
Gestational diabetes (controlled by diet)	195 (2.2)	16 (2.8)	179 (2.2)	0.396[b]
Gestational diabetes (controlled by insulin)	83 (0.9)	3 (0.5)	80 (1.0)	0.373[e]
IUGR[a]	179 (2.0)	12 (2.1)	167 (2.0)	1.000[b]
Admission to NICU/SCBU	1,217 (13.7)	88 (15.4)	1,129 (13.6)	0.271[b]
Gestational age – mean (sd)	39.5 (2.1)	39.3 (2.4)	39.5 (2.0)	0.035[c]
Birthweight				
<1,500g	67 (0.8)	6 (1.0)	61 (0.7)	
1500g-2500g	401 (4.5)	38 (6.6)	362 (4.4)	
2500g-4500g	8,178 (92.3)	509 (88.8)	7,670 (92.6)	
>4500g	212 (2.4)	20 (3.5)	192 (2.3)	0.0138[e]
Breastfed on discharge - yes	4,131 (46.6)	300 (52.4)	4,427 (53.4)	0.648[b]
Smoker in hsd smoking during pregnancy - yes	2,801 (31.6)	225 (39.3)	2,576 (31.1)	<0.001[b]
Maternal				
Age – mean (sd)	32.0 (5.2)	30.6 (5.7)	32.1 (5.1)	<0.001[c]
BMI – mean (sd)	25.5 (4.7)	25.9 (5.0)	25.5 (4.7)	0.072[c]
Self-rated health				
Excellent	2,798 (31.6)	159 (27.7)	2,639 (31.9)	
Very good	3,505 (39.6)	223 (38.9)	3,282 (39.6)	
Good	2,020 (22.8)	150 (26.2)	1,870 (22.6)	
Fair	475 (5.4)	35 (6.1)	440 (5.3)	
Poor	60 (0.6)	6 (1.0)	54 (0.7)	0.097[b]
Current smoker - yes	2,028 (22.9)	174 (30.4)	1,854 (22.4)	<0.001[b]
Alcohol intake				
Never	1,457 (16.4)	192 (33.5)	1,265 (15.3)	
<1/month	2,409 (27.2)	178 (31.1)	2,231 (26.9)	
1-2/month	2,398 (27.1)	119 (20.8)	2,279 (27.5)	
1-2/week	2,176 (24.6)	73 (12.7)	2,103 (25.4)	
3-4/week	353 (4.0)	10 (1.7)	343 (4.1)	
5-6/week	49 (0.6)	<5	48 (0.6)	
Daily	16 (0.2)	<5	16 (0.2)	<0.001[e]
Maternal depression – median (IQR)	1 (3)	1 (3)	1 (3)	0.0384[d]

Chronic illness - yes	1,019 (11.5)	55 (9.6)	964 (11.6)	0.159[b]
Highest level of education				
Primary or less	186 (2.1)	43 (7.5)	143 (1.7)	
Lower secondary	739 (8.3)	77 (13.4)	662 (8.0)	
Upper secondary	1,608 (18.2)	131 (22.9)	1,477 (17.8)	
Technical/vocational	1,194 (13.5)	114 (19.9)	1,080 (13.0)	
Non Degree	1,787 (20.2)	74 (12.9)	1,713 (20.7)	
Degree or equivalent	1,617 (18.3)	68 (11.9)	1,549 (18.7)	
Post degree cert/diploma	1,047 (11.8)	36 (6.3)	1,011 (12.2)	
Post degree masters or PhD	680 (7.7)	30 (5.2)	650 (7.8)	<0.001[b]
Infant				
Infant gender - male	4,479 (50.6)	347 (60.6)	4,132 (49.9)	<0.001[b]
Socioeconomic				
Household income – median (IQR)	48,000 (33,600)	31,200 (24,963)	48,000 (33,330)	<0.001[d]
Household social class				
Professional workers	1,711 (19.3)	52 (9.1)	1,659 (20.0)	
Managerial and technical	2,863 (32.3)	109 (19.0)	2,754 (33.2)	
Non-manual	1,472 (16.6)	85 (14.8)	1,387 (16.7)	
Skilled manual	1,168 (13.2)	122 (21.3)	1,046 (12.6)	
Semi-skilled	638 (7.2)	72 (12.6)	566 (6.8)	
Unskilled	119 (1.3)	15 (2.6)	104 (1.3)	
All others gainfully occupied	37 (0.4)	8 (1.4)	29 (0.4)	
Never worked	850 (9.6)	110 (19.2)	740 (8.9)	<0.001[b]
Accommodation occupancy				
Owner occupied	6,315 (71.3)	223 (38.9)	6,091 (73.5)	
Local authority housing	623 (7.0)	85 (14.8)	538 (6.5)	
Private rental	1,622 (18.3)	249 (43.5)	1,374 (16.6)	
Living with parents	146 (1.6)	8 (1.4)	138 (1.7)	
Occupied rent free	152 (1.7)	8 (1.4)	144 (1.7)	<0.001[b]
Number of bedrooms – mean (sd)	3.5 (0.9)	3.1 (0.9)	3.5 (0.9)	<0.001[c]
Connectedness in community				
Strongly agree	3,777 (42.6)	186 (32.5)	3,591 (43.3)	
Agree	4,292 (48.5)	320 (55.8)	3,972 (47.9)	
Disagree	618 (7.0)	54 (9.4)	564 (6.8)	
Strongly disagree	171 (1.9)	13 (2.3)	158 (1.9)	<0.001[b]
Early environment				
Siblings - yes	5,410 (61.1)	355 (62.0)	5,055 (61.0)	0.687[b]
Books in home				
None	29 (0.3)	6 (1.0)	23 (0.3)	
<10	585 (6.6)	111 (19.4)	474 (5.7)	
10-20	1,626 (18.4)	169 (29.5)	1,457 (17.6)	
21-30	1,562 (17.6)	105 (18.3)	1,457 (17.6)	
>30	5,056 (57.1)	182 (31.8)	4,874 (58.8)	<0.001[b]

Days/week learning activities done at home– mean (sd)	4.9 (1.2)	4.5 (1.4)	4.9 (1.2)	<0.001[c]
Level of support				
Get enough	6,446 (72.8)	344 (60.0)	6,102 (73.7)	
Don't get enough	856 (9.7)	49 (8.6)	807 (9.7)	
Don't get any	482 (5.4)	36 (6.3)	446 (5.4)	
Don't need any	458 (5.2)	51 (8.9)	407 (4.9)	
Family not in country	616 (7.0)	93 (16.2)	523 (6.3)	<0.001[b]
Quality of parental attachment – median (IQR)	44 (4)	44 (4)	44 (4)	0.529[d]
Parental sleep hours/night – median (IQR)	11 (2)	10 (1)	11 (2)	<0.001[d]
English maternal native language – yes	7,512 (84.8)	309 (53.9)	7,203 (86.9)	<0.001[b]
Verbal interaction – talk to baby while busy doing things				
Never	201 (2.3)	16 (2.8)	185 (2.2)	
Rarely	33 (0.4)	<5	30 (0.4)	
Sometimes	658 (7.4)	50 (8.7)	608 (7.3)	
Often	2,130 (24.0)	143 (25.0)	1,987 (24.0)	
Always	5,836 (65.9)	361 (63.0)	5,475 (66.1)	0.385[e]
Partner in household	7,929 (89.5)	491 (85.7)	7,438 (89.8)	0.003[b]

[a]Intrauterine growth restriction; [b]Pearson's Chi-squared test; [c]Welch two sample t-test; [d]Wilcoxon rank sum test ; [e]Fisher's Exact Test (p-value simulated where cell number too small);

Table A12.5 Average accuracy of machine learning algorithms across 10-fold cross validation on training data calculated at a decision threshold of 0.5

	Model 1 (8 features)	Model 2 (15 features)	Model 3 (24 features)
Random forest	0.85 (0.02)	0.87 (0.02)	0.87 (0.03)
Logistic regression	0.70 (0.01)	0.75 (0.01)	0.76 (0.01)
Support vector machine	0.78 (0.02)	0.81 (0.03)	0.80 (0.03)

Table A12.6 Performance of Model 2 across different decision thresholds on independent test set

Decision threshold	Sensitivity	Specificity	Accuracy
0.50	0.40	0.91	0.87
0.45	0.47	0.87	0.84
0.40	0.55	0.82	0.80
0.35	0.61	0.76	0.75
0.30	0.69	0.70	0.70
0.28[a]	0.72	0.66	0.68

[a]Optimal cut point chosen using Youden's J statistic where the optimal cut-off is the threshold that maximises the distance to the diagonal line)

Table A12.7 Contingency table showing how model would perform across 60,000 births at a decision threshold of 0.5

		Standardised cognitive test age 5		
		Low	'Normal'	
Model predictions	Low cognitive ability	1,560	5,610	7,170
	'Normal cognitive ability'	2,340	50,490	52,830
		3,900	56,100	60,000

Figure A12.1 Flowchart of study population in Growing Up in Ireland (GUI) study

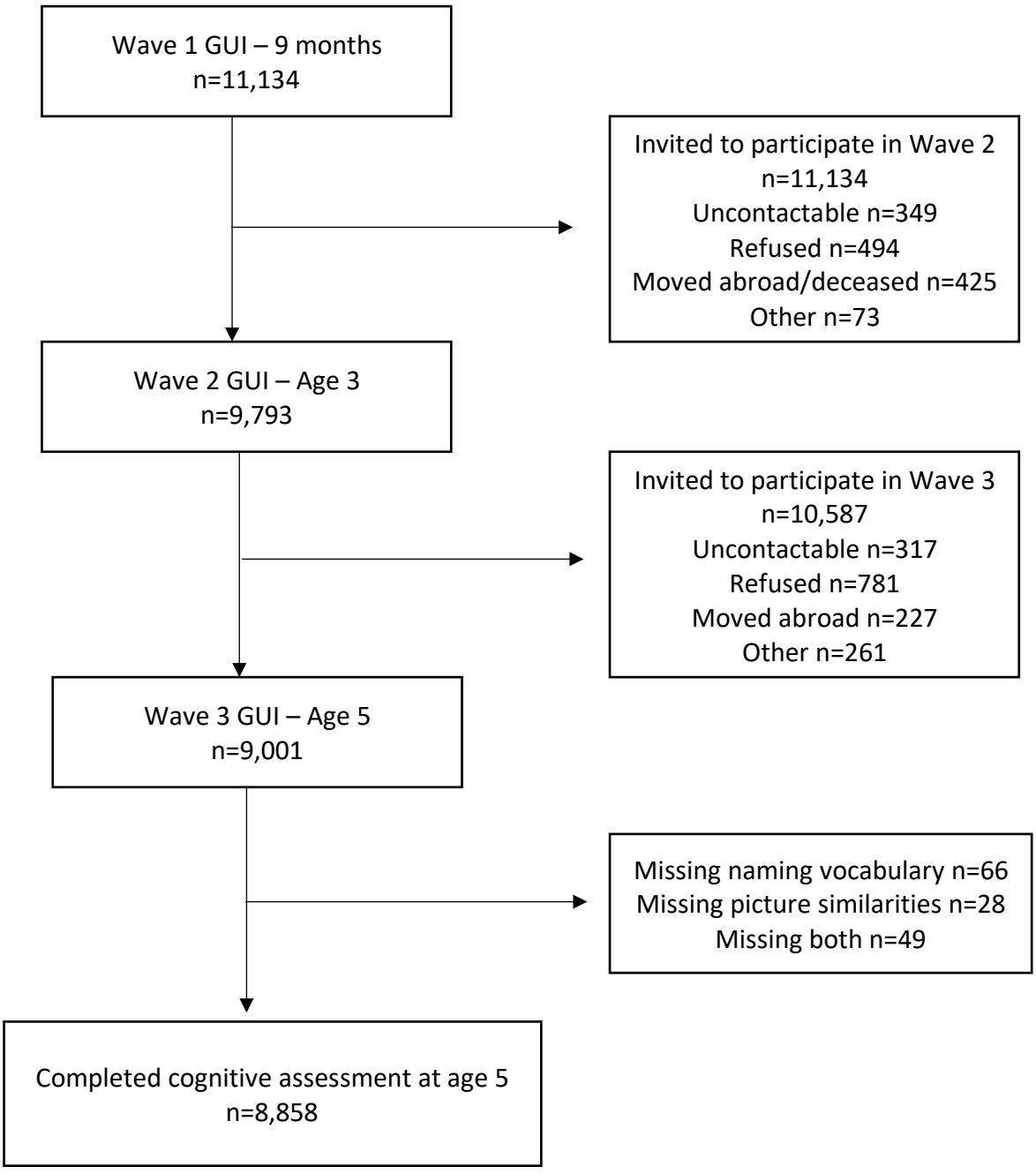


Figure A12.1 Legend
This figure shows the number of participants at each wave of the GUI study. Shown in the boxes to the right are the number invited to participate at each wave, and the reasons for non-participation.

