Design Approaches for Developing Quality Checklists in Healthcare Organizations: A Scoping Review.

Elizabeth Kwong[1*], Amy Cole[1], Dorothy Sippo[2], Fei Yu[3], Karthik Adapa[1], Christopher M. Shea[4], Carlton Moore[5], Shiva Das[6], Lukasz Mazur[3,6]

[1]Carolina Health Informatics Program, University of North Carolina at Chapel Hill, North Carolina, United States of America
[2]Department of Radiology, School of Medicine, University of North Carolina at Chapel Hill, North Carolina, United States of America
[3]School of Information and Library Science, University of North Carolina at Chapel Hill, North Carolina, United States of America
[4]Department of Health Policy and Management, Gillings School of Public Health, University of North Carolina at Chapel Hill, North Carolina, United States of America
[5]Division of Hospital Medicine, School of Medicine, University of North Carolina at Chapel Hill, North Carolina, United States of America
[6]Department of Radiation Oncology, School of Medicine, University of North Carolina at Chapel Hill, North Carolina, United States of America

*Corresponding author: Elizabeth Kwong
E-mail: ekwong@unc.edu (EK)

Keywords: design methods; quality checklist; safety checklist; healthcare; scoping review

Abstract

Quality checklists have demonstrated benefits in healthcare and other high-reliability organizations, but there remains a gap in the understanding of design approaches and levels of stakeholder engagement in the development of these quality checklists. This scoping review aims to synthesize the current knowledge base regarding the use of various design approaches for developing quality checklists in healthcare. Secondary objectives are to explore theoretical frameworks, design principles, stakeholder involvement and engagement, and characteristics of the design methods used for developing quality checklists. The review followed the Preferred Reporting Items for Systematic Reviews 2020 checklist. Seven databases (PubMed, APA PsycInfo, CINAHL, Embase, Scopus, ACM Digital Library, and IEEE Xplore) were searched for studies using a comprehensive search strategy developed in collaboration with a health sciences librarian. Search terms included "checklist" and "user-centered design" and their related terms. The IAP2 Spectrum of Participation Framework was used to categorize studies by level of stakeholder engagement during data extraction. Twenty-nine studies met the inclusion criteria for this review. Twenty-three distinct design methods were identified that were predominantly non-collaborative in nature (e.g., interviews, surveys, and other methods that involved only one researcher and one participant at a given time). Analysis of the levels of stakeholder engagement revealed a gap in studies that empowered their stakeholders in the quality checklist design process. Highly effective, clear, and standardized methodology are needed for the design of quality checklists. Future work needs to explore how stakeholders can be empowered in the design process, and how different levels of stakeholder engagement might impact implementation outcomes.

Introduction

Quality Checklists in Healthcare

Researchers have used "quality checklist" or "safety checklist" interchangeably to refer to the safety and quality management checklists that outline current evidence-based practices, serve as mnemonic devices, focus on evaluation or performance measurement, assist in maintaining or improving the safety of an organization, or take the form of a cognitive aid or goal sheet (1). A safety and quality management checklist (hereafter referred to as "quality checklist") is an algorithmic listing of actions to verify if an action has taken place and is used to manage and control the quality of deliverables and ensure no step is forgotten (2,3). Quality checklists have been used informally across various industries, with their widespread adoption following a notable aircraft incident in 1935. In this incident, two expert pilots were killed, and engineers on board were injured when the gust-lock was not released prior to takeoff, rendering the elevators inoperable (4,5). After the incident, Boeing developed a series of checklists for the pilots to ensure critical tasks like this were completed. Since then, checklists have become essential for aviation regulation and safety, used by pilots consistently and mandated by the Federal Aviation Administration (FAA) and other regulators internationally. This practice was subsequently adopted more broadly by the military (4,5).

