Approaches to Mobile Health Evaluation: A Comparative Study

A mHealth Evaluation Comparison Study

Ms Samantha Dick; Business Information Systems, Cork University Business School, University College Cork, Cork, Ireland.

Corresponding author: samantha.dick@ucc.ie

Dr Yvonne O'Connor; Business Information Systems, Cork University Business School, University College Cork, Cork, Ireland.

Dr Ciara Heavin; Business Information Systems, Cork University Business School, University College Cork, Cork, Ireland.

This is an Accepted Manuscript of an article published by Taylor & Francis in Information Systems Management on 27th November 2019 at:https://www.tandfonline.com/doi/full/10.1080/10580530.2020.1696550

Abstract

A major challenge faced by mobile health (mHealth) is identifying an evaluation technique which provides a rigorous evaluation while capturing the unique characteristics of the intervention. This study investigates traditional and emerging methods of mHealth evaluation, identifying existing gaps. This research is a useful first step towards developing an evaluation technique which will facilitate implementation and enable mHealth to reach its potential in accelerating socio-economic development, particularly in Low and Middle Income countries (LMICs).

Keywords

mHealth (mobile health), evaluation, randomized controlled trials (RCTs), comparison, socio-economic development.

Introduction

Mobile health (mHealth) is defined as medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices (World Health Organization, 2011). mHealth can be used in both public health and clinical medicine to improve healthcare systems, support healthcare professionals, and provide better health outcomes for patients (Burns, Keating, & Free, 2016; Davis, DiClemente, & Prietula, 2016). mHealth involves the use of mobile devices capable of supporting audio, photography, geolocation, sensors, internet access, and third-party apps (Davis et al., 2016). However, there is little or no quality control or regulations ensuring the usability, accuracy or safety of these mHealth interventions (Boudreaux, Waring, Hayes, Sadasivam, Mullen, & Pagoto, 2014).

The success of a mHealth intervention is dependent not only on the intervention components being safe and effective, but also on the end-user's willingness and ability to use it, and the context within which it is being used (Chib, van Velthoven, & Car, 2014). As a result, clinical evaluations where the primary goal is to determine effectiveness in a controlled research environment are often lacking consideration of the many other, complex variables which are required in order for the intervention to be successful. A robust mHealth evaluation should examine multiple criteria, such as user feedback on the mHealth intervention, the robustness of the technology, user engagement strategies, and user interaction, in addition to ensuring safety and accuracy (White, Burns, Giglia, & Scott, 2017).

A rigorous, mixed-methods approach is required to untangle the "why" and the "how" of mHealth interventions (Hatt, Chatterji, Miles, Comfort, Bellows, & Okello, 2015). Hatt et al. (2015) believe that global public health practitioners should use the most rigorous systematic approach to answer questions and make decisions. The Randomized Controlled Trial (RCT) has long been considered

the “gold standard” approach to pharmaceutical evaluations and it is perceived by many to be the best mechanism for mHealth evaluation (Pham, Wiljer, & Cafazzo, 2016b). However in recent years, evaluation methods have been proposed that may serve as alternatives to the RCT but they are, as yet, unsuccessful in changing the perception that the RCT is the most suitable approach to mHealth evaluation. Indeed, Pham et al. (2016b) noted that at no time throughout their study were alternative methodologies to the RCT mentioned as being more suitable for mHealth evaluation.

Medical research also produces non-medical social effects such as increased productivity, greater competitiveness and economic growth. As a consequence, health-related research thereby contributes indirectly to a country’s Gross Domestic Product (Roback, Dalal, & Carlsson, 2011). In order for mHealth research to have the greatest impact on socio-economic development, it is important that fit-for-purpose techniques are used to adequately evaluate interventions. Effective evaluations aid in minimising resource wastage and maximising the development potential of mHealth (van Velthoven, Car, Zhang, & Marušić, 2013). With this in mind, the research question posed for this study is: *“How can we characterize novel and existing approaches to mHealth and what gaps exist in these approaches?”* In order to satisfy this question, we examine the traditional method, and identify three newly emerging methods of mHealth evaluation, outlining a comprehensive comparison between the characteristics of each methodology, and the challenges posed by the unique field of mHealth. The evaluation considers the Randomized Controlled Trial (RCT), the Continuous Evaluation of Evolving Behavioral Interventions Technologies (CEEBIT), the Multiphase Optimization Strategy (MOST) and the Sequential Multiple Assignment Randomized Trial (SMART). By gaining a greater understanding of these evaluation approaches, we endeavor to identify the gaps in the existing approaches and highlight the requirements for the development of a truly fit-for-purpose approach to mHealth evaluation.

This article is structured as follows. The next section presents a review of the literature including the benefits of employing mHealth, and the impact this has on socio-economic development. The following section considers the literature search methodology, and the development of the comparison criteria. Subsequently, a brief overview of each evaluation methodology is presented, followed by a table outlining the comparison of each. Finally, a comprehensive discussion of the methodology comparison is considered, followed by the conclusions.

Literature Review

mHealth Benefits

With the rapidly shifting demographic and health profile of high-income countries to an ageing population with increasing chronic diseases, care is moving from hospital to community settings (Health Service Executive, 2017). There is an increasing focus on making care more patient-centric and empowering patients to manage their own care (Shankar, Prasad, Ankur, Talwar, & Jain, 2013). This patient empowerment movement has seen an exponential rise in the development of mHealth interventions, many of these are freely available to download (IMS Institute for Healthcare Informatics, 2015). These interventions allow the patient to take charge of their own healthcare, supporting them in many areas of their health including; medication reminders, women's health and pregnancy support, fitness and lifestyle, diet and nutrition, mental health and stress management (IMS Institute for Healthcare Informatics, 2015).

Health gains in society may result in improved diet and living conditions, safer work environments and healthier lifestyles (Roback et al., 2011). mHealth has the potential to have a profound impact on socio-economic development in several ways, including influencing patient behaviour, enabling remote treatment of chronic diseases and equipping healthcare workers to make better

clinical decisions. It is estimated that mHealth could save €99 billion in healthcare costs in the European Union alone (Shankar et al., 2013). In order to achieve this, demonstrable efficacy, cost-effectiveness and accessibility must be prioritized alongside technical software development (The Lancet, 2017). Yet, the challenge remains in identifying an evaluation technique which can achieve these aspects, while keeping pace with the technology industry. Consequently, without a robust evidence base, mHealth will not become part of government policy at a fast pace, nor will policy-makers be aware of its possible application (World Health Organization, 2011).

Many mHealth interventions are being used in Low-to-Middle Income Countries (LMICs) or developing countries (Chib, 2018), where the potential for socio-economic development is far greater than in more developed regions. Although there is vast potential for mHealth to have meaningful benefits for socio-economic development such as improving disease management, health outcomes, and reducing disease burden, these are yet to be adequately evaluated (Nilsen, Kumar, Shar, Varoquiers, Wiley , Riley, Pavel, & Atienza, 2012), especially in LMICs (Chib et al., 2014; Hamine, Gerth-Guyette, Faulx, Green, & Ginsburg, 2015). Countries moving from a developing to developed status typically have low standards of living, a weak industrial and commercial base and poor infrastructure (Kowal & Paliwoda-Pękosz, 2017). Investment in Information and Communication Technology (ICT) correlate strongly with indicators of economic growth (Kowal et al., 2017). Most mHealth interventions considered successful in LMICs are based in Non-Governmental Organizations and not integrated into mainstream public health services (Mechael, Batavia, Kaonga, Searle, Kwan, Goldberger, Fu, & Ossman, 2010).

mHealth and Socio-Economic Development

With both developed and developing countries leveraging mHealth, developing countries have a unique opportunity to accelerate developed nations in the provision of quality, affordable and

accessible healthcare services. mHealth interventions have the potential to contribute to socio-economic development in LMICs. Madon (2000) discusses four main dimensions of socio-economic development; 1. Social wellbeing, 2. Physical environment, 3. Economic growth, and 4. Political wellbeing. In the following sections, we outline how mHealth interventions which are well designed, effectively evaluated, and successfully implemented and sustained, can advance the socio-economic development of developing countries.

Social Wellbeing

The most obvious merits of mHealth lie in the area of health improvement. Many interventions have the potential to contribute to improvements in a variety of health areas including; mental health (Burns, Begale, Duffecy, Gergle, Karr, Giangrande, & Mohr, 2011; Naslund, Marsch, McHugo, & Bartels, 2015), sexual health (Burns et al., 2016), and maternal health (Stephani, Opoku, & Quentin, 2016). mHealth contributes to patient empowerment by allowing self-management of chronic diseases such as diabetes (Holmen, Torbjornsen, Wahl, Jenum, Smastuem, Arsand, & Ribu, 2014) and HIV/AIDS (Déglise, Suggs, & Odermatt, 2012), and connecting communities of support for people with similar experiences, such as breastfeeding families (White et al., 2017). mHealth also allows the removal of physical and financial barriers to healthcare, by allowing remote access to health information and contact with healthcare professionals (Opoku, Scott, & Quentin, 2015; Rutledge, Kott, Schweickert, Poston, Fowler, & Haney, 2017).

Economic Growth

Patients who are empowered are more likely to be motivated to manage their disease and adhere to medications (Brown & Bussell, 2011). This in turn contributes to an improvement in overall health outcomes and an increase in productivity due to fewer days lost at work due to illness (Khan & Socha-Dietrich, 2018).

mHealth as a diagnostic tool can result in a cheaper and faster health assessment process (Matthews, Kulkarni, Whitesides, Sarrafzadeh, Gerla, & Massey, 2009; Bourouis, Zerdazi, Feham, & Bouchachia, 2013), resulting in higher quality data recording than paper-based tools (Hall, Fottrell, Wilkinson, & Byass, 2014). The use of mHealth can also contribute to streamlined processes for Health Care Workers (HCWs), and improved communication between HCWs (Hall et al., 2014).

Physical Environment

mHealth interventions can facilitate instant data transfer allowing faster diagnosis, and can provide safer online or cloud storage of sensitive health data (Steinhubl, Muse, & Topol, 2015). Supply chain management is a problem in developing countries, with long delays in reporting and subsequent restocking of stock-outs, something that could be overcome by the implementation of mHealth interventions (Shieshia, Noel, Andersson, Felling, Alva, Agarwal, Lefevre, Misomali, Chimphanga, & Nsona, 2014). Particularly in developing nations with poor infrastructure, physical access to a health facility or contact with a HCWs can take many hours and requires traveling long distances. For example in Niger, less than a quarter of the population are within a 1-hour walk of a health center during the wet season (Blanford, Kumar, Luo, & MacEachren, 2012). mHealth can bridge this physical distance by connecting patients with HCWs, and with larger health facilities providing remote advice and care to people in need (Mahmud, Rodriguez, & Nesbit, 2010).

