



Systematic Review

Artificial Intelligence Implementation in Healthcare: A Theory-Based Scoping Review of Barriers and Facilitators

Taridzo Chomutare ^{1,*}, Miguel Tejedor ¹, Therese Olsen Svenning ¹, Luis Marco-Ruiz ¹, Maryam Tayefi ¹, Karianne Lind ¹, Fred Godtlielsen ^{1,2}, Anne Moen ^{1,3}, Leila Ismail ^{4,5,6}, Alexandra Makhlysheva ¹ and Phuong Dinh Ngo ¹

¹ Norwegian Centre for E-Health Research, 9019 Tromsø, Norway

² Department of Mathematics and Statistics, Faculty of Science and Technology, UiT The Arctic University of Norway, 9037 Tromsø, Norway

³ Institute for Health and Society, Faculty of Medicine, University of Oslo, 0318 Oslo, Norway

⁴ Department of Computer Science and Software Engineering, College of Information Technology, United Arab Emirates University, Al Ain 15551, United Arab Emirates

⁵ National Water and Energy Center, United Arab Emirates University, Al Ain 15551, United Arab Emirates

⁶ School of Computing and Information Systems, Faculty of Engineering and Information Technology, The University of Melbourne, Parkville, VIC 3010, Australia

* Correspondence: taridzo.chomutare@ehealthresearch.no



Citation: Chomutare, T.; Tejedor, M.; Svenning, T.O.; Marco-Ruiz, L.; Tayefi, M.; Lind, K.; Godtlielsen, F.; Moen, A.; Ismail, L.; Makhlysheva, A.; et al. Artificial Intelligence Implementation in Healthcare: A Theory-Based Scoping Review of Barriers and Facilitators. *Int. J. Environ. Res. Public Health* **2022**, *19*, 16359. <https://doi.org/10.3390/ijerph192316359>

Academic Editor: George Crooks

Received: 18 October 2022

Accepted: 2 December 2022

Published: 6 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: There is a large proliferation of complex data-driven artificial intelligence (AI) applications in many aspects of our daily lives, but their implementation in healthcare is still limited. This scoping review takes a theoretical approach to examine the barriers and facilitators based on empirical data from existing implementations. We searched the major databases of relevant scientific publications for articles related to AI in clinical settings, published between 2015 and 2021. Based on the theoretical constructs of the Consolidated Framework for Implementation Research (CFIR), we used a deductive, followed by an inductive, approach to extract facilitators and barriers. After screening 2784 studies, 19 studies were included in this review. Most of the cited facilitators were related to engagement with and management of the implementation process, while the most cited barriers dealt with the intervention's generalizability and interoperability with existing systems, as well as the inner settings' data quality and availability. We noted per-study imbalances related to the reporting of the theoretic domains. Our findings suggest a greater need for implementation science expertise in AI implementation projects, to improve both the implementation process and the quality of scientific reporting.

Keywords: artificial intelligence; machine learning; CFIR; AI implementation; eHealth; healthcare; deep learning; diagnosis; prognosis

1. Introduction

Nowadays, artificial intelligence (AI) has become ubiquitous, and much more advanced and user-friendly than it was two decades ago. In many respects, AI has become reliable and permeates many aspects of our daily lives, such as face and speech recognition apps. Yet, only recently have we seen a corresponding rate of adoption in the healthcare services. AI systems can emerge as a smart solution to reduce clinical staff workload in a world with increasingly saturated healthcare systems. AI is different from simple technology interventions in the sense that AI does not just manage data, but it provides suggestions and recommendations directly shaping the clinical decision process [1,2].

There exists a wide body of literature on the barriers to and facilitators of implementing AI in healthcare [3–5]. However, much of what we know about these barriers and facilitators comes from anecdotal evidence [6], narrative commentaries [7] and reviews [8–11], mostly without any empirical support or sound theoretical basis. As a result, the determinants

of AI implementation success in healthcare are still poorly understood [12]. We lack a complete overview of all the factors that are relevant to implementing AI in clinical settings. In this study, we turn to implementation science [13] to analyze the facilitators and barriers, based on accounts from existing implementations.

Implementation science is a fairly new field, whose emerging theories, models and frameworks have the potential to inform our understanding of AI implementation in a more widely accessible and systematic way. This multidisciplinary approach, combining AI and implementation science, transcends the traditional boundaries of each of the fields. Blending these two disparate, yet complementary, fields is key to our understanding of AI implementation in healthcare. However, there is a need to reconcile the methodological differences and conflicting domain-specific jargon. In the next two subsections, we explore the fundamental aspects of each of these two fields.

1.1. Artificial Intelligence

AI is not a new concept, but renewed interest in the field is widely attributed to the increasing abundance of digital data and the advancements in data analytic approaches. AI comprises many different areas that range from logic-based models to machine learning (ML). Logic-based models [14] have been successfully used in areas such as biomedical ontologies management (e.g., SNOMED-CT automatic concept classification [15]) and decision support (e.g., SAGUE, Arden syntax, GLIF, etc. [16]). Conversely, ML has had a less prominent role, partially due to the lack of health data availability for training data-driven algorithms. Data-driven methods have the capacity to unveil patterns in data that otherwise would remain hidden [17]. They stand a comparatively better chance at dealing with subpopulations, where one clinical guideline may not suffice to provide the optimal treatment (e.g., multimorbid patients).

In the scope of this study, we refer to AI as systems that are used to solve healthcare problems of interest and are powered by ML. Witten et al. [18] define ML as “a family of statistical and mathematical modeling techniques that use a variety of approaches to automatically learn and improve the prediction of a target state, without explicit programming”. This definition precludes most expert systems and other basic knowledge-based AI systems that use simple rule-based processes or Boolean rules.

1.2. Implementation Science

In a seminal paper, Eccles and Mittman [13] define implementation science as “. . . the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practice into routine practice. . .”. In contrast, AI, comprised mostly of computing sciences, defines implementation as generally referring to development of software components according to a specification, for example, implementing an algorithm. To the extent that these two fields define implementation in significantly different ways, their focus as academic fields will also diverge markedly. Computing sciences focus more on developing artefacts rather than systematically studying how the artefacts are put into routine use. This lack of shared meaning will inevitably have serious consequences for search strategies to find relevant articles in academic databases. For the purpose of discussion in this study, we use the definition of implementation from implementation science.

1.3. Pilot Study vs. Implementation Trial

In the context of screening for implementation trials, there is a thin line between a pilot study and an implementation trial. Pilot studies and feasibility studies are necessary components in the path to implementation. Curran et al. [19] describe a progressive path from efficacy studies, followed by effectiveness studies and then proceeding to implementation research. Pearson et al. [20] distinguish between studies conducted for testing effectiveness and studies intended to evaluate implementation strategies, using three conceptualizations named Hybrid Type 1, Type 2 and Type 3. These conceptualizations are based on Curran et al.’s [19] work on combining both effectiveness studies and implementation sci-

ence elements. Major distinctions are made between the purpose of the study and the methods used. The primary purpose of Hybrid Type 1 is for testing the clinical or public health effectiveness of an intervention. Hybrid Type 2 considers both the clinical effectiveness and evaluation of an implementation strategy. The primary goal of Hybrid Type 3 is to evaluate the effectiveness of the implementation strategies, with a secondary goal to observe other data such as health outcomes. In the current study, we focus on studies evaluating implementation strategies (Hybrid Type 2 and 3), where a full implementation already exists or where the organization is committed to a full roll-out, and the smaller implementation trial forms part of a risk minimization strategy. Thus, pilot studies that fall under Hybrid Type 1 are excluded.

1.4. Objectives

The goal of this scoping review is to characterize the barriers and facilitators influencing the implementation of ML methods in the healthcare setting. This study differs from the existing reviews in at least two major ways. First, whereas the existing reports on barriers and facilitators are fragmented, this study analyzes these barriers and facilitators in a more systematic and theoretic way, which allows us to identify reporting problems and knowledge gaps. Second, the existing reviews do not discriminate based on the phases of implementation. Therefore, most of the studies include algorithm development, efficacy and effectiveness studies. The current study, on the other hand, focuses on empirical observations from the late phases of implementation and roll-out.

2. Methodology

This scoping review follows standard reporting, based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) extension for scoping reviews [21] (see Additional file S4—PRISMA checklist). A scoping review is an appropriate methodology for exploring new areas of research [22]. There are only a few implementations, and reviewing auxiliary information sources, such as reports or websites of the implementation, adds value to our overall understanding of the implementation context. The authors are a multidisciplinary team of statisticians, data scientists, computer scientists and clinicians. The authors have had experience in the implementation of data-driven ML methods and their performance evaluation.

2.1. Protocol

Since this review is scoping in nature and required an additional search phase, no protocol was published in advance. However, the data extraction form was designed before starting the search.

2.2. Eligibility Criteria

The screening goal was to exclude articles that do not study actual or real-world implementations. To identify and understand the barriers and facilitators, based on empirical observations and real experiences with production systems, we included implementation trials and excluded early pilot testing of algorithms. Table 1 shows the eligibility criteria based on the population, intervention, comparator, outcomes and study type (PICOS). We searched for papers published between 2015 and 2021. Only publications in English were included.

Table 1. Study eligibility criteria based on PICOS.

	Inclusion	Exclusion
Population	<ul style="list-style-type: none"> • Humans • Any disease groups 	<ul style="list-style-type: none"> • Animals
Intervention	<ul style="list-style-type: none"> • Use ML techniques • Implemented in healthcare settings or approved by the FDA, EMA or equivalent body 	<ul style="list-style-type: none"> • Systems without ML component • Purely logic-based systems
Comparator	<ul style="list-style-type: none"> • Not relevant 	<ul style="list-style-type: none"> • Not relevant
Outcomes	<ul style="list-style-type: none"> • Implementation results or evaluation • Implementation barriers • Implementation facilitators • Economic benefits, process improvement, treatment outcomes, etc. 	<ul style="list-style-type: none"> • Papers without implementation results (other than technical and algorithmic performance)
Study type	<ul style="list-style-type: none"> • Prospective validation in clinical settings • Retrospective implementation evaluation • Published between 2015 and 2021 	<ul style="list-style-type: none"> • Review papers, expert opinions • Non-English papers • Pilot and experimental studies without real-world implementation

2.3. Information Sources and Selecting Sources of Evidence

The databases and indexes that we searched included PubMed, IEEE, ACM, Google Scholar, and the Web of Science. These sources represent the major indices of scientific articles related to both AI–ML and the healthcare sciences.

In addition to scientific publication databases, we used other sources of information, such as the database of FDA-approved AI systems [23]. We performed an additional Google search on the Internet to gain a better understanding of both the functions of the system and the implementation context. We used website information and any available reports relevant to the specific implementation. These auxiliary information sources are appropriate for use in a scoping review and helped us screen the studies. However, the data extraction was only based on scientific articles.

2.4. Search Query and Two-Phase Search

As a multidisciplinary team of researchers, we knew about the conflicting definitions of implementation. However, we could not anticipate the extent of the problem or how it would affect our search results. We defined an iterative search process with two phases. In the first phase, we searched only the title with terms such as “implement*” and “practice”. Through a limited screening of the title and abstract, we quickly realized that many potentially relevant papers were missed, and most of the papers were about implementing algorithms. In phase two of the search, we had to define our search more broadly; include both title and abstract, and more synonyms. This iterative approach to a search strategy is supported by the literature [24].

We identified a broad spectrum of studies, and, given the lack of a unified vocabulary for indexing relevant articles, we had to make a subjective judgement regarding where a study fell on a continuum: (i) algorithm implementation, (ii) efficacy, effectiveness or algorithm validation, (iii) implementation trial, or (iv) full implementation. An overwhelming majority of the search hits fell within (i) and (ii). Only the studies identified as class (iii) or (iv) were included in this study.

The basic structure of the search query was «Artificial intelligence AND implementation AND healthcare». Synonyms and terms related to AI were then added using the

logical disjunction operator (OR). The initial abstract screening was done using Rayyan [25]. All the search strings are available in [Additional File S1—search string]. An example of the search in PubMed was as follows:

(«machine learning»[Title/Abstract] OR machine learning[mesh] OR «artificial intelligence»[Title/Abstract] OR artificial intelligence[mesh] OR «deep learning»[Title/Abstract] OR deep learning[mesh] OR «neural network»[Title/Abstract] OR «image analysis»[Title/Abstract] OR «deep neural networks»[Title/Abstract] OR «supervised learning»[Title/Abstract] OR «unsupervised learning»[Title/Abstract] OR «reinforcement learning»[Title/Abstract] OR «automated algorithms»[Title/Abstract] OR «adaptive algorithms» [Title/Abstract]) AND (implement* [Title] OR practice [Title] OR approved [Title]) AND (y_10[Filter])

2.5. Data Extraction and Items

The data extraction variables were developed through weekly brainstorming sessions. At least four co-authors (TC, TOS, MT, PDN) participated in each brainstorming session, defining the list of topics relevant for extraction. The initial sessions were focused on the free definition of the topics and variables useful for extraction. As the brainstorming sessions advanced, the categories of variables were inductively defined, leading to the final list of agreed variables for data extraction as shown in Table 2. The final list of extraction items was calibrated through limited tests by four co-authors.

Table 2. Data extraction items.

Data Type	Examples
Study	authors, year, title, journal
Description	country of implementation, product name, company/research group, timeline for implementation, implementation phase
Role of AI	patient group, primary users, training required, medical specialty, medical task
Technology	AI methods, algorithms, hardware, transparency, interpretability, explainability
Data	type of input, sample size for training
Ethics	security and privacy, bias, other ethics issues
Clinical Validation	type, sample size
Legal	process for approval, approval status, other legal issues/processes
Barriers Facilitators	qualitative methods used to extract the barriers and facilitators

2.6. Critical Appraisal of Individual Sources of Evidence

We used the Mixed Methods Appraisal Tool (MMAT) [26] to critically assess the quality of the included studies. MMAT was an appropriate tool because the nature of relevant studies varied widely between qualitative, quantitative and mixed methods. Three co-authors (TC, MATH, LMR) assessed the quality of the studies and disagreements were resolved by discussion.

2.7. Synthesis of Results

Qualitative methods were used to synthesize the extracted facilitators and barriers based on the Consolidated Framework for Implementation Research (CFIR) (see Figure 1 and the codebook in Additional file S5). CFIR is a framework used by many implementation research studies. It provides an index of constructs for organizing findings in a consistent and understandable manner [27]. It naturally invites us to follow a deductive strategy in the synthesis of results. However, due to the many technical and organizational details found

in AI implementations, we considered that a more granular presentation was convenient in the synthesis of results, as previous studies in the field of clinical decision support (CDS) had shown [28]. To that end, we opted for a mixed approach, aiming to join the proven coherency of the CFIR constructs, for the general classification of barriers and facilitators, with the detailed approach that open inductive coding provided for defining items about the specific context under examination. In this way, we broke down the details of each CFIR construct in the framework's codebook into more granular sub-constructs that were easily mappable to specific barriers and facilitators in AI implementations. With this rationale in mind, we split the analysis of results in two stages and performed a mixed inductive-deductive approach.