In healthcare, the use of checklists was influenced by the Michigan Health and Hospital Association (MHA) Keystone Center for Patient Safety and Quality Keystone ICU project conducted between 2003 and 2005. During this project, a checklist was used to ensure adherence to evidence-based, infection-control practices, successfully reducing the risk of central line-associated bloodstream infections in intensive care unit patients (6). The development and implementation of the World Health Organization (WHO) surgical safety checklist in 2007-2008 further promoted the use of checklists in healthcare. This 19-item checklist sought to reduce medical errors and adverse events during surgery while improving the consistency of care (5,7). Subsequent healthcare research has explored the development and application

of checklists in various areas, including inpatient care (8), obstetrics (9), hospital discharge (10,11), chemotherapy treatment (12), COVID-19 prevention and control (13,14), and other care areas and procedures. Checklists in healthcare have demonstrated several benefits: they serve as memory aids, ameliorating the effects of fatigue, stress, and distraction, standardize task performance, ensure adherence to best practices, and promote team communication (5,15). Notably, checklists in healthcare have also been crucial for safety management, the improvement of care processes, and the reduction of mortality and morbidity (16–18). However, despite these benefits, barriers to effective design, adoption, and implementation quality checklist design persist. These barriers include slow development and adoption, inconsistent use, and a lack of effective, standardized methodology for quality checklist development, despite research and evidence highlighting the benefits of checklists (1,19).

Design of Quality Checklists

Although the concept of "design" has existed before the Industrial Revolution, the emergence and study of design methods originated in the 1950s and 1960s with the application of novel and "scientific" methods to problem-solving after World War II and the recognition of increased complexity in industrial products (20,21). Since then, design methods have evolved with the publication of books and articles on design methods across various industries, as well as the introduction of artificial intelligence and automation. Influential design researchers including Horst Rittel, Nigel Cross, Herbert Simon, Don Norman, and more along with design organizations such as IDEO have been credited with transforming the field and introducing concepts such as design thinking, solutions-focused problem-solving, user-centered design, and human-centered design (22). These design methods and concepts have been adopted and adapted in engineering, architecture, aviation, technology, product design, and healthcare for a variety of tools, systems, and purposes, including quality checklists.

Several studies have reported considerations, recommendations, and frameworks for the design of quality checklists, which individual organizations can adapt to their respective procedures and processes (1,23–26). These studies recognize that the content and format of quality checklists depend on their specific context, so they often provide only broad guidelines or basic methods. Nevertheless, some established methods for designing checklists include literature reviews, focus groups, Delphi consensus, task analyses, heuristic evaluation, interviews, surveys, and personal experience (19,23). Guidance for the design of quality checklists often involves a review of existing literature, consideration of the user skills, and the experience, context, systems, and environment in which the checklist is intended to be used (1,19). These guidelines advocate for the involvement of potential users and stakeholders in the design process or design team, but do not provide detailed instruction on how to engage stakeholders effectively.

Stakeholder engagement, or the process of incorporating the stakeholder in the design and development process, is being increasingly used and promoted in health research (27,28), but is incompletely described in the design and development of quality checklists in healthcare. While gathering input and insights from stakeholders has demonstrated impact, it also presents challenges, leading to literature-based recommendations on how to best engage stakeholders (25,27–30). Existing studies that have utilized varying levels of stakeholder engagement to design quality checklists acknowledge the importance of stakeholder involvement. However, more research is needed to elucidate the gaps and challenges with user engagement specifically in quality checklists design and enable researchers to develop, evaluate, and implement quality checklists more effectively in their own organizations.

While the utilization of quality checklists has been documented in various settings, there is a shortage of published guidance, rigorous design methodology, and understanding of stakeholder engagement for quality checklists, particularly in healthcare (1). Despite growing evidence and various methods for quality checklist design, a lack of a highly effective, standardized methodology for the design of quality checklists in healthcare and medicine has contributed to inconsistent use, adoption barriers, and gaps in

implementation (1,5). Due to the nature and purpose of quality checklists, these issues, if not addressed, may negatively impact stakeholder satisfaction, usability, quality of care and safety. Therefore, this scoping review aims to synthesize design approaches in the development of quality checklists in healthcare organizations to better understand gaps in design methodology and stakeholder engagement that can be addressed in future studies and ultimately improve quality of care.

Aim and Research Questions

In this review, we scope and synthesize the current state of knowledge regarding the design approaches and associated levels of stakeholder engagement for developing quality checklists in healthcare organizations. More specifically, our review addresses the following questions:

1. What are the characteristics of the design methods adopted to develop quality checklists in healthcare organizations, and how were the design methods defined and measured?
2. What theoretical frameworks and design principles were used to develop quality checklists in healthcare organizations?
3. Who is involved in the development of quality checklists in healthcare organizations, and what is the level of stakeholder engagement in the process?
4. What knowledge gaps exist in the literature in the utilization of design approaches to develop quality checklists in healthcare organizations?