Political Wellbeing

Many African and Middle-Eastern countries continue to criminalize homosexuality (Cameron & Berkowitz, 2016), with severe penalties, including beatings, public humiliation, and life

imprisonment (Beyrer, 2014). This has been attributed to exacerbating the HIV epidemic in developing countries (Hagopian, Rao, Katz, Sanford, & Barnhart, 2017). As a result of this political standpoint, men who have sex with men living in these countries are perhaps the most vulnerable population in terms of contraction of HIV, subject to stigmatization and inability to access appropriate care. The use of mHealth can enable this vulnerable population to access confidential information and advice, test results, appointment scheduling and medication reminders (Catalani, Philbrick, Fraser, Mechael, & Israelski, 2013). mHealth interventions have shown successes in both the prevention and management of HIV in resource-poor settings (Catalani et al., 2013; Devi, Syed-Abdul, Kumar, Iqbal, Nguyen, Li, & Jian, 2015).

mHealth in LMICS

While the use of mHealth can provide profound opportunities (Nilsen et al., 2012; Davey & Davey, 2014), research into the assessment of mHealth from a developing country perspective has been limited (Tariq & Akter, 2011; Peiris, Praveen, Johnson, & Mogulluru, 2014). The proliferation of lightweight mHealth interventions which fail to translate or scale into health systems has led to subsequent criticisms of "pilotitis" which plague the mHealth field (Labrique, Vasudevan, Chang, & Mehl, 2013, p. 468). The novelty of the field of mHealth in LMICs may partly explain why there has been a focus on pilot studies, many of which have not been followed up with rigorous evaluation, or taken to scale (Hall et al., 2014). Additionally, premature scale-up of an unevaluated mHealth intervention could harm the entire field (Chib et al., 2014). We need to reduce the plague of "pilotitis" and ensure the integrity of mHealth as a discipline. This can be achieved through the use of rigorous evaluation methodologies which will achieve a fuller understanding of the sustainability, scalability and usability of mHealth beyond early phases (Labrique et al., 2013; Franz-Vasdeki, Pratt, Newsome, & Germann, 2015).

Ineffective Evaluation Impacting Potential

The potential for mHealth to contribute to socio-economic development is clear. In order for this to be realised, mHealth interventions must be adequately evaluated using techniques which include research, design and analysis planning which is fit-for-purpose (Istepanian & Woodward, 2017). Pham et al. (2016b) argue that most mHealth interventions that make it to the evaluation stage are evaluated in a RCT. This has lead to an "all or nothing" situation, resulting in a grey area surrounding the efficacy and safety of publicly available mHealth interventions. Although many health ICT usability studies have been conducted to explore usability requirements, discover usability problems and design solutions, few of the studies reported have evaluated the usability of mobile technologies (Brown III, Yen, Rojas, & Schnall, 2013). Additionally, it is argued that mHealth evaluation is not a solitary activity but instead, should involve a range of stakeholders including developers, consumers, policy makers, and physicians to ensure the clinical viablity of the application (The Lancet, 2017). This is imperative as another challenge facing the evaluation of mHealth lies in the subjectivity of evaluation. mHealth projects often involve a large number of stakeholders from different backgrounds and disciplines. It is likely that individual goals for the project may differ between stakeholders and may lead to discrepancies in what constitutes a "successful" intervention (Heeks, 2002). A rigorous and reliable evaluation technique is required that will counter this challenge.

Healthcare systems already face excessive costs and poor outcomes, so the adoption of unevaluated mHealth interventions may have a detrimental effect, exacerbating the problems (Kumar, Nilsen, Abernethy, Atienza, Patrick, Pavel, Riley, Shar, Spring, Spruijt-Metz, Hedeker, Honavar, Kravitz, Lefebvre, Mohr, Murphy, Quinn, Shusterman, & Swendeman, 2013). The literature attributes the lack of evidence on the effectiveness of mHealth interventions to the lack

of appropriate evaluation frameworks (Déglise et al., 2012; Kumar et al., 2013; Hall et al., 2014; Pham et al., 2016b). In the context of patient safety and security, the absence of effective evaluation has the potential to damage the field of mHealth by allowing sub-standard technologies and applications into the public domain which may cause substantial harm and conclude in future litigation (Rahman, 2015). The lack of evaluation across the mHealth field as a whole is a major weakness and threatens the credibility of mHealth as a concept (Hall et al., 2014). Rigorous evaluation of mHealth interventions is essential not only to quantify their effect, but to ensure that they do no harm (Pagoto & Bennett, 2013).

Governments and health departments are becoming more interested in mHealth, with eHealth and mHealth being included in strategic plans (Leon, Schneider, & Daviaud, 2012; Njoroge, Zurovac, Ogara, Chuma, & Kirigia, 2017). This is a positive move for mHealth, so it is more important than ever for policy makers to make informed decisions and allocate resources to only the highest quality mHealth interventions. mHealth interventions have an enormous potential to accelerate socio-economic development in developing countries. However, so little evidence is available on the feasibility, acceptability, and effectiveness of mHealth, that many have not been taken to scale, remaining at the pilot phase (Labrique et al., 2013; Hall et al., 2014; Shuchman, 2014). Without an evaluation framework that is fit-for-purpose, the field cannot move forward and fulfil its potential to impact on social wellbeing, economic growth, physical environment, and political wellbeing (Cohen, Bancilhon, & Grace, 2018).

mHealth evaluations have been criticized for often lacking the scientific rigor of the RCT, (Stephani et al., 2016) and the current evidence on their effectiveness is not convincing enough for policy-makers (Chib et al., 2014). Of note, the Food and Drug Administration (FDA) only began to monitor mHealth interventions classified as a medical device in 2013 (Herron, 2016) but despite

this, many mHealth interventions still fell outside of the remit of FDA monitoring and were not subject to regulation, whereas many others were subject to stringent regulation which hampered the development of potentially useful software (Malvey & Slovensky, 2017). Recently, the FDA have made the decision to demarcate the boundaries of applications which pose a low risk and do not meet their definition of medical devices (Taylor, 2017; US Food and Drug Adminstration, 2017). This change of regulation by the FDA may also indicate that mHealth interventions no longer qualifying as a medical device, may not require a RCT evaluation. This may contribute to changing the perception that the RCT is best for mHealth interventions. However, a problem still remains as no rigorous and reliable alternative to the RCT has yet been identified.

The RCT is considered to be the "gold standard" for the evaluation of interventions, and has subsequently been adopted into the mHealth field (Burns et al., 2016; Pham et al., 2016b; Stephani et al., 2016). However, much of the literature around the topic has suggested that this methodology may be incompatable with mHealth, identifying weaknesses such as the lack of qualitative analysis, high resource cost, and lengthy timeframe (Nilsen et al., 2012; Kumar et al., 2013; Ben-Zeev, Scheuller, Begale, Duffecy, Kane, & Mohr, 2015; Hatt et al., 2015; Pham et al., 2016b; White et al., 2017). From these weaknesses, a comparison criteria was derived against which to compare the RCT with novel evaluation technologies; 1. Data collection and analysis, 2. Standard execution protocol, 3. Sample size, 4. Cost, 5. Protocol, and 6. Time (Kaplan, 2001; Nilsen et al., 2012; Kumar et al., 2013; Ben-Zeev et al., 2015; Pham et al., 2016b). These criteria are used to compare the evaluation methodologies and identify gaps in their processes in the context of mHealth evaluation.

The evaluation of mHealth interventions presents several unique challenges; particularly the rapid pace at which technology evolves (Nilsen et al., 2012; Kumar et al., 2013; Ben-Zeev et al., 2015),

and the complexity of mHealth interventions, requiring a mixed-method evaluation to capture all of the socio-technical aspect of the intervention (Kaplan, 2001; Hatt et al., 2015; Pham et al., 2016b). Additionally, Pham et al. (2016b) identified three weaknesses of the RCT in relation to mHealth evaluation. These include 1. The high cost of trial implementation, 2. The rigid protocol of the RCT and 3. The issue of blinding. Together, these mHealth challenges and RCT weaknesses were used to develop the comparison criteria for the methodologies used in this study.

This study aims to examine the mHealth evaluation literature to compare the characteristics of each technique, and the criteria identified for successful mHealth evaluation and to identify novel evaluation techniques which may provide an alternative to the RCT.

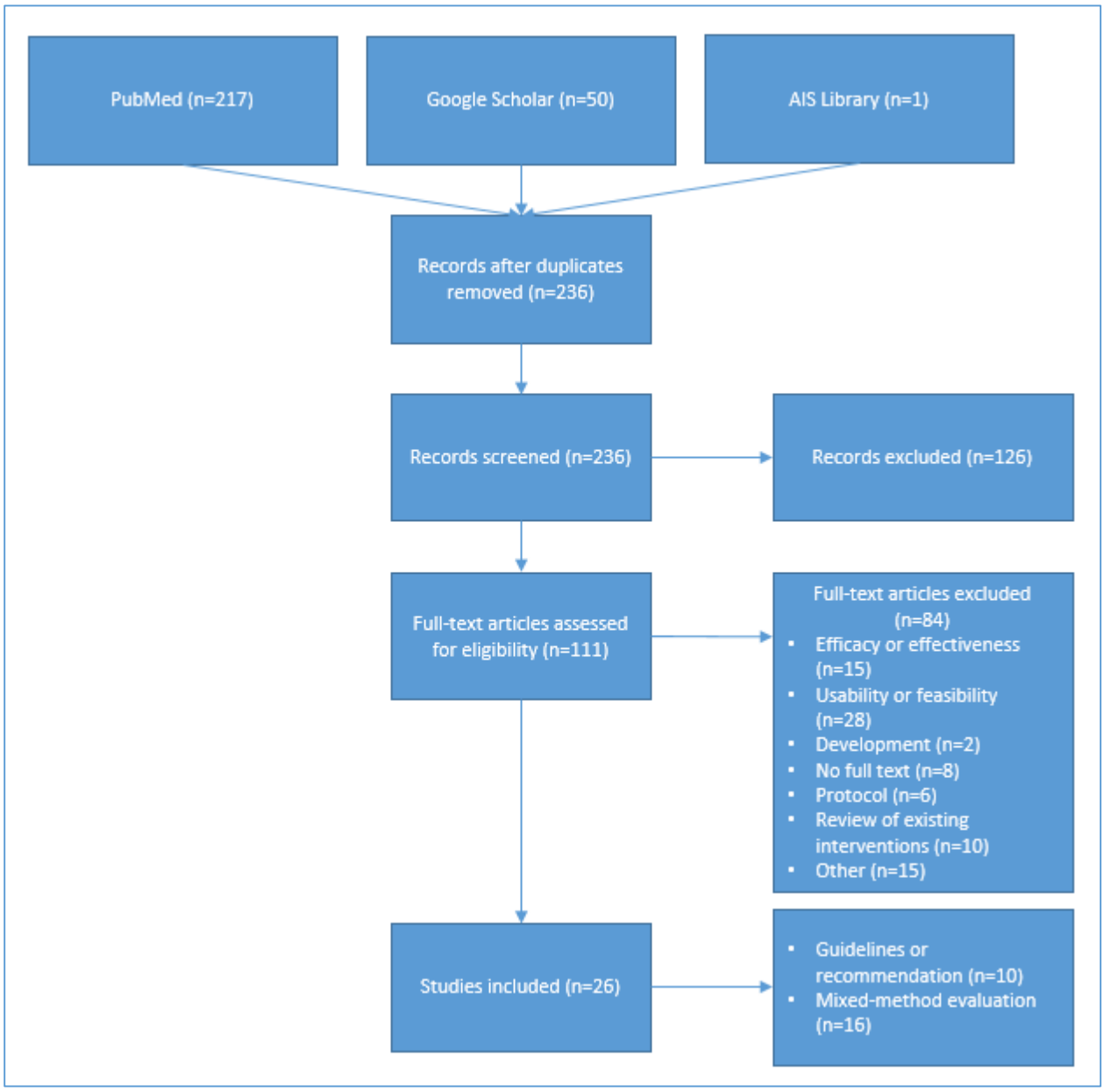
Methodology

This study was conducted as part of a larger research project. PubMed, Google Scholar, and the AIS library were searched in January 2017 using search terms under the pillars of "mHealth" and "evaluation." The mHealth literature has indicated that the most robust evalution of mHealth should conduct mixed-method evaluation (Kaplan, 2001; Hatt et al., 2015; Pham et al., 2016b) therefore, records were included if they: 1. Utilised a mixed-method evaluation, 2. Described a single application or intervention, 3. Described a pilot evaluation, 4. Provided guidance for evaluation. Records were excluded if they: 1. Used a single method evaluation (e.g. examined only effectiveness/efficacy, or only usability/acceptability), 2. Conducted a review of existing mHealth interventions, 3. Described a protocol, 4. Described health data, 5. Conducted post-hoc reviews or assessments of quality (i.e. after implementation). Due to the large number of records, only the first 5 pages of Google Scholar were retrieved. Saturation point was quickly reached as the majority of records identified up to this point were duplications of records found in PubMed, with no new records being identified after 50 records. When screening the papers, the authors only