Figure A12.2 Framework for feature selection

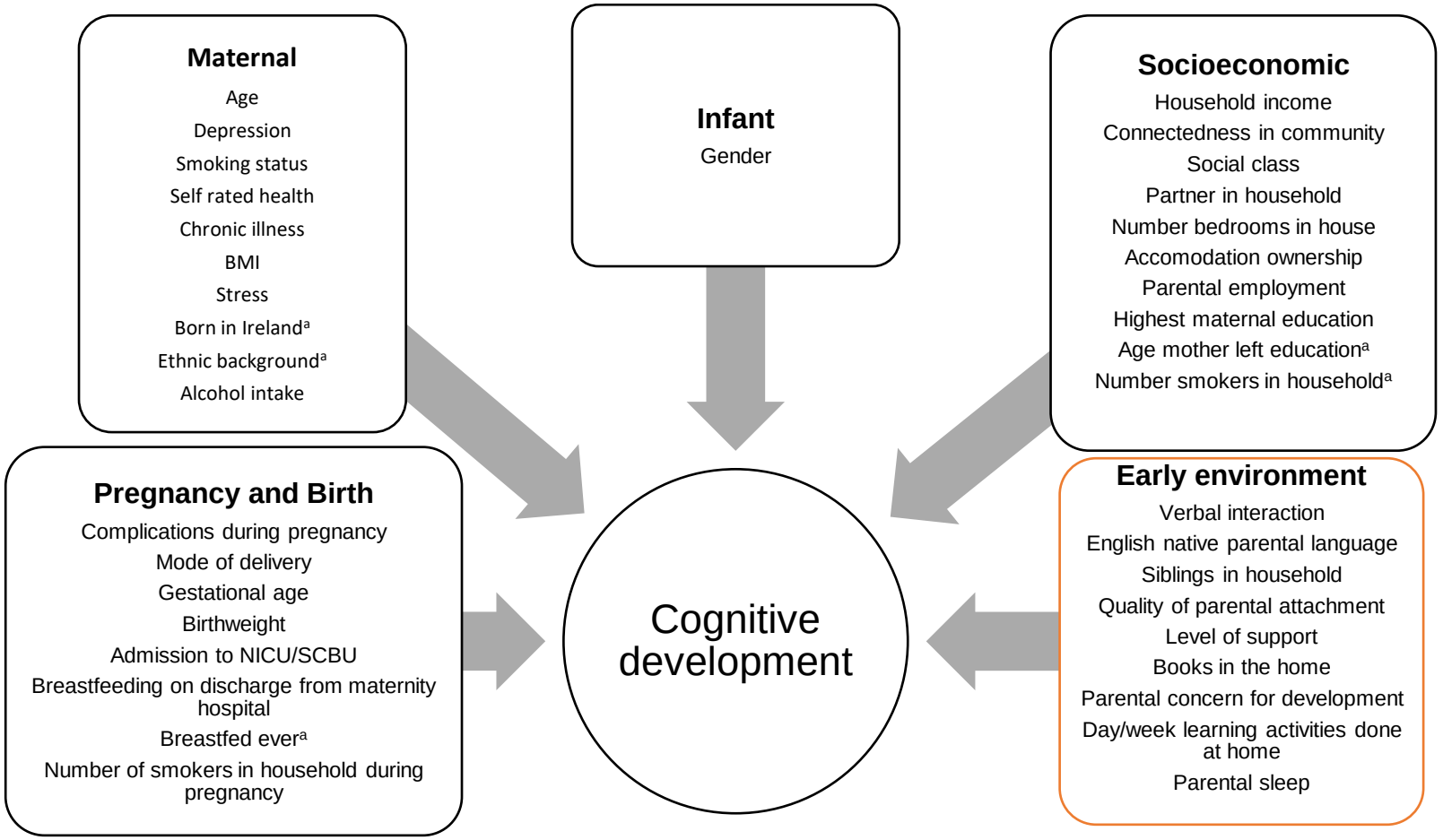


Figure A12.2 Legend

This figure demonstrates the theoretical framework used to identify potential features in the Growing Up in Ireland dataset. All features considered in the feature selection process are shown. [a]Feature removed after examining correlations for redundancy

Figure A12.3 Receiver operating characteristic curves of Models 1-3

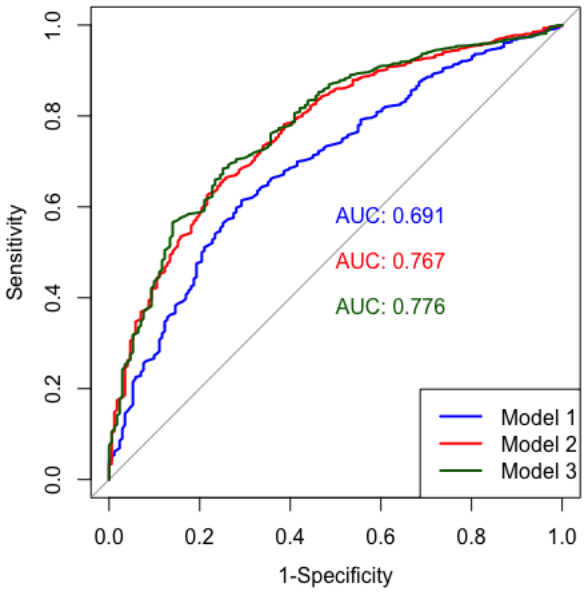


Figure A12.3 Receiver operating characteristic curves of Models 1-3
This figure shows the receiver operating characteristic curves for prediction of low cognitive ability at age 5 for Random Forest Model 1 (8 features), Model 2 (15 features), Model 3 (23 features). Model 3 achieved the highest area under the curve (AUC) of 0.776.

Supplementary Text

Evaluation Metrics

Sensitivity = $\frac{TP}{TP+FN}$

Specificity = $\frac{TN}{TN+FP}$

$$\frac{TN}{TN + FP}$$

Accuracy = $\frac{TP+TN}{TP+FP+TN+FN}$

Supplementary analyses using a cognitive ability score cut-off of 1 standard deviation below the mean to define low cognitive ability.

Figure A12.4 Overview of modelling process for cognitive ability > 1SD below mean (Below average cognitive ability (BACA)

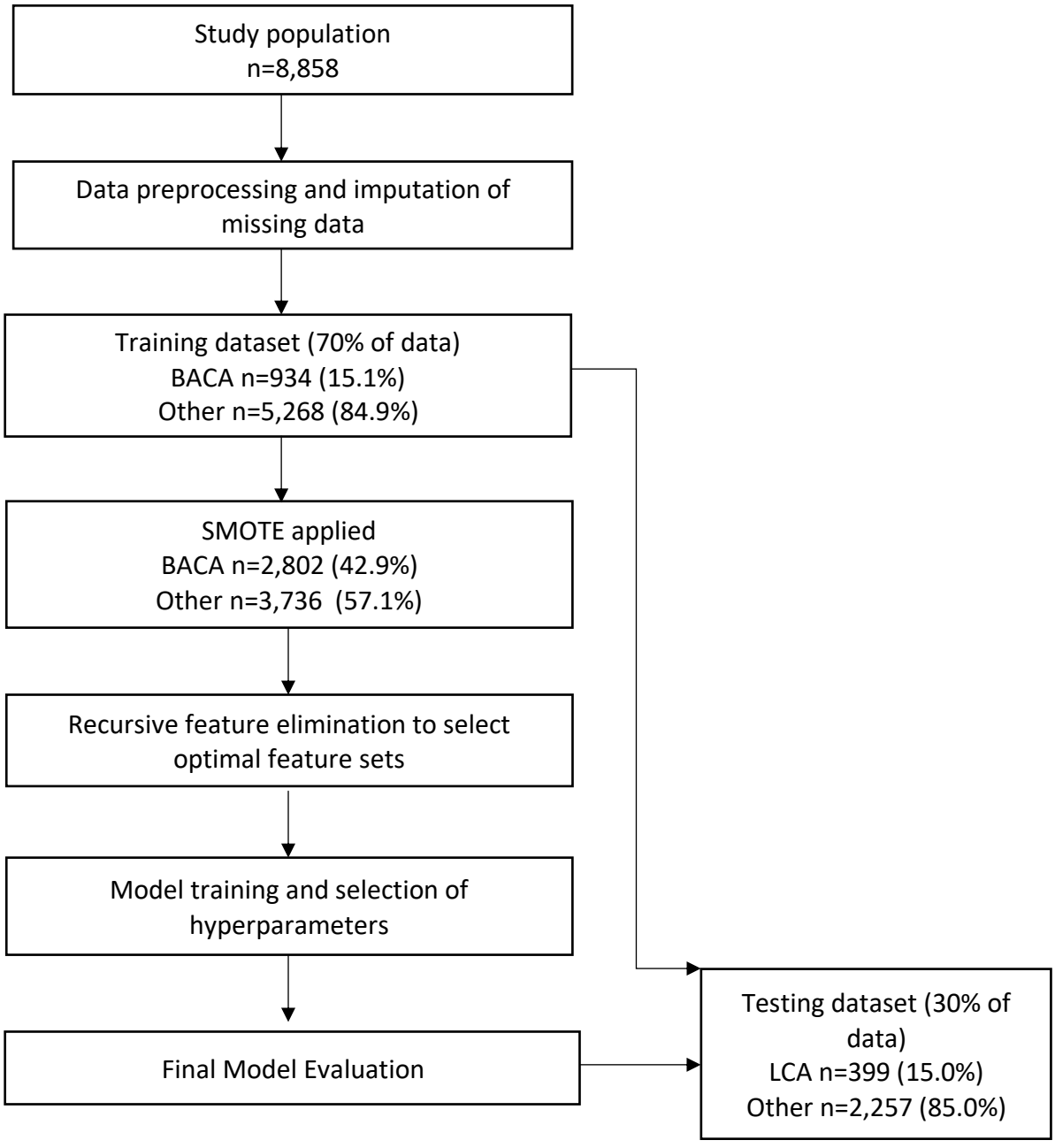


Figure A12.4 Legend
This figure shows the modelling process and numbers involved at each stage of modelling using an alternative cut off point of a general cognitive ability score ≥ 1SD below mean. Those below the cut-off were categorised as below average cognitive ability (BACA).

Table A12.8 Accuracy of different machine learning algorithms on 10-fold cross-validation for predicting cognitive ability more than 1 standard deviation below mean at age 5

	Model 1 (9 features)	Model 2 (18 features)	Model 3 (31 features)
Accuracy[a]			
Random forest	0.86 (0.01)	0.87 (0.01)	0.88 (0.01)
Logistic regression	0.67 (0.01)	0.71 (0.01)	0.73 (0.01)
SVM	0.77 (0.01)	0.81 (0.02)	0.82 (0.02)

[a]Calculated at a decision threshold of 0.5

Table A12.9 Performance metrics of final models tested on independent test set for predicting cognitive ability > 1 standard deviation below mean at age 5

	Model 1 (9 features)	Model 2 (18 features)	Model 3 (31 features)
Accuracy[a]	0.82 (0.80-0.83)	0.82 (0.81-0.84)	0.82 (0.81-0.84)
Sensitivity[a]	0.22	0.29	0.35
Specificity[a]	0.92	0.92	0.91
AUROC[b] (95% confidence interval)	0.63 (0.60-0.66)	0.68 (0.65-0.71)	0.70 (0.66-0.72)
Positive predictive value[a]	0.33	0.39	0.40
Negative predictive value[a]	0.87	0.88	0.89

[a]Calculated at a decision threshold of 0.5, [b] Area under receiver operating characteristic curve

Table A12.10 Performance metrics of random forest models trained with the five most important socioeconomic features alone.

Feature	AUROC (95% CI)	Accuracy[a]	Sensitivity[a]	Specificity[a]
Household social class	0.56 (0.52-0.59)	0.94 (0.93-0.94)	0.00	1.00
PCG Highest education	0.55 (0.52-0.58)	0.92 (0.91-0.93)	0.06	0.98
Household income	0.64 (0.59-0.68)	0.80 (0.78-0.81)	0.33	0.83
Number of bedrooms in home	0.52 (0.50-0.54)	0.94 (0.93-0.94)	0.00	1.00
PCG alcohol intake	0.62 (0.58-0.66)	0.81 (0.79-0.82)	0.28	0.84

[a] Calculated at a decision threshold of 0.5

Appendix 13 Supplementary material for Chapter 8: Predicting 2-year cognitive outcomes in very preterm infants using machine learning

Supplemental text

Rebalancing dataset

An imbalanced dataset is one where the classification categories are not approximately equally represented.[1] In registry data and cohort data this is often the case, where datasets are composed largely of those without the condition or disease of interest. Most machine learning algorithms assume a roughly balanced class distribution. When class imbalance is present the decision boundary is biased toward the majority class.[2] An algorithm which learns from class-imbalanced data tends to have poor predictive accuracy for samples in the minority class when it is shown new data. It tends to classify most new samples as being in the majority class.[1,3] In our original dataset there were n=231 infants with cognitive delay at 2 year follow up and n= 831 infants without delay. Training the algorithms on this class imbalanced data would likely result in a classifier bias towards the majority class (i.e. typical cognitive development). To address this, the synthetic minority oversampling technique (SMOTE) was applied to the training dataset.[1]

The aim of the SMOTE algorithm is to create a balanced dataset where the outcome classes, in this case 'cognitive delay' and 'typical cognitive development', have approximately equal observations. It achieves this through a combination of oversampling the minority class and undersampling the majority class. Unlike other oversampling techniques which simply duplicate minority class cases, SMOTE creates synthetic examples in the minority class.[1] Each minority class sample can be

considered as operating in a 'feature space'.[1] The algorithm randomly selects *k* minority class samples from a sample's nearest minority class neighbours (*k* is a parameter set by the user). Then, synthetic examples are created at random locations along the lines joining the *k* minority class neighbours. This is repeated until the desired amount of oversampling, set by the user using the 'perc.over' parameter, is achieved. In our study, oversampling was set to perc.over = 200 and the number of nearest neighbours used to generate new samples was set to *k*=5. For each existing sample categorised as 'cognitive delay' in the training dataset (n=162) two new samples were generated using information from the 5 nearest neighbours of each 'cognitive delay' sample, resulting in n=486 samples categorised as 'cognitive delay' in the rebalanced dataset.[1,4] Undersampling of the majority class, if required, is achieved by random removal of samples from the majority class. In our study SMOTE was applied only to the training dataset. This ensured that the final model was tested on completely unseen and untouched real life data.

Feature Effect

Feature effect was explored using partial dependence plots (PDPs) created with the 'iml' package.[5] The PDPs shows the average predicted probability of cognitive delay (CD) for a given feature value on the x-axis. For every value *x* of the feature of interest, model predictions are calculated and averaged across every observation, based on the observation having *x* as the value for that feature.[5] A lower value on the y-axis suggests that CD is less likely at that value of the feature on x-axis. Conversely, a higher value on the y-axis suggests a higher probability of CD. PDPs ignore whether the value *x* is plausible or even possible for that observation, therefore making plot

estimations unreliable in the presence of correlated features. Plots should also be interpreted with caution for values with few or no observations. The feature effect for the ten most important features in Model B and Model D are shown in eFigure 1 and eFigure 2. In Model B head circumference was strongly correlated with gestational age (Pearson correlation coefficient (r) = 0.80), birthweight (r = 0.82) and duration of hospitalisation (r = -0.63). Therefore, for the plot showing the effect of head circumference, the correlated features were removed from the model. Similarly, the plot for birthweight excludes the effect of gestational age (r = 0.71) and head circumference (r=0.82). The same procedure was performed for correlated features in Model D. The PDPs show how a machine learning model, such as the gradient boosted machine model in eFigure 3, can model non-linear relationships.