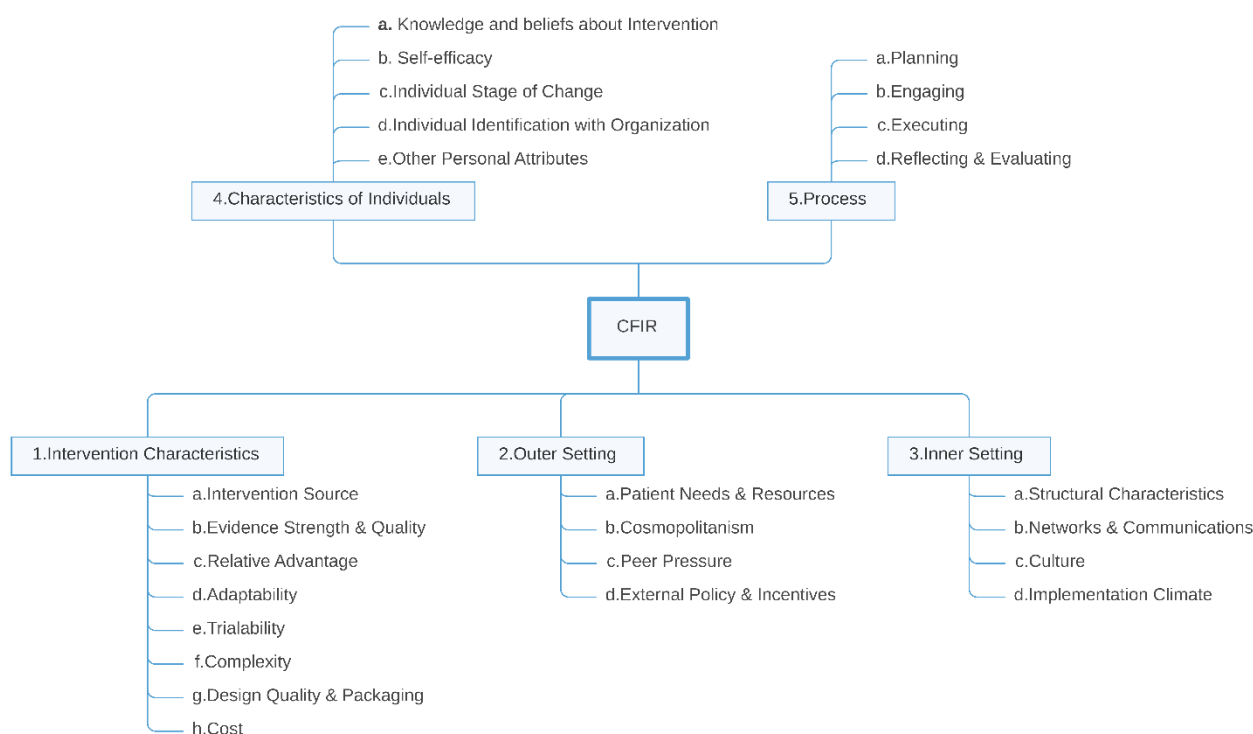


Figure 1. Theoretical constructs of the Consolidated Framework for Implementation Research (CFIR).

2.8. Open Inductive Coding and Mapping onto the CFIR Framework

Five co-authors (LMR, TOS, MATH, MT, PDN) read the full papers, extracting any section that pointed to a possible barrier or facilitator. Free comments (e.g., observations and interpretations) from the reviewers were allowed. All the papers were reviewed by at least two co-authors. Both text segments and free comments were imported into the qualitative analysis software, MaxQDA [29], for further analysis.

Two co-authors (TOS, LMR) went through the extracted segments of the selected papers independently. Initially, a deductive approach to code the segments into CFIR constructs was used. After one iteration, the constructs were considered not granular enough. Then, two reviewers (TOS, LMR) proceeded with an inductive approach, with no predefined code list. The reviewers marked all the segments of text that indicated a barrier or a facilitator for AI adoption.

Once all the papers had been coded, the reviewers met with three other members of the team, who had read all the papers but had not coded the texts. Iterative meetings were performed to go through all the coded texts and crosscheck the results. Equivalent labels were merged into one single concept when agreements were found. Any disagreements were discussed until all the members agreed on the optimal concept to code a specific fragment of the text by checking the full text and re-reading the section of interest. The

usual sources of disagreement were the scope of one concept and the specific barriers and facilitators that one concept should encompass.

The same concept could be described as a barrier or as a facilitator by different studies (for instance, data quality was described as a barrier with “insufficient data quality” and a facilitator as “availability of high-quality data”). This process resulted in an index of concepts that fully categorized all the barriers and facilitators found in the full texts. The index of concepts evolved iteratively until the end of this inductive analysis, refining the semantics of each concept and its scope.

The index of concepts was analyzed by the team and mapped into the constructs of the CFIR framework. Any disagreements about which CFIR construct was the most appropriate for the concept were resolved by discussing the possible options until an agreement was reached. With regards to coverage, the CFIR fully covered the concepts defined in our index, and all of the index concepts could be mapped to CFIR constructs.

3. Results

3.1. Selection of Sources of Evidence

Phase one of the search resulted in a total of 607 articles, while the second phase resulted in 2177 articles, after the removal of duplicates. Four co-authors (TC, TOS, MT, PDN) independently screened the titles and abstracts according to the inclusion and exclusion criteria. This resulted in the removal of 2668 articles, leaving 116 relevant articles. A full-text assessment was conducted on these 116 relevant articles, which resulted in 19 included articles, as shown in Figure 2, 11 of which were published in 2020.

As shown in Table 3, about half of the included studies were conducted in the USA and Canada (47%, $n = 9/19$); about a quarter (26%, $n = 5/19$) were conducted in North-eastern Asia, and only three in Europe. Several medical fields were represented: sepsis (16%, $n = 3/19$), diabetes (11%, $n = 2/19$), cardiology (11%, $n = 2/19$), mental health (11%, $n = 2/19$), emergency care (11%, $n = 2/19$) and palliative care (5%, $n = 1/19$), and the rest were for all patients (16%, $n = 3/19$). The most common medical task (an AI-use case) was screening (79%, $n = 15/19$). Only two of the studies had AI systems targeted towards use by patients, while the rest were meant to be used by clinicians or healthcare staff (89%, $n = 17/19$). In terms of AI algorithms, the majority of the studies applied deep learning (63%, $n = 12/19$).

In terms of the appraisal, three co-authors (TC, LMR, MATH) used the MMAT template to independently appraise the included studies (see Additional file S2—MMAT), and any disagreements were reconciled through discussion. Except for one study, all the other studies had well-defined research questions and sufficient data to address the questions they posed. We categorized the studies into quantitative non-randomized (58%, $n = 11/19$), qualitative (37%, $n = 7/19$) and quantitative descriptive (5%, $n = 1/19$). In all the quantitative studies, the participants were representative of the target population. For all but one of the qualitative studies, the methods used were appropriate to answer the posed questions.

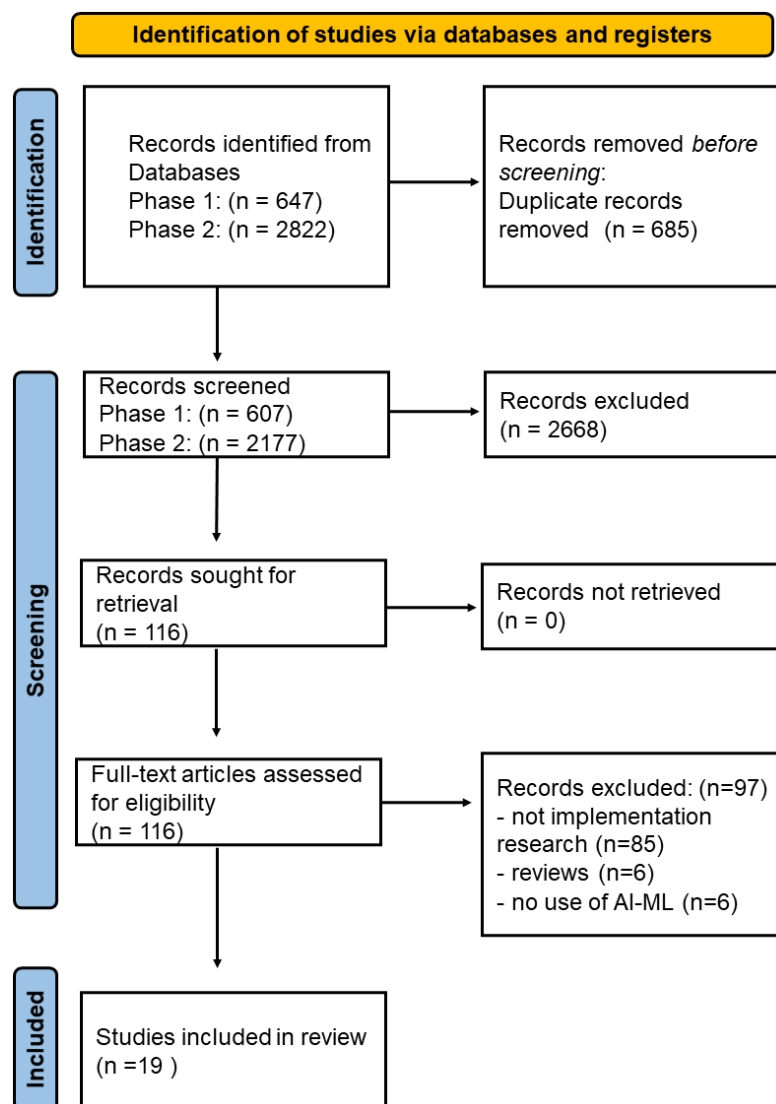


Figure 2. PRISMA flow diagram based on the template by Page et al. [30]. Study characteristics and critical appraisal.

Table 3. Properties of the included studies.

Study (Year)	Country	Medical Field	Medical Tasks (Problem)	Primary Users	AI Techniques
Lee [31] (2015)	USA	Emergency Dept. patients	Screening	Clinicians, nurses, planners	Machine learning
McCoy [32] (2017)	USA	Sepsis	Screening	Clinicians and nurses	Machine learning
Moon [33] (2018)	Korea	Delirium	Screening	Clinicians	Logistic regression
van der Heijden [34] (2018)	Netherlands	Diabetes/retinopathy	Screening	Clinicians	Deep Learning
Schuh [35] (2018)	Austria	All patients	Screening	Clinicians	Deep learning, fuzzy logic, decision tree

Table 3. Cont.

Study (Year)	Country	Medical Field	Medical Tasks (Problem)	Primary Users	AI Techniques
Guo [36] (2019)	China	All patients	Screening	Patients	Deep learning
Cruz [37] (2019)	Spain	Cardiology, Gastroenterology, Psychiatry	Quality improvement	Clinicians (GPs, Pediatricians)	Deep learning
Joerin [38] (2019)	USA/Canada	Psychology	Treatment	Staff, patients and family caregivers	Natural language processing
Gonçalves [39] (2020)	Brazil	Sepsis	Screening	Nurses	Deep learning
Sendak [40] (2020)	USA	Sepsis	Screening	Clinicians	Deep Learning
Gonzalez-Briceno [41](2020)	Mexico	Diabetes/retinopathy	Screening	Clinicians	Deep Learning
Xu [42] (2020)	China	All patients	Screening	Nurses and clinicians	Deep learning
Cho2020 [20]	Korea	Cardiology	Screening	Nurses and clinicians	Deep learning
Romero-Brufau [43] (2020)	USA	All patients	Screening, prognosis, treatment	Clinicians, outpatient care coordinators	Decision tree
Scheinker [44] (2020)	USA	Chronic kidney disease, diabetes	Screening, prognosis, treatment	Clinicians	Deep learning
Davis [45] (2020)	USA	Radiology	Screening	Clinicians	Deep learning
Petitgand [46] (2020)	Canada	Emergency Dept.	Diagnose	Clinicians	Deep learning
Betrian [47] (2021)	Japan	Mental health	Treatment	Patients (receiver) nurse (controller)	Not specified
Murphree [48] (2021)	USA	Palliative care	Screening	Palliative care team (clinicians)	Gradient Boosting Machine (GBM)

3.2. Results of Individual Sources of Evidence

The concepts identified through the inductive process were further classified into ten broader themes: evaluation and testing, background, management and engagement, data quality and management, trust and transparency, clinical workflow, interoperability, finance and resources, technical design and AI policy and regulation. As illustrated in Table 4, most of the facilitators were based on the management and engagement theme (47%, $n = 27/57$). None of the reviewed articles reported barriers related to management and engagement. The second most common facilitators were related to the theme of evaluation and testing (14%, $n = 8/57$), while a third were related to technical design (12%, $n = 7/57$). For the barriers, the most common were interoperability issues (19%, $n = 7/36$), data quality and management (17%, $n = 6/36$) and trust and transparency (14%, $n = 5/36$).

Table 4. Inductive extraction of concepts and themes.

Theme	Facilitators	Barriers	Concept
Evaluation and testing	8	3	-
Background	5	2	Experiences and prior knowledge, Prior evidence, Healthcare demand
Management and engagement	27	-	External collaboration, Planning, Feedback incorporation, Communication, Involvement, Motivation, Leadership, Education of workforce, Patient needs, Champions
Data quality and management	1	6	Data availability, Data quality
Trust and transparency	1	5	Interpretability, Trust
Clinical workflow	4	4	Integration, Disruptiveness (alert fatigue)
Interoperability	2	7	Model Interoperability, Data interoperability, Generalizability
Finance and resources	1	3	Available Resources, Cost
Technical design	7	4	Usability, Documentation and presentation of results, Adaptability, Innovation, Complexity, Trialability
AI policy and regulation	1	2	Organizational policy and culture, Regulation and law
Totals	57	36	

Table 5 shows the concepts extracted from each study. Three studies had no easily discernible barriers or facilitators [34,36,41]. Three studies that reported facilitators did not report any barriers [38,39,47], and two studies that reported barriers did not report any facilitators [33,35].

Table 5. Facilitators and barriers based on the concepts.

Study	Facilitators	Barriers
Lee [31]	Healthcare demand, Evaluation and testing, Generalizability, Data availability, Available Resources, Trialability, Motivation	Regulation and law
Betrianana [47]	Healthcare demand, Planning, Education of workforce, Involvement, Evaluation and testing	-
Cho [49]	Generalizability, Evaluation and testing	Evaluation and testing, Interpretability, Model interoperability
Cruz [37]	Evaluation and testing, Integration, Leadership, Usability	Data availability
Davis [45]	Integration, Usability	Evaluation and testing, Trust
Gonçalves [39]	Motivation, Experiences and prior knowledge	-
Joerin [38]	Involvement, Evaluation and testing, Patient needs, Adaptability	-
McCoy [32]	Healthcare demand, Communication, Feedback incorporation, Education of workforce	Disruptiveness (alert fatigue)
Moon [33]	-	Model Interoperability, Data quality
Murphree [48]	Involvement, Communication	Generalizability
Petitgand [46]	Involvement, Organizational policy and culture	Data interoperability, Usability, Documentation and presentation of results, Trust

Table 5. Cont.

Study	Facilitators	Barriers
Romero-Brufau [43]	Planning, Involvement, Education of workforce, Adaptability	Usability, Data quality, Data availability, Generalizability, Evaluation and testing
Scheinker [44]	Prior evidence, Involvement, Planning, Evaluation and testing	Trust, Complexity, Disruptiveness
Schuh [35]	–	Data quality, Experiences and prior knowledge, Cost, Regulation and law, Data interoperability
Sendak [40]	Involvement, Planning, External collaboration, Leadership, Integration, Interpretability, Evaluation and testing, Champions, Education of workforce	Cost, Trust, Available Resources, Generalizability, Prior evidence, Integration
Xu [42]	Education of workforce, Evaluation and testing, Innovation, Usability, Integration	Data availability, Integration
Gonzalez-Briceno [41]	–	–
Guo [36]	–	–
van der Heijden [34]	–	–

3.3. Mapping Extracted Concepts to CFIR

The mapping between the coded concepts and the corresponding CFIR constructs is available in [Additional file S3—CFIR mappings]. In total, 69 facilitators and 46 barriers were identified and coded following the CFIR framework. The result of this mapping is summarized in Figure 3. Most of the studies reported V. Process as the main facilitator for AI implementation in healthcare (35%, $n = 24/69$). The second most popular facilitators were based on III. Inner setting (29%, $n = 20/69$), followed by I. Intervention characteristics (27%, $n = 19/69$). Most of the barriers were reported for I. Intervention characteristics (41%, $n = 19/46$), followed by III. Inner setting (33%, $n = 15/46$) and II. Outer setting, which had the least number of both barriers and facilitators.

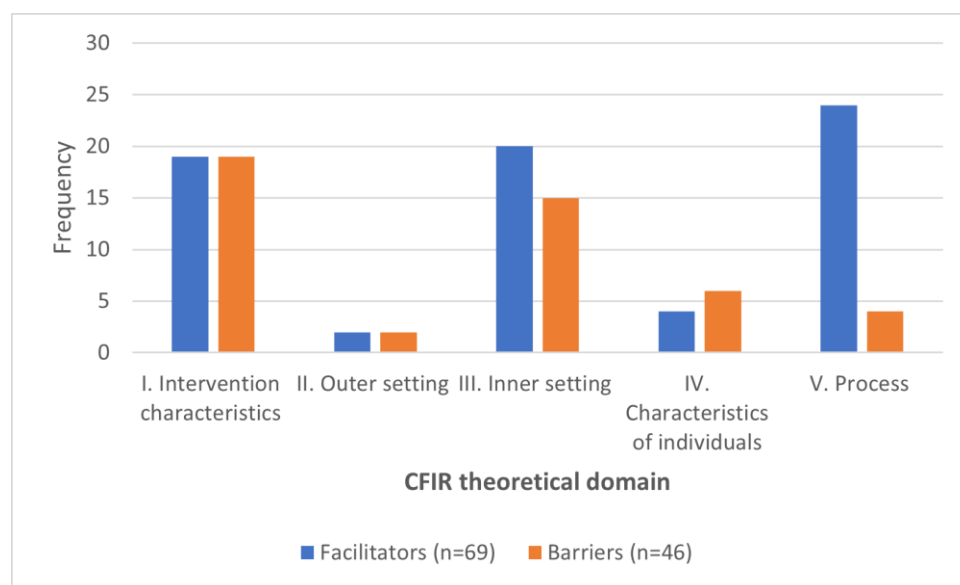


Figure 3. Frequency of facilitators versus barriers.

4. Discussion

Viewed in total, the reporting of facilitators and barriers related to the I. Intervention characteristics, III. Inner setting and V. Process appear somewhat balanced, with some under-reporting of the II. Outer setting and the IV. Characteristics of individuals. Viewed per study, however, the reporting imbalances are more apparent, and we highlight two kinds of imbalance. The first case relates to how a study can concentrate on a single theoretic domain and neglect the rest, and the second is where, regardless of the theoretic domain, a study focuses on either one of facilitators or barriers.

In the first case of theoretic domain imbalance, some studies focused on the characteristics of the intervention [33,49], while others focused on the process [40,47]. Consequently, we end up with an incomplete picture of the implementation for any single study, and it is difficult to compare findings across studies [50]. In the second case, which was typically the case, the studies focused more on the facilitating factors than the barriers. In extreme cases, a study might focus on facilitators alone [38,39,47] and completely neglect the barriers, or the other way round [33,35]. We attribute this poor reporting and imbalance to a lack of implementation science expertise.

We further describe some of the salient highlights in each of the five CFIR domains. These highlights are based on the frequency of discussion they generated in the included studies. The major facilitating factors were related to the implementation process itself and the involvement of users, and this finding is consistent with the literature [51]. In contrast, barriers were mostly associated with the intervention characteristics and inner setting, specifically interoperability, trust and transparency and non-availability of high-quality data.

4.1. Intervention Characteristics

4.1.1. Evidence Strength and Quality

ML algorithms need to continuously learn from new data. As data change and as methodological techniques advance, so must the models, and this presents several challenges. One of challenges is the continuous need to validate algorithms and test whether their specificity and sensitivity have deteriorated. This partially explains why many of the included studies conducted fresh validation tests.

In the validation process, it is essential to make sure the training data represents the population to which the AI system is applied. In practice, results may not be representative across populations [43,48], and experiences may not be generalizable to a new setting [40]. Projects that have been properly evaluated and tested are more likely to succeed in the implementation process. In this regard, [49] recognized clinical trials and multi-center studies are a necessary part of implementation [33,49].

4.1.2. Design Quality and Complexity

Technical design decisions might affect the implementation by facilitating or hindering this process. Usability was cited as both an important facilitator and a barrier. For instance, Romero-Brufau et al. [43] faced problems related to the documentation and presentation of results, and reported difficulties understanding patient information from the decision support system. They dedicated two months to refining the interface of the system and adapting it to the workflow, reflecting the importance of customization in the implementation process. Intuitive, unintrusive, and easy-to-use systems have better chances to succeed in the implementation process [42,45].

4.1.3. Interoperability, Adaptability and Generalizability

In order to successfully implement an AI system in a clinical workflow, the system must interoperate with the targeted hospital systems. In [46], the lack of data interoperability was exposed when overworked nurses were asked to print and deliver medical histories in paper form. This resulted in a situation where medical histories were often not printed, and

thus were not provided to physicians. Data interoperability issues led to poor integration in the clinical workflow.

4.1.4. Integration with Clinical Workflow

A lack of integration with the clinical workflow can be a barrier to the implementation process. For example, Sendak et al. [40] reported workflow issues with model retraining and updating, which are intrinsic ML processes. In addition, projects that are too technically complex or disruptive are at risk of hindering the implementation process:

“Models that require additional work, even if it is as little as looking at another screen and clicking a few more times, are much less likely to be implemented or sustained” [44].

As a solution, some studies showed that ML-based methods compatible with logic-based CDS methods are easier to integrate in the clinical workflow. An example is the use of neural networks for knowledge discovery during the development stage, where results have been later discretized as Arden syntax ECA rules in the production stage [35].

4.2. Outer Setting

External Policies and Incentives

The outer setting was discussed by only one study [35], mostly from the perspective of the legislative environment as a barrier, and the study was conducted in Europe, where AI algorithms used in healthcare are considered Software as a Medical Device (SaMD) and require CE-certification by law. This certification is expensive and time-consuming. However, an exemption allows AI software under clinical evaluation to be used without CE conformity. This requires only an approval from an ethical board and a study protocol adhered to for auditing. This exemption is generally utilized due to costs related to certification [35].

4.3. Inner Setting

Resource Availability

The availability of high-quality data resources within the organization was discussed as an important determinant factor. Most of the studies used electronic health records (EHR) as the primary source of data, and they reported their complexity and inadequate use as a barrier [37,42]. Due to the complex nature of the EHR, key data that can be used to predict the outcome of interest is not always available or ready in the structured format for AI algorithms [31,37,43].

“... key data that reliably predict the outcome of interest may not be readily available as structured, discrete data inputs from the EHR...” [43]

Missing data, noisy data, or data without proper labels and identifiers were among the main factors that lowered data quality and were consequently reported as barriers. Besides data quality, the frequency of data updates is another important issue in maintaining the validity of predictive models [33]. Although data quality and management were usually seen as a barrier by most of the studies, Lee et al. [31] mentioned that rich data availability was a facilitator of the implementation process.

4.4. Characteristics of Individuals

Knowledge, Beliefs and Other Personal Attributes

This domain speaks to the perceptions and beliefs of the individuals involved in the implementation. For instance, a lack of trust among clinicians might hinder the implementation process. The clinicians must trust that the system maintains good sensitivity and specificity and provides trustworthy suggestions in line with evidence-based practice and clinical judgement. As Sendak et al. [40] note:

“Clinical leaders prioritized positive predictive value as a performance measure and were willing to trade-off model interpretability for performance gains”.

At the beginning of the implementation process in [46], the physicians showed interest in the use of an AI-based decision support system that improves diagnostics. However, two of them reported errors in the medical histories, which led them to a wrong diagnosis. As a consequence of sharing those reports among the physicians, the decision support system was perceived as prone to error, generating persistent distrust, and so undermining the usefulness of the system.

Another factor is explainability, a characteristic that directly conditions the transparency and trust of the AI implementation, which, in turn, are precursors of privacy and fairness [52]. No explainability technique is a one-size-fits-all solution for every intervention. Each AI system needs to adapt its explainability to the context and the audience using the model. For instance, a CDS based on a logistic regression model is perfectly understandable by clinicians, but it may be opaque in the context of a patient-oriented app. Other models, such as neural networks, are generally opaque and could be complemented with recent discoveries in explainability techniques such as feature relevance or visualization [53–55].

4.5. Process

Champions and Key Stakeholders

User involvement ranked as the most reported facilitator, followed by the education of key stakeholders. In the very beginning of a project, it is useful to have a common justification [47] and an early mapping of the workflow [43]. In order to attain this, it is necessary to get the relevant participants on board as early as possible [43,44,48]. The stakeholders' feedback and involvement, especially from the leadership, clinicians and users, are also necessary throughout the implementation process [32,38,46,48]. In many instances, the projects strongly supported by the leadership have a higher probability to succeed. Senior leadership support can be crucial to achieve a shared vision among different stakeholders to reach the desired impact.

4.6. Implication of the Results and Recommendations for the Future

The barriers and facilitating factors emerging from this study are not surprising, since they are widely reported in the literature. The included studies presumably have overcome many of the barriers since the studies are based on the late stages of implementation. We expected that insight into the determinants of their successes would shed new light on our basic understanding of AI implementation in clinical settings. However, what we uncovered was insufficient and imbalanced reporting of some key theoretical domains, which suggests a lack of implementation science expertise in the reporting of relevant projects.

The traditional recommendation for e-health implementation processes is to involve both ICT and clinical domain experts. All the included studies seem to have followed this basic recommendation, but our findings suggest there still is a missing piece of the puzzle—socio-organizational considerations. Considering the successes of implementation science as a field, perhaps it is time we looked beyond these traditional recommendations in order to uncover additional synergies based on new modes of inquiry native to implementation science, integrating insights from social science theories and abstractions.

We also showed that domain differences between AI and implementation science have an impact on multidisciplinary research. Since implementation has become a key aspect of AI in healthcare, it is important to unify the vocabulary to make relevant research more accessible to both fields. This could start with annotating relevant publications with an appropriate keyword indicating the implementation stage or purpose of the study, for example, using Curran et al.'s [19] Hybrid Types or research pipeline model (*ibid.*). Classifying implementation stages is an important problem [56] and may reduce the ambiguity of terminology and bridge the gap between data science and implementation science.

4.7. Limitations

Perhaps one of the major limitations of this study is the uncertainty regarding coverage of the relevant literature, which was conditioned by multiple factors. First, we noted that some AI implementations might not have been subject to rigorous scientific study or evaluation, while other implementations were only reported locally in internal reports. This made the implementations essentially inaccessible. In addition, ambiguity related to terminology was a huge factor in successfully identifying all the relevant studies. We allude to the difficulties of defining implementation and the consequences it had on our search strategy and screening.

It is possible that attentional bias is a factor in our findings. Since we set out to identify advanced implementations, it is conceivable these implementations faced comparatively fewer challenging barriers than those of a typical implementation. This might partially explain why there were many more facilitating factors than barriers. In looking at successful implementations, it is quite possible we missed many important barriers from failed implementations.

5. Conclusions

This study exemplifies a theory-based approach to synthesizing determinants of AI implementation success and formalizes known gaps and biases related to how AI implementations are reported. In addition to highlighting the major facilitators and barriers, we noted a widespread imbalance and insufficient reporting of AI implementations in clinical settings. We single out the II. Outer setting and IV. Characteristics of individuals as two key theoretical domains, which were not fully explored in the included studies. As a result, we know very little about the knowledge and beliefs, self-efficacy and other personal attributes of the people involved in the implementations. Similarly, any policies, incentives, collaborative networks or competitive pressures that helped or hindered these implementations are largely unknown. These factors represent an important knowledge gap and require further inquiry before AI implementation in healthcare can be more fully understood.

Further, we recommend two remedial actions based on our findings: (i) implementation science expertise should be a part of every AI implementation project in healthcare in order to improve both the implementation process and the quality of scientific reporting, and (ii) scientific publications involving AI implementations in clinical settings should be annotated with an implementation stage or purpose to make relevant research more easily accessible.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijerph192316359/s1>, Additional File S1—Search strings; Additional File S2—MMAT critical appraisal; Additional File S3—CFIR mappings; Additional File S4—PRISMA checklist; Additional File S5—CFIR codebook.

Author Contributions: P.D.N. conceived the idea as project manager; T.C. designed the approach for the study; M.T. (Miguel Tejedor), T.O.S., T.C., P.D.N. screened, T.O.S., L.M.-R., M.T. (Maryam Tayefi), M.T. (Miguel Tejedor) extracted, T.O.S., L.M.-R. analyzed and M.T. (Maryam Tayefi), P.D.N. interpreted the data; K.L. made a substantial contribution to the search strategy; A.M. (Anne Moen), F.G., A.M. (Alexandra Makhlysheva), L.I., K.L. reviewed the manuscript and substantively revised it. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Norwegian Centre for E-health Research.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ismail, L.; Materwala, H.; Karduck, A.P.; Adem, A. Requirements of Health Data Management Systems for Biomedical Care and Research: Scoping Review. *J. Med. Internet Res.* **2020**, *22*, e17508. [[CrossRef](#)] [[PubMed](#)]
2. Ismail, L.; Materwala, H.; Tayefi, M.; Ngo, P.; Karduck, A.P. Type 2 Diabetes with Artificial Intelligence Machine Learning: Methods and Evaluation. *Arch. Comput. Methods Eng.* **2022**, *29*, 313–333. [[CrossRef](#)]
3. Victor Mugabe, K. Barriers and facilitators to the adoption of artificial intelligence in radiation oncology: A New Zealand study. *Tech. Innov. Patient Support Radiat. Oncol.* **2021**, *18*, 16–21. [[CrossRef](#)] [[PubMed](#)]
4. Strohm, L.; Hehakaya, C.; Ranschaert, E.R.; Boon, W.P.C.; Moors, E.H.M. Implementation of artificial intelligence (AI) applications in radiology: Hindering and facilitating factors. *Eur Radiol.* **2020**, *30*, 5525–5532. [[CrossRef](#)] [[PubMed](#)]
5. Morrison, K. Artificial intelligence and the NHS: A qualitative exploration of the factors influencing adoption. *Future Healthc. J.* **2021**, *8*, e648–e654. [[CrossRef](#)] [[PubMed](#)]
6. Warsavage, T., Jr.; Xing, F.; Barón, A.E.; Feser, W.J.; Hirsch, E.; Miller, Y.E. Quantifying the incremental value of deep learning: Application to lung nodule detection. *PLoS ONE* **2020**, *15*, e0231468. [[CrossRef](#)]
7. Davenport, T.; Kalakota, R. The potential for artificial intelligence in healthcare. *Future Healthc. J.* **2019**, *6*, 94–98. [[CrossRef](#)] [[PubMed](#)]
8. Capobianco, E. High-dimensional role of AI and machine learning in cancer research. *Br. J. Cancer* **2022**, *126*, 523–532. [[CrossRef](#)] [[PubMed](#)]
9. Mazaheri, S.; Loya, M.F.; Newsome, J.; Lungren, M.; Gichoya, J.W. Challenges of Implementing Artificial Intelligence in Interventional Radiology. *Semin. Intervent. Radiol.* **2021**, *38*, 554–559. [[CrossRef](#)]
10. Fischer, U.M.; Shireman, P.K.; Lin, J.C. Current applications of artificial intelligence in vascular surgery. *Semin. Vasc. Surg.* **2021**, *34*, 268–271. [[CrossRef](#)] [[PubMed](#)]
11. Ben Ali, W.; Pesaranghader, A.; Avram, R.; Overtchouk, P.; Perrin, N.; Laffite, S. Implementing Machine Learning in Interventional Cardiology: The Benefits Are Worth the Trouble. *Front Cardiovasc. Med.* **2021**, *8*, 711401. [[CrossRef](#)] [[PubMed](#)]
12. Nilsen, P. Making sense of implementation theories, models and frameworks. *Implement. Sci.* **2015**, *10*, 53. [[CrossRef](#)] [[PubMed](#)]
13. Eccles, M.P.; Mittman, B.S. Welcome to Implementation Science. *Implement. Sci.* **2006**, *1*, 1. [[CrossRef](#)]
14. Van de Velde, S.; Heselmans, A.; Delvaux, N.; Brandt, L.; Marco-Ruiz, L.; Spitaels, D.; Cloetens, H.; Kortteisto, T.; Roshanov, P.; Kunnamo, I. A systematic review of trials evaluating success factors of interventions with computerised clinical decision support. *Implement. Sci.* **2018**, *13*, 114. [[CrossRef](#)] [[PubMed](#)]
15. Gaudet-Blavignac, C.; Foufi, V.; Bjelogrić, M.; Lovis, C. Use of the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) for Processing Free Text in Health Care: Systematic Scoping Review. *J. Med. Internet Res.* **2021**, *23*, e24594. [[CrossRef](#)]
16. Soares, A.; Jenders, R.A.; Harrison, R.; Schilling, L.M. A Comparison of Arden Syntax and Clinical Quality Language as Knowledge Representation Formalisms for Clinical Decision Support. *Appl. Clin. Inform.* **2021**, *12*, 495–506. [[CrossRef](#)]
17. Ismail, L.; Materwala, H.; Znati, T.; Turaev, S.; Khan, M.A. Tailoring time series models for forecasting coronavirus spread: Case studies of 187 countries. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 2972–3206. [[CrossRef](#)]
18. Witten, I.H.; Frank, E.; Hall, M.A.; Pal, C.J. *Data Mining: Practical Machine Learning Tools and Techniques*; Morgan Kaufmann: Burlington, MA, USA, 2016.
19. Curran, G.M.; Bauer, M.; Mittman, B.; Pyne, J.M.; Stetler, C. Effectiveness-implementation hybrid designs: Combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med. Care* **2012**, *50*, 217–226. [[CrossRef](#)]
20. Pearson, N.; Naylor, P.-J.; Ashe, M.C.; Fernandez, M.; Yoong, S.L.; Wolfenden, L. Guidance for conducting feasibility and pilot studies for implementation trials. *Pilot Feasibility Stud.* **2020**, *6*, 167. [[CrossRef](#)]
21. Tricco, A.C.; Lillie, E.; Zarin, W.; O'Brien, K.K.; Colquhoun, H.; Levac, D.; Moher, D.; Peters, M.D.; Horsley, T.; Weeks, L.; et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann. Intern. Med.* **2018**, *169*, 467–473. [[CrossRef](#)]
22. Arksey, H.; O'Malley, L. Scoping studies: Towards a methodological framework. *Int. J. Soc. Res. Methodol.* **2005**, *8*, 19–32. [[CrossRef](#)]
23. Benjamins, S.; Dhunoo, P.; Meskó, B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: An online database. *NPJ Digit. Med.* **2020**, *3*, 118. [[CrossRef](#)]
24. Levac, D.; Colquhoun, H.; O'Brien, K.K. Scoping studies: Advancing the methodology. *Implement. Sci.* **2010**, *5*, 69. [[CrossRef](#)] [[PubMed](#)]
25. Ouzzani, M.; Hammady, H.; Fedorowicz, Z.; Elmagarmid, A. Rayyan—A web and mobile app for systematic reviews. *Syst. Rev.* **2016**, *5*, 210. [[CrossRef](#)] [[PubMed](#)]
26. Hong, Q.N.; Gonzalez-Reyes, A.; Pluye, P. Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the Mixed Methods Appraisal Tool (MMAT). *J. Eval. Clin. Pract.* **2018**, *24*, 459–467. [[CrossRef](#)] [[PubMed](#)]
27. Damschroder, L.J.; Aron, D.C.; Keith, R.E.; Kirsh, S.R.; Alexander, J.A.; Lowery, J.C. Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implement. Sci.* **2009**, *4*, 50. [[CrossRef](#)] [[PubMed](#)]

28. Van de Velde, S.; Kunnamo, I.; Roshanov, P.; Kortteisto, T.; Aertgeerts, B.; Vandvik, P.O.; Flottorp, S. The GUIDES checklist: Development of a tool to improve the successful use of guideline-based computerised clinical decision support. *Implement. Sci.* **2018**, *13*, 86. [[CrossRef](#)]
29. Kuckartz, U.; Rädiker, S. *Analyzing Qualitative Data with MAXQDA*; Springer: Berlin/Heidelberg, Germany, 2019.
30. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [[CrossRef](#)]
31. Lee, E.K.; Atallah, H.Y.; Wright, M.D.; Post, E.T.; Thomas, C., IV; Wu, D.T.; Haley, L.L., Jr. Transforming hospital emergency department workflow and patient care. *Interfaces* **2015**, *45*, 58–82. [[CrossRef](#)]
32. McCoy, A.; Das, R. Reducing patient mortality, length of stay and readmissions through machine learning-based sepsis prediction in the emergency department, intensive care unit and hospital floor units. *BMJ Open Quality* **2017**, *6*, e000158. [[CrossRef](#)]
33. Moon, K.J.; Jin, Y.; Jin, T.; Lee, S.M. Development and validation of an automated delirium risk assessment system (Auto-DelRAS) implemented in the electronic health record system. *Int. J. Nurs Stud.* **2018**, *77*, 46–53. [[CrossRef](#)] [[PubMed](#)]
34. Van der Heijden, A.A.; Abramoff, M.D.; Verbraak, F.; van Hecke, M.V.; Liem, A.; Nijpels, G. Validation of automated screening for referable diabetic retinopathy with the IDx-DR device in the Hoorn Diabetes Care System. *Acta Ophthalmol.* **2018**, *96*, 63–68. [[CrossRef](#)] [[PubMed](#)]
35. Schuh, C.; de Bruin, J.S.; Seeling, W. Clinical decision support systems at the Vienna General Hospital using Arden Syntax: Design, implementation, and integration. *Artif. Intell. Med.* **2018**, *92*, 24–33. [[CrossRef](#)] [[PubMed](#)]
36. Guo, P.; Deng, W. (Eds.) Design and Implementation of Intelligent Medical Customer Service Robot Based on Deep Learning. In Proceedings of the 2019 16th International Computer Conference on Wavelet Active Media Technology and Information Processing, Chengdu, China, 13–15 December 2019.
37. Cruz, N.P.; Canales, L.; Muñoz, J.G.; Pérez, B.; Arnott, I. Improving Adherence to Clinical Pathways through Natural Language Processing on Electronic Medical Records. In *MEDINFO 2019: Health and Wellbeing e-Networks for All*; IOS Press: Amsterdam, The Netherlands, 2019; pp. 561–565.
38. Joerin, A.; Rauws, M.; Ackerman, M.L. Psychological artificial intelligence service, Tess: Delivering on-demand support to patients and their caregivers: Technical report. *Cureus* **2019**, *11*, e3972. [[CrossRef](#)]
39. Gonçalves, L.S.; Amaro, M.L.M.; Romero, A.L.M.; Schamne, F.K.; Fressatto, J.L.; Bezerra, C.W. Implementation of an Artificial Intelligence Algorithm for sepsis detection. *Rev. Bras Enferm.* **2020**, *73*, e20180421. [[CrossRef](#)]
40. Sendak, M.P.; Ratliff, W.; Sarro, D.; Alderton, E.; Futoma, J.; Gao, M. Real-world integration of a sepsis deep learning technology into routine clinical care: Implementation study. *JMIR Med. Inform.* **2020**, *8*, e15182. [[CrossRef](#)]
41. Gonzalez-Briceno, G.; Sanchez, A.; Ortega-Cisneros, S.; Contreras, M.S.G.; Diaz, G.A.P.; Moya-Sanchez, E.U. Artificial intelligence-based referral system for patients with diabetic retinopathy. *Computer* **2020**, *53*, 77–87. [[CrossRef](#)]
42. Xu, H.; Li, P.; Yang, Z.; Liu, X.; Wang, Z.; Yan, W.; He, M.; Chu, W.; She, Y.; Li, Y.; et al. Construction and application of a medical-grade wireless monitoring system for physiological signals at general wards. *J. Med. Syst.* **2020**, *44*, 1–15. [[CrossRef](#)]
43. Romero-Brufau, S.; Wyatt, K.D.; Boyum, P.; Mickelson, M.; Moore, M.; Cognetta-Rieke, C. Implementation of artificial intelligence-based clinical decision support to reduce hospital readmissions at a regional hospital. *Appl. Clin. Inform.* **2020**, *11*, 570–577. [[CrossRef](#)]
44. Scheinker, D.; Brandeau, M.L. Implementing analytics projects in a hospital: Successes, failures, and opportunities. *INFORMS J. Appl. Anal.* **2020**, *50*, 176–189. [[CrossRef](#)]
45. Davis, M.A.; Rao, B.; Cedeno, P.; Saha, A.; Zohrabian, V.M. Machine Learning and Improved Quality Metrics in Acute Intracranial Hemorrhage by Non-Contrast Computed Tomography. *Curr. Probl. Diagn. Radiol.* **2020**, *51*, 556–561. [[CrossRef](#)] [[PubMed](#)]
46. Petitgand, C.; Motulsky, A.; Denis, J.-L.; Régis, C. Investigating the Barriers to Physician Adoption of an Artificial Intelligence-Based Decision Support System in Emergency Care: An Interpretative Qualitative Study. In *Digital Personalized Health and Medicine*; IOS Press: Amsterdam, The Netherlands, 2020; pp. 1001–1005.
47. Betriana, F.; Tanioka, T.; Osaka, K.; Kawai, C.; Yasuhara, Y.; Locsin, R.C. Interactions between healthcare robots and older people in Japan: A qualitative descriptive analysis study. *Jpn. J. Nurs. Sci.* **2021**, *18*, e12409. [[CrossRef](#)] [[PubMed](#)]
48. Betriana, F.; Tanioka, T.; Osaka, K.; Kawai, C.; Yasuhara, Y.; Locsin, R.C. Improving the delivery of palliative care through predictive modeling and healthcare informatics. *J. Am. Med. Inform. Assoc.* **2021**, *28*, 1065–1073.
49. Cho, K.J.; Kwon, O.; Kwon, J.M.; Lee, Y.; Park, H.; Jeon, K.H.; Kim, K.H.; Park, J.; Oh, B.H. Detecting patient deterioration using artificial intelligence in a rapid response system. *Crit. Care Med.* **2020**, *48*, e285–e289. [[CrossRef](#)]
50. Dovigi, E.; Kwok, E.Y.L.; English, J.C., 3rd. A Framework-Driven Systematic Review of the Barriers and Facilitators to Tele dermatology Implementation. *Curr. Dermatol. Rep.* **2020**, *9*, 353–361. [[CrossRef](#)]
51. Servaty, R.; Kersten, A.; Bruckamp, K.; Möhler, R.; Mueller, M. Implementation of robotic devices in nursing care. Barriers and facilitators: An integrative review. *BMJ Open* **2020**, *10*, e038650. [[CrossRef](#)]
52. Arrieta, A.B.; Díaz-Rodríguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; García, S.; Gil-López, S.; Molina, D.; Benjamins, R. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [[CrossRef](#)]
53. Cortez, P.; Embrechts, M.J. Using sensitivity analysis and visualization techniques to open black box data mining models. *Inf. Sci.* **2013**, *225*, 1–17. [[CrossRef](#)]

54. Strumbelj, E.; Kononenko, I. An efficient explanation of individual classifications using game theory. *J. Mach. Learn. Res.* **2010**, *11*, 18.
55. Henelius, A.; Puolamäki, K.; Boström, H.; Asker, L.; Papapetrou, P. A peek into the black box: Exploring classifiers by randomization. *Data Min. Knowl. Discov.* **2014**, *28*, 1503–1529. [[CrossRef](#)]
56. Wang, D.; Ogihara, M.; Gallo, C.; Villamar, J.A.; Smith, J.D.; Vermeer, W.; Cruden, G.; Benbow, N.; Brown, C.H. Automati-classification of communication logs into implementation stages via text analysis. *Implement. Sci.* **2016**, *11*, 119. [[CrossRef](#)] [[PubMed](#)]