included papers which focused on the pilot stage of implementation. Implementation refers to the "effort directed toward diffusing appropriate Information Technology (IT) within a user community" (Cooper and Zmud 1990, p.124). There is a lack of consensus, however, on the number of stages involved with IT implementation (Thompson, 1969; Pierce & Delbecq, 1977; McFarlan & McKenney, 1982; Cooper et al., 1990). For the purpose of this study, a pilot implementation is defined as: "a field test of a properly engineered, yet unfinished system, in its intended environment, using real data and aiming – through real-use experience – to explore the value of the system, improve or assess its design, and reduce implementation risk." (Hertzum, Bansler, Havn and Simonsen, 2012, p.2). Focusing on this stage of implementation (i.e. pilot study) is integral to allowing us to characterize and identify the gaps in approaches to mHealth evaluation, to fulfill the requirements of sustainable, scalable and usable mHealth interventions in LMICs.

Following the literature search, the titles and abstracts of 236 records were screened for eligibility. 111 records met the eligibility criteria, and were retrieved and reviewed. 84 studies were excluded as they were deemed irrelevant, and 26 studies were included. Of these, 16 utilised a mixed-method evaluation, and 10 provided guidance for evaluation. Full details of the search process are outlined in Figure 1.

Figure 1 – Literature Search Process



The studies utilising mixed-method evaluations were examined to determine the methodology used. The evaluation techniques used in each of these studies is listed in Table 1. We found that there appears to be very little consistency in the way the mHealth interventions were evaluated.

This is reflective of the lack of standardised, widely accepted evaluation frameworks for mHealth evaluation as identified in the literature (Déglise et al., 2012; Kumar et al., 2013; Hall et al., 2014).

Table 1 – Mixed Method Evaluation Techniques of Included Studies

Mixed Method Evaluation		
Pre-post study with interviews or focus groups	5	(Cafazzo, Casselman, Hamming, Katzman, & Palmert, 2012; Lemay, Sullivan, Jumbe, & Perry, 2012; Battle, Farrow, Tibaijuka, & Mitchell, 2015; Ho, Newton, Boothe, & Novak-Lauscher, 2015; Vu, Nguyen, Tran, & Muhajarine, 2016)
Cluster RCT with interviews	2	(Chang, Kagaayi, Arem, Nakigozi, Ssempijja, Serwadda, Quinn, Gray, Bollinger, & Reynolds, 2011; Jamison, Karlan, & Raffler, 2013)
Post-task questionnaire with interviews	2	(Lim, Cloete, Dunsmuir, Payne, Scheffer, von Dadelszen, Dumont, & Ansermino, 2015; Braun, Lasway, Agarwal, L'Engle, Layer, Silas, Mwakibete, & Kudrati, 2016)
Uncontrolled trial with usability questionnaire	2	(Mohr, Montague, Stiles-Shields, Kaiser, Brenner, Carty-Fickes, Palac, & Duffecy, 2015; Ha, Tesfalul, Littman-Quinn, Antwi, Green, Mapila, Bellamy, Ncube, Mugisha, & Ho-Foster, 2016)
RCT with user evaluation	1	(Pham, Khatib, Stansfeld, Fox, & Green, 2016a)
Randomized crossover with acceptability questionnaire	1	(Dahlberg, Jaensson, Eriksson, & Nilsson, 2016)
Multiphase Optimization Strategy	1	(Buman, Epstein, Gutierrez, Herb, Hollingshead, Huberty, Hekler, Vega-López, Ohri-Vachaspati, & Hekler, 2015)
Randomized trial with feedback	1	(McClure, Anderson, Bradley, An, & Catz, 2016)
Comparative analysis with focus groups	1	(Neupane, Odendaal, Friedman, Jassat, Schneider, & Doherty, 2014)

We then analysed the ten papers which provided guidelines and recommendations to identify novel mHealth evaluation methodologies. The papers were coded by one researcher, and ten distinct methodologies were identified. These are presented in Table 2. The evaluation methdologies mentioned most frequently were the MOST (n=5), SMART (n=4), and the CEEBIT (n=4). Due to time and resource constraints, the less frequently mentioned methodologies (e.g. studies using

principles of existing study design and the Mobile App Rating Scale (MARS)) were not included in the comparison.

Table 2 – Guidelines and Recommendations for Alternative mHealth Evaluation

Guidelines & Recommendations for mHealth Evaluation		
Multiphase Optimization Strategy (MOST)	5	(Nilsen et al., 2012; Whittaker, Merry, Dorey, & Maddison, 2012; Kumar et al., 2013; Pham et al., 2016b; Buscemi, Janke, Kugler, Duffecy, Mielenz, George, & Gorin, 2017)
Sequential Multiple Assignment Randomized Trial (SMART)	4	(Nilsen et al., 2012; Kumar et al., 2013; Pham et al., 2016b; Buscemi et al., 2017)
Continuous Evaluation of Evolving Behavioral Intervention Technologies (CEEBIT)	4	(Kumar et al., 2013; Mohr, Schueller, Riley, Brown, Cuijpers, Duan, Kwasny, Stiles-Shields, & Cheung, 2015; Pham et al., 2016b)
Evaluation utilizing principles of existing study designs (e.g. n-of-1, stepped-wedge RCT)	3	(Nilsen et al., 2012; Whittaker et al., 2012; Pham et al., 2016b; White et al., 2017)
Mobile Application Rating Scale (MARS)	2	(Stoyanov, Hides, Kavanagh, & Wilson, 2016; White et al., 2017)
Integrate, Design, Assess and Share (IDEAS) Framework	1	(Mummah, Robinson, King, Gardner, & Sutton, 2016)
ResearchKit	1	(Pham et al., 2016b)
Micro-randomized trial	1	(Pham et al., 2016b)
Meaningful Use Criteria	1	(Song, Kim, & Yi, 2013)
Collaborative Adaptive Interactive Technology	1	(White et al., 2017)

The following section presents an overview of the methodology of each of the four evaluation techniques, and is followed by Table 3 which compares each evaluation technique against the criteria required for successful mHealth evaluation.

Findings

Randomized Controlled Trial (RCT)

The RCT is a trial in which subjects are randomly assigned to one of two groups: one (the experimental group) receiving the intervention that is being tested and the other (the comparison

group or control) receiving an alternative (conventional or placebo) treatment (Kendall, 2003). All intervention groups are treated identically except for the experimental treatment (Sibbald & Roland, 1998). The two groups are then followed up to see if there are any differences between them in outcome. The results and subsequent analysis of the trial are used to assess the effectiveness of the intervention, which is the extent to which a treatment, procedure or service does more good than harm (Kendall, 2003).

RCTs are considered to be the "gold standard" for examining the effectiveness of a medical intervention in a clinical domain due to their ability to control for confounding factors and bias (Kendall, 2003). Calls for greater rigor in evaluation has increased the number of mHealth RCTs conducted in developed and developing countries (Burns et al., 2016). The majority of mHealth researchers are continuing to use the RCT for evaluating mHealth interventions, suggesting that researchers view this design to be the "gold-standard" for any clinical trial evaluating intervention efficacy (Pham et al., 2016b).

Continuous Evaluation of Evolving Behavioral Intervention Technologies (CEEBIT)

Behavioral Intervention Technologies (BITs) are web-based and mobile interventions intended to support patients and consumers in changing behaviors related to health, mental health and well-being (Mohr, Cheung, Schueller, Brown, & Duan, 2013). CEEBIT is an evaluation method involving the deployment of substantively new versions of an intervention along with the previous version, with users randomized to available versions. The most efficacious version, based on a priori criteria, is retained (Kumar et al., 2013). This framework was proposed by Mohr et al. (2013) as an alternative to the RCT.

This methodology addresses the current weak evidence base and lack of discussion addressing how to evaluate interventions effectively and efficiently, and provides a solution to the challenge of rapid change, evolution and expanding expectations (Mohr et al., 2013). CEEBIT is suitable for the ongoing evaluation of interventions as they go to scale, adapting to the changing technological landscape and allowing for intervention improvement over time (Kumar et al., 2013). The method is statistically powered to continuously evaluate intervention efficacy and accounts for updated versions through a sophisticated elimination process (Pham et al., 2016b).

Multiphase Optimization Strategy (MOST)

MOST uses a principled method for identifying which components are active in an intervention and which levels of each component lead to the best outcomes. Its underlying principles are drawn from engineering and emphasize efficiency (Collins, Murphy, & Strecher, 2007). Promising components of an intervention are identified in a screening phase through either factorial or fractional factorial analysis of variance design (Kumar et al., 2013). MOST uses three phases as a replacement for the cycle of confirmatory trial, exploratory analysis, revision and subsequent confirmatory trial (Collins, Murphy, Nair, & Strecher, 2005). The final, optimized intervention is then evaluated in a standard RCT in the confirming phase to evaluate efficacy (Collins et al., 2007).

The traditional approach to intervention development has involved constructing an intervention a priori and then evaluating it in a standard RCT, after which, post-hoc analyses are done and adjustments are made (Collins et al., 2007). MOST is a system aimed at creating an optimal version of a multicomponent intervention (Clough & Casey, 2015), which can then be evaluated in a RCT.

Sequential Multiple Assignment Randomized Trial (SMART)

The SMART approach is a randomized experimental design that has been developed especially for building time-varying adaptive interventions (Collins et al., 2007). It allows investigators to evaluate the timing, sequencing, and adaptive selection of treatments in a principled fashion by use of randomized data (Almirall, Compton, Gunlicks-Stoessel, Duan, & Murphy, 2012). All questions within the SMART trial are addressed by means of randomized experiments and the end goal of the SMART approach is the development of evidence-based adaptive intervention strategies which are then evaluated in a subsequent RCT (Collins et al., 2007).

The SMART approach considers the order of components in an intervention, as opposed to considering each component in isolation (Collins et al., 2007). Researchers decide which aspects of treatment require investigation and then individuals are randomly assigned to various intervention choices over time (Kumar et al., 2013). With this approach, a number of important treatment questions can be answered; the optimal length of the intervention, the best approach to take for treatment of non-responders and the level of support required for individuals (Clough et al., 2015). The SMART technique can be integrated into the MOST procedure, or be used as a stand-alone technique (Collins et al., 2007).

The following section presents a comprehensive comparison of the four evaluation techniques outlined previously. This is followed by an in-depth discussion of how each evaluation technique meets the unique characteristics of mHealth.

Comparison of Evaluation Techniques

Table 3 outlines the main characteristics of each evaluation technique in a comparison of their suitability to mHealth evaluation. These criteria have been determined as important in terms of

choosing an evaluation technique as they examine both protocol-related factors, including standard execution and data collection techniques as well as logistical, resource requirement factors such as time and cost.

Table 3 - Comparison of Evaluation Techniques

	RCT Characteristics	CEEBIT Characteristics	MOST Characteristics	SMART Characteristics	mHealth Challenges
Data Collection & Analysis	Quantitative analysis - The analysis is focused on estimating the size of difference in predefined outcomes (Sibbald & Roland 1998).	Collects outcome and use data in real time. CEEBIT capitalizes on data generates by BITs to continuously evaluate efficacy in a manner consistent with the current socio-technologic environment (Mohr et al., 2013).	A key feature of MOST is that each new intervention produced will have been engineered, and empirically demonstrated, to be an improvement over the previous version (Collins, Nahum-Shani, & Almirall, 2011b).	SMART designs provide a framework to empirically determine the most powerful version of an intervention (Clough et al., 2015). However, the SMART design does not compare the intervention to a control or comparative treatment condition (Clough et al., 2015).	mHealth literature suggests the need for mixed-method evaluation to accurately capture the socio-technical reasons for using an inititative (Chib et al., 2014).
Standard Execution Protocol	Double blinding (usually investigator and participants) is the usual standard and will eliminiate any confounding factors occuring after randomisation (Kendall, 2003).	In instances when non-randomized assignment methods are warranted, statistical methods can mitigate overt bias when all confounding factors are observed (Mohr et al., 2013).	Based on randomized experimentation, meaning that a high degree of confidence can be placed on the results (Collins et al., 2007).	Based on randomized experimentation, placing a high degree of confidence on the results (Collins et al., 2007). To avoid information bias, the use of a blinded, independent evaluator is suggested (Almirall, Nahum-Shani, Sherwood, & Murphy, 2014).	It is difficult to blind participants receiving an mHealth intervention due to the physical presence of the device (Eysenbach, 2002; Stroux, 2012).
Sample Size	The sample size must be large enough to eliminate chance (Kendall, 2003).	The sample size required in a CEEBIT methodology is considerably reduced due to a more liberal Type I error rate of 50% (Mohr et al., 2013).	Interaction effect sizes tend to be small, making it important to ensure that there is sufficient statistical power to test any interactions that are of particular interest (Collins et al., 2007).	The sample size required will depend on the primary aim for the trial and the level of analyses. In a longitudinal comparison of two groups, the sample size requirement is identical to that of a two-group, longitudinal RCT (Almirall et al., 2014).	Recruiting adequate numbers may be challenging in developing countries where cultural and religious barriers may resist technological change (Tariq et al., 2011).

	RCT Characteristics	CEEBIT Characteristics	MOST Characteristics	SMART Characteristics	mHealth Challenges
Cost	RCTs are expensive to carry out (Comstock, 2012) often due to the large sample size and length of follow-up time required (Rosen, Manor, Engelhard, & Zucker, 2006).	The reduced sample size and rapid, real-time evaluation may contribute to lower financial costs than other methodologies.	MOST does not directly assess the overall effectiveness of the intervention to a comparative treatment or control condition but the process does ensure that the most efficacious version of the intervention goes forward to the final testing stage, thereby making for a more efficient use of time and resources (Clough et al., 2015).	It would be more cost-effective both in terms of dollars spent, and in terms of the value of scientific information gained to use the SMART methodology, than to use a RCT to evaluate each question in the SMART stages (Almirall et al., 2014).	In low-income countries where there may not be the resources to carry out expensive trials (Rosen et al., 2006; World Health Organisation, 2012) and also in high-income countries where the sheer volume of mHealth interventions available, may mean it is not feasable to carry out resource-intensive evaluations.
Protocol	Rigid protocol, designed for the elimination of bias and confounding factors (Pham et al., 2016b).	Protocol is fluid, allowing for consumer choice to be incorporated into the evaluation as a fully observed pre-randomization factor (Mohr et al., 2013).	Protocol is fluid, exact details about its implementation depend on the application (Collins et al., 2007). It is a "general approach," rather than an off-the-shelf procedure (Collins, Baker, Mermelstein, Piper, Jorenby, Smith, Christiansen, Schlam, Cook, & Fiore, 2011a).	SMART designs are not, as per common misconceptions, "adaptive trial designs," they are a fixed study design (Almirall et al., 2012; Almirall et al., 2014).	Software is meant to evolve, change and progress over time at a rapid pace (Ben-Zeev et al., 2015).
Time	RCTs are notoriously long (Pham et al., 2016b), with the entire process, including funding proposals and publication, taking upto 17 years (Mohr et al., 2013; Pagoto & Bennett, 2013).	CEEBIT can support the rapid evaluation of BITs in near-real time through deployment sites located in care-providing organizations or commercial market places with the aim of protecting consumers from ineffective or inferior BITs (Mohr et al., 2013).	A challenge is whether a full cycle of MOST can be completed within the five-year duration of the typical National Institutes of Health (NIH) funding cycle (Collins et al., 2011a).	The overarching aim of a SMART is to construct a high quality adaptive intervention based on data (Almirall et al., 2014). This may save resources in the long run as the end intervention will be already optimized.	In the mHealth field, the fast pace at which technology evolves, may make lengthy study designs unsuitable for evaluation as in the time it takes to design and evaluate an intervention, the mHealth space may move so fast that the intervention is obsolete before it has even been implemented (Kumar et al., 2013).

Discussion

There are a number of factors which must be taken into account when selecting an evaluation method for a mHealth intervention. This study has suggested six of these factors, identified as important from recent literature surrounding the challenges and unique characteristics of mHealth evaluations. The mHealth literature widely suggests that a mixed-methods approach should be used when evaluating a mHealth intervention (Chib et al., 2014). Quantitative data is important to determine if, and to what extent an intervention is functional and beneficial, but because of the many socio-technical aspects of mHealth, failing to include a qualitative evaluation may mean that the intervention will fail to be implemented and sustained (Chib et al., 2014). These sociotechnical factors include the social, cultural, religious and behavioral interactions of the end user, as well as other technological issues (Chib et al., 2014), such as adequate cellular service and charging points, an issue which is particularly important in LMICs. The data collection of a standard RCT is quantitative, and therefore is unable to isolate the socio-technical aspects of mHealth which are so important for their successful implementation (Pham et al., 2016b). In addition, the inclusion/exclusion criteria employed in RCTs can produce high internal validity but may produce poor external validity, impacting the effectiveness of interventions in a "real-world" setting (Clough et al., 2015). The CEEBIT method has the potential to include a qualitative aspect but the selection of outcome measures will depend on whether the research question primarily pertains to the efficacy or the effectiveness of the intervention (Clough et al., 2015). However, if a researcher is primarily concerned with the efficacy of an intervention, it could be possible that a qualitative evaluation will be absent and compromise the implementation of the intervention. Similarly, the MOST method consists of three phases, each of which addresses a different set of questions about the intervention by means of randomized, quantitative experiments and ending with the optimized

intervention being evaluated in a RCT. Although the SMART trial follows a quantitative methodology, the pilot SMART trial can include qualitative aspects such as focus groups or interviews to help uncover new and potentially important tailoring variables (Almirall et al., 2012).

All of the methodologies outlined in Table 1 include a randomization process. Randomization is used to eliminate certain biases and confounding factors and therefore allows a high level of confidence to be placed on the results. The randomizations in SMART allow unbiased comparisons between treatment components at each decision stage in their development (Almirall et al., 2014). As outlined, there is a difficulty in blinding recipients of a mHealth intervention due to the physical presence of the device but the SMART trial suggests the use of an independent evaluator who is blind to treatment assignment to eliminate any information bias which may result (Almirall et al., 2014). This is important because a lack of blinding in a study design could lead to an over-estimation of the effects of an intervention, as was illustrated by Colditz, Miller, and Mosteller (1989) who found that medical interventions evaluated within randomized trials that did not use a double-blind design reported a significantly greater likelihood of success on average than the studies that used double blinding.

A large sample size may be problematic in the area of mHealth, particularly in low- and middle-income countries where there may be cultural and religious barriers to technology, resisting change and creating challenges with recruitment (Tariq et al., 2011) as well as technological barriers which may prevent usage such as intermittent reception and unreliable electricity for charging of devices. RCTs require a relatively large sample size in order to eliminate chance (Kendall, 2003), MOST requires a large enough sample size to detect small variations in interaction effects with sufficient statistical power (Collins et al., 2007) and SMART claims to require a sample size similar to that of a RCT (Almirall et al., 2014), but the use of a fractional factorial design can reduce the required

sample size. However, CEEBIT claims to require a much smaller sample size due to its Type 1 error rate of 50%, compared to the standard Type 1 error rate of 5%.

RCTs are notoriously expensive (Comstock, 2012), rendering them potentially unsuitable for LMICs where financial resources are scarce (Rosen et al., 2006; Frethein, Witter, Lindahl, & Olsen, 2012), or even in a developed country, given the sheer volume of mHealth interventions currently requiring evaluation, as outlined earlier. The MOST and SMART methodologies both suggest greater cost-effectiveness than RCTs as they ensure that the most efficient, fully optimized version of the intervention is going forward to be tested in a RCT (Almirall et al., 2012; Clough et al., 2015) and the use of MOST does not require more resources than the classical approach (Kugler, Balantekin, Birch, & Savage, 2016). Although it could be argued that because these methodologies still require RCT evaluation, they are equally as expensive, but the typical cycle of intervention-RCT-post hoc analyses-revision of intervention-RCT considerably increases the length of time the intervention spends in development (Collins et al., 2007). By presenting an optimized intervention for RCT, both financial and time resources can be saved because the traditional RCT evaluates the intervention only as a whole, as opposed to evaluating the individual components in isolation (Collins et al., 2007). In contrast to the RCT, which aims to evaluate an already developed intervention, SMART aims to develop an adaptive intervention based on data (Almirall et al., 2014).

Timing is an important factor in the evaluation of mHealth interventions given how fast technology evolves and develops. This is illustrated in the exponential growth and change to the field of mHealth which has occurred in the last decade. The RCT process is typically lengthy, taking 5.5 years on average to complete (Pham et al., 2016b) and there is the risk that a mHealth intervention may become obsolete in the time it takes to evaluate it. The CEEBIT methodology deals with this

issue, evaluating interventions in near real-time through constant deployment (Mohr et al., 2013). The MOST methodology faces challenges as to whether a full cycle can be completed within the typical NIH funding cycle (Collins et al., 2011a). Although as Collins et al. (2011a) argue, this five-year funding cycle is merely an administrative necessity with no intrinsic scientific meaning or merit. However, it may still be incompatible with the speed at which the field of mHealth develops. As mentioned previously, a relatively short period of time in the context of drug development and trialing equates to a very long period of time in the mHealth field and the MOST methodology may be, similarly to the RCT, regarded as too long for mHealth.

This research aims to answer the research question: *"How can we characterize popular evaluation approaches to mHealth evaluation and what gaps exist in these approaches ?"* We found that of the four evaluation techniques presented in this paper, each has many benefits in the context of a mHealth evaluation. However there are several gaps in their strengths, particularly in terms of time and cost, both of which are critical in LMICs with limited resources. Although none of the methodologies include mixed-method components, there is the potential for qualitative evaluation to be included in their design. The RCT is hindered by its high cost, and large sample size, but MOST and SMART go some way in filling this gap by allowing for the use of smaller sample sizes using fractional factorial designs. Similarly, the cost of a RCT is high, and although the MOST and SMART methods do recommend eventual evaluation in a RCT, the presentation of a final product that is optimized may save costs in the long run. Time appears to be a common issue across the evaluation techniques, except CEEBIT, which allows for fast deployment of interventions, coupled with a much smaller sample size. We present three examples of the application of MOST, SMART and CEEBIT in the evaluation of interventions, as an indication of how these may be applied to mHealth evaluation. Since each methodology is relatively novel, very

few applied examples of these methodologies were identified in the mHealth literature. The three following examples were chosen as they present constructive feedback on the application of each methodology, in mHealth or similar fields. We present these in an attempt to help us to understand how the alternative methodologies may be used in a mHealth evaluation.

The Prevention of Substance Use in College Students - MOST

The MOST methodology was applied during the development of an online intervention for the prevention of substance use among college student-athletes (Wyrick, Rulison, Fearnow-Kenney, Milroy, & Collins, 2014). The authors applied MOST after running an initial pilot of their program using the classic "treatment package" approach. They used MOST to optimize the lesson components of the intervention. The final optimized version of the intervention was then evaluated in a RCT with the original version as the comparison group. The authors reported the primary challenge as having to break the existing intervention into components that could be separated and tested. They argue that the optimization phase does not require an unusually large sample size, as factorial and fractional factorial experimental designs are used. However, they acknowledged that because they used an iterative approach involving three experiments, that the total sample size did amount to a larger sample than that required by a RCT. Given the time restraints, they were unable to follow students for an extended period of time, suggesting that the approach may be unsuitable for outcomes that cannot be captured in a short time-frame. However, the longer term RCT evaluation will be able to measure these. The author's state that MOST efficiently uses scarce financial and participant resources, but the iterative process of intervention improvement will always be time-constrained (Wyrick et al., 2014).

Evaluating the Effectiveness of Antipsychotic Treatment in Patients – SMART

An applied example of the SMART methodology for mHealth evaluation was unavailable, but we identified one case study comparing a SMART study with traditional RCTs. Moodie, Karran, and Shortreed (2016) compared the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) SMART study with traditional RCTs. CATIE was a multisite study designed to evaluate the effectiveness of antipsychotic treatment strategies for patients with schizophrenia. They found that SMARTs allow the possibility of studying treatment interactions and delayed effects so may provide a "real-world" assessment of treatment sequences than RCTs. The study also had a very low risk of bias, and the retention rate was higher than those typically observed in RCTs. However, they did not find evidence to suggest greater generalizability when using SMART (Moodie et al., 2016). Although there is a lack of literature documenting the application of SMART for mHealth intervention, their use in multi-component interventions is promising, and could be applied to mHealth (Kumar et al., 2013).

Continual Evaluation of a Smartphone Application for Anxiety in Adults - CEEBIT

A hypothetical application of the use of CEEBIT is presented in an evaluation of the PsychAssist app by Clough et al. (2015). PsychAssist aims to assist in the treatment of anxiety in adults. The authors use a 50% alpha rate based on the assumption that there will be symmetry of preference among the deployed versions. This sample size results in a smaller required sample size, enabling faster testing, dissemination and elimination of inferior versions. Consumers are protected against prolonged use of an inferior version as a-priori rules are established to determine which versions to eliminate. The authors chose a clinical outcome to determine inferiority, but acknowledged that other outcomes such as usage or satisfaction could also be used. The authors argue that CEEBIT is well designed to keep pace with the evolving nature and speed of development of mHealth interventions (Clough et al., 2015).

Conclusion

To determine the success of a mHealth intervention, evaluation should examine user feedback, the robustness of the technology, user engagement strategies, user interaction, safety and accuracy (White et al., 2017). Additionally, to ensure that mHealth interventions have the maximum impact on healthcare worldwide, evaluation must be cost-effective, using an evaluation technique which is fit-for-purpose and allows resources to be used in the most efficient way. The comparison of the RCT, CEEBIT, MOST and SMART methodologies shows that each approach is capable of examining a number of these factors. However, it is apparent that the methodologies discussed are unable to examine all elements simultaneously within the strict time constraints imposed by the rapidly evolving field of mHealth. This may mean that, particularly in LMICs, that mHealth interventions are not able to fulfil their potential impact on socio-economic development. mHealth has the potential to overcome physical and social barriers to healthcare access, empowering patients, and allowing them to take a more active role in their health. However, without thorough and timely evaluations, the true potential of mHealth across healthcare, social and economic environments cannot be realized and harnessed.

The identification and use of an effective tool for mHealth evaluation will have a far-reaching impact, benefitting those in mHealth development and research, strengthening system developer's ability to adequately plan and project the costs, and the length of time required to evaluate their intervention. Additionally, as the regulatory landscape of mHealth evolves, those responsible will have the tools to make more informed decisions regarding the safety and effectiveness of publicly available interventions. Funders will be able to prioritize projects which have clear, effective evaluation plans in place and policy makers will be supported in their decision-making (Mechael et al., 2010). Policy makers need to make important decisions on the use of public funds – to target

which disease areas, which populations and which interventions (Baltussen & Niessen, 2006; World Health Organization, 2017). In the context of mHealth, the challenge lies in the vast number of new interventions available for a myriad of health problems. Healthcare policy makers and health researchers have different goals, languages and attitudes towards information (Choi, Pang, Lin, Puska, Sherman, Goddard, Ackland, Sainsbury, Stachenko, Morrison, & Clottey, 2005). Healthcare policy makers are concerned with maximizing population health with interventions that are evidence-based, cost effective, and equitable. The evidence for which is produced by health researchers (Clancy, Glied, & Lurie, 2012). However, each of these aspects are based on a single criteria, making it difficult to prioritize one agenda. Policy makers need to make choices taking into account multiple criteria simultaneously. There is a documented lack of economic evaluations on mHealth interventions, particularly in LMICs, and this is a barrier to implementation and limited policy investment (Iribarren, Cato, Falzon, & Stone, 2017). Studies which have taken an economic approach consistently indicate that the socio-economic benefits of health research broadly exceed the research costs (Roback et al., 2011). However, conducting mHealth evaluations which are ineffective may result in wasted resources, thereby diluting the socio-economic development impact of mHealth

This research serves as a useful first step toward developing an evaluation technique which incorporates a multi-criteria decision making tool for policy makers to use in prioritizing mHealth interventions. The use of such a tool would provide a focus on providing data which is most valuable to decision-makers and ensure that social preferences, epidemiological priorities and ethical values are not neglected in the decision making process (Drake, de Hart, Monleón, Toro, & Valentim, 2017). It would offer a structured and transparent approach to identify an intervention by allowing for clear consideration of the importance of different criteria and improve the quality,

transparency, consistency and accountability of decision making (Adunlin, Diaby, & Xiao, 2015; Thokala, Devlin, Marsh, Baltussen, Boysen, Kalo, Longrenn, Mussen, Peacock, Watkins, & Ijzerman, 2016).

This investigation has limitations; the evaluation methodologies have been described in their simplest, pure state and the observations made apply only to those. Evaluation methodologies are adapted to suit the context in which they are being carried out and these context-dependent adaptations have not been taken into account in the above comparison. It was possible to examine only the most common evaluation methodologies and it is possible that this comparison of evaluation methodologies against mHealth criteria is not exhaustive. Further investigation should examine the contextual adaptations applied to the RCT, CEEBIT, MOST and SMART methodologies and the potential of these adaptations to produce a better fit for mHealth evaluation. Additionally, due to time and resources constraints, we were unable to compare other, less frequently mentioned methodologies against the mHealth criteria. There is further scope to assess the suitability of adapted methodologies, such as the n-of-1, and the stepped wedge trial, as well as the MARS scale for evaluating mHealth interventions.

Contributions and Future Recommendations

This study has identified a set of criteria, important to mHealth evaluation, by which existing evaluation methodologies can be compared against. These criteria have been determined as important in terms of choosing an evaluation technique as they examine both protocol-related factors and logistical, resource requirement factors such as time and cost. Further, we have demonstrated that the methodologies discussed in this study are unable to simultaneously examine all elements of mHealth evaluation, within the strict time constraints imposed by the rapidly growing field of mHealth. We have identified a mismatch between the requirements of mHealth

evaluations and the methodologies currently available. These mismatches will inform future research into the development of a fit-for-purpose evaluation technique. An evaluation technique that is fit-for-purpose will allow for timely and economical evaluation of mHealth interventions.

The development of a fit-for-purpose evaluation technique for mHealth will ease the strain on regulating bodies as they struggle to deal with the volume of untested mHealth interventions. A fast, reliable and cost-effective method of evaluation will allow for high-quality mHealth interventions to reach beyond the pilot stage, and lessen the plague of "pilotitis" which is hampering progress of the field. Meeting the complex criteria for effective mHealth evaluation will strengthen the ability of policy makers to make confident funding decisions when faced with the vast number of mHealth interventions being proposed. Additionally, planning for the use of a fit-for-purpose technique in the evaluation in a mHealth project will allow research teams to submit stronger proposals to funders.

There is a need for further research into the area of mHealth evaluation to build upon the methodologies currently available for the assessment of mHealth interventions. The mismatches between the mHealth criteria and the evaluation methodologies must be addressed to allow for the development of a holistic approach that will allow evaluations of mHealth interventions to provide the most robust and thorough results and contribute to timely, successful and long-term mHealth implementation.

A potential next step toward the development of a fit-for-purpose mHealth evaluation technique could be the development of a detailed decision model. Such a model could be used in practice to assess which of the currently available evaluation techniques are most suitable for the mHealth intervention under evaluation. As previously discussed, a multi-criteria decision model for mHealth evaluation would have far reaching benefits for the mHealth field by providing a standard

of evaluation, increasing confidence in the field. The development of such a model would allow regulatory bodies to more clearly outline which types of mHealth fall under their remit, and as such, which evaluation technique they require to meet standards. Additionally, mHealth project teams following a decision model will be able to strengthen applications for funding by complying with standards. Finally, policy makers and governments will be able to make strong, confident decisions about the funding of mHealth interventions that have been evaluated using a decision model.

Acknowledgements

This work was supported by the Supporting LIFE project (305292) which is funded by the Seventh Framework Programme for Research and Technological Development of the European Commission www.supportinglife.eu.

Author Biographies

Samantha Dick is a researcher at University College Cork (UCC), Ireland. She holds a BSc in Public Health and earned an MSc by research in the area of mobile health (mHealth) evaluation from the Health Information Systems Research Centre (HISRC) at UCC.

Dr Yvonne O'Connor is a lecturer in Cork University Business School, University College Cork (UCC), Ireland. She is also a senior researcher within the Health Information Systems Research Centre (HISRC) at UCC. Her research focuses on research focuses on the introduction, sustained use and ethical implications of information systems in healthcare.

Dr Ciara Heavin is a Senior Lecturer in Business Information Systems at Cork University Business School, University College Cork, Ireland. Her research focuses on opportunities for using information systems (IS) in the global healthcare ecosystem and in digital transformation. Dr Heavin has directed funded research in the investigation, development, and implementation of innovative technology solutions in the health/healthcare domain. She has published articles in several top international IS journals and conference proceedings. Dr Heavin has co-authored two books with Daniel J. Power: Decision Support, Analytics, and Business Intelligence and Data-Based Decision Making and Digital Transformation.

References

Adunlin, G., Diaby, V., & Xiao, H. (2015). Application of multicriteria decision analysis in health care: a systematic review and bibliometric analysis. *Health expectations: an international journal of public participation in health care and health policy, 18*(6), 1894-1905. doi:10.1111/hex.12287

Almirall, D., Compton, S. N., Gunlicks-Stoessel, M., Duan, N., & Murphy, S. A. (2012). Designing a Pilot Sequential Multiple Assignment Randomised Trial for Developing an Adaptive Treatment Strategy. *Statistics in Medicine, 31*(17), 1887-1902. doi:10.1002/sim.4512.

Almirall, D., Nahum-Shani, I., Sherwood, N. E., & Murphy, S. A. (2014). Introduction to SMART designs for the development of adaptive interventions: with application to weight loss research. *Translational Behavioral Medicine, 4*(3), 260-274. doi:10.1007/s13142-014-0265-0

Baltussen, R., & Niessen, L. (2006). Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost Effectiveness and Resource Allocation, 4*, 14-14. doi:10.1186/1478-7547-4-14

Battle, J. D., Farrow, L., Tibaijuka, J., & Mitchell, M. (2015). mHealth for Safer Deliveries: A mixed methods evaluation of the effect of an integrated mobile health intervention on maternal care utilization. *Healthc (Amst), 3*(4), 180-184. doi:10.1016/j.hjdsi.2015.10.011

Ben-Zeev, D., Scheuller, S. M., Begale, M., Duffecy, J., Kane, J. M., & Mohr, D. C. (2015). Strategies for mHealth Research: Lessons from 3 Mobile Intervention Studies. *Administration and Policy in Mental Health and Mental Health Services Research, 42*(2), 157-167. doi:10.1007/s10488-014-0556-2

Beyrer, C. (2014). Pushback: The Current Wave of Anti-Homosexuality Laws and Impacts on Health. *PLOS Medicine, 11*(6), e1001658. doi:10.1371/journal.pmed.1001658

Blanford, J. I., Kumar, S., Luo, W., & MacEachren, A. M. (2012). It's a long, long walk: accessibility to hospitals, maternity and integrated health centers in Niger. *International journal of health geographics, 11*(1), 24. doi:10.1186/1476-072X-11-24

Boudreaux, E. D., Waring, M. E., Hayes, R. B., Sadasivam, R. S., Mullen, S., & Pagoto, S. (2014). Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. *Translational Behavioral Medicine, 4*(4), 363-371. doi:10.1007/s13142-014-0293-9

Bourouis, A., Zerdazi, A., Feham, M., & Bouchachia, H. (2013). M-Health: Skin Disease Analysis System Using Smartphone's Camera. *Procedia Computer Science, 19*, 1116-1120. doi:10.1016/j.procs.2013.06.157

Braun, R., Lasway, C., Agarwal, S., L'Engle, K., Layer, E., Silas, L., Mwakibete, A., & Kudrati, M. (2016). An evaluation of a family planning mobile job aid for community health workers in Tanzania. *Contraception, 94*(1), 27-33. doi:10.1016/j.contraception.2016.03.016

Brown III, W., Yen, P.-Y., Rojas, M., & Schnall, R. (2013). Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. *Journal of Biomedical Informatics, 46*(6), 1080-1087. doi:10.1016/j.jbi.2013.08.001

Brown, M. T., & Bussell, J. K. (2011). Medication adherence: WHO cares? *Mayo Clinic Proceedings, 86*(4), 304-314. doi:10.4065/mcp.2010.0575

Buman, M. P., Epstein, D. R., Gutierrez, M., Herb, C., Hollingshead, K., Huberty, J. L., Hekler, E. B., Vega-López, S., Ohri-Vachaspati, P., & Hekler, A. C. (2015). BeWell24: development and process evaluation of a smartphone "app" to improve sleep, sedentary, and active behaviors in US Veterans with increased metabolic risk. *Translational Behavioral Medicine, 6*(3), 438-448. doi:10.1007/s13142-015-0359-3

Burns, K., Keating, P., & Free, C. (2016). A systematic review of randomised control trials of sexual health interventions delivered by mobile technologies. *BMC Public Health, 16*(1), 778. doi:10.1186/s12889-016-3408-z

Burns, M. N., Begale, M., Duffecy, J., Gergle, D., Karr, C. J., Giangrande, E., & Mohr, D. C. (2011). Harnessing context sensing to develop a mobile intervention for depression. *Journal of Medical Internet Research, 13*(3), e55. doi:10.2196/jmir.1838

Buscemi, J., Janke, E. A., Kugler, K. C., Duffecy, J., Mielenz, T. J., George, S. M. S., & Gorin, S. N. S. (2017). Increasing the public health impact of evidence-based interventions in

behavioral medicine: new approaches and future directions. *Journal of behavioral medicine, 40*(1), 203-213. doi:10.1007/s10865-016-9773-3

Cafazzo, J. A., Casselman, M., Hamming, N., Katzman, D. K., & Palmert, M. R. (2012). Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. *Journal of Medical Internet Research, 14*(3). doi:10.2196/jmir.2058

Cameron, D., & Berkowitz, B. (2016). The state of gay rights around the world. Retrieved from https://www.washingtonpost.com/graphics/world/gay-rights/ [Accessed: 11th September 2017] Archived by WebCite® at http://www.webcitation.org/77zmUiYBq

Catalani, C., Philbrick, W., Fraser, H., Mechael, P., & Israelski, D. M. (2013). mHealth for HIV Treatment and Prevention: A Systematic Review of the Literature. *The Open AIDS Journal, 7*, 17-41. doi:10.2174/1874613620130812003

Chang, L. W., Kagaayi, J., Arem, H., Nakigozi, G., Ssempijja, V., Serwadda, D., Quinn, T. C., Gray, R. H., Bollinger, R. C., & Reynolds, S. J. (2011). Impact of a mHealth intervention for peer health workers on AIDS care in rural Uganda: a mixed methods evaluation of a cluster-randomized trial. *AIDS and Behavior, 15*(8), 1776. doi:10.1007/s10461-011-9995-x

Chib, A. (2018). Afterword: Reflections on a Decade of mHealth Innovation in Asia. In E. Baulch, J. Watkins, & A. Tariq (Eds.), *mHealth Innovation in Asia: Grassroots Challenges and Practical Interventions* (pp. 123-131). Dordrecht: Springer Netherlands.

Chib, A., van Velthoven, M. H., & Car, J. (2014). mHealth Adoption in Low-Resource Environments: A Review of the Use of Mobile Healthcare in Developing Countries. *Journal of Health Communication, 20*(1), 4-34. doi:10.1080/10810730.2013.864735

Choi, B., Pang, T., Lin, V., Puska, P., Sherman, G., Goddard, M., Ackland, M., Sainsbury, P., Stachenko, S., Morrison, H., & Clottey, C. (2005). Can scientists and policy makers work together? *Journal of Epidemiology and Community Health, 59*(8), 632-637. doi:10.1136/jech.2004.031765

Clancy, C. M., Glied, S. A., & Lurie, N. (2012). From research to health policy impact. *Health services research, 47*(1pt2), 337-343. doi:10.1111/j.1475-6773.2011.01374.x

Clough, B. A., & Casey, L. M. (2015). Smart designs for smart technologies: Research challenges and emerging solutions for scientist-practitioners within e-mental health. *Professional Psychology-Research and Practice, 46*(6), 429-436. doi:10.1037/pro0000053

Cohen, J., Bancilhon, J. M., & Grace, T. (2018). Digitally connected living and quality of life: An analysis of the Gauteng City-Region, South Africa. *The Electronic Journal of Information Systems in Developing Countries, 84*(1), e12010. doi:10.1002/isd2.12010

Colditz, G. A., Miller, J. N., & Mosteller, F. (1989). How Study Design Affects Outcomes in Comparisons of Therapy. I: Medical. *Statistics in Medicine, 8*, 441-454.

Collins, L. M., Baker, T. B., Mermelstein, R. J., Piper, M. E., Jorenby, D. E., Smith, S. S., Christiansen, B. A., Schlam, T. R., Cook, J. W., & Fiore, M. C. (2011a). The Multiphase Optimization Strategy for Engineering Effective Tobacco Use Interventions. *Annals of Behavioral Medicine, 41*(2), 208-226. doi:10.1007/s12160-010-9253-x

Collins, L. M., Murphy, S. A., Nair, V. N., & Strecher, V. J. (2005). A Strategy for Optimizing and Evaluating Behavioural Interventions. *Annals of Behavioural Medicine, 30*(1), 65-73. doi:10.1207/s15324796abm3001_8

Collins, L. M., Murphy, S. A., & Strecher, V. (2007). The Multiphase Optimization Strategy (MOST) and the Sequential Multiple Assignment Randomized Trial (SMART) New

Methods for More Potent eHealth Interventions. *American Journal of Preventative Medicine, 32*(5S), S112-S118. doi:10.1016/j.amepre.2007.01.022

Collins, L. M., Nahum-Shani, I., & Almirall, D. (2011b). Optimization of behavioral dynamic treatment regimens based on sequential, multiple assignment, randomized trial (SMART). *Clin Trials., 11*(4), 426-434. doi:10.1177/1740774514536795

Comstock, J. (2012). Is the RCT too slow for mobile health? Retrieved from http://www.mobihealthnews.com/19378/is-the-rct-too-slow-for-mobile-health [Accessed: 22nd December 2017] Archived by WebCite® at http://www.webcitation.org/77zmJVQfJ

Cooper, R. B., & Zmud, R. W. (1990). Information technology implementation research: a technological diffusion approach. *Management science, 36*(2), 123-139. doi:10.1287/mnsc.36.2.123

Dahlberg, K., Jaensson, M., Eriksson, M., & Nilsson, U. (2016). Evaluation of the Swedish Web-Version of Quality of Recovery (SwQoR): secondary step in the development of a mobile phone app to measure postoperative recovery. *JMIR research protocols, 5*(3). doi:10.2196/resprot.5881

Davey, S., & Davey, A. (2014). Mobile-health technology: Can it strengthen and improve public health systems of other developing countries as per Indian strategies? A systematic review of the literature. *International Journal of Medicine and Public Health, 4*(1), 40-45. doi:10.4103/2230-8598.127121

Davis, T. L., DiClemente, R., & Prietula, M. (2016). Taking mHealth Forward: Examining the Core Characteristics. *JMIR Mhealth Uhealth, 4*(3), e97. doi:10.2196/mhealth.5659

Déglise, C., Suggs, S., & Odermatt, P. (2012). Short Message Service (SMS) Applications for Disease Prevention in Developing Countries. *Journal of Medical Internet Research, 14*(1), e3. doi:10.2196/jmir.1823

Devi, B. R., Syed-Abdul, S., Kumar, A., Iqbal, U., Nguyen, P.-A., Li, Y.-C., & Jian, W.-S. (2015). mHealth: An updated systematic review with a focus on HIV/AIDS and tuberculosis long term management using mobile phones. *Computer Methods and Programs in Biomedicine, 122*(2), 257-265. doi:10.1016/j.cmpb.2015.08.003

Drake, J. I., de Hart, J. C. T., Monleón, C., Toro, W., & Valentim, J. (2017). Utilization of multiple-criteria decision analysis (MCDA) to support healthcare decision-making FIFARMA, 2016. *Journal of Market Access & Health Policy, 5*(1), 1360545. doi:10.1080/20016689.2017.1360545

Eysenbach, G. (2002). Issues in evaluating health websites in an Internet-based randomized controlled trial. *Journal of Medical Internet Research, 4*(3), e17. doi:10.2196/jmir.4.3.e17

Franz-Vasdeki, J., Pratt, B. A., Newsome, M., & Germann, S. (2015). Taking mHealth solutions to scale: enabling environments and successful implementation. *Journal of Mobile Technology in Medicine, 4*(1), 35-38. doi:10.7309/jmtm.4.1.8

Frethein, A., Witter, S., Lindahl, A., & Olsen, I. (2012). Performance-based financing in low- and middle-income countries: still more questions than answers. *Bulletin of the World Health Organization, 90*(8), 559-559A. doi:10.2471/BLT.12.106468

Ha, Y. P., Tesfalul, M. A., Littman-Quinn, R., Antwi, C., Green, R. S., Mapila, T. O., Bellamy, S. L., Ncube, R. T., Mugisha, K., & Ho-Foster, A. R. (2016). Evaluation of a mobile health approach to tuberculosis contact tracing in Botswana. *Journal of Health Communication, 21*(10), 1115-1121. doi:10.1080/10810730.2016.1222035

Hagopian, A., Rao, D., Katz, A., Sanford, S., & Barnhart, S. (2017). Anti-homosexual legislation and HIV-related stigma in African nations: what has been the role of PEPFAR? *Global Health Action, 10*(1), 1306391. doi:10.1080/16549716.2017.1306391

Hall, C. S., Fottrell, E., Wilkinson, S., & Byass, P. (2014). Assessing the impact of mHealth interventions in low- and middle-income countries - what has been shown to work? *Global Health Action, 7*. doi:10.3402/gha.v7.25606

Hamine, S., Gerth-Guyette, E., Faulx, D., Green, B. B., & Ginsburg, A. S. (2015). Impact of mHealth Chronic Disease Management on Treatment Adherence and Patient Outcomes: A Systematic Review. *J Med Internet Res, 17*(2), e52. doi:10.2196/jmir.3951

Hatt, L. E., Chatterji, M., Miles, L., Comfort, A. B., Bellows, B. W., & Okello, F. O. (2015). A False Dichotomy: RCTs and Their Contributions to Evidence-Based Public Health. *Global Health: Science and Practice, 3*(1), 138-140. doi:10.9745/Ghsp-D-14-00245

Health Service Executive. (2017). ICP for Prevention and Management of Chronic Disease. Retrieved from https://www.hse.ie/eng/about/who/cspd/icp/chronic-disease/ [Accessed: 27th October 2017] Archived by WebCite® at http://www.webcitation.org/77zl8Mjus

Heeks, R. (2002). Information Systems and Developing Countries: Failure, Success, and Local Improvisations. *The Information Society, 18*(2), 101-112. doi:10.1080/01972240290075039

Herron, J. (2016). Bad Apps: mHealth Apps Doubling as Medical Devices. *Journal of Electronic Resources in Medical Libraries, 13*(4), 177-181. doi:10.1080/15424065.2016.1256800

Hertzum, M., Bansler, J. P., Havn, E., & Simonsen, J. (2012). Pilot implementation: Learning from field tests in IS development. *Communications of the Association for Information Systems, 30*(1), 313-328. doi:10.17705/1CAIS.03020

Ho, K., Newton, L., Boothe, A., & Novak-Lauscher, H. (2015). *mobile Digital Access to a Web-enhanced Network (mDAWN): assessing the feasibility of mobile health tools for self-management of type-2 diabetes.* Paper presented at the AMIA Annual Symposium Proceedings, San Francisco. PMID:26958197

Holmen, H., Torbjornsen, A., Wahl, A. K., Jenum, A. K., Smastuem, M. C., Arsand, E., & Ribu, L. (2014). A Mobile Health Intervention for Self-Management and Lifestyle Change for Persons With Type 2 Diabetes, Part 2: One-Year Results From the Norwegian Randomized Controlled Trial RENEWING HEALTH. *JMIR mHealth and uHealth, 2*(4), e57. doi:10.2196/mhealth.3882

IMS Institute for Healthcare Informatics. (2015). *Patient Adoption of mHealth*. Retrieved from https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/patient-adoption-of-mhealth.pdf?la=en&hash=B3ACFA8ADDB143F29EAC0C33D533BC5D7AABD689

Iribarren, S. J., Cato, K., Falzon, L., & Stone, P. W. (2017). What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. *PLoS ONE, 12*(2). doi:10.1371/journal.pone.0170581

Istepanian, R. S. H., & Woodward, B. (2017). *mHealth Fundamentals and Application*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Jamison, J. C., Karlan, D., & Raffler, P. (2013). *Mixed method evaluation of a passive mHealth sexual information texting service in Uganda*. Retrieved from https://www.nber.org/papers/w19107.pdf

Kaplan, B. (2001). Evaluating informatics applications - some alternative approaches: theory, social interactionism, and call for methodological pluralism. *International journal of medical informatics, 64*(1), 39-56. doi:10.1016/S1386-5056(01)00184-8

Kendall, J. M. (2003). Designing a research project: randomised controlled trials and their principles. *Emergency Medicine Journal, 20*(2), 164-168. doi:10.1136/emj.20.2.164

Khan, R., & Socha-Dietrich, K. (2018). Investing in medication adherence improves health outcomes and health system efficiency: Adherence to medicines for diabetes, hypertension, and hyperlipidaemia *OECD Health Working Papers, 105*. doi:10.1787/18152015

Kowal, J., & Paliwoda-Pękosz, G. (2017). ICT for Global Competitiveness and Economic Growth in Emerging Economies: Economic, Cultural, and Social Innovations for Human Capital in Transition Economies. *Information Systems Management, 34*(4), 304-307. doi:10.1080/10580530.2017.1366215

Kugler, K. C., Balantekin, K. N., Birch, L. L., & Savage, J. S. (2016). Application of the multiphase optimization strategy to a pilot study: an empirical example targeting obesity among children of low-income mothers. *BMC Public Health, 16*(1181). doi:10.1186/s12889-016-3850-y

Kumar, S., Nilsen, W. J., Abernethy, A., Atienza, A., Patrick, K., Pavel, M., Riley, W. T., Shar, A., Spring, B., Spruijt-Metz, D., Hedeker, D., Honavar, V., Kravitz, R., Lefebvre, C. R., Mohr, D. C., Murphy, S. A., Quinn, C., Shusterman, V., & Swendeman, D. (2013). Mobile Health Technology Evaluation: The mHealth Evidence Workshop. *American Journal of Preventive Medicine, 45*(2), 228-236. doi:doi.org/10.1016/j.amepre.2013.03.017

Labrique, A., Vasudevan, L., Chang, L. W., & Mehl, G. (2013). H_pe for mHealth: More "y" or "o" on the horizon? *International journal of medical informatics, 82*(5), 467-469. doi:10.1016/j.ijmedinf.2012.11.016

Lemay, N. V., Sullivan, T., Jumbe, B., & Perry, C. P. (2012). Reaching remote health workers in Malawi: baseline assessment of a pilot mHealth intervention. *Journal of Health Communication, 17*(sup1), 105-117. doi:10.1080/10810730.2011.649106

Leon, N., Schneider, H., & Daviaud, E. (2012). Applying a framework for assessing the health system challenges to scaling up mHealth in South Africa. *BMC Medical Informatics and Decision Making, 12*, 123-123. doi:10.1186/1472-6947-12-123

Lim, J., Cloete, G., Dunsmuir, D. T., Payne, B. A., Scheffer, C., von Dadelszen, P., Dumont, G. A., & Ansermino, J. M. (2015). Usability and feasibility of PIERS on the Move: an mHealth app for pre-eclampsia triage. *JMIR mHealth and uHealth, 3*(2). doi:10.2196/mhealth.3942

Madon, S. (2000). The Internet and socio-economic development: exploring the interaction. *Information technology & people, 13*(2), 85-101. doi:10.1108/09593840010339835

Mahmud, N., Rodriguez, J., & Nesbit, J. (2010). A text message-based intervention to bridge the healthcare communication gap in the rural developing world. *Technology and Health Care, 18*(2), 137-144. doi:10.3233/THC-2010-0576

Malvey, D. M., & Slovensky, D. J. (2017). Global mHealth policy arena: status check and future directions. *Mhealth, 3*, 41. doi:10.21037/mhealth.2017.09.03

Matthews, J., Kulkarni, R., Whitesides, G., Sarrafzadeh, M., Gerla, M., & Massey, T. (2009). *A light-weight solution for real-time dengue detection using mobile phones.* Paper presented at the Proceedings of the International Conference on Mobile Computing, Applications, and Services, MobiCase

McClure, J. B., Anderson, M. L., Bradley, K., An, L. C., & Catz, S. L. (2016). Evaluating an Adaptive and Interactive mHealth Smoking Cessation and Medication Adherence Program: A Randomized Pilot Feasibility Study. *JMIR Mhealth Uhealth, 4*(3), e94. doi:10.2196/mhealth.6002

McFarlan, F. W., & McKenney, J. L. (1982). The Information Archipelago-Gaps and Bridges *Harvard Business Review, 60*(5), 109-119.

Mechael, P., Batavia, H., Kaonga, N., Searle, S., Kwan, A., Goldberger, A., Fu, L., & Ossman, J. (2010). *Barriers and Gaps Affecting mHealth in Low and Middle Income Countries: Policy White Paper*. Retrieved from Columbia: http://www.globalproblems-globalsolutions-files.org/pdfs/mHealth_Barriers_White_Paper.pdf

Mohr, D. C., Cheung, K., Schueller, S. M., Brown, C. H., & Duan, N. (2013). Continuous Evaluation of Evolving Behavioural Intervention Technologies. *American Journal of Preventative Medicine, 45*(4). doi:10.1016/j.amepre.2013.06.006

Mohr, D. C., Montague, E., Stiles-Shields, C., Kaiser, S. M., Brenner, C., Carty-Fickes, E., Palac, H., & Duffecy, J. (2015). *MedLink: a mobile intervention to address failure points in the treatment of depression in general medicine.* Paper presented at the Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare. PMID:26640740

Mohr, D. C., Schueller, S. M., Riley, W. T., Brown, C. H., Cuijpers, P., Duan, N., Kwasny, M. J., Stiles-Shields, C., & Cheung, K. (2015). Trials of intervention principles: evaluation methods for evolving behavioral intervention technologies. *Journal of Medical Internet Research, 17*(7). doi:10.2196/jmir.4391

Moodie, E. E. M., Karran, J. C., & Shortreed, S. M. (2016). A case study of SMART attributes: a qualitative assessment of generalizability, retention rate, and trial quality. *Trials, 17*(1), 242. doi:10.1186/s13063-016-1368-3

Mummah, S. A., Robinson, T. N., King, A. C., Gardner, C. D., & Sutton, S. (2016). IDEAS (Integrate, Design, Assess, and Share): a framework and toolkit of strategies for the development of more effective digital interventions to change health behavior. *Journal of Medical Internet Research, 18*(12). doi:10.2196/jmir.5927

Naslund, J. A., Marsch, L. A., McHugo, G. J., & Bartels, S. J. (2015). Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. *J Ment Health, 24*(5), 321-332. doi:10.3109/09638237.2015.1019054

Neupane, S., Odendaal, W., Friedman, I., Jassat, W., Schneider, H., & Doherty, T. (2014). Comparing a paper based monitoring and evaluation system to a mHealth system to support the national community health worker programme, South Africa: an evaluation. *BMC Medical Informatics and Decision Making, 14*(1), 69. doi:10.1186/1472-6947-14-69.

Nilsen, E., Kumar, S., Shar, A., Varoquiers, C., Wiley , T., Riley, W., Pavel, M., & Atienza, A. (2012). Advancing the science of mHealth. *Journal of Health Communication, 17*(Suppl 1), 5-10. doi:10.1080/10810730.2012.677394

Njoroge, M., Zurovac, D., Ogara, E. A. A., Chuma, J., & Kirigia, D. (2017). Assessing the feasibility of eHealth and mHealth: a systematic review and analysis of initiatives implemented in Kenya. *BMC Research Notes, 10*(1), 90. doi:10.1186/s13104-017-2416-0

Opoku, D., Scott, P., & Quentin, W. (2015). Healthcare professionals' perceptions of the benefits and challenges of a teleconsultation service in the amansie-west district of Ghana. *Telemedicine and e-Health, 21*(9), 748-755. doi:10.1089/tmj.2014.0210

Pagoto, S., & Bennett, G. G. (2013). How behavioural science can advance digital health. *Translational Behavioural Medicine, 3*(3), 271-276. doi:10.1007/s13142-013-0234-z

Peiris, D., Praveen, D., Johnson, C., & Mogulluru, K. (2014). Use of mHealth Systems and Tools for Non-Communicable Diseases in Low- and Middle-Income Countries: a Systematic Review. *Journal of Cardiovascular Translational Research, 7*(8), 677-691. doi:10.1007/s12265-014-9581-5

Pham, Q., Khatib, Y., Stansfeld, S., Fox, S., & Green, T. (2016a). Feasibility and efficacy of an mHealth game for managing anxiety:"Flowy" randomized controlled pilot trial and design evaluation. *Games for health journal, 5*(1), 50-67. doi:10.1089/g4h.2015.0033

Pham, Q., Wiljer, D., & Cafazzo, J. A. (2016b). Beyond the Randomized Controlled Trial: A Review of Alternatives in mHealth Clinical Trial Methods. *JMIR mHealth and uHealth, 4*(3), e107. doi:10.2196/mhealth.5720

Pierce, J. L., & Delbecq, A. L. (1977). Organization structure, individual attitudes and innovation. *Academy of management review, 2*(1), 27-37. doi:10.2307/257602

Rahman, M. (2015). Experts of the Future: 4 Areas of Emerging Litigation. Retrieved from https://www.theexpertinstitute.com/experts-future-4-areas-emerging-litigation/ [Accessed: 21st December 2017] Archived by WebCite® at http://www.webcitation.org/77zls8yi5

Roback, K., Dalal, K., & Carlsson, P. (2011). Evaluation of Health Research: Measuring Costs and Socioeconomic Effects. *International Journal of Preventative Medicine, 2*(4), 203-215. PMID:22174959

Rosen, L., Manor, O., Engelhard, D., & Zucker, D. (2006). In Defense of the Randomized Controlled Trial for Health Promotion Research. *American Journal of Public Health, 96*(7), 1181-1186. doi:10.2105/AJPH.2004.061713

Rutledge, C. M., Kott, K., Schweickert, P. A., Poston, R., Fowler, C., & Haney, T. S. (2017). Telehealth and eHealth in nurse practitioner training: current perspectives. *Advances in medical education and practice, 8*, 399. doi:10.2147/AMEP.S116071

Shankar, R., Prasad, S., Ankur, B., Talwar, A., & Jain, N. (2013). *Socio-economic impact of mHealth. An assessment report for the European Union*. Retrieved from https://www.gsma.com/iot/wp-content/uploads/2013/06/Socio-economic_impact-of-mHealth_EU_14062013V2.pdf

Shieshia, M., Noel, M., Andersson, S., Felling, B., Alva, S., Agarwal, S., Lefevre, A., Misomali, A., Chimphanga, B., & Nsona, H. (2014). Strengthening community health supply chain performance through an integrated approach: Using mHealth technology and multilevel teams in Malawi. *Journal of global health, 4*(2), 117-128. doi:10.7189/jogh.04.020406

Shuchman, M. (2014). The promise and pitfalls of global mHealth. *Canadian Medical Association Journal, 186*(15), 1134-1135. doi:10.1503/cmaj.109-4861

Sibbald, B., & Roland, M. (1998). Understanding controlled trials: Why are randomised controlled trials important? *BMJ, 316*(7126), 201. doi:10.1136/bmj.316.7126.201

Song, J. H., Kim, I. K., & Yi, B.-K. (2013). Preliminary comparison of mHealth architectures based on MU2 criteria. *Studies in health technology and informatics, 192*, 1127-1127. doi:10.3233/978-1-61499-289-9-1127

Steinhubl, S. R., Muse, E. D., & Topol, E. J. (2015). The emerging field of mobile health. *Science translational medicine, 7*(283), 283rv283. doi:10.1126/scitranslmed.aaa3487

Stephani, V., Opoku, D., & Quentin, W. (2016). A systematic review of randomized controlled trials of mHealth interventions against non-communicable diseases in developing countries. *BMC Public Health, 16*(572), 572. doi:10.1186/s12889-016-3226-3

Stoyanov, S. R., Hides, L., Kavanagh, D. J., & Wilson, H. (2016). Development and validation of the user version of the Mobile Application Rating Scale (uMARS). *JMIR mHealth and uHealth, 4*(2). doi:10.2196/mhealth.5849

Stroux, L. (2012). *Evidence for mHealth*. Retrieved from https://www.gsma.com/iot/wp-content/uploads/2012/08/Stroux_CDT_Market_Research_Project_2012_v2.pdf

Tariq, A., & Akter, S. (2011). *An Assessment of m-Health in Developing Countries Using Task Technology Fit Model*, Detroit, Michigan.

Taylor, N. P. (2017). FDA proposes freeing apps, software from burden of medical device definition. Retrieved from https://www.fiercebiotech.com/medtech/fda-proposes-freeing-apps-software-from-burden-medical-device-definition [Accessed: 19th December 2017] Archived by WebCite® at http://www.webcitation.org/77zkvH5Jl

The Lancet. (2017). Does mobile health matter? *The Lancet, 390*(10109), 2216. doi:doi.org/10.1016/S0140-6736(17)32899-4

Thokala, P., Devlin, N., Marsh, K., Baltussen, R., Boysen, M., Kalo, Z., Longrenn, T., Mussen, F., Peacock, S., Watkins, J., & Ijzerman, M. (2016). Multiple Criteria Decision Analysis for Health Care Decision Making—An Introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value in Health, 19*(1), 1-13. doi:10.1016/j.jval.2015.12.003

Thompson, V. A. (1969). *Bureaucracy and Innovation*. Alabama: University of Alabama Press University.

US Food and Drug Adminstration. (2017). Statement on advancing new digital health policies to encourage innovation, bring efficiency and modernization to regulation. Retrieved from https://www.prnewswire.com/news-releases/statement-from-fda-commissioner-scott-gottlieb-md-on-advancing-new-digital-health-policies-to-encourage-innovation-bring-efficiency-and-modernization-to-regulation-300568421.html [Accessed: 10th September 2018] Archived by WebCite® at http://www.webcitation.org/77zlPXcry

van Velthoven, M. H., Car, J., Zhang, Y., & Marušić, A. (2013). mHealth series: New ideas for mHealth data collection implementation in low- and middle-income countries. *J Glob Health, 3*(2), 020101-020101. doi:10.7189/jogh.03.020101

Vu, L. T. H., Nguyen, N. T. K., Tran, H. T. D., & Muhajarine, N. (2016). mHealth information for migrants: an e-health intervention for internal migrants in Vietnam. *Reproductive health, 13*(1), 55. doi:10.1186/s12978-016-0172-6

White, B. K., Burns, S. K., Giglia, R. C., & Scott, J. A. (2017). Designing evaluation plans for health promotion mHealth interventions: a case study of the Milk Man mobile app. *Health Promotion Journal of Australia, 27*(3), 198-203. doi: 10.1071/HE16041

Whittaker, R., Merry, S., Dorey, E., & Maddison, R. (2012). A development and evaluation process for mHealth interventions: examples from New Zealand. *Journal of Health Communication, 17*(sup1), 11-21. doi:10.1080/10810730.2011.649103

World Health Organisation. (2012). Performance-based financing in low- and middle-income countries: still more questions than answers. Retrieved from http://www.who.int/bulletin/volumes/90/8/12-106468/en/ [Accessed:

World Health Organization. (2011). *mHealth New horizons for health through mobile technologies*. Retrieved from Switzerland: http://www.who.int/goe/publications/goe_mhealth_web.pdf

World Health Organization. (2017). *Aligning public financial management and health financing: a process guide for identifying issues and fostering dialogue* (9241513071). Retrieved from Geneva: http://apps.who.int/iris/bitstream/handle/10665/254680/9789241512039-eng.pdf?sequence=1

Wyrick, D. L., Rulison, K. L., Fearnow-Kenney, M., Milroy, J. J., & Collins, L. M. (2014). Moving beyond the treatment package approach to developing behavioral interventions: addressing questions that arose during an application of the Multiphase Optimization Strategy (MOST). *Translational Behavioral Medicine, 4*(3), 252-259. doi:10.1007/s13142-013-0247-7