In this study, a quality checklist is defined as a list of items to be performed to complete a task to verify if an action has taken place (or not), and gives information with regard to quality assurance activities and helps control the quality of deliverables (2,3). This scoping review contributes to the growing literature examining how stakeholder engagement affects the design and development of quality checklists, adds to the existing knowledge base of design approaches and frameworks, and highlights gaps and challenges for future research.

Methods

Data Sources and Search Strategy

A comprehensive search strategy was developed in collaboration with a health sciences librarian and executed in seven databases were included in the scoping review and searched on May 3, 2023, including PubMed, APA PsycInfo, CINAHL, Embase, Scopus, ACM Digital Library, and IEEE Xplore.

The search strategy consists of relevant search terms identified based on our research aim and research questions. For example, our search terms included "checklist," its term variations, and related terms (e.g., "cognitive aid", "care pathway", "clinical pathway", "care map", critical path", "clinical path") in combination with "user-centered design" and related terms ("community-based participatory research", "human-centered design", "co-create", "co-created", "co-creation", "co-design", "user centered design", "participatory design", "design thinking", "rapid prototyping", "rapid design", "participatory research", "community-based research", "community based research", "action research", "design sprint", "agile design", "agile method", "agile methodology", "agile project management", "scrum framework", and "agile" + "scrum").

Following the search, retrieved citations were exported to SciWheel (reference management tool; sciwheel.com) and then uploaded into Covidence (systematic review management tool; covidence.org) for manual screening. Both SciWheel and Covidence removed duplicated citations from the aggregated search results.

Eligibility Criteria

Eligibility criteria was developed according to the PICOT Framework (31). To be eligible for inclusion, studies had to be an empirical study involving or describing the design, development, or evaluation of a quality checklist in a healthcare organization (e.g., hospital, clinic, cancer center). Studies needed to involve a safety and quality management checklist of any format (e.g., paper or digital) as the intervention but could involve any comparator or outcome. Studies could be from any publishing year or a healthcare organization in any state, country, or location. We excluded studies that did not provide information on the design, development, or evaluation of the quality checklist intervention, as well as those that only used the checklist as part of another study or to assess another intervention. Additionally, studies published in a language other than English were excluded. Table 1 provides details on the inclusion and exclusion criteria that were used to screen search results.

Table 1. Inclusion and exclusion criteria.

Selection Criteria	Inclusion	Exclusion
Population	• Study is a primary research article	• Exclude studies that are not primary research articles (e.g., do not include editorials, book chapters, opinion pieces) • Exclude systematic / literature reviews • Exclude protocols
Intervention	• Include studies that involve a safety and quality management checklist	• Exclude studies that do NOT involve a safety and quality management checklist
Comparators	• N/A	• N/A
Outcomes	• N/A	• N/A
Timing	• Studies published anytime	• N/A
Setting	• Include studies that take place in a healthcare organization • Include healthcare organizations in any state/country/location	• Exclude studies that do NOT take place in a healthcare organization
Study Design	• Include studies that provide or describe information on the design, development, or evaluation of a quality checklist	• Exclude studies that do NOT provide or describe information on the design, development, or evaluation of a quality checklist

		(e.g., exclude studies that just use a checklist for an experiment)
Language	• Studies available in English (includes publications in another language, as long as an English-language version is available)	• Studies in which publications are not in the English language

Study Selection

Two investigators (EK and AC) independently screened titles and abstracts against the inclusion and exclusion criteria. Discrepancies were resolved by a third investigator (LM). Two investigators (EK and AC) then independently screened full-text articles for inclusion, with discrepancies resolved by a third investigator (LM).

Data Extraction and Synthesis

Data extraction was performed by one investigator (EK) by using the data extraction template provided in Covidence. After customizing the template based on our research questions, it was pilot tested with three included studies and revised to ensure all relevant data were captured. For each study, the following information was extracted: DOI, title, lead author and their contact details, year, country in which the study was conducted, healthcare organization type, study aim, study design (i.e., the framework of research methods to conduct the study), start date, end date, study funding sources, possible conflicts of interest for study authors, checklist type, theoretical frameworks or design principles used, level(s) of stakeholder engagement used, characteristics of design methods used, implementation methods or frameworks used (if applicable), population description, inclusion criteria, exclusion criteria, method of participant recruitment, total number of participants, baseline population characteristics, intervention characteristics, outcomes (e.g., mean, standard deviation, p-value), study strengths, and study limitations. Extracted data were synthesized to address each research question. Specifically, the IAP2 Spectrum of Participation framework and Vaughn's definitions for levels of participation based on the IAP2 Spectrum of Participation framework (32,33) was used to guide data synthesis for Research Question 3 (Who is

involved in the development of quality checklists in healthcare organizations, and what is the level of stakeholder engagement in the process?), classify the level of stakeholder engagement in each study, and qualify what level of participation was used by each study. Stakeholder engagement was defined as one of five levels of engagement according to this framework:

- Inform: Provide stakeholders with balanced and objective information
- Consult: Obtain feedback from stakeholders on analysis, issues, decisions, etc.
- Involve: Work with stakeholders to make sure concern and aspirations are considered and understood
- Collaborate: Partner with stakeholders in each aspect of the decision-making
- Empower: Place final decision making in the hands of the stakeholders

Quality Assessment

While the scoping review was not focused on reported outcomes and may not have required a quality assessment, we utilized the Agency for Healthcare Research and Quality (AHRQ) Evidence-Based Practice Center (EPC) approach to interpret levels of evidence for each study (34). Using this approach, one investigator (EK) performed a quality assessment of each study within the Covidence system and a second investigator (AC) validated the quality assessment. Studies were assessed using the five EPC domains: study limitations, consistency, directness, precision, and reporting bias.

Results

Study Selection

The database search yielded 2551 records, of which 518 were duplicates. After excluding the duplicate articles, a total of 2033 records were identified for screening. After initial screening of article titles and abstracts, 1940 records were excluded, leaving 93 articles for full-text retrieval and review. After full-text

screening, 32 articles were retained and synthesized. These 32 articles reported on 29 studies (6 articles reported on 3 studies, i.e., 2 articles were published about 1 study with different cuts of information or data presented, and were combined into 1 record, for 3 different studies). Inter-rater reliability statistics were calculated for the title and abstract review stage and the full text review stage, generating a Cohen's Kappa of 0.632 and 0.745, respectively. The article selection and review process is detailed in the PRISMA flow diagram (35) in Fig 1.

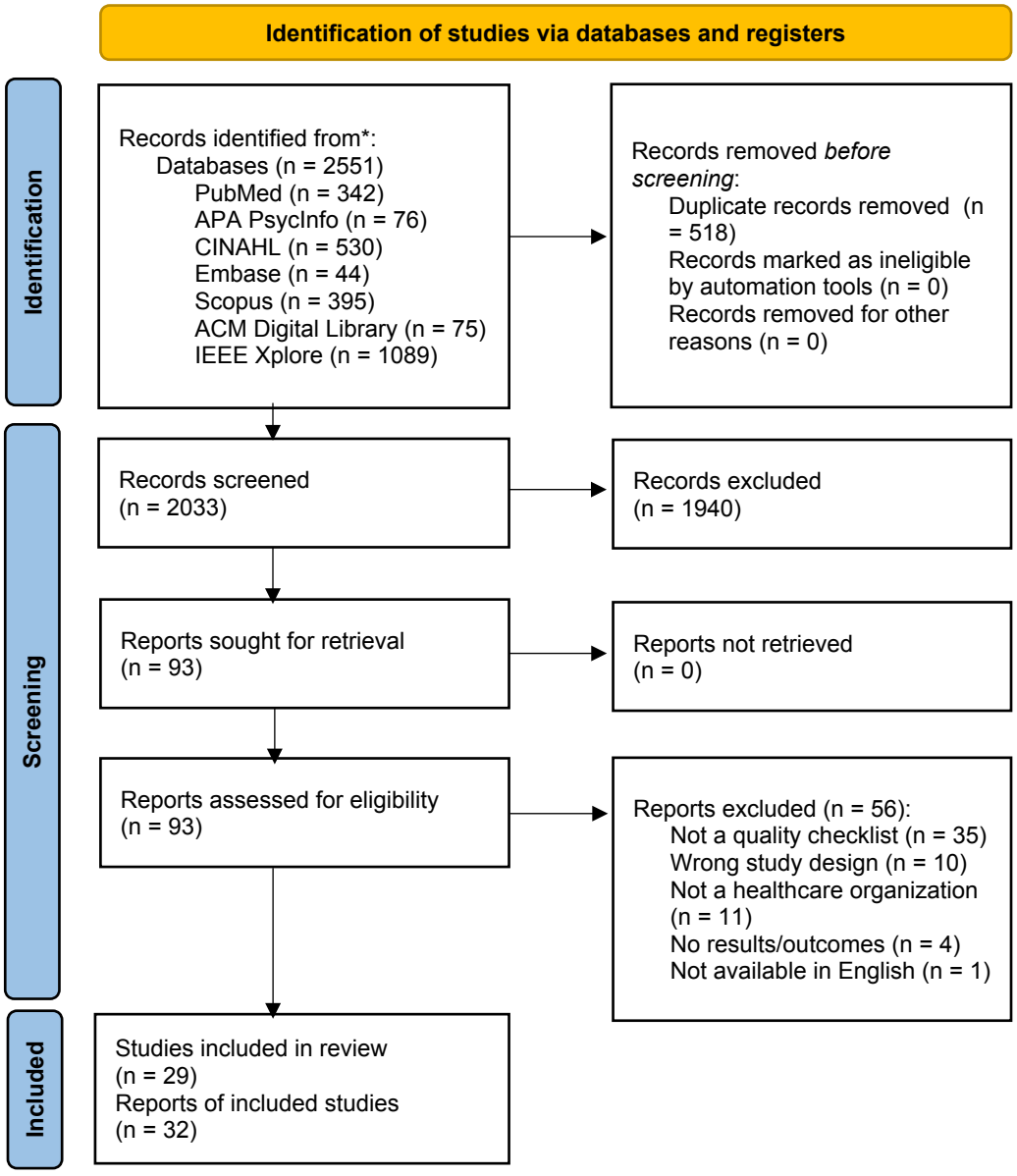


Fig 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram. Modified from Page *et al.*, 2021 (35).

Study Quality Assessment

Included studies had a moderate strength of evidence overall, as they were observational in nature and not randomized controlled trials. The majority of studies (86%) were scored as "direct" (study has evidence that links interventions directly to a health outcome); remaining studies did not have measurable outcome data, intermediate data, or were only available for proxy respondents. One study (3%) (36) provided details of its power analysis, including effect size calculations and optimal information size, demonstrating a "consistent" level (same direction or similar magnitude of effect) and "precise" level (degree of certainty to reach clinically useful conclusion) for consistency and precision domains, respectively. Seventeen of the 29 studies (59%) had "low" levels of study limitations, indicating low protection against bias. These limitations predominantly resulted from small sample size and studies being conducted at one site or within one group. Reporting bias was rated as "suspected" for 66% of studies and "undetected" for 34% of studies. The "suspected" reporting bias resulted from incomplete reporting of the full study, with only a portion of the outcomes disclosed, and limitations to tested tasks and outcomes.

Characteristics of Included Studies

Among the 29 studies identified, 34% were conducted in the United States; studies are representative of 11 different countries or regions. Studies span approximately 23 years from 2001 to 2023. Studies had between 4 and 138 participants, with a mean of 29 participants (Note: one study (37) did not provide a population number). Recruitment of participants included voluntary recruitment, convenience sampling, snowball approaches, word of mouth, social media and regular media (e.g., posters and recruitment flyers), email, and phone. Several studies cited the design, development, or evaluation of the checklist as

part of a quality improvement project in the organization or system. Additionally, the majority of the study designs used were pre- and post- studies (59%).

While not a primary or secondary research question, one exploratory section of extracted evidence included documentation on whether or not the study included implementation of the checklist, and if so, what implementation methods or frameworks were included. The majority (55%) of the studies did not involve implementation while 31% of the studies discussed the implementation of the checklist, but did not describe an implementation method or framework. Additionally, 50% of the studies assessed a digital/electronic quality checklist, 12% evaluated a paper-based quality checklist, and 12% evaluated both a digital- and paper-based checklist. For additional information, see Tables 2 and 3.

305 **Table 2. Overview of included studies.** Summary characteristics for the 34 included studies are provided

306 in alphabetical order by first author's last name.

First Author's Last Name	Year of Publication	Country	Study design	Design Methods	Theoretical frameworks / design principles	Level of stakeholder engagement	Total # of participants
Ashiru-Oredope (38)	2022	Tanzania, Zambia, Uganda, and Ghana (Sub-Saharan Africa)	Before and after	Modified Delphi	Modified Delphi process	Involve	14
Bartlett Ellis (39,40)	2020, 2021	United States	Before and after	Expert panel, use case	APA's App Evaluation Model	Consult	7
Bearsley-Smith (41)	2008	Australia	Before and after	Discussion / focus group	Action research	Consult	18
Benton (42)	2020	United States	Before and after	Literature review, interviews, survey, observations, simulated case	None described	Consult	34
Bentvelsen (43)	2021	The Netherlands	Interrupted time series	Contextual inquiry, questionnaires	CeHRes Roadmap, Minimum Viable Product (MVP) Strategy	Involve	28
Bergerød (44)	2021	Norway	Before and after	Reflective writing, consensus generating session	Consensus method based on a modified NGT; Organizing for Quality model	Involve	20
Bowie (45)	2015	UK	Controlled before and after	Mini-Delphi, content validity index exercise, literature and practice policy document reviews, expert steering group	Participatory design approach; SEIPS work system model; mini-Delphi	Collaborate	18
Campbell-Yeo (46)	2021	Canada	Before and after	Agile co-design sessions, expert consensus, interviews, focus groups	Agile, collaborative co-design process; Theoretical Domains Framework (TDF)	Collaborate	20
Carayon (36)	2020	United States	Controlled before and after	Scenario-based simulation, surveys	None described	Inform	32
Ceulemans (47)	2022	Belgium	Before and after	Semi-structured interviews, focus groups, pilot testing	Experience-based co-design approach	Consult	15
Chudleigh (48,49)	2021, 2022	UK	Controlled before and after	Interviews, group feedback, co-design workshops	Normalisation Process Theory (NPT); experience-based co-design	Involve	42

Flohr (50)	2018	Canada	Before and after	Observation, interviews, simple scenarios, mockups	Work domain analysis (WDA), participatory design approach	Involve	15
Higby (37)	2009	UK	Before and after	Process mapping	Action research process	Consult	Not described
Kerwin (51)	2017	United States	Before and after	Literature review, modified Delphi approach	Participatory design, modified Delphi approach	Collaborate	14
Kirkham (52)	2021	UK	Before and after	Delphi process, surveys	Delphi method	Involve	30
Kulp (53)	2017	United States	Before and after	Use case	In-the-wild co-design approach	Consult	5
Kuo (54)	2019	United States	Interrupted time series	Literature review, modified Delphi, interviews, think-aloud, scenario-based design/ activities, focus groups	Activity Theory (AT)	Collaborate	42
Li (55)	2014	China	Case report	Co-design workshop, literature review	Co-design methodology based on organisational semiotics (OS)	Inform	4
Mastrianni (56)	2023	United States	Before and after	Literature/ case review, co-design sessions, near-live simulation sessions	In-the-wild co-design sessions	Collaborate	8
McIntosh (57)	2017	Australia	Cluster RCT	Simulation-based testing, interviews, usability testing	User-centered design	Consult	40
Moody (58)	2001	Australia	Before and after	Literature review, focus groups, questionnaires	Action research	Collaborate	30
Østergaard (59,60)	2019, 2023	Denmark	Controlled before and after	Participatory design workshops	Cross's research through design; participatory design (PD)	Collaborate	85 nurses (approx.)
Rebic (61)	2021	Canada	Before and after	System Usability Scale (SUS) usability testing, interviews	User-centered design	Consult	12
Rose (62)	2022	Canada	Interrupted time series	Systematic review, interviews, touchpoint video, Delphi	Experience-based co-design, participatory research approach; three-round modified Delphi	Involve	138

Sarcevic (63)	2016	United States	Before and after	Literature review, usability testing with scenarios, focus groups, questionnaire	Iterative, user-centered approach	Consult	4 (approx.)
Schild (64)	2019	Germany	Controlled before and after	Literature search, usability testing, questionnaires, think-aloud, brainstorming session, context interviews, personas, storyboarding	User-centered design	Consult	43 (approx.)
Tarola (65)	2018	United States	Case report	Delphi consensus method, simulated scenarios, surveys	Three-iterative modified Delphi consensus method	Consult	9
Tyler (66)	2021	UK	Before and after	Literature review, co-design workshop, interviews, observations	Co-design approach	Collaborate	40 (approx.)
Wu (67)	2014	United States	Controlled before and after	Observations, narrative simulation	Participatory design approach	Inform	37

Table 3. Characteristics of included studies. Characteristics include: country of publication, study design, year published, total number of participants, and implementation methods/frameworks used (where applicable).

Characteristics of Studies	Number of Studies	Percentage
Country		
United States	10	34%
UK	5	17%
Canada	4	14%
Australia	3	10%
China	1	3%
Belgium	1	3%
Denmark	1	3%
Germany	1	3%
Norway	1	3%
The Netherlands	1	3%
Tanzania, Zambia, Uganda, and Ghana (Sub-Saharan Africa)	1	3%
Study Design		
Before and after study	17	59%
Controlled before and after	6	21%
Interrupted time series	3	10%
Case report	2	7%
Cluster RCT	1	3%
Year Published		
2000-2009	3	12%
2010-2019	11	44%
2020-2023	15	44%
Total number of participants		
1-9 participants	6	21%
10-19 participants	7	24%
20-29 participants	3	10%
30-39 participants	5	17%
40-49 participants	5	17%
50+ participants	2	7%
Not provided	1	3%
Implementation Methods / Frameworks		
Not applicable (did not involve implementation)	16	55%
Involved implementation of the checklist, but did not describe an implementation framework	9	31%

Action Research Cycle / Model of Implementation of Action Research	1	3%
Behavior Change Wheel (BCW) and Theoretical Domains Framework (TDF)	1	3%
Integrated pathway management approach	1	3%
Normalization Process Theory (NPT)	1	3%
Checklist Type		
Digital / Electronic	17	50%
Paper-based	4	12%
Both digital and paper-based	4	12%
Not described / discussed	4	12%

Design Methods

Twenty-three distinct design methods were identified across the 29 studies, with many studies often employing more than one design method. Of these 23 design methods, the most common ones were interviews (41%), literature reviews (34%), simulated sessions/scenarios (31%), Delphi consensus (24%), and co-design / participatory design / consensus generating workshops (24%), which were a mix of collaborative (involve the cooperation of multiple parties working together) and non-collaborative approaches (involve one researcher and one stakeholder at a given time) (Table 4).

Design methods were used in various ways for the design and development of checklists within the studies, including generating checklist content, providing feedback on checklist design and format, collecting data from testing, and improving the design through iterations. Studies also varied in their quality of information provided on design methods. For example, the use of a literature review ranged from an informal literature review (e.g., uncomprehensive literature search) to a full systematic review, with the level of detail in the reported information often minimal. For a list of design methods in each study, refer to Table 2.

Table 4. Overview of design methods used.

Specific Design Methods	Description / Summary	Number of Studies	Percentage

Interviews^	Meeting with users to ask questions and get information	12	41%
Literature review^	Overview/summary of published works in a specific field of study	10	34%
Simulation sessions / simulated scenarios^	Technique to replace and amplify experiences with guided ones that replicate aspects of real world settings (68)	9	31%
Delphi consensus / modified Delphi / mini Delphi+	Systematic process of using the collective opinion of panel members (69)	7	24%
Co-design workshop / participatory design workshop / consensus generating workshop+	Collaborative approach where designers work together with non-designers (70)	7	24%
Surveys / questionnaires^	Set of questions for the purpose of gathering information through survey or study	6	21%
Focus groups+	Small group of people assembled to discuss, respond, and provide opinions and feedback (71)	6	21%
Observation^	Observing participants in their natural environment to gather information	4	14%
Usability testing (e.g., SUS)^	Testing how easy a design is to use with a group of presentative users (72)	4	14%
Use case^	List of tasks or events to identify and organize system requirements	2	7%
Expert panel / expert steering group+	Group of experts and lay members who provide advice, input, and direction (71)	2	7%
Think-aloud activities^	Verbalizing thoughts when engaging with a task	2	7%
Contextual inquiries^	In-depth observation and intervews iusers' workplace while they work (73)	1	3%
Narrative simulation^	Presents a consistently unfolding scenario to participants (67)	1	3%
Pilot testing / prototype testing^	Small-scale test performed before larger-scale study	1	3%
Content validity index exercise^	Quantify the strength of agreement on aspects of the content (45)	1	3%
Touchpoint video^	Composite film representing all key touch points (crucial moments) in the experience (62)	1	3%
Mockups^	Model or replica of a design/device	1	3%
Reflective writing^	Participants reflect on articles in writing (44)	1	3%
Process mapping^	Technique to visually map out workflows and processes	1	3%
Brainstorming session+	Group problem-solving technique to generate new ideas or solutions	1	3%
Personas^	Fictional characters created to represent different user types (74)	1	3%
Storyboarding^	"Freeze-frame" sketches showing scenarios of how people work with the system(73)	1	3%

*some studies may use multiple design methods
+predominantly collaborative method; ^predominantly non-collaborative method

Theoretical Frