Type of paper: Review

Evidence-based evaluation of eHealth interventions: A systematic literature review

Amia Enam\textsuperscript{1}, Johanna Torres-Bonilla\textsuperscript{1} and Henrik Eriksson\textsuperscript{1}

\textsuperscript{1} Centre for Healthcare Improvement (CHI), Chalmers University of Technology, Gothenburg, Sweden.

Amia Enam, Centre for Healthcare Improvement (CHI), Chalmers University of Technology, 412 96 Gothenburg, Sweden, Phone: +1(860)9449850, Email: amiaenam@gmail.com

\textsuperscript{1}Corresponding author.
Abstract

Background
Until now, the use of technology in healthcare is driven mostly by the assumptions about the benefits of eHealth rather than the evidence of it. It is noticeable that the magnitude of evidence of effectiveness and efficiency of eHealth is not proportionate to the number of interventions that are regularly conducted. Reliable evidence generated through comprehensive evaluation of eHealth interventions may accelerate the growth of eHealth for long-term successful implementation and help to experience the benefits of eHealth in an enhanced way.

Objective
The objective of this article was to understand how the evidence of effectiveness and efficiency of eHealth can be generated through evaluation. Hence, we aim to discern, (1) the phases of eHealth intervention as depicted by eHealth literature, (2) the aspects of eHealth intervention that are supposed to be evaluated according to eHealth literature, (3) how empirical studies about eHealth are evaluated, and (4) how evidence can be generated while evaluating eHealth interventions.

Methods
A systematic literature review was conducted to explore the evaluation methods for eHealth interventions. PRISMA guidelines were followed. We searched Google Scholar and Scopus for the published articles that addressed the evaluation of eHealth or described an eHealth intervention study. A qualitative analysis of the selected articles was conducted in several steps.

Results
We identified two distinct branches of conceptual framework that discussed the evaluation of eHealth interventions. The first branch identified evaluation as one of the phases of the intervention. The other phases identified through the literature review included design, pre-testing, pilot study, pragmatic trial, and post-intervention. The second branch provided the aspects of evaluation during intervention including clinical, human/social, organizational, technological, cost, ethical/legal, and transferability. The analysis of the empirical studies revealed that evaluation varied greatly and it may not be plausible to employ a standard evaluation procedure. Lastly, we proposed a conceptual model based on our findings that shows how the evidence of effectiveness and efficiency can gradually and reliably be built by evaluating specific aspects during specific intervention phases.

Conclusions
The importance of evidence in eHealth has not been discussed as rigorously as have the diverse research approaches and evaluation frameworks. Further research directed towards evidence-based evaluation can not only improve the quality of intervention studies, but also facilitate successful long-term implementation of eHealth in general. We conclude that the development of more robust and comprehensive eHealth intervention studies or an improved validation of evaluation methods could ease the transferability of results among similar studies. Thus, the resources can be used to supplementary research in eHealth.

**Keywords**: eHealth intervention, evaluation, evidence, systematic review

**Introduction**

The use of eHealth is still driven by assumptions about the benefits of eHealth rather than evidence of it [1]. With time, the trustworthiness and robustness of eHealth to facilitate safe and cost-efficient care are being questioned due a lack of evidence [2]. This may trigger reluctance to investing and developing policies related to eHealth in organizations as well as in countries [3].

The term “eHealth” was introduced in the 1990s [4]; however, it was hardly in use until 1999 [5]. According to Eysenbach [5], “e-health is an emerging field in the intersection of medical informatics, public health and business, referring to health services and information delivered or enhanced through the Internet and related technologies” (p. 1). Eysenbach believes eHealth stands for more than Internet and medicine [5]. In our study, eHealth was used as the broadest umbrella encompassing everything that comes within information and communication technology and healthcare including telemedicine, mobile health, health informatics.

Systematic evaluation can capture the evidence and criteria that evaluative judgment is based on and curtail the sources of biases [6]. The quality of an evaluation is assessed by the credibility of evidence assembled through it and using evidence in refining the policies and programs [7]. Evaluation of eHealth intervention is complex due to several reasons (e.g., the need for multidisciplinary collaboration [8], context-dependency [9], and differences in epistemological beliefs considering the interventions in clinical studies or including social aspects as well [10-14]. Therefore, variety exists concerning how the evaluation of eHealth interventions are performed and presented. Garnering robust evidence through evaluation becomes difficult due to these circumstances.