Table A13.1 Features with >25% missing values

Feature	Missing n (%)
Umbilical artery P02	648 (61.0%)
Umbilical artery PCO2	621 (58.5%)
Umbilical artery pH	608 (57.3%)
Venous P02	576 (54.2%)
Venous PC02	534 (50.3%)
Venous pH	521 (49.1%)
Birth length	269 (25.3%)

Table A13.2 Missing data for 90 features considered in modelling process

Feature	Description	Feature type	N (%) missing
Parent education[a]	0-3 years 3-6 years 7-9 years 10-12 years >12 years	Categorical	203 (19.1)
Head circumference	Measured in cm	Numeric	179 (16.9)
Rupture of membrane timing[a]	0-12 hours 12-24 hours >24 hours – 1 week >1 week – 4 weeks	Categorical	178 (16.8)
Receiving breastmilk at discharge	Completely Partially None	Categorical	149 (14.0)
Family language	Swedish or other Scandinavian language Non Scandinavian language	Categorical	148 (14.0)
Foetal presentation	Breech Cephalic Transverse Other	Categorical	117 (11.0)
Maximum grade of retinopathy of prematurity (ROP)	No ROP Grade 1 Grade 2 Grade 3 Grade 4 Grade 5	Categorical variable	62 (5.8)
ROP treated	Yes No	Categorical	62 (5.8)
Antenatal steroids[a]	No Yes, <8 hours before delivery Yes, 8-24 hours before delivery Yes, >24 hours before delivery Yes, missing time data	Categorical	61 (5.7)
Apgar 1 minute		Numeric	34 (3.2)
Apgar 10 minutes		Numeric	29 (2.7)
Number of erythrocyte transfusions		Numeric	28 (2.6)
Apgar 5 minutes		Numeric	26 (2.4)
Duration of continuous	Duration in days	Numeric	24 (2.3)

positive airway pressure (CPAP)			
Highest confirmed grade of intraventricular haemorrhage (IVH)	No confirmed IVH Grade 1 Grade 2 Grade 3 Grade 4	Categorical (derived)	17 (1.6)
Number of plasma transfusions		Numeric	12 (1.1)
Mode of delivery	Elective c-section Emergency c-section Vacuum Vaginal	Categorical (derived)	10 (0.9)
Cardiac compressions	Yes No	Categorical	9 (0.8)
CPAP during resuscitation	Yes No	Categorical	9 (0.8)
Ventilation mask during resuscitation	Yes No	Categorical	9 (0.8)
Adrenaline during resuscitation	Yes No	Categorical	9 (0.8)
Correction of acidosis during resuscitation	Yes No	Categorical	9 (0.8)
Intubation during resuscitation	Yes No	Categorical	9 (0.8)
Maternal illness	Yes No	Categorical	7 (0.7)
Pregnancy complication	Yes No	Categorical	7 (0.7)
Placental abruption	Yes No	Categorical	7 (0.7)
Chorioamnionitis	Yes No	Categorical	7 (0.7)
Preterm contractions	Yes No	Categorical	7 (0.7)
Intrauterine growth restriction	Yes No	Categorical	7 (0.7)
Preeclampsia	Yes No	Categorical	7 (0.7)
Rupture of membranes before 37 weeks	Yes No	Categorical	7 (0.7)
Rhesus immunisation	Yes No	Categorical	7 (0.7)

Birthweight	Birthweight in grams (20g intervals)	Numeric	7 (0.6)
Z.score for birthweight	Gestational-age specific birthweight z-score	Numeric	6 (0.6)
Maternal Epilepsy	Yes No	Categorical	5 (0.5)
Maternal type 1 diabetes	Yes No	Categorical	5 (0.5)
Maternal chronic lung disease or asthma	Yes No	Categorical	5 (0.5)
Maternal age	Age in years	Numeric	1 (0.09)

Categorizations of parental education, timing of rupture of membranes, and timing of antenatal corticosteroids reflect how these variables were reported to the Swedish Neonatal Quality Register.

Table A13.3 Characteristics of study population

Variable	Valid n	Total n (%)	Typical cognitive development n = 831	Cognitive delay n = 231	P-value
Pregnancy					
No. of foetuses	1,062				
Singleton		751 (70.7)	589 (70.9)	162 (70.1)	
Twin		286 (26.9)	222 (26.7)	64 (27.7)	
Triplet		25 (2.4)	20 (2.4)	5 (2.2)	0.898[a]
Maternal illness	1,055	459 (43.5)	351 (42.5)	108 (47.0)	0.264[b]
Epilepsy	1,057	8 (0.8)	7 (0.8)	1 (0.4)	1.000[a]
Type 1 diabetes mellitus	1,057	6 (0.6)	4 (0.5)	2 (0.9)	0.616[a]
Chronic lung disease or asthma	1,057	51 (4.8)	42 (5.1)	9 (3.9)	0.578[b]
Other	1,057	348 (32.9)	264 (31.9)	84 (36.5)	0.217[b]
Pregnancy complication	1,055	1,003 (95.1)	778 (94.3)	225 (97.8)	0.023[a]
Placental abruption	1,055	148 (14.0)	115 (13.9)	33 (14.3)	0.960[b]
Chorioamnionitis	1,055	77 (7.3)	53 (6.4)	24 (10.4)	0.054[b]
Gestational diabetes	1,055	14 (1.3)	8 (1.0)	6 (2.6)	0.094[a]
IUGR	1,055	178 (16.9)	145 (17.6)	33 (14.3)	0.291[b]
Preeclampsia	1,055	161 (15.3)	125 (15.2)	36 (15.7)	0.153[b]
ROM <37 weeks (no contractions)	1,055	169 (16.0)	118 (14.3)	51 (22.2)	0.005[b]
Antenatal steroids	1,001				
No		67 (6.7)	50 (6.4)	17 (7.8)	
< 8 hours before delivery		148 (14.8)	123 (15.7)	25 (11.4)	
8-24 hours before delivery		152 (15.2)	119 (15.2)	33 (15.1)	
>24 hours before delivery		617 (61.6)	478 (61.1)	139 (63.5)	
Yes, timing unknown		17 (1.7)	12 (1.5)	5 (2.3)	0.461[a]
Birth					
Delivery hospital	1,062				
Level 1/2		186 (17.5)	154 (18.5)	32 (13.9)	
Level 3/4		871 (82.0)	673 (81.0)	198 (85.7)	
Other		5 (0.5)	4 (0.5)	1 (0.4)	0.251[a]
Time of birth	1,062				
8am-5pm		552 (52.0)	426 (51.3)	126 (54.5)	
5pm-8am		510 (48.0)	405 (48.7)	105 (45.5)	0.419[b]
Presentation	945				
Breech		288 (30.5)	227 (30.8)	61 (29.3)	
Cephalic		565 (59.8)	439 (59.6)	126 (60.6)	
Transverse		38 (4.0)	41 (5.6)	13 (6.2)	
Other		54 (5.7)	30 (4.1)	8 (3.8)	0.957[a]
Maternal age at delivery	1,061				
N, mean (sd)		31.8 (5.6)	831, 31.4 (5.5)	230, 31.2 (5.8)	0.739[c]
Delivery	1,052				
Forceps or vacuum		5 (0.5)	5 (0.6)	0 (0.0)	

Elective Section		29 (2.8)	24 (2.9)	5 (2.2)	
Emergency Section		679 (64.5)	524 (63.8)	155 (67.1)	
Vaginal		339 (32.2)	268 (32.6)	71 (30.7)	0.675[a]
Gestational age	1,062				
N, mean (sd)		26.5 (2.2)	831, 26.7 (2.2)	231, 26.0 (2.2)	<0.001[c]
Birthweight - g	1,055				
N, median (IQR)		880 (380)	824, 900 (365)	231, 800 (320)	<0.001[d]
Z score for birthweight	1,054				
N, median (IQR)		-1.06 (2.00)	825, -0.99 (1.98)	231, -1.19 (2.06)	0.157[d]
Head circumference	883				
N, mean (sd)		24.7 (2.5)	685, 24.9 (2.5)	198, 24.2 (2.5)	0.001[c]
Sex	1,062				
Boy		566 (53.3)	414 (49.8)	152 (65.8)	
Girl		496 (46.7)	417 (50.2)	79 (34.2)	<0.001[b]
Apgar 1 min	1,028				
N, median (IQR)		6 (4)	805, 6 (4)	223, 5 (3.5)	0.013[d]
Apgar 5 min	1,036				
N, median (IQR)		8 (3)	809, 8 (3)	227, 7 (4)	0.001[d]
Apgar 10 min	1,033				
Median (IQR)		9 (2)	808, 9 (2)	225, 9 (3)	0.010[d]
Neonatal Resuscitation					
Extra oxygen	1,053	835 (79.3)	650 (78.7)	185 (81.5)	0.406[b]
CPAP	1,053	777 (73.8)	621 (75.2)	156 (68.7)	0.061[b]
Chest compressions	1,053	11 (1.0)	10 (1.2)	1 (0.4)	0.473[a]
Adrenaline	1,053	3 (0.3)	2 (0.2)	1 (0.4)	0.518[a]
Ventilation mask	1,053	683 (64.9)	518 (62.7)	165 (72.7)	0.007[b]
Correction acidosis	1,053	4 (0.4)	4 (0.5)	0 (0.0)	0.583[a]
Intubation	1,053	371 (35.2)	265 (32.1)	106 (46.7)	<0.001[b]
Respiratory					
TTN any	1,062	1 (0.1)	1 (0.1)	0 (0.0)	1.000[a]
RDS	1,062	887 (83.5)	680 (81.8)	207 (89.6)	0.007[b]
Pneumothorax	1,062	33 (3.1)	23 (2.8)	10 (4.3)	0.320[b]
BPD[e]	1,062	637 (60.0)	475 (57.2)	162 (70.1)	<0.001[b]
Severe BPD	1,062	106 (10.0)	71 (8.5)	35 (15.2)	0.005[b]
PPHN	1,062	81 (7.6)	56 (6.7)	25 (10.8)	0.054[b]
Extra O2 total days	1,062				
Median (IQR)		49 (65)	831, 45 (62)	231, 67 (65)	<0.001[d]
Surfactant given	1,062	697 (65.6)	530 (63.8)	167 (72.3)	0.020[b]
Any mechanical ventilation	1,062	690 (65.0)	556 (62.5)	134 (77.5)	<0.001[b]
Days mechanical vent total	1,062				
N, Median (IQR)		4 (17)	831, 3 (14)	231, 12 (26)	<0.001[d]
Days CPAP	1,038				
N, median (IQR)		26 (30)	809, 25 (30)	229, 30 (30)	<0.001[d]
Days conventional mechanical ventilation	1,062				
N, median (IQR)		1 (6)	831, 1 (5)	231, 3 (9)	<0.001[d]
Days HFOV	1,062				

N, median (IQR)		0 (8)	831, 0 (5)	231, 3 (15)	<0.001[d]
Inhaled NO	1,062	45 (4.2)	29 (3.5)	16 (6.9)	0.035[b]
Postnatal steroids	1,062	341 (32.1)	238 (28.6)	103 (44.6)	<0.001[b]
Infection					
Early onset sepsis	1,062	14 (1.3)	6 (0.7)	8 (3.5)	0.004[a]
Late onset sepsis	1,062	206 (19.4)	151 (18.2)	55 (23.8)	0.068[b]
Antibiotic treatment	1,062	979 (92.2)	762 (91.7)	217 (93.9)	0.325[a]
Days of antibiotics	1,062				
N, median (IQR)		14 (20)	831, 12 (19)	231, 20 (25)	<0.001[d]
Neurology					
Non IVH intracranial haemorrhage	1,062	12 (1.1)	6 (0.7)	6 (2.6)	0.028[a]
Seizures	1,062	16 (1.5)	9 (1.1)	7 (3.0)	0.059[a]
EEG or aEEG	1,062	69 (6.5)	49 (5.9)	20 (8.7)	0.175[b]
PVL - any	1,062	17 (1.6)	8 (1.0)	9 (3.9)	0.004[a]
IVH - confirmed	1,045	300 (28.7)	217 (26.5)	83 (36.7)	0.001[b]
Highest grade IVH	1,045				
No confirmed IVH		745 (71.3)	602 (73.5)	143 (63.3)	
Grade 1		129 (12.3)	100 (12.2)	29 (12.8)	
Grade 2		97 (9.3)	73 (8.9)	24 (10.6)	
Grade 3		36 (3.4)	24 (2.9)	12 (5.3)	
Grade 4		38 (3.6)	20 (2.4)	18 (8.0)	<0.001[b]
Other diseases					
Hyperbilirubinemia	1,062	946 (89.1)	738 (88.8)	208 (90.0)	0.680[b]
NEC	1,062	78 (7.3)	53 (6.4)	25 (10.8)	0.032[b]
PDA – medical tx	1,062	388 (36.5)	288 (34.7)	100 (43.3)	0.020[b]
PDA – surgical tx	1,062	153 (14.4)	104 (12.5)	49 (21.2)	0.001[b]
Max Grade ROP					
None	1,052	542 (51.5)	444 (54.1)	98 (42.4)	
Grade 1		142 (13.5)	115 (14.0)	27 (11.7)	
Grade 2		167 (15.9)	125 (15.2)	42 (18.2)	
Grade 3		198 (18.8)	136 (16.6)	62 (26.8)	
Grade 4		3 (0.3)	1 (0.1)	2 (0.9)	<0.001[a]
Other treatments					
Insulin	1,062	131 (12.3)	88 (10.6)	43 (18.6)	0.002[b]
ROP treated	1,000	124 (12.4)	84 (10.8)	40 (17.9)	0.007[b]
Inotrope support	1,062	200 (18.8)	136 (16.4)	64 (27.7)	<0.001[b]
Inotrope duration	1,062				
N, median (IQR)		0 (0)	831, 0 (0)	231, 0 (1.5)	<0.001[d]
Any erythrocyte transfusion	1,062	911 (85.8)	705 (84.8)	206 (89.2)	0.118[b]
Erythrocyte number transfusions	1,034				
N, median (IQR)		13 (28)	808, 13 (27)	226, 14 (26)	0.034[d]
Any plasma transfusion	1,062	548 (51.6)	408 (49.1)	140 (60.6)	<0.001[b]
Plasma number transfusion	1,050				
M, median (IQR)		2 (11)	822, 1 (11)	228, 2 (19)	<0.001[d]
Hospital stay and discharge					

Length hospital stay	1,062				
N, median (IQR)		2 (11)	831, 72 (40)	231, 83 (42.5)	<0.001[d]
Length hospital and homecare	1,062				
N, median (IQR)		75 (40)	831, 73 (40)	231, 85 (43.5)	<0.001[d]
Discharged to home	1,062	927 (87.3)	746 (89.8)	181 (78.4)	<0.001[b]
Breastfeeding on discharge	942				
None		421 (44.7)	311 (41.9)	110 (55.3)	
Partial		454 (48.2)	373 (50.2)	81 (40.7)	
Completely		67 (7.1)	59 (7.9)	8 (4.0)	0.002[a]
Receiving breastmilk on discharge	913				
None		320 (35.0)	235 (32.2)	85 (46.2)	
Partially		355 (38.9)	284 (39.0)	71 (38.6)	
Completely		238 (26.1)	210 (28.8)	28 (15.2)	<0.001[b]
Sociodemographic					
Family language	914				
Swedish or other Scandinavian language		641 (70.1)	541 (74.8)	100 (52.5)	<0.001[b]
Smoking in family	1,062	131 (12.3)	98 (11.8)	33 (14.3)	0.365[b]
Parental education					
0-3 years		8 (0.9)	5 (0.7)	3 (1.6)	
3-6 years		3 (0.3)	2 (0.3)	1 (0.5)	
7-9 years		50 (5.8)	31 (4.6)	19 (10.4)	
10-12 years		322 (37.5)	243 (35.9)	79 (43.2)	
>12 years		476 (55.4)	395 (58.4)	81 (44.3)	<0.001[a]
BSID-III[f]					
Corrected age in years at BSID-III	1,024				
N, mean (sd)		2.2 (0.2)	804, 2.2 (0.2)	220, 2.2 (0.2)	0.593[c]
Cognition index score	1,062				
N, mean (sd)		96.8 (16.1)	831, 103 (10.5)	231, 73.5 (10.0)	<0.001[c]

Abbreviations: BSID-III, Bayley Scales of Infant and Toddler Development, Third Edition; CPAP, continuous positive airway pressure; cPVL, cystic periventricular leukomalacia; EEG, electroencephalogram; aEEG, amplitude-integrated electroencephalogram; GA, gestational age; HFOV, high frequency oscillatory ventilation; IVH, intraventricular haemorrhage; IUGR, Intrauterine growth restriction; NEC, necrotising enterocolitis; RDS, respiratory distress syndrome; ROM, rupture of membranes; ROP, retinopathy of prematurity; TTN, transient tachypnoea of the newborn.

[a]Fishers exact test, [b]Pearson's Chi-squared test, [c]Welch two sample t test, [d]Wilcoxon rank sum test

[e]Defined as a registered ICD-10 code for BPD (P27.1) or a registration of supplemental oxygen use at 36 weeks of postmenstrual age

[f]Reasons for not completing: Child declined or would not participant n=25, Inattention, hyperactivity or fatigue n=7, Language barrier n=1, Parent declined n=5, Resources or administrative reason n=6, child unable to complete n=14, reason unclear or not recorded n=6.

Table A13.4 Final 26 features included in model

1	Gestational diabetes
2	Gestational age
3	Birthweight
4	Head circumference
5	Sex
6	Apgar 10 minutes
7	Intubation during neonatal resuscitation
8	Umbilical artery catheter
9	Bronchopulmonary dysplasia
10	Duration of CPAP (days)
11	Duration of conventional mechanical ventilation (days)
12	Duration of high frequency oscillatory ventilation (days)
13	Postnatal steroids
14	Duration antibiotics (days)
15	CNS bleed (excluding IVH)
16	Cystic periventricular leukomalacia (cPVL)
17	Grade IVH
18	Insulin treatment
19	Retinopathy of prematurity treated
20	Duration of inotropes (days)
21	Number of plasma transfusions
22	Duration of hospitalisation (days)
23	Discharge to home
24	Receiving breastmilk
25	Family language
26	Parental education

Table A13.5 Results of internal validation using ten-fold cross validation for Models A-D

Model Name	Algorithm	Features	Hyperparameters tuned in grid search (values)	Accuracy
Model A	Random forest	26	Mtry (1:26)	0.86
Model B	Logistic Regression	26		0.69
Model C	Support vector machine	26	C (0.25,0.50,1,2,4,8,16,32,64,128)[a] Sigma (0.01, 0.02, 0.05, 0.1, 0.2, 0.3, 0.4)	0.82
Model D	Gradient boosting machine	26	Ntrees (400,500,600,700,800,900)[a] Interaction depth (4,5,6,7,8,9,10,11,12,13,14,15) Shrinkage (0.1) n.minobsinnode (10)	0.87

[a]An initial random gridsearch was performed first to inform the range of tuning parameters chosen for training process.

Figure A13.1 Feature selection using Boruta algorithm

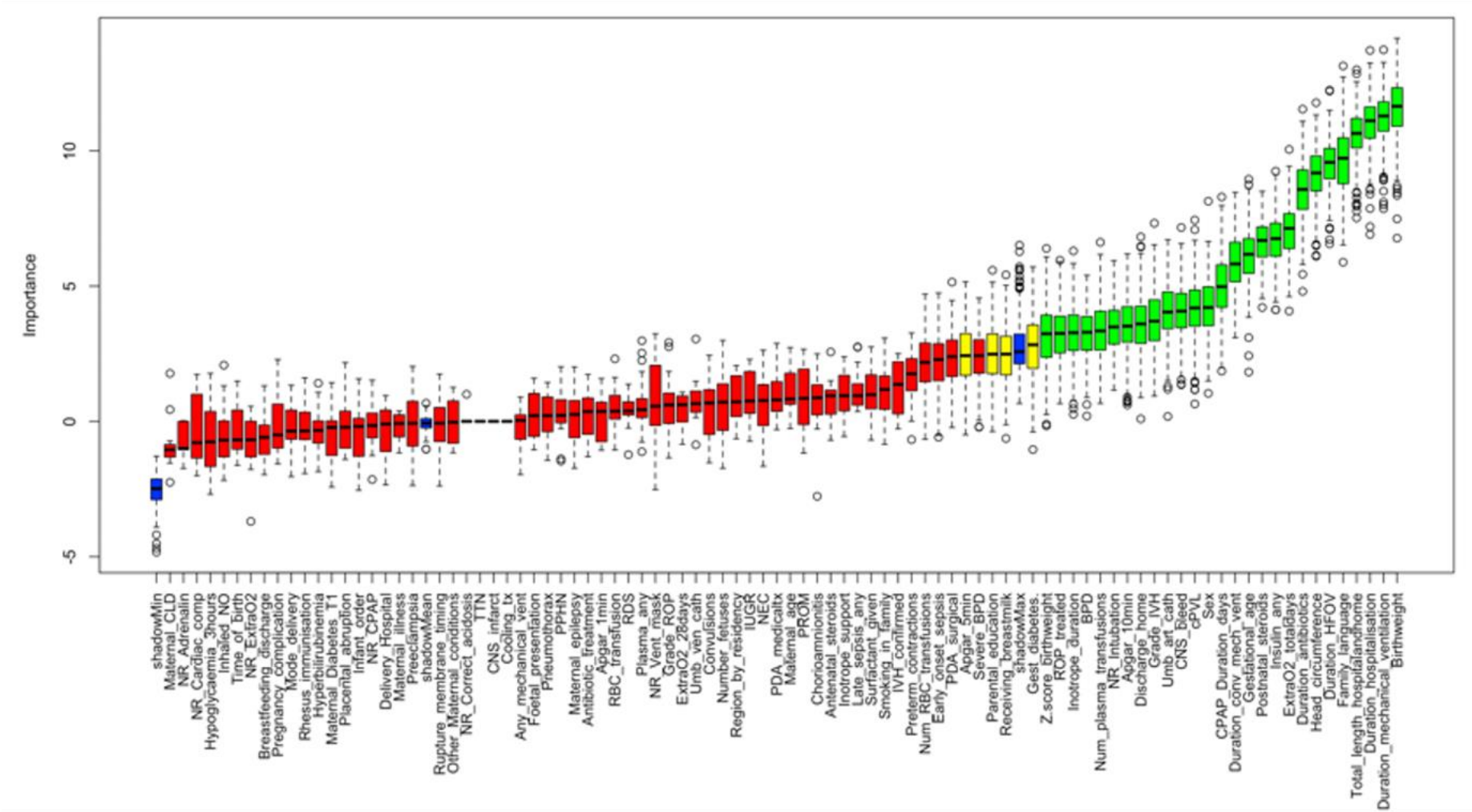


Figure A13.1 Legend

Abbreviations: CLD, Maternal chronic lung disease; NR, Neonatal resuscitation; cardiac_comp, cardiac compressions; CPAP, Continuous positive airway pressure; TTN, Transient tachypnoea of the newborn; CNS, Central nervous system; vent, ventilation; PPHN, Persistent pulmonary hypertension of the newborn; RBC, Red blood cell; RDS, Respiratory distress syndrome; ROP, Retinopathy of prematurity; Umb_ven_cath, Umbilical venous catheter; Umb_art_cath, Umbilical arterial catheter; IUGR, Intrauterine growth restriction; NEC, Necrotising enterocolitis; PDA, Patent ductus arteriosus; PROM, Preterm rupture of membranes; IVH, Intraventricular haemorrhage; BPD, Bronchopulmonary dysplasia; Gest_diabetes, Gestational diabetes; cPVL, Cystic periventricular leukomalacia; Duration_conv_mech_vent, Duration of conventional mechanical ventilation; HFOV, High frequency oscillatory ventilation.

A standardized feature importance score is shown in the y-axis. The blue boxplots represent the minimum, mean, and maximum z-scores of the shadow attributes which are created by randomly shuffling the values of the original features. The green, yellow, and red boxplots represent the z-scores of the confirmed, tentative, and rejected features respectively.

Figure A13.2. Features selected by Boruta algorithm with correlation coefficients >0.70

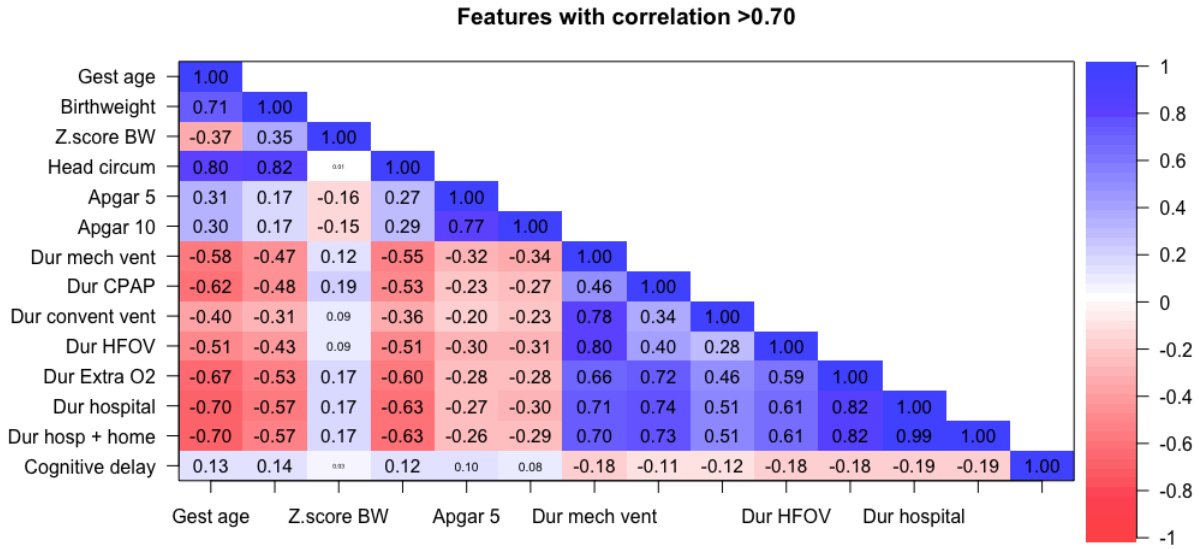


Figure A13.2 Legend
Abbreviations: Gest age, Gestational age; Z.score BW, Z score for birthweight; Head circum, Head circumference; Dur mech vent, Duration of all mechanical ventilation in days; Dur CPAP, Duration of continuous positive airway pressure in days; Dur convent vent, Duration of conventional mechanical ventilation in days; Dur HFOV, Duration of high frequency oscillatory ventilation in days; Dur Extra O2, Duration of extra oxygen; Dur hospital, Duration of hospitalisation; Dur hosp + home, Duration of hospital and home care.

Figure A13.3 Receiver operating characteristic curves of Models A-D

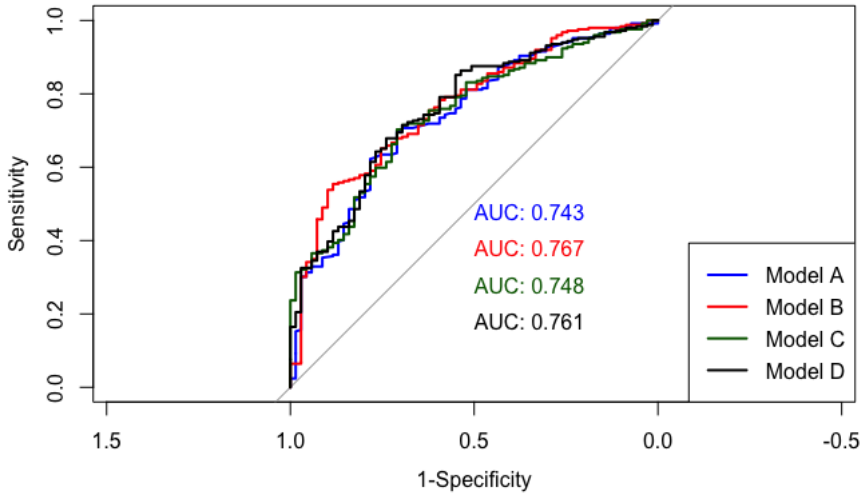


Figure A13.3 Legend

The receiver operating characteristic (ROC) curves for Models A-D tested on the unseen dataset are shown. The sensitivity is shown on the y-axis and 1-Specificity on the x-axis.

To ensure the performance of Model B (selected as our optimal model) was not affected by bias introduced by a single 70-30 split of the dataset, we performed ten random simulations of the 70-30 split and tested the AUROC across each unseen dataset. The mean AUROC of Model B across the ten simulations of the 70-30 split was 0.76 (95% Confidence Interval 0.69 – 0.82).

Figure A13.4 Feature effect plots for Model B (Logistic Regression)

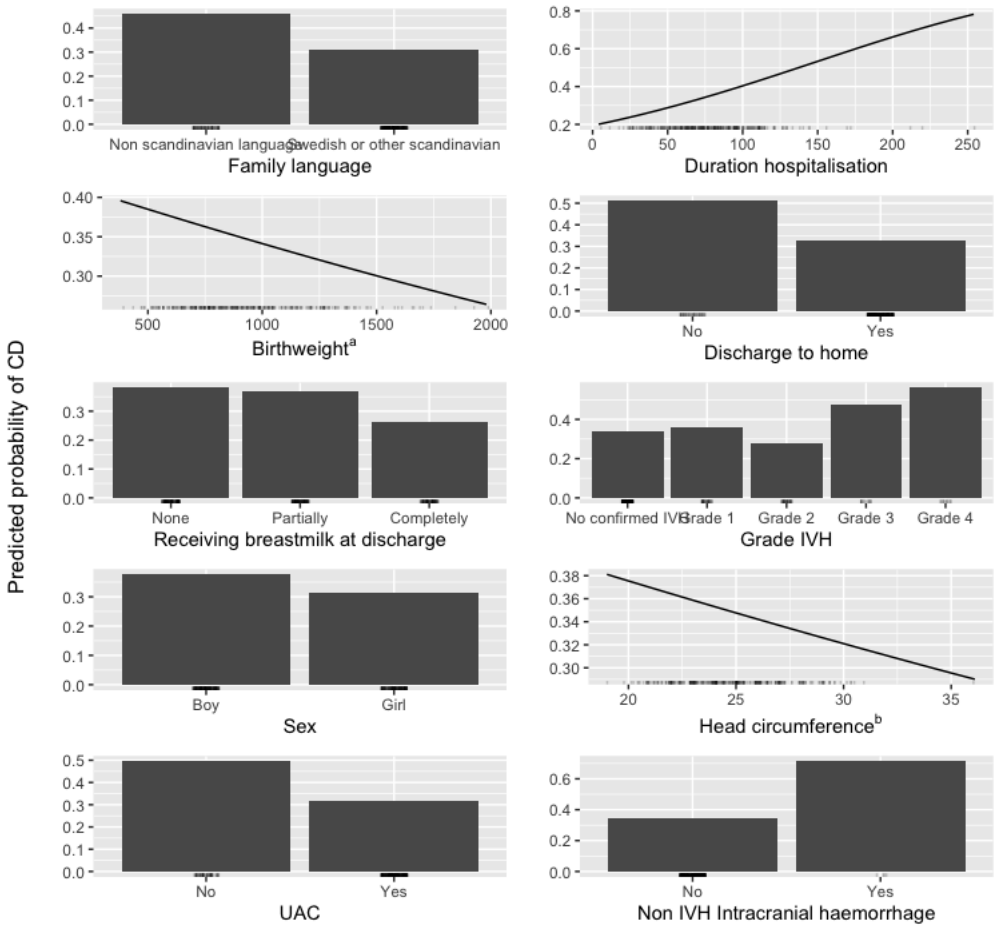


Figure A13.4 Legend

[a]Model excludes gestational age and head circumference. [b]Model excludes gestational age, birthweight and duration of hospitalisation. On the y-axis of the plots is the probability of cognitive delay and on the x-axis is either the categorical feature label or the continuous feature value. The distribution of feature values is shown along the x-axis. The predicted probability of CD was higher for families with who reported a non-Scandinavian family language, for those who were not discharged to home, for boys, for those who received no breastmilk at discharge, for those with Grade 4 IVH and for those with Non IVH intracranial haemorrhage. As the duration of hospitalisation increased, the predicted probability of CD increased. As head circumference and birthweight increased, the predicted probability of CD decreased.

Figure A13.5 Feature effect plots for Model D (Gradient Boosted Machine)

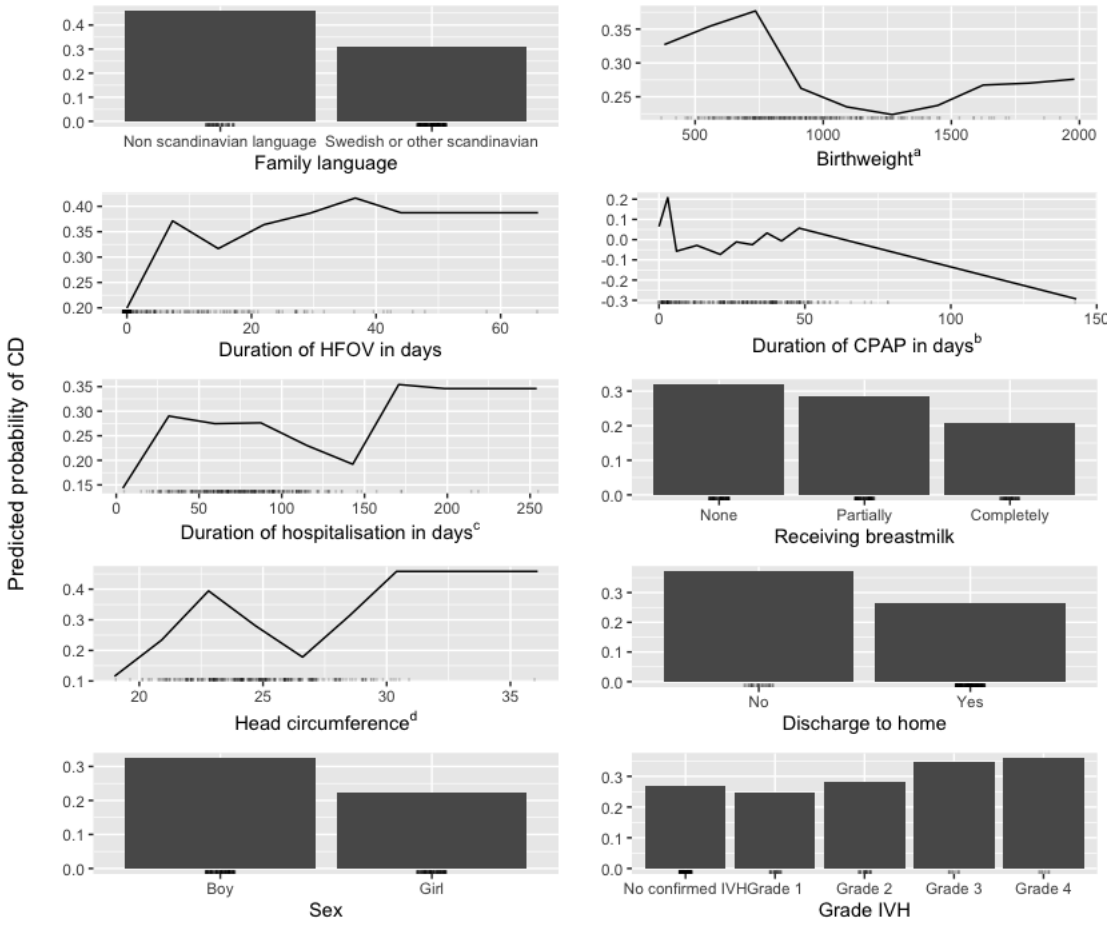


Figure A13.5 Legend

[a]Model excludes gestational age and head circumference, [b]Model excludes duration of hospitalisation, [c]Model excludes Duration of CPAP in days, [d]Model excludes birthweight and gestational age. On the y-axis of the plots is the probability of cognitive delay and on the x-axis is either the categorical feature label or the continuous feature value. The distribution of feature values is shown along the x-axis. For plots examining continuous features, for values where there are few or no observations the plots should be interpreted with caution as estimates may be unreliable.

Supplemental References

1 Chawla N, Bowyer K, Hall L, Kegelmeyer W. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res* 2002; 16: 341–348.

2 Kim A, Jung I. Optimal selection of resampling methods for imbalanced data with high complexity. PLoS One. 2023 Jul 27;18(7):e0288540. doi: 10.1371/journal.pone.0288540. PMID: 37498823; PMCID: PMC10374143.

3 Blagus R, Lusa L. SMOTE for high-dimensional class-imbalanced data. *BMC Bioinformatics* 2013; 14: 106.

4 RDocumentation. SMOTE: SMOTE algorithm for unbalanced classification problems.[Internet] [cited 9th November 2023]. Available from https://www.rdocumentation.org/packages/DMwR/versions/0.4.1/topics/SMOTE

5 Molnar C, Bischl B, Casalicchio G. Iml: An R package for Interpretable Machine Learning. *The Journal of Open Source Software* 2018; 3: 786.

Appendix 14 Big Data, Machine Learning, and Population Health: Predicting Cognitive Outcomes in Childhood

Andrea K. Bowe[1], Gordon Lightbody[1,2], Anthony Staines[3], Deirdre M. Murray[1,4]

[1] INFANT Research Centre, University College Cork, Cork, Ireland

[2] Department of Electrical and Electronic Engineering, University College Cork, Cork, Ireland

[3] School of Nursing, Psychotherapy, and Community Health, Dublin City University, Ireland

[4] Department of Paediatrics, Cork University Hospital, Cork, Ireland

Published in Paediatric Research

Annual Review Issue "The emerging role of artificial intelligence, big data analysis and precision medicine in Pediatrics"

Abstract

The application of machine learning to address population health challenges has received much less attention than its application in the clinical setting. One such challenge is addressing disparities in early childhood cognitive development - a complex public health issue rooted in the social determinants of health, exacerbated by inequity, characterised by intergenerational transmission, and which will continue unabated without novel approaches to address it. Early life, the period of optimal neuroplasticity, presents a window of opportunity for early intervention to improve cognitive development. Unfortunately for many, this window will be missed, and intervention may never occur or occur only when overt signs of cognitive delay manifest. In this review we explore the potential value of machine learning (ML) and big data analysis in the early identification of children at risk for poor cognitive outcome, an area where there is an apparent dearth of research. We compare and contrast traditional statistical methods with ML approaches, provide examples of how ML has been used to date in the field of neurodevelopmental disorders, and present a discussion of the opportunities and risks associated with its use at a population level. The review concludes by highlighting potential directions for future research in this area.

Early Life and Neurodevelopmental Plasticity

The phrase "the first 1000 days", a period spanning conception to three years of age, is one increasingly seen in policy documents relating to child health.[1] The phrase has its origins in the work of Professor David Barker, a physician and epidemiologist, whose trio of seminal articles published in the Lancet between 1986 and 1993 prompted the development of the 'foetal origins hypothesis'.[2–5] A fundamental concept in Barker's work was developmental plasticity, formally defined as "the ability of a single genotype to produce more than one alternative form of structure, physiological state, or behaviour in response to environmental conditions".[6]

It is widely accepted that early life is a unique period, differing from adulthood, during which the brain shows greater potentiality toward plasticity.[7] This is explained in part by the highly dynamic dendritic spines, which are predominant during early development, and display huge potential for adaptability but, which over time and in adulthood are much more stable.[8] Early life is the most crucial period for shaping the developing brain, and represents a window of both vulnerability and opportunity.[9] A failure to achieve early foundational cognitive skills may result in a permanent loss of opportunity to achieve full cognitive potential. Conversely, with appropriate intervention, it is a period during which there is opportunity to disrupt the intergenerational transmission of inequity, particularly with regard to the developing brain.[9]

Cognitive Function and The Importance of Social Determinants

Cognitive function is a broad construct consisting of multiple domains, which can differ across literature and disciplines. Most definitions consistently refer to domains of learning, understanding, reasoning, problem solving, memory, language,

attention, and decision making.[10–12] There has been considerable debate in the literature regarding the measurement of cognitive functioning, but also regarding its representativeness of overall human functioning.[13] While a thorough discussion of this debate is beyond the scope of this article, it is important to clearly acknowledge that childhood cognitive function is only one domain of overall child development. Other domains, including motor development, the development of positive social skills and emotional intelligence, and the establishment of appropriate behaviours are both intimately intertwined and of commensurate importance.[14]

For many children, particularly those exposed to adverse environmental conditions in early life, the true potential of their cognitive ability will never be reached. This gap between potential and actual ability brings with it significant implications for outcomes across multiple domains throughout the life course, including educational attainment[15], social mobility[16], financial well-being[17], and survival.[18] Birth and pregnancy cohort studies have to date provided an evidence-base illustrating the important early life determinants of cognitive development. Consistently, socioenvironmental factors have been shown to be among the strongest determinants of cognitive function in childhood.[19–25] The effect of socioenvironmental disadvantage on cognitive function can be seen as early as 2 years of age, but if unabated, the adverse effect of disadvantage on cognitive function appears to be amplified over time.[22]

Evidence for Early Intervention

Across epidemiological, animal, and human research there is consensus agreement that early life is a period of fundamental importance in cognitive development, and strategies for intervention should ideally be targeting this unique period.[26–33]

One of the most widely cited examples supporting early cognitive intervention is the Abecedarian Study, a small randomised controlled trial begun in 1972, and involving infants at risk of poor cognitive outcomes based on an individual risk index which combined multiple factors including parental education, family income, welfare payments and the school history of family members.[34] The trial involved 111 infants, of whom 57 were randomly assigned to an intensive preschool programme, which commenced at a mean age of 4.4 months. The intervention consisted of high quality early education delivered with high teacher to pupil ratios, five days per week for fifty weeks of the year, with free transportation and a focus on family involvement.[34,35] At age 3, significantly higher IQ scores were observed in the intervention group, and these persisted into adulthood. At age 21, the intervention group still had significantly higher IQ scores and academic achievement scores, as well as significantly more years of education completed compared to the control group.[36]

This pre-emptive approach to early cognitive intervention may also benefit other high risk cohorts such as pre-term infants, who on average have childhood IQ scores approximately 10 points lower than their full-term peers.[37,38] A Cochrane review examining the effects of early intervention provided in the first 12 months of life, for infants identified as at risk based on a gestational age <37 weeks or a birth weight <2,500 grams, concluded that early interventions can improve cognitive outcomes in infancy, with benefits persisting into preschool age.[39]

Early pre-emptive approaches targeting high risk individuals have also been shown to alter neurodevelopmental trajectories for other outcomes such as autism spectrum disorders (ASD). Whitehouse et al demonstrated, in a randomized clinical

trial of infants with very early signs of potential ASD, that pre-emptive intervention commenced around the first year of life reduced the odds of ASD classification at age 3. [40]

These studies illustrate the effectiveness of early pre-emptive intervention, commenced prior to overt signs of cognitive difficulty and at the time of optimal neuroplasticity. Outside of the controlled trial setting, real world programmes, such as the federally sponsored Early Head Start programme in the United States, demonstrate that at a population level such interventions are both feasible, achievable, and effective with appropriate resourcing.[41,42] However, achieving optimal individual outcomes from early intervention is dependent on the accurate identification and involvement of infants and families who are likely to benefit most.

Current Population-based Developmental Screening

Internationally, many countries rely on universal developmental screening programmes to identify children who may benefit from early intervention programmes, with the majority of developmental screening assessment being based on the presence of a delay in reaching developmental milestones.[43] Population-based screening programmes based on the principle of universalism apply structured developmental screening, independent of pre-existing risk factors, to all children. While this approach has many merits, there are significant limitations. Firstly, the opportunity for pre-emptive intervention is lost as these infants are already behind their typically developing counterparts at the time of detection, with valuable opportunities in the period of optimal neuroplasticity being lost. Secondly, population based screening programmes depend on voluntary presentation of parents and children, and tend to exemplify the inverse care law in practice, whereby

the most vulnerable and most at risk of developmental delay, are least likely to attend.[44-46] Thirdly, such screening programmes do not provide a systematic, robust assessment of the environmental, social, and demographic factors, to which the infant will be exposed in early life which are the strongest and most consistent predictors of cognitive outcomes in childhood.[47,48]

The birth of a child is unique in many ways, not least because there is almost universal interaction between a mother and a healthcare professional. Therefore, identification of high risk mother-infant dyads in the perinatal period could provide a unique opportunity to engage families with early intervention programmes**.** Machine learning and artificial intelligence (AI), which has shown great promise in prediction across other fields of medicine, may have potential to improve population-based strategies for the early identification of children at risk of poor cognitive outcomes, and this deserves further exploration.

Machine Learning and Predicting Risk

Adverse cognitive outcomes in childhood are complex, heterogenous, and result from highly interactive relationships between biological, environmental, and social factors, the mechanisms of which are often poorly understood. This real-world complexity, which may be difficult to model, particularly with more traditional statistical methods, may have to date acted as a deterrent for researchers investigating risk prediction tools for this purpose. However, there is now increasing potential to statistically model such complex interactions with machine learning. With the huge recent advances in both the quantity and quality of data available through large birth cohorts, electronic health records, imaging data, genomics, video, medical imaging, and electronic devices, the application of artificial intelligence and

machine learning may assist in finding the optimal predictive patterns to enable early interventions. [49]

The 'exposome' is a term used for the totality of life course environmental exposures from the prenatal period onwards. [50] Adopting this lens in evaluating the role of environmental factors may enable a more complete understanding of the exposures involved in childhood cognitive development. [51,52] Prospective birth cohorts and linked electronic medical records provide a unique opportunity for evaluating such exposures as they can capture a wealth of valuable information on prenatal, perinatal, and early childhood exposures. Following an individual from the intrauterine period or from birth enables a longitudinal assessment of early life exposures and later life outcomes.[53] There has been an increase in both the number and the size of birth cohort studies in recent years, and in Europe alone more than 110 birth cohorts representing 27 different countries have been formed.[54] Additionally, the capacity to collect large volumes of data, to biobank serum samples, to perform genomics, and to use technology such as wearable devices has resulted in an explosion in the quantity, complexity, and depth of data now collected in such cohorts, providing huge opportunity to utilise such data to advance both our understanding of the aetiology of disease, and our ability to predict outcomes. To truly harness the value of these cohorts and data sources, methods of data linkage and harmonisation require further research. Through amalgamating data sources, large data repositories could be created which would be particularly valuable for all fields of research, but especially for AI. Further work is needed in this area to address the issues around data protection, consent, and data sharing.

Generally two classes of statistical model are recognised – those where the aim is to understand the causal pathways leading to a certain outcome, and those where the aim is to predict as accurately as possible a certain outcome.[55] These two classes have much in common, and a tangled, but often constructive relationship.[56] Statistical models such as those intending to investigate the causal pathways between early life determinants and cognitive function in childhood estimate the effect size between a preceding factor and the outcome, adjusted for the confounding factors in this relationship, based on a formal modelling process.[57] The inclusion of additional variables in the model is based on previous evidence, changes in the size of the effect between the causative factor and the outcome, and goodness of fit. [56,57]

Prediction models however combine multiple predictors based on their ability to predict the occurrence of a particular event in the future (prognostic model) or their ability to predict the presence of a particular health state (diagnostic model).[58] Prediction model studies aim to develop and validate multivariable prediction models which combine individual predictors to predict an individual's probability of either having or developing a particular outcome.[59]

Published prediction models examining the prediction of cognitive outcomes among general paediatric populations are scarce compared to the published literature on prediction models for high risk groups such as pre-term infants.[57] A summary of three recently published multivariable models predicting cognitive outcomes in general paediatric populations is shown in Table 1.[20-22]

A major limitation in the traditional statistical regression methods used in these models is their reliance on strong assumptions, many of which are unrealistic

in complex real word data. Traditional regression methods are often used in ways which assume that the effect of a predictor on the outcome increases or decreases uniformly throughout the range of the prediction (assumption of linearity), however in truth many relationships are non-linear. It is often assumed that the residuals (the difference between the observed value of the dependent variable and the model predicted value of the dependent variable) have a constant variance at each level of the predictor (assumption of homoscedasticity), and that the residuals have a normal distribution.[60] Methods exist to deal with all of these issues, but may not be properly or commonly used. Further limitations are the constraints around interaction terms in traditional statistical models. Traditional models assume predictors affect the outcome in an additive manner unless otherwise specified. Interaction terms must be determined a-priori and entered into the model, a process not commonly performed despite a wealth of evidence demonstrating strong interactions between many of the predictors in the models, such as that between prematurity and maternal education, or between preterm brain injury and socioeconomic status.[61-63]

There is considerable debate regarding the true differences between traditional statistical approaches and machine learning, often compounded by the fact that many machine learning techniques, such as linear and logistic regression, are drawn from traditional statistical methods.[64] While traditional statistical models and machine learning models may use similar means to reach their goal, the key difference tends to lie between the objectives of the models. Traditional statistical methods, while they can be used for prediction, tend to focus on inferring relationships between variables, while following a defined set of assumptions and rules based on the data to which the model is fitted.[65–67] The purpose of ML is to

make a repeatable prediction by learning from patterns within the data, without prior assumptions or rules governing the process.[66,67] ML does not make inferences regarding the causal relationship between features and outcomes. Using designated training data, with a defined set of potential predictors and a defined outcome, a ML algorithm can learn a set of rules from the data which can then be applied to novel, unseen data to make a prediction on the outcome. [68] ML approaches can handle enormous quantities of data, and also can combine multiple different forms of data from multiple different sources including epidemiological data, imaging data, environmental and genomic data. ML techniques can efficiently deal with missing data, and driven by the data and generalisation performance, are adept at handling feature redundancy (collinearity) and feature selection.[68]

A major advantage of machine learning is the ability to handle large numbers of highly interactive predictors and to combine them in nonlinear relationships. With both the quantity and diversity of data growing, non-linear and interactive relationships within data are increasingly important and may be the key to improving predictive accuracy. ML techniques can, to a point, automatically uncover such interactive relationships in the data without the need for a priori specification of interactions.[69,70] For the purposes of predicting cognitive development, this is likely to be important. It is often the case that the effect of one predictor will change given the value of another predictor. For example, the combined effect of adverse socioenvironmental factors and preterm brain injury results in significantly lower childhood IQ than either alone.[63]

It should not be assumed, however, that ML techniques will always result in improved predictive performance compared to traditional methods.[71] This is likely to

depend on the outcome being predicted, the number of features involved in training the prediction, the type and variability of the features, and the presence of interactions between features. The optimal predictive algorithm will depend on the predictive problem, and prior to moving to more complex predictive models, traditional methods with logistic and linear regression modelling should always be trialled and reported for comparison.[71]

Applications of Machine Learning in Neurodevelopmental Pediatrics

The past decade has seen a rapid expansion in the application of ML and AI, for the purposes of diagnosis and prognosis, in clinical medicine.[68] Significant progress has been made specifically in the areas of radiology and imaging, pathology, and cancer. [72] There is growing interest in the application of machine learning in the field of paediatric neurodevelopmental disorders (NDDs), a heterogenous group of disorders impacting different domains of function including cognition, language, social, motor and behavioural, but progress has been slower.[73] The term NDD includes conditions such as intellectual disability, autism spectrum disorder (ASD), and attention deficit hyperactivity disorder (ADHD). [73]

Diverse data, ranging from MRI data to motor skills data collected via drag-and-drop gaming devices, have been combined with machine learning techniques for both diagnostic and prognostic purposes in the area of NDDs. Examples are provided in Table 2. Accurate, objective, diagnostic tools could be a particularly valuable aid to clinicians in this field, where many of the conditions currently rely on subjective, labour intensive, expert clinical assessments. In many countries, this results in lengthy waiting lists for assessment in developmental paediatric clinics and delays in diagnosis for those affected.

One example is ADHD, the diagnosis of which relies upon a detailed clinical assessment based upon observed behaviour and reported symptoms.[74] Pupillometric data (data based on dilation of the pupil) from 50 participants age 10-12 years used in a support vector machine (SVM) algorithm achieved a sensitivity of 0.773 and a specificity of 0.753 in the detection of ADHD, based on the results of nested 10-fold cross validation.[75] This study however is limited by the small sample size of 50, and the absence of model evaluation on an independent test set raises questions over the generalisability of their model to wider populations. This is a limitation seen in many of the studies in this field.

MRI data has also been used in ML algorithms for the purpose of diagnosing ADHD. Sen et al trained a SVM using structural and functional MRI data from 558 participants. The final model, which produced an accuracy of 0.673, sensitivity of 0.454, and specificity of 0.851 on an independent test set of 171 participants, is unlikely to be clinically useful in its current form, but suggests there may be signals within such data that could assist in the diagnosis of ADHD.[76]

For ASD, another neurodevelopmental disorder which benefits from early intervention but for which diagnosis is often delayed, ML has been increasingly studied as a potential tool for early detection.[77] A 2019 systematic review and meta-analysis examining the accuracy of ML algorithms based on brain MRI data for detecting ASD reported a pooled Area Under the Receiver Operating Curve (AUROC) of 0.90, a sensitivity of 0.83, and a specificity of 0.84 for ML models trained with structural MRI data. The results should however be interpreted with caution, as the majority of included studies performed internal validation only, and did not externally validate their models in independent test sets.[78]

While the examples outlined thus far have largely focussed on ML for diagnostic purposes, ML methods have also been deployed for prediction of later NDDs.[79,80] A deep learning model, trained on white matter connectomes from brain MRI data obtained at time of birth, and tested on an independent test set, correctly classified 84% of infants as having below or above the median level of cognitive ability at 2 years of age.[80] The results, which are based on a relatively small sample size of 37 in the independent test set, are promising but the clinical utility at a population-level is likely limited.

Novel methods of neurodevelopmental assessment, outside of traditional medical investigations, are also being analysed using ML algorithms. For example, a deep learning model trained on data from a touch-based computer game, aimed to harness this measure of motor skill to detect developmental disabilities in children. An AUROC of 0.817 based on internal cross-validation provides a basis for further research in this area.[81]

Moving from Clinical to Population-based Prediction

Acknowledging the limitations of these studies, overall the results indicate that artificial intelligence and machine learning may have an important role in the early prediction of certain neurodevelopmental disorders. However, many of the prediction studies in this field to date are based on data, such as MRI and EEG, which is not routinely collected at birth at a population level. Therefore, in clinical practice these studies may provide methods of prediction for higher risk populations who undergo these investigations, but are likely to be largely unsuitable for implementation at a population-based level due to resource constraints and cost. At a population level, machine learning would appear to be more advantageous when

used with data that is readily available, cost effective to collect, and available early in life. This however may come at the cost of reduced predictive ability compared with models based on direct physiological and anatomical measures.

In the context of population health, the application of machine learning to address population health issues and mitigate health disparities has received much less attention than its clinical applications. Prediction models at a population level tend to use existing electronic medical records, large cohort studies, environmental databases, government databases , and internet based data.[82] Population based data sources have the potential to incorporate social, environmental, and other structural determinants in prediction models, potentially improving accuracy.[83] Social determinants of health have long been studied in epidemiological research, but findings of these studies emphasise the need for methods that are better able to capture the complex, interactive, and flexible relationships between these predictors and health outcomes.[83]

The majority of well-established predictors of cognitive function in childhood, including poverty, gestational age, drug, alcohol and tobacco use during pregnancy, birthweight, admission to the neonatal intensive care unit, parental education and socioeconomic status, are predictors that are readily available in the perinatal period. To date, the accuracy of algorithms applying traditional statistical methods to such data for the purpose of identifying children at risk of low cognitive function in childhood have shown mixed results with regards to their accuracy (Table 1). The question of how we can best harness this information for identifying high risk children remains. Application of machine learning methods to large population based

datasets, using readily available, routinely collected information may assist in this objective but remains relatively unexplored in the literature.

Risks of Individual Risk Predictions and Machine Learning Algorithms

There are undoubtedly benefits to the early identification of children at risk of poor cognitive outcomes, namely providing early interventions during the period of optimal neuroplasticity and potentially altering the cognitive trajectory for the individual child. However, this approach is not without risk.

Early identification of those at risk is not synonymous with labelling, yet even the use of terms such as “at risk” or “higher risk” could have detrimental effects on both the child and their family. Labelling a child at a very early stage of life, as being at risk of a poor cognitive outcome, can perpetuate a self-fulfilling prophecy due to well documented effects on parent and teacher expectations, as well as the potentially detrimental effect on a child’s own self-concept.[84–86] Furthermore, using a label based largely on social factors, which are transmitted through generations and rooted in inequities generated by social and political policy, may shift the blame of societal failures onto the individual child or family. These risks are especially pertinent for children and families who may be incorrectly identified as at risk on a screening or risk prediction tool and who are unlikely to benefit from early intervention. For those correctly identified, the question is whether these risks are outweighed by the risk of overlooking a child until their cognitive difficulties manifest in overt signs such as academic failures, a point at which intervention may be too late.[87]

There are also broader ethical considerations with the use of machine learning algorithms in healthcare in general.[88–90] There is a growing body of literature

examining algorithmic fairness, a topic warranting urgent attention following examples of ML algorithms exacerbating health inequities, raising serious ethical concerns.[91] Each step of development of a ML algorithm is fraught with potential for bias – from the initial decision to use ML for the specific problem, to how it is deployed in practice.[92] A major threat to algorithmic fairness is the inherent bias in the data on which many of these models are built. For example, healthcare utilisation data represents those who attend for healthcare, but not necessarily those who need it; cohort data tends not to represent marginalised populations such as the homeless who may be unable to participate in longitudinal follow up. A further threat is the use of "black box" models which lack transparency, cannot be interpreted by the user, and therefore have inherent risks for producing bias. There are opportunities to harness the power of ML in addressing population health challenges, but stringent assessment of potential for bias, quality, and applicability are required prior to dissemination in practice.[93]

Conclusions and Future Research

There is an optimal window in early life during which the trajectory of cognitive development is more amenable to change.[94] Infants at risk of poor cognitive outcomes due to prenatal, birth-related, or socioenvironmental risk factors, whom we fail to identify and provide with appropriate interventions, may never meet their true cognitive potential. For children who begin their educational journey lacking basic cognitive skills, a downward spiral of repeated failure, grade retention, and behavioural difficulties may culminate in school dropout, and its associated future implications which include unemployment and substance abuse.[95] Equally it is the case that these children at higher risk are likely to live in poorer households, in more

deprived areas, and be members of marginalised groups, so population level interventions including educational policy, planning policy, and economic policy, also have a central role to play.

It is estimated that less than half of children with developmental delay are detected prior to school entrance, with the vast majority of those detected receiving no intervention in the very early years.[96,97] It is abundantly clear that the current practice of waiting for delays to present prior to intervention is inadequate and goes against the large body of evidence highlighting the optimal period of neuroplasticity in early life.[98] The risk factors for poor cognitive outcomes in childhood are well established in the literature, are amenable to identification in the perinatal period, and are often routinely collected. Yet, to date we have been largely unable to harness this information in a personalised and meaningful way to identify the high-risk mother-infant dyads and provide appropriate interventions.

Our improved capacity to collect large volumes of rich data and to apply novel statistical methods particularly suited to prediction provide an opportunity for researchers and clinicians to investigate alternative methods of identifying, at an earlier stage, children at risk of developmental and cognitive delay. Utilising, to the best of their predictive ability, currently available data in the form of birth cohorts, electronic health records, and population registries should be a first step.

Addressing health and social inequity is an age old problem crying out for novel solutions. In this review we have focussed primarily on the domain of cognitive development. This is not to put undue emphasis on cognitive ability as a measure of success, nor to endorse a meritocratic society where a child's emotional, social, or creative skills are viewed as less important. The purpose is to strive towards creating

a society whereby every child is given the opportunity to meet their cognitive potential, irrespective of the circumstances into which they are born. To achieve this, high risk infants and their families should be supported with appropriate, evidence-based interventions at the earliest possible stage, prior to the onset of cognitive delay. This requires early, accurate identification of those who will benefit. Combining the strengths of machine learning, big data analysis, and paediatric public health for a potential solution deserves thorough exploration for the individual, community, economic, and societal benefits which could result.

Author contributions

AB – Conceptualization, methodology, writing – original draft preparation, writing – review and editing

GL – methodology, writing – review and editing

AS – methodology, writing – review and editing

DM – Conceptualization, methodology, writing – review and editing

Table 14.1 Study Characteristics of Multivariable Prediction Models for Cognitive Outcomes in Childhood

Authors	Camargo-Figuera et al [20]	Camacho et al[21]	Eriksen et al[22]
Journal	BMC Pediatrics	Paediatric Perinatal Epidemiology	PLOS One
Continent	South America	Europe	Europe
Sample	Pelotas Birth Cohort	Millennium Cohort Study	Lifestyle During Pregnancy Study (sampled from the Danish National Birth Cohort)
Design	Prospective cohort	Prospective Cohort	Prospective cohort
Sample size	3,312	9,487	1,782
Year of Recruitment	2004	2000-2002	1997-2003[a]
Exclusion Criteria	Conditions associated with very low IQ e.g. severe mental retardation	Nil	Multiple pregnancies, language barrier, impaired hearing or vision, congenital disabilities implying mental retardation

Age at Cognitive Assessment	6	3	5
Cognitive Assessment	Wechsler Primary and Preschool Scales of Intelligence -III	Bracken School Readiness Assessment	Wechsler Primary and Preschool Scales of Intelligence - Revised
Cognitive outcome variable	Binary. Low IQ defined by a z-score < −1.	Binary. Not school ready defined by score <1 standard deviation below mean	Continuous
Number of risk factors at outset	32	29	27
Rationale given for candidate variables	Yes - selected based on previous literature and availability	Yes – selected based on previous literature and availability	No – but broad range (>20) selected
Statistical Model	Multivariable logistic regression	Multivariable logistic regression	Multivariable linear regression

Method of initial screening of candidate variables	Forward and backward stepwise selection	Forward and backward stepwise selection	Univariable association p ≥ 0.10
Interaction terms fitted	No	No	No
Multicollinearity addressed/discussed	No	Yes	Yes
No. Predictors in Final Model	13	13	9
Validation performed	Yes. Internal and external validation performed.	Yes. Internal validation only.	No
Predictive value measured	External validation Area Under Receiver Operating Curve (AUROC) 0.75	Internal validation AUROC 0.80 Sensitivity 72%	R squared 0.29

	Sensitivity 70.3% Specificity 68%	Specificity 74%.	
Predictors in Final Model	Child – gender, height-for-age deficit; head circumference-for-age deficit Parental – breastfeeding, parental smoking, maternal perception of child's health, skin colour Socioenvironmental – parental employment status, maternal education, income, number of siblings, number of persons per room	Child – gender, ethnicity, developmental milestones Parental – maternal age, maternal mental health, breastfeeding Socioenvironmental – socioeconomic class, maternal education, income, number of children, employment status, housing type	Child - gender, birth weight, height, head circumference Parental – maternal BMI, breastfeeding Socioenvironmental – Maternal IQ, Parental education

[a] In 2003, a prospective follow-up of 1750 mother-child pairs was initiated, sampled on the basis of maternal alcohol drinking patterns from The Danish National Birth Cohort (DNBC)

Table 14.2 Examples of Machine Learning Methods Applied in Neurodevelopmental Paediatrics

Authors	Das and Khanna [75]	Sen et al [76]	Emerson et al [79]	Girault et al [80]	Kim et al [81]
Year Published	2019	2018	2017	2019	2021
Purpose	Diagnostic	Diagnostic	Prognostic	Prognostic	Diagnostic
Outcome	ADHD	ADHD	ASD at 24-months of age	Cognitive ability at age 2 years	Developmental disability
Outcome Measure	Diagnosed by child neurologist using DSM-IV criteria	Diagnostic criteria differed across study sites	DSM-IV Criteria	Mullens Scales of Early Ability. Outcome categorised as above or below the median	Diagnosis made by paediatric psychologist
Features	Pupillometric data	Structural and functional MRI	MRI at 6-months of age	MRI at birth	Drag and drop data from computer based game
Participants	50 (28 cases, 22 controls)	558 (279 cases, 279 controls)	59 infants with high familial risk of ASD (11 cases, 48 controls)	115 infants	370 children (147 cases, 233 controls)

Algorithms trained	LR, SVM, KNN, NB, DT, RF	SVM	SVM	DL	DL
Evaluation method	Nested 10-fold cross validation	Independent test set (77 cases, 94 controls)	Cross-validation with nested leave one out procedure	Independent test set (22 cases, 15 controls)	10-fold cross validation
Evaluation Result (best performing models)	NB: AUROC 0.871, sensitivity 0.656, specificity 0.863;	Accuracy 0.6725, sensitivity 0.4545, specificity 0.8510	Accuracy 0.966, sensitivity 0.818, specificity 1.000	Accuracy 0.838, Sensitivity 0.864, Specificity 0.800	AUROC 0.817, Sensitivity 0.757, Specificity 0.740

LR – Logistic Regression, SVM – Support Vector Machine, KNN – K-Nearest Neighbours, NB – Naïve Bayes, DT – Decision Tree, RF- Random Forest, ADHD – Attention Deficit Hyperactivity Disorder, ASD – Autism Spectrum Disorder, DSM-IV – Diagnostic and Statistics Manual of Mental Disorders

References

1 Moore TG, Arefadib N, Deery A, West S. The First Thousand Days: An Evidence Paper. Parkville, Victoria, 2017.

2 Barker D et al. Fetal nutrition and cardiovascular disease in adult life. Lancet 1993; 341: 938–941.

3 Barker D, Osmond C. Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales. Lancet 1986; 1(8489): 1077–1081.

4 Barker D et al. Weight in infancy and death from ischaemic heart disease. Lancet 1989; 9.

5 Wadhwa PD, Buss C, Entringer S, Swanson JM. Developmental origins of health and disease: brief history of the approach and current focus on epigenetic mechanisms. Seminars in reproductive medicine 2009; 27: 358–368.

6 Barker D. Developmental origins of adult health and disease. Journal of Epidemiology and Community Health 2004; 58: 114 LP – 115.

7 Fu M, Zuo Y. Experience-dependent structural plasticity in the cortex. Trends in neurosciences 2011; 34: 177–187.

8 Cioni G, Inguaggiato E, Sgandurra G. Early intervention in neurodevelopmental disorders: underlying neural mechanisms. Developmental Medicine & Child Neurology 2016; 58: 61–66.

9 Spencer N, Raman S, O'Hare B, Tamburlini G. Addressing inequities in child health and development: towards social justice. BMJ pediatrics open 2019; 3: e000503–e000503.

10 American Psychological Association. American Psychological Association. APA Dictionary of Psychology. 2020.

11 Nouchi R, Kawashima R. Improving Cognitive Function from Children to Old Age: A Systematic Review of Recent Smart Ageing Intervention Studies. Advances in Neuroscience 2014; 2014: 235479.

12 National Research Council (US) Panel to Review the Status of Basic Research on School-Age Children. National Research Council (US) Panel to Review the Status of Basic Research on School-Age Children. Development During Middle Childhood: The Years From Six to Twelve. National Academies Press (US): Washington (DC), 1984.

13 Ganuthula VRR, Sinha S. The Looking Glass for Intelligence Quotient Tests: The Interplay of Motivation, Cognitive Functioning, and Affect. Frontiers in psychology 2019; 10: 2857.

14 Drigas AS, Papoutsi C. A New Layered Model on Emotional Intelligence. Behav Sci (Basel). 2018 May 2;8(5):45. doi: 10.3390/bs8050045. PMID: 29724021; PMCID: PMC5981239

15 Lager A, Bremberg S, Vågerö D. The association of early IQ and education with mortality: 65 year longitudinal study in Malmö, Sweden. BMJ (Clinical research ed) 2009; 339: b5282.

16 Forrest LF, Hodgson S, Parker L, Pearce MS. The influence of childhood IQ and education on social mobility in the Newcastle Thousand Families birth cohort. BMC public health 2011; 11: 895.

17 Furnham A, Cheng H. Childhood Cognitive Ability Predicts Adult Financial Well-Being. Journal of Intelligence . 2017; 5. doi:10.3390/jintelligence5010003.

18 Whalley LJ, Deary IJ. Longitudinal cohort study of childhood IQ and survival up to age 76. BMJ (Clinical research ed) 2001; 322: 819.

19 Tong S, Baghurst P, Vimpani G, McMichael A. Socioeconomic position, maternal IQ, home environment, and cognitive development. The Journal of pediatrics 2007; 151: 284–8, 288.e1.

20 Camargo-Figuera FA, Barros AJD, Santos IS, Matijasevich A, Barros FC. Early life determinants of low IQ at age 6 in children from the 2004 Pelotas Birth Cohort: a predictive approach. BMC pediatrics 2014; 14: 308.

21 Camacho C, Straatmann VS, Day JC, Taylor-Robinson D. Development of a predictive risk model for school readiness at age 3 years using the UK Millennium Cohort Study. BMJ open 2019; 9: e024851.

22 Eriksen H LF et al. Predictors of intelligence at the age of 5: family, pregnancy and birth characteristics, postnatal influences, and postnatal growth. PloS one 2013; 8: e79200.

23 Lawlor DA et al. Early life predictors of childhood intelligence: findings from the Mater-University study of pregnancy and its outcomes. Pediatric and perinatal epidemiology 2006; 20: 148–162.

24 von Stumm S, Plomin R. Socioeconomic status and the growth of intelligence from infancy through adolescence. Intelligence 2015; 48: 30–36.

25 Schoon I, Jones E, Cheng H, Maughan B. Family hardship, family instability, and cognitive development. Journal of epidemiology and community health 2012; 66: 716–722.

26 Bugental DB, Corpuz R, Schwartz A. Preventing children's aggression: outcomes of an early intervention. Developmental psychology 2012; 48: 1443–1449.

27 Vinen Z, Clark M, Paynter J, Dissanayake C. School Age Outcomes of Children with Autism Spectrum Disorder Who Received Community-Based Early Interventions. Journal of autism and developmental disorders 2018; 48: 1673–1683.

28 Yoshinaga-Itano C, Sedey AL, Mason CA, Wiggin M, Chung W. Early Intervention, Parent Talk, and Pragmatic Language in Children With Hearing Loss. Pediatrics 2020; 146: S270–S277.

29 Freitag H, Tuxhorn I. Cognitive function in preschool children after epilepsy surgery: rationale for early intervention. Epilepsia 2005; 46: 561–567.

30 Ramey CT, Ramey SL. Prevention of intellectual disabilities: early interventions to improve cognitive development. Preventive medicine 1998; 27: 224–232.

31 Gillette Y. Family-centered early intervention: an opportunity for creative practice in speech-language pathology. Clinics in communication disorders 1992; 2: 48–60.

32 Pungello EP et al. Early educational intervention, early cumulative risk, and the early home environment as predictors of young adult outcomes within a high-risk sample. Child Development 2010; 81: 410–426.

33 Watanabe K, Flores R, Fujiwara J, Tran LTH. Early childhood development interventions and cognitive development of young children in rural Vietnam. The Journal of nutrition 2005; 135: 1918–1925.

34 Masse L, Barnett S. A Benefit Cost Analysis of the Abecedarian Early Childhood Intervention. 2002.

35 Campbell FA, Ramey CT. Effects of early intervention on intellectual and academic achievement: a follow-up study of children from low-income families. Child Dev 1994; 65: 684–698.

36 Campbell FA, Ramey CT, Pungello E, Sparling J, Miller-Johnson S. Early Childhood Education: Young Adult Outcomes From the Abecedarian Project. Applied Developmental Science 2002; 6: 42–57.

37 Sejer EPF, Bruun FJ, Slavensky JA, Mortensen EL, Schiøler Kesmodel U. Impact of gestational age on child intelligence, attention and executive function at age 5: A cohort study. BMJ Open 2019; 9. doi:10.1136/bmjopen-2019-028982.

38 Turpin H et al. The interplay between prematurity, maternal stress and children's intelligence quotient at age 11: A longitudinal study. Sci Rep 2019; 9: 450.

39 Spittle AJ, Orton J, Doyle LW, Boyd R. Early developmental intervention programs post hospital discharge to prevent motor and cognitive impairments in preterm infants. Cochrane Database Syst Rev 2007; : CD005495.

40 Whitehouse AJO et al. Effect of Preemptive Intervention on Developmental Outcomes Among Infants Showing Early Signs of Autism: A Randomized

Clinical Trial of Outcomes to Diagnosis. JAMA Pediatrics 2021; 175: e213298–e213298.

41 Love JM, Chazan-Cohen R, Raikes H, Brooks-Gunn J. What makes a difference: Early Head Start evaluation findings in a developmental context. Monographs of the Society for Research in Child Development 2013; 78: vii–viii, 1–173.

42 Morris PA et al. New Findings on Impact Variation From the Head Start Impact Study: Informing the Scale-Up of Early Childhood Programs. AERA Open 2018; 4: 2332858418769287.

43 Kerstjens JM et al. Support for the global feasibility of the Ages and Stages Questionnaire as developmental screener. Early Human Development 2009; 85: 443–447.

44 Hirai AH, Kogan MD, Kandasamy V, Reuland C, Bethell C. Prevalence and Variation of Developmental Screening and Surveillance in Early Childhood. JAMA pediatrics 2018; 172: 857–866.

45 Wolf ER et al. Gaps in Well-Child Care Attendance Among Primary Care Clinics Serving Low-Income Families. Pediatrics 2018; 142: e20174019.

46 Edwards K et al. Improving Access to Early Childhood Developmental Surveillance for Children from Culturally and Linguistically Diverse (CALD) Background. International journal of integrated care 2020; 20: 3.

47 Flensborg-Madsen T, Falgreen Eriksen H-L, Mortensen EL. Early life predictors of intelligence in young adulthood and middle age. PloS one 2020; 15: e0228144.

48 Bradley RH, Corwyn RF. Socioeconomic status and child development. Annual review of psychology 2002; 53: 371–399.

49 Dwyer D, Koutsouleris N. Annual Research Review: Translational machine learning for child and adolescent psychiatry. Journal of Child Psychology and Psychiatry 2022; n/a. doi:https://doi.org/10.1111/jcpp.13545.

50 Santos S et al. Applying the exposome concept in birth cohort research: a review of statistical approaches. European journal of epidemiology 2020; 35: 193–204.

51 Vrijheid M et al. Early-life environmental exposures and childhood obesity: An exposome-wide approach. Environmental Health Perspectives 2020; 128: 1–14.

52 Steer CD, Bolton P, Golding J. Preconception and prenatal environmental factors associated with communication impairments in 9 Year old children using an exposome-wide approach. PLoS ONE 2015; 10. doi:10.1371/journal.pone.0118701.

53 Canova C, Cantarutti A. Population-Based Birth Cohort Studies in Epidemiology. International journal of environmental research and public health 2020; 17: 5276.

54 Pansieri C, Pandolfini C, Clavenna A, Choonara I, Bonati M. An Inventory of European Birth Cohorts. International journal of environmental research and public health 2020; 17: 3071.

55 Ramspek CL et al. Prediction or causality? A scoping review of their conflation within current observational research. European Journal of Epidemiology 2021; 36: 889–898.

56 Blakely T, Lynch J, Simons K, Bentley R, Rose S. Reflection on modern methods: when worlds collide-prediction, machine learning and causal inference. International journal of epidemiology 2021; 49: 2058–2064.

57 Linsell L, Malouf R, Morris J, Kurinczuk JJ, Marlow N. Risk Factor Models for Neurodevelopmental Outcomes in Children Born Very Preterm or With Very Low Birth Weight: A Systematic Review of Methodology and Reporting. American Journal of Epidemiology 2017; 185: 601–612.

58 Wolff RF et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Annals of Internal Medicine 2019; 170: 51–58.

59 Steyerberg EW et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS medicine 2013; 10: e1001381.

60 Barker LE, Shaw KM. Best (but oft-forgotten) practices: checking assumptions concerning regression residuals. The American Journal of Clinical Nutrition 2015; 102: 533–539.

61 Benavente-Fernández I, Siddiqi A, Miller SP. Socioeconomic status and brain injury in children born preterm: modifying neurodevelopmental outcome. Pediatric Research 2020; 87: 391–398.

62 Patra K, Greene MM, Patel AL, Meier P. Maternal Education Level Predicts Cognitive, Language, and Motor Outcome in Preterm Infants in the Second Year of Life. American journal of perinatology 2016; 33: 738–744.

63 Benavente-Fernández I et al. Association of Socioeconomic Status and Brain Injury With Neurodevelopmental Outcomes of Very Preterm Children. JAMA Network Open 2019; 2: e192914–e192914.

64 Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. Nature Methods 2018; 15: 233–234.

65 Azzolina D et al. Machine learning in clinical and epidemiological research: Isn't it time for biostatisticians to work on it? Epidemiology Biostatistics and Public Health 2019; 16. doi:10.2427/13245.

66 Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. Nature Methods 2018; 15: 233–234.

67 Rajula HSR, Verlato G, Manchia M, Antonucci N, Fanos V. Comparison of Conventional Statistical Methods with Machine Learning in Medicine: Diagnosis, Drug Development, and Treatment. Medicina (Kaunas, Lithuania) 2020; 56: 455.

68 Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. The New England journal of medicine 2016; 375: 1216–1219.

69 Touw WG et al. Data mining in the Life Sciences with Random Forest: a walk in the park or lost in the jungle? Briefings in bioinformatics 2013; 14: 315–326.

70 Mooney SJ, Westreich DJ, El-Sayed AM. Commentary: Epidemiology in the era of big data. Epidemiology (Cambridge, Mass) 2015; 26: 390–394.

71 Christodoulou E et al. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. Journal of Clinical Epidemiology 2019; 110: 12–22.

72 Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nature Medicine 2022. doi:10.1038/s41591-021-01614-0.

73 Uddin M, Wang Y, Woodbury-Smith M. Artificial intelligence for precision medicine in neurodevelopmental disorders. npj Digital Medicine 2019; 2: 112.

74 National Institute for Health and Care Excellence (NICE). Attention deficit hyperactivity disorder: diagnosis and management. [Internet]. [London]. NICE; 2018. Available from https://www.nice.org.uk/guidance/ng87/chapter/Recommendations#diagnosis (accessed 12 Apr2022).

75 Das W, Khanna S. A Robust Machine Learning Based Framework for the Automated Detection of ADHD Using Pupillometric Biomarkers and Time Series Analysis. Scientific Reports 2021; 11: 16370.

76 Sen B, Borle NC, Greiner R, Brown MRG. A general prediction model for the detection of ADHD and Autism using structural and functional MRI. PLoS One 2018; 13: e0194856–e0194856.

77 Zwaigenbaum L, Penner M. Autism spectrum disorder: advances in diagnosis and evaluation. BMJ 2018; 361: k1674.

78 Moon SJ, Hwang J, Kana R, Torous J, Kim JW. Accuracy of Machine Learning Algorithms for the Diagnosis of Autism Spectrum Disorder: Systematic Review and Meta-Analysis of Brain Magnetic Resonance Imaging Studies. JMIR Ment Health 2019; 6: e14108.

79 Emerson RW et al. Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age. Sci Transl Med 2017; 9: eaag2882.

80 Girault JB et al. White matter microstructural development and cognitive ability in the first 2 years of life. Hum Brain Mapp 2019; 40: 1195–1210.

81 Kim HH, An J il, Park YR. A Prediction Model for Detecting Developmental Disabilities in Preschool-Age Children Through Digital Biomarker-Driven Deep Learning in Serious Games: Development Study. JMIR Serious Games 2021; 9: e23130–e23130.

82 Morgenstern JD et al. Predicting population health with machine learning: a scoping review. BMJ Open 2020; 10: e037860.

83 Mhasawade V, Zhao Y, Chunara R. Machine learning and algorithmic fairness in public and population health. Nature Machine Intelligence 2021; 3: 659–666.

84 Jussim L, Harber KD. Teacher expectations and self-fulfilling prophecies: knowns and unknowns, resolved and unresolved controversies. Personality and social psychology review : an official journal of the Society for Personality and Social Psychology, Inc 2005; 9: 131–155.

85 Shifrer D. Stigma of a label: educational expectations for high school students labeled with learning disabilities. J Health Soc Behav 2013; 54: 462–480.

86 Taylor LM, Hume IR, Welsh N. Labelling and self-esteem: the impact of using specific vs. generic labels. Educational Psychology 2010; 30: 191–202.

87 Leigh JE. Early labelling of children: Concerns and alternatives. Topics in Early Childhood Special Education 1983; 3: 1–6.

88 Char DS, Shah NH, Magnus D. Implementing Machine Learning in Health Care - Addressing Ethical Challenges. N Engl J Med 2018; 378: 981–983.

89 McCradden MD, Joshi S, Mazwi M, Anderson JA. Ethical limitations of algorithmic fairness solutions in health care machine learning. Lancet Digit Health 2020; 2: e221–e223.

90 Mhasawade V, Zhao Y, Chunara R. Machine learning and algorithmic fairness in public and population health. Nature Machine Intelligence 2021; 3: 659–666.

91 Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019; 366: 447–453.

92 Striving for health equity with machine learning. Nature Machine Intelligence 2021; 3: 653.

93 de Hond AAH et al. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. npj Digital Medicine 2022; 5: 2.

94 Almond D, Currie J. Human capital development before age five. In: Card D, Ashenfelter OBT-H of LE (eds). Handbook of Labor Economics. Elsevier, 2011, pp 1315–1486.

95 Lansford JE, Dodge KA, Pettit GS, Bates JE. A Public Health Perspective on School Dropout and Adult Outcomes: A Prospective Study of Risk and Protective Factors From Age 5 to 27 Years. The Journal of adolescent health : official publication of the Society for Adolescent Medicine 2016; 58: 652–658.

96 Morelli DL et al. Challenges to implementation of developmental screening in urban primary care: a mixed methods study. BMC Pediatrics 2014; 14: 16.

97 Dearlove J, Kearney D. How good is general practice developmental screening? BMJ (Clinical research ed) 1990; 300: 1177–1180.

98 Klebanov P, Brooks-Gunn J. Cumulative, human capital, and psychological risk in the context of early intervention: links with IQ at ages 3, 5, and 8. Annals of the New York Academy of Sciences 2006; 1094: 63–82.

Appendix 15 Modules and courses completed during PhD

Modules

Module code	**Module name**	**Credits**
PG6030	Reproducible Research Practices using R	5
PG7049	The PhD Journey: Research Skills for Doctoral Researchers	10

Courses

Provider	**Course title**
Insight Statistical Consulting	R for Data Science
Statistical Horizons	Directed Acyclic Graphs for Causal Inference
Statistical Horizons	Machine learning

Appendix 16 Additional publications during PhD

The following publications were achieved during the time I was registered as a PhD student in UCC but were unrelated to this thesis

1. **Bowe AK**, Barrett PM. Bridging the academic practice gap in public health-The role of the clinical academic in public health. Public Health. 2024 Feb;227:e1-e2. doi: 10.1016/j.puhe.2023.12.026. Epub 2024 Jan 18. PMID: 38242829.
2. **Bowe A**, Marron L, Devlin J, Kavanagh P. An Evaluation of the Impact of a Multicomponent Stop Smoking Intervention in an Irish Prison. Int J Environ Res Public Health. 2021 Nov 15;18(22):11981. doi: 10.3390/ijerph182211981. PMID: 34831737; PMCID: PMC8624287.
3. Carr S, O'Donoghue PJ, **Bowe A**, O'Ceallaigh B, Siney E, Kelly JL. Central Slip Repair using Trans-articular K-wires: A Comparative Study. JPRAS Open. 2023 Jun 2;37:163-170. doi: 10.1016/j.jpra.2023.05.004. PMID: 37593586; PMCID: PMC10427979.
4. White P, Crowe C, **Bowe A**, Brennan A, Bruton O, O'Sullivan MB, O'Mahony MT, Sheahan A, Barrett P. An audit of COVID-19 death reporting in counties Cork and Kerry, Ireland, winter 2021-2022. Ir J Med Sci. 2023 Aug;192(4):1589-1594. doi: 10.1007/s11845-022-03211-7. Epub 2022 Nov 16. PMID: 36383325; PMCID: PMC9666975.