It is relevant to understand evidence-based medicine (EBM) while discussing the importance of evidence in eHealth interventions. A common query is how EBM can help generate evidence for eHealth interventions. “Evidence-based medicine is the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients. The practice of EBM means integrating individual clinical expertise with the best available external clinical evidence from systematic
research” [15] (p. 71). As per this definition, “evidence” in EBM is conspicuously related to the clinical aspect. Although it is argued whether EBM is only about randomized controlled trials [16] or not [15], it is quite explicit that it usually does not contemplate anything outside clinical practices. However, eHealth interventions have more aspects to evaluate besides the clinical aspects. An extensive assessment of the aspects including socio-technical aspects is needed through each phase of the technology’s life cycle while evaluating eHealth interventions [10-11,17]. Hence, to gather evidence from eHealth intervention, the evaluation process requires a distinct approach than what is usually put forward within EBM.

Our current objective was to elucidate how the evidence of effectiveness and efficiency of eHealth can be generated through evaluation. A systematic literature review was conducted to answer the following research questions:

1. What are the phases of eHealth intervention as depicted by eHealth literature?
2. What aspects of eHealth intervention are supposed to be evaluated according to eHealth literature?
3. How have the empirical studies of eHealth been evaluated?
4. How can evidence be generated in eHealth evaluation?

**Methods**

**Systematic review**

**Identification and screening**

A systematic search of relevant literature was conducted following PRISMA guidelines [18]. Google Scholar and Scopus were used to search the following identified terms, “research methods” and “eHealth interventions”, “study design” and “eHealth interventions”, “evaluation methods” and “eHealth interventions”, “eHealth interventions” and “evaluation framework”, “evidence based” and “evaluation” and “eHealth interventions”. We included articles published between 1990 and 2016. As the term eHealth evolved during the 1990s [4], we deemed it reasonable to consider the literature published on eHealth interventions since then. One-thousand six-hundred twenty-four records were found.

The screening of the articles was conducted in three steps. For the first two steps, the screening was based on the title of the manuscripts using a predefined set of exclusion and inclusion criteria. The exclusion criteria for the first step were books; patents; citations; literature reviews; meta-analyses; or publications solely addressing behavior change theory, ergonomics, drugs, sedentary issues, or physical activity intervention. The number of records was reduced to 813 after the first elimination. At this point, all the records were listed together, and duplicate records were removed. During the second step of screening, only those records that included either a conceptual discussion about an eHealth intervention, and/or discussion about an eHealth intervention that focused on adult patients and caregivers were selected. The third step of the screening process
started with 279 records; this time, the abstracts and the methodology of the articles were read by two of the authors individually. At the end of step three, the authors discussed, and 81 articles were selected for thorough reading and analysis. Besides the articles selected through systematic search, 10 records were added for further analysis. These records were found from the citation of the articles selected using the systematic literature search. All 10 articles were included in this study because of their relevance to our objective.

**Eligibility and inclusion**

To extract and record useful information from the articles, an Microsoft Excel spreadsheet was created with the criteria shown in Appendix A. While most of the criteria were adopted from Cochrane review [19], criterion such as learning point was included by us. During these phases, the articles \(N = 91\) were read meticulously, which led to the final screening, where 46 articles were selected for the qualitative analysis. The flow diagram of the article selection process is presented in Figure 1.

![PRISMA flow diagram of the article selection process](image)

**Figure 1. PRISMA flow diagram of the article selection process**

**Qualitative analysis**

Based on the summary of articles mentioned in previous section, the articles \(N = 46\) were classified into two categories: (1) conceptual exploration of eHealth interventions \(n = 21\), and (2) empirical studies of eHealth interventions \(n = 25\). Most of the articles from the first category presented the outcome of the studies as frameworks and a thematic analysis of those frameworks revealed two distinct patterns defined as Group A: phases of an eHealth intervention \(n = 10\), and Group B: aspects of evaluation in eHealth interventions \(n = 11\).
The articles categorized as empirical studies of eHealth interventions were analyzed based on several characteristics (i.e., duration of the intervention, number of participants, used a framework/pre-defined theory for evaluation or designing the intervention, aspects assessed for evaluation, phases involved in the evaluation, data collection method, and presentation of intervention results). The purpose of the analysis was to compare the characteristics of empirical studies with the conceptual articles.

Results

The results section is divided into four sections according to the research questions. The first three parts are the direct outcome of the analysis of the variation of evaluating eHealth interventions (i.e., phases of an eHealth intervention, aspects of evaluation in eHealth interventions, and evaluation reported in empirical studies of eHealth interventions) found in the literature. In the last part, we developed a conceptual model showing the building of evidence in eHealth evaluation.

The phases of eHealth intervention as depicted by eHealth literature

The articles assigned to Group A (n = 10) provide frameworks to design an eHealth intervention. Depending on the characteristics of the intervention (i.e., type of intervention, objective of the intervention, starting point of the intervention), the frameworks included distinct phases. A spectrum of phases of an eHealth intervention was identified including design, pre-testing, pilot study, pragmatic trial, evaluation, and post-intervention. Table 1 provides a compilation of the frameworks assigned to each phase of an eHealth intervention.

Table 1

Description of identified phases to design an eHealth intervention

<table>
<thead>
<tr>
<th>Phase</th>
<th>Area of focus</th>
<th>Key activities</th>
</tr>
</thead>
</table>
| Design phase        | Conceptualization | – Gather theoretical foundations and empirical evidence to detect existing problems and identify viable solutions [20,21-23,25].  
|                     |                   | – Define the objectives of the to-be-developed technology [20,24].              |
| Contextual Inquiry  |                   | – Identify the end-users and stakeholders to define and analyze the characteristics of the context the technology is going to be implemented [20,21,23-25]. |
| Value specification |                   | – Prioritize the critical values of the technology derived from the end-users and stakeholders’ needs [12,22-25]. |
| Requirements        |                   | – Translate the values into functional and technical requirements that frame the final design and the technology development [12,23-25]. |
| specification       | Conduct short-term trials | – Provide evidence of efficacy of the technology [12,25,26]. |
Measure factors such as optimal intensity, timing, safety, feasibility, usability, intervention content, and logistic issues [21,22,26]. Evaluate the correspondence between technology capabilities and technology requirements [25].

<table>
<thead>
<tr>
<th>Pilot study</th>
<th>Strategic plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Define the preliminary plan of the pilot study (i.e., objective, timeline, budget, sponsors, and team members [27,28].</td>
<td></td>
</tr>
<tr>
<td>Identify related ethical and legal issues [20,28].</td>
<td></td>
</tr>
</tbody>
</table>

Study design

Define the study type, duration, and participants [20,26,28], as well as data collection methods [28]. Design the recruitment process to conform to statistical validity and minimize selection bias [27].

Evaluation

Evaluate simultaneously the technology and its impact [27]. Evaluate the effectiveness of the intervention [22].

Pragmatic trial phase

Define the study type, duration, and participants [20,26,28], as well as data collection methods [28]. Design the recruitment process to conform to statistical validity and minimize selection bias [27].

Evaluation

Formative and summative evaluation (discussed in the evaluation phase) Internal and external evaluation (discussed in the evaluation phase)

Evaluation phase

Formative evaluation

Generate measures that provide timely feedback [12,23]. Perform an evaluative iterative process, since the findings from each step are used to inform subsequent steps [21].

Summative evaluation

Provide generalizable knowledge and benefits of the intervention [12,23].

Internal evaluation

Perform an evaluative process intrinsic to information and communication technology implementations and conducted by the implementation team [12].

External evaluation

Conduct the evaluation by external evaluators to provide expertise where is needed and minimize the bias of in-house evaluators [12].

Post-intervention phase

Conduct post-marketing or surveillance studies to follow-up the technology once scaled up and used by a wider audience [22,26].

Aspects of eHealth intervention supposed to be evaluated according to eHealth literature

The articles categorized under Group B (n = 11) provide frameworks for evaluating eHealth interventions. The articles offer various dimensions of an eHealth intervention that are supposed to be evaluated. We classified these dimensions into seven aspects of evaluation: organizational aspect, technological aspect, human/social aspect, clinical aspect, cost and economic aspect, ethical and legal aspect, transferability aspect. Table 2 exhibits the compilation of the frameworks under these categories.
<table>
<thead>
<tr>
<th>Aspects of assessment</th>
<th>Key areas of measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Organizational aspect</strong></td>
<td>– Organizational setting where the intervention is taking place; it can differ depending on the scale of the intervention (e.g., health center(s), region, country) [29].</td>
</tr>
<tr>
<td></td>
<td>– All type of individuals or groups in the healthcare system that participate in the eHealth intervention, their characteristics, and expectations [30].</td>
</tr>
<tr>
<td></td>
<td>– Organizational performance and professional practice standards [30].</td>
</tr>
<tr>
<td></td>
<td>– Changes in the functions of the healthcare provider, skills and resource demands, and the roles of the professionals in the organization [29,31-33].</td>
</tr>
<tr>
<td></td>
<td>– Representativeness and participation rates of the healthcare professionals during the intervention [34].</td>
</tr>
<tr>
<td></td>
<td>– Capability of the organization to implement the intervention [30,34-38] and the extent that the technology fits the organizational strategy, operations, culture, processes [30].</td>
</tr>
<tr>
<td></td>
<td>– Sustainability or the degree that the technology becomes accustomed in the daily practice of an organization [29,32,34].</td>
</tr>
<tr>
<td><strong>Technological aspect</strong></td>
<td>– Ensure trust [38], effectiveness, and contribution of quality of care [30,36] of the technology implemented.</td>
</tr>
<tr>
<td></td>
<td>– System performance: hardware and software requirements, correct functioning of the components [29,38], and system capability to meet users’ needs and fit the work patterns of the healthcare system’s professionals [30,39,40].</td>
</tr>
<tr>
<td></td>
<td>– Usability: broad experience of the users with the system. [29,33,37,40].</td>
</tr>
<tr>
<td></td>
<td>– Privacy and security: safety and reliability of the technology [29], and security of the data managed in the technology [37,38].</td>
</tr>
<tr>
<td></td>
<td>– Technical accuracy: quality of the transfer of data [41].</td>
</tr>
<tr>
<td></td>
<td>– Information Quality: relates to accuracy, completeness and availability of the information produced by the system (e.g., patients’ records, reports, images, and prescriptions), and it depends on users’ subjectivity. [30,39].</td>
</tr>
<tr>
<td></td>
<td>– Service quality: measures the support and follow-up service delivered by the technology provider. [39].</td>
</tr>
<tr>
<td></td>
<td>– Triability: the ability of the innovation to be tested on a small scale before the final implementation [40].</td>
</tr>
<tr>
<td></td>
<td>– Maturity: whether the system has been used on a sufficient number of patients to address all the technical problems [36].</td>
</tr>
<tr>
<td></td>
<td>– Interoperability: communication between the technology and the pre-existing systems, the fit between the technology and the existing work practices [37].</td>
</tr>
<tr>
<td><strong>Human/social aspect</strong></td>
<td>– Acceptance and usability satisfaction of the technology used in the intervention [30,31,33,36,38,39,41] where the user can be physicians, nurses and other staff, and patients depending on the type of the participants in the intervention [41].</td>
</tr>
<tr>
<td></td>
<td>– System use: volume of use, who is using, purpose of use, and motivation to use the technology [39].</td>
</tr>
<tr>
<td></td>
<td>– User satisfaction: perceived usefulness, enjoyment, decision-making satisfaction, and overall satisfaction for the technology [30,39].</td>
</tr>
<tr>
<td></td>
<td>– Psychological aspects such as satisfaction, wellbeing, and other psychological variables, and social aspects such as accessibility to the technology, the social relationships evolving over the transmission of care or activities of the patients under the intervention [31].</td>
</tr>
</tbody>
</table>
| Clinical aspect | – Benefits and unanticipated negative effects of the intervention, biologic outcomes including disease risk factors, behavioral outcomes of the participants, staff who deliver the intervention, and the sponsors, and quality of life outcomes to evaluate participants’ mental health and satisfaction [34].  
– Long-term measurements of the diagnostic and clinical effectiveness [41, 35, 36], safety of care [33, 35, 36], and quality of care [33]. |
| Cost and economic aspect | – Cost analysis methods to compare the intervention with relevant alternatives in terms of costs and consequences [36].  
– Diverse cost analysis methods can be considered (e.g., cost minimization analysis, cost-effectiveness analysis, cost-benefit analysis, cost-utility analysis, and cost-consequence analysis) [30, 31, 41], and are conducted from several perspectives such as societal, third-party payers, healthcare providers, or patient [31].  
– Diverse costs can be included such as investment cost, monthly user charge of equipment, costs of used communication line, education of the technology, costs of patients and their close relatives [41], wages of doctor and other staff [30, 41], expenditure and revenue for the healthcare organization adopting the technology [36], and resource utilization and opportunity cost of the eHealth intervention [34]. |
| Ethical/legal aspect | – Ethical concerns of the application itself and its implementation, including all the stakeholders’ viewpoints on using the technology and the key ethical principles associated with the context that the intervention is conducted within [35, 36].  
– Legal aspect identifies and analyzes the legislative documents and legal obligations that may exist in each context involved in the intervention [30, 35, 36]. |
| Transferability aspect | – Participation and representativeness of the intervention, percentage of persons who receive or are affected by the program, and the characteristics of participants and nonparticipants to investigate the extent that participants are representative and what population group should be a priority for future research [34].  
– Transferability of results from studies of eHealth from one setting to another, and the assessment of validity and reliability of the study [36]. |

**Evaluation reported in empirical studies of eHealth interventions**

The articles categorized as Empirical Studies of eHealth intervention show substantial variation in the approaches taken to evaluate the interventions. The use of standardized frameworks and theories for evaluating the interventions was hardly noticed in these articles. Table 3 demonstrates the result of the analysis.

<p>| Table 3 |
| Evaluation reported in empirical studies of eHealth interventions |</p>
<table>
<thead>
<tr>
<th>Reference</th>
<th>Used a framework/pre-defined theory for evaluation or designing the intervention</th>
<th>Intervention duration</th>
<th>Number of participants</th>
<th>Data collection method</th>
<th>Presentation of result of the intervention</th>
<th>Aspects for evaluation</th>
<th>Phases involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>[42]</td>
<td>No</td>
<td>1-2 weeks</td>
<td>17</td>
<td>Ql: Thematic analysis of text messages.</td>
<td>Descriptive interpretation of data.</td>
<td>Clinical aspect: Mental health symptoms; mental health coping strategies; mental health treatment and management; lifestyle behaviors. Human/social aspect: Social relationships and leisure activities; motivation and personal goal setting; and independent living.</td>
<td>Pilot study</td>
</tr>
<tr>
<td>[43]</td>
<td>No</td>
<td>6 weeks and 12 weeks</td>
<td>101</td>
<td>Qn: Survey with standardized questionnaire (PHQ-9).</td>
<td>Statistically</td>
<td>Clinical aspect: Depressive symptom severity.</td>
<td>Pilot study</td>
</tr>
<tr>
<td>[45]</td>
<td>No</td>
<td>2 years</td>
<td>49</td>
<td>Qn: Survey Ql: Behavioral analysis, semi-structure interviews.</td>
<td>Descriptive interpretation of data.</td>
<td>Technological aspect: Content, layout, program design and usability testing of a web-based program.</td>
<td>Design and pre-testing</td>
</tr>
<tr>
<td>[46]</td>
<td>No</td>
<td>12 weeks</td>
<td>225</td>
<td>Qn: Survey (online questionnaire).</td>
<td>Statistically</td>
<td>Clinical aspect: Disorder severity rating, number of</td>
<td>Pilot study</td>
</tr>
<tr>
<td>No.</td>
<td>Design</td>
<td>Follow-up</td>
<td>Sample Size</td>
<td>Qn:</td>
<td>Data Collection Method</td>
<td>Statistical Analysis</td>
<td>Pilot Study Type</td>
</tr>
<tr>
<td>-----</td>
<td>--------</td>
<td>-----------</td>
<td>-------------</td>
<td>-----</td>
<td>------------------------</td>
<td>----------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>[48]</td>
<td>No</td>
<td>73.2 days on average</td>
<td>2632</td>
<td>Qn: Survey with standardized questionnaire (KPD-38).</td>
<td>Statistically</td>
<td>Clinical Aspect: Physical and psychological impairment. Human/social aspect: Social problems.</td>
<td>Pragmatic trial</td>
</tr>
<tr>
<td>[49]</td>
<td>No</td>
<td>11 months</td>
<td>241</td>
<td>Qn: Survey with standardized questionnaires (PSS, COPE-index, GSE, WHO-5) and a single question for QoL.</td>
<td>Statistically</td>
<td>Clinical aspect: Stress level. Human/social aspect: Social situation, well-being, QoL, self-efficacy.</td>
<td>Pilot study</td>
</tr>
<tr>
<td>[50]</td>
<td>No</td>
<td>12 months</td>
<td>Patients from 20 community nursing services</td>
<td>Qn: Log-in to the system data, questionnaire. QI: In-depth interview.</td>
<td>Statistically and in descriptive format.</td>
<td>Organizational aspect: Efficiency and quality of care.</td>
<td>Evaluation study</td>
</tr>
<tr>
<td>[52]</td>
<td>No</td>
<td>6-months</td>
<td>46</td>
<td>Qn: Survey (BDI-II and others).</td>
<td>Statistically</td>
<td>Clinical aspects: Depressive symptom severity, medication</td>
<td>Pilot study</td>
</tr>
<tr>
<td>#</td>
<td>No.</td>
<td>Duration</td>
<td>Sample</td>
<td>Qn:</td>
<td>Ql:</td>
<td>Statistically</td>
<td>Clinical aspect:</td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>----------</td>
<td>--------</td>
<td>-----</td>
<td>-----</td>
<td>---------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>54</td>
<td>No</td>
<td>12 months</td>
<td>43</td>
<td>Survey with standardized questionnaire (FIM).</td>
<td></td>
<td>Statistically</td>
<td>Human/social aspect:</td>
</tr>
<tr>
<td>55</td>
<td>No</td>
<td>8 weeks</td>
<td>8</td>
<td>Survey with standardized questionnaire (QIDS-C, PHQ-9, GAD-7).</td>
<td></td>
<td>Statistically</td>
<td>Clinical aspect:</td>
</tr>
</tbody>
</table>


**[58]** | No | 6 weeks | 86 in Study-I and 39 in Study-II | Likert scale questionnaire. | Semi-structured interview. | Statistically and in descriptive format (this paper includes the descriptive part only). | | | Technological aspect: | System usefulness (expected and perceived), impact of reminders, ease of use. | Pilot study aspect: | | |

**[59]** | No | 6 weeks | 7 | Survey with | | | Clinical aspect: | Depressive | Pilot study |
<table>
<thead>
<tr>
<th>Study</th>
<th>Framework</th>
<th>Duration</th>
<th>Sample Size</th>
<th>Study Design</th>
<th>Study Duration</th>
<th>Sample Size</th>
<th>Questionnaire</th>
<th>Data Analysis</th>
<th>Human/social aspect</th>
<th>Technological aspect</th>
<th>Clinical aspect</th>
<th>Pilot study</th>
</tr>
</thead>
<tbody>
<tr>
<td>[60]</td>
<td>CONSORT</td>
<td>12 weeks</td>
<td>562</td>
<td>Qn: Survey</td>
<td>Statistically</td>
<td></td>
<td></td>
<td></td>
<td>Treatment adherence, perceived usefulness.</td>
<td></td>
<td></td>
<td>Pilot study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>with</td>
<td>and in</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>standardized</td>
<td>descriptive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>questionnaire</td>
<td>format.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[61]</td>
<td>Framework 6 weeks</td>
<td>20</td>
<td>Qn: Survey</td>
<td>Statistically</td>
<td>Clinical aspect</td>
<td></td>
<td></td>
<td></td>
<td>Design and pre-testing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>by Campbell and colleagues 2000</td>
<td></td>
<td></td>
<td>with</td>
<td>and in</td>
<td>Psychiatric rating scale, DSM disorder.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>standardized</td>
<td>descriptive</td>
<td>Human/social aspect: Perception of the system.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>questionnaire</td>
<td>format (this</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[62]</td>
<td>No</td>
<td>4 weeks</td>
<td>27</td>
<td>Qn: Survey</td>
<td>Technological aspect: System use and acceptability.</td>
<td></td>
<td></td>
<td></td>
<td>Pilot study</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>with</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>standardized</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>questionnaire</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[63]</td>
<td>No</td>
<td>4 months</td>
<td>42</td>
<td>Qn: Semi-structured and written interview.</td>
<td>Human/social aspect: Empowerment and assistance in decision making.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Technological aspect: Acceptance and willingness to use the system.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[64]</td>
<td>No</td>
<td>8 weeks</td>
<td>108</td>
<td>Qn: Survey</td>
<td>Clinical aspect: Adherence to medication.</td>
<td></td>
<td></td>
<td></td>
<td>Pilot study</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>with</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>standardized</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>questionnaire</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Questionnaire Details:**

- **MADRS-S, BDI-II, ISI, WSAS, PTQ, AS-18** for standardized questionnaire in [60].
- **PHQ-9, MOS, GAD-7, IPAQ, WHODAS-II** for standardized questionnaire in [60].
- **BPRS** and **Structured clinical interview (SCID)** for standardized questionnaire in [61].
- **BDI-2, QIDS-SR, EQ-5D-5L, DAS-SF12** for standardized questionnaire in [62].
- **Positive and Negative Syndrome Scale and** for standardized questionnaire in [63].

**Human/social aspect details:**

- Treatment adherence, perceived usefulness in [60].
- Psychiatric rating scale, DSM disorder in [61].
- Perception of the system in [61].
- Empowerment and assistance in decision making in [63].
- Acceptance and willingness to use the system in [63].
- Adherence to medication in [64].

**Technological aspect details:**

- System use and acceptability in [62].
- Acceptance and willingness to use the system in [63].
**QI:** Structured interview.

<table>
<thead>
<tr>
<th>No</th>
<th>3 months</th>
<th>40</th>
<th>Qn: Survey with standardized questionnaires (WHO-QOL-Brief, Dutch empowerment questionnaire, Pearlin mastery scale, social network questionnaire, client satisfaction questionnaire, SUS).</th>
<th>Statistically and in descriptive format.</th>
<th>Clinical aspect: Mastery. Human/social aspect: QoL, empowerment, social cohesion, satisfaction with care.</th>
<th>Pragmatic trial</th>
</tr>
</thead>
</table>

*Note.* AS-18: affective scale-rating scale-18; ASI-3: anxiety sensitivity index-3; BADS-SF: behavioral activation for depression-short form; BDI-II: beck depression inventory-II; BPRS: psychiatric rating scale; CES-D: center for epidemiologic studies depression; COPE-index: careres of older people in Europe index; DASS-21: depression anxiety stress scale-21; DSM: diagnostic and statistical manual (for mental disorder); Eheals: eHealth literacy scale; EQ-5D-5L: EuroQol; FIM: functional independence measure; gad-7: generalized anxiety disorder-7; GSE: general self-efficacy; HADS: hospital anxiety and depression scale; HRSD-24: hamilton rating scale for depression-24; IMR: illness management and recovery; IPAQ: international physical activity questionnaires; ISI: insomnia severity index; KPD-38: klinisch psychologisches diagnosesystem-38; MADRS-S: self-reported Montgomery-Åsberg depression rating scale; MHSES: mental health self-efficacy scale; MOS: medical outcome study; OASIS: outcome and assessment information set; PDSS-SR: panic disorder severity scale; PHQ-9: patient health questionnaire-9; PSS: perceived stress scale; PTQ: perseverative thinking questionnaire; QIDS-CR: quick inventory of depressive symptomatology; Ql: qualitative; QLES-Q-SF: quality of life enjoyment and satisfaction-short form; QoL: quality of life; Qn: quantitative; SANS: assessment of negative symptoms; SCID: structured clinical interview for disorder; SF-12: short form-12; WHO-5: wellbeing index; WHODAS-II: world health organization disability assessment schedule-II; WSAS: work and social adjustment scale.
Generation of evidence in eHealth intervention: Evidence in eHealth Evaluation (EeHE) model

The variety of frameworks found in the systematic literature review led us to a conceptual model—the EeHE (see Figure 2). It was noticeable from the review that while some frameworks elaborate the aspects of evaluation for eHealth intervention and some organize the intervention into certain phases, no visible interaction has been made so far between these two groups of frameworks (i.e., what to assess during what phase). Most of the frameworks from Group A consider evaluation to be one of the phases of intervention [12,21,23] and some of them break down the evaluation into summative and formative [12,23]. The frameworks from group B provide various aspects of evaluation without mentioning the time of evaluation. The EeHE model exhibits the accumulation of evidence by assessing certain aspects of evaluation in distinct intervention phases.

![Evidence in eHealth Evaluation Model](image)

**Figure 2. The Evidence in eHealth Evaluation model**

Here, an eHealth intervention having all six phases (i.e., design, pre-testing, pilot study, pragmatic trial, evaluation, and post-intervention) is conceived to be a comprehensive intervention. We propose that the generation of robust evidence of effectiveness and efficiency would be plausible when the evaluation is conducted through all the phases of an intervention. Moreover, the aspects of evaluation (i.e., organizational aspect, technological aspect, human/social aspect, clinical aspect, cost aspect, ethical and legal aspect, and transferability aspect) would vary in each phase depending on
activities of the phases. When an eHealth intervention initiates with the design phase, the decisions are made based on the evaluation of the technological aspect and cost of technology development. The formal evaluation of the intervention begins in succeeding phases. The evaluation of technological, human/social, and cost aspects occurs in the pre-testing phase. During the pilot study phase, the focus of evaluation shifts primarily to clinical aspect followed by human/social, technological, and ethical and legal. Depending on the evidence garnered in the pilot study, the intervention may proceed to the next phase or go back to the design phase. Since the intervention is scaled up in the pragmatic trial, the evaluation is conducted to identify whether the technology enabled care can be executed within the realistic layout of an organization. Hence, the key areas of evaluation in this phase are organizational and cost aspect along with other aspects such as clinical, human/social, technological, and ethical and legal. The last phase of gathering evidence is summative evaluation where all the aspects are assessed including transferability. The comprehensive evaluation process gradually accumulates the evidence that reaches its peak in summative evaluation phase and is used in post-intervention phase to make future decisions. The model also exhibits how the involvement of patients increases continuously from the design phase to the pragmatic trial escalating the complexity of the evaluation process.

The EeHE model proposes to extend the evaluation process throughout the six phases of eHealth intervention. The underlying idea is to assess specific aspects in each phase instead of evaluating all aspects in a single phase. This way of evaluating eHealth intervention can capture comprehensive evidence for eHealth interventions, which usually are dynamic and complex in nature.

Discussion

From the articles reviewed in this study, we revealed that numerous approaches to conceptualize and conduct eHealth intervention co-exist. Several attributes of evaluation of eHealth intervention have become known through this review. There are vivid differences between how evaluation is conducted in empirical studies and how it is discussed in the conceptual articles. Moreover, a wide range of variety prevails within each group. Evaluation has been depicted as both static action performed at the end of the intervention [20,25-28] and dynamic action dividing it further into summative and formative evaluation [12,21,23]. Depending on the evaluators, evaluation can also be classified into internal and external assessment [12]. However, concerning empirical studies, all conducted evaluation at the end of the intervention. Although several aspects of evaluation have been found in conceptual articles [32-39,41], the empirical studies only evaluated clinical and human/social aspects [45,51,61,65].

Although analyzing standardization of eHealth evaluation was not an objective of the review, the variability found in the articles compels us to think whether it hinders the sharing of evidence among eHealth interventions. Scarcity of evidence, in turn, could delay the growth of eHealth. It is noticeable that the need for evidence is not clearly
stated in any of the articles. The evaluation of the empirical studies typically focuses on the success or failure of the technology (eHealth) in that intervention. It seems that the numerous efforts taken in eHealth research are still quite disconnected, and they are thus unable to create a synergic effect on the growth of eHealth.

The findings also reveal that comprehensive evaluation enables the creation of better evidence. The inclusion of relevant information regarding other aspects besides the clinical (e.g., organizational aspect and cost aspect) allows creating reusable knowledge to facilitate the transfer of results to other settings [36], and to obtain useful insights for long-term implementations. It can be assumed that assessing all the aspects in a single study might end up with a confounding result since all the aspects are interrelated and inferior performance in an aspect can affect the performance in other aspects, which may create a misleading result. Therefore, to gather evidence of effectiveness and efficacy we have suggested the EeHE model. The model encourages evaluating certain aspects during phases of the intervention, and its main objective is to gather evidence gradually and credibly. We acknowledge the fact that an eHealth intervention including all the phases presented in the model will become cumbersome regarding resource consumption. This model is not a prescription, but just a way to show the progression of evidence in eHealth intervention in a reliable manner.

**Conclusion**

To date, the importance of evidence has not been discussed as rigorously as have been the diverse research approaches and evaluation frameworks. In this article, the EeHE model was developed to exhibit how evidence can be generated evaluating certain aspects in each intervention phase. Assessing distinct aspects during distinct phases is a novel concept discussed in this article and requires further analysis. Moreover, this study implies an inconsistency between the literary concepts and practices of eHealth intervention, which has not been noted until now.

Since health interventions are context-specific, the transferability of results from eHealth studies may be difficult. Moreover, neither the conceptual nor the empirical studies suggested the long-term implementation of specific technology into the healthcare settings where it has been tested. We believe that this may be caused by a lack of or insufficient availability of preliminary evidence of the effectiveness and efficiency after conducting the micro-trials or short-term tests on the effects of the technology. Consequently, it appears that lack of evidence hinders the growth of eHealth. Further research directed towards evidence-based evaluation can not only improve the quality of that intervention study, but also facilitate long-term implementation of eHealth in general. We conclude that the development of more robust and comprehensive eHealth intervention studies or an improved validation of evaluation methods could ease the transferability of results among similar studies. Thus, the resources can be used to supplementary research in eHealth.
Acknowledgements: This work was partly funded by the European Union’s Horizon 2020 research and innovation programme, grant agreement No 643588.

Conflicts of Interest: There are no conflicts of interest to declare.

References


Appendix A

Criteria for summarizing the articles found in the systematic literature review

<table>
<thead>
<tr>
<th>Summarizing criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Objective</strong></td>
</tr>
<tr>
<td><strong>2. Methods</strong></td>
</tr>
<tr>
<td><strong>2.1. Types of studies</strong></td>
</tr>
<tr>
<td><strong>2.2. Background of the authors</strong></td>
</tr>
<tr>
<td><strong>2.3. Types of participants</strong></td>
</tr>
<tr>
<td><strong>2.4. Types of outcome measures</strong></td>
</tr>
<tr>
<td><strong>2.5. Data collection method and data analysis</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>3. Learning points</strong></td>
</tr>
</tbody>
</table>

**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EBM</td>
<td>Evidence-based medicine</td>
</tr>
<tr>
<td>EeHE</td>
<td>Evidence in eHealth evaluation</td>
</tr>
<tr>
<td>Cl</td>
<td>Clinical</td>
</tr>
<tr>
<td>Ql</td>
<td>Qualitative</td>
</tr>
<tr>
<td>Qn</td>
<td>Quantitative</td>
</tr>
</tbody>
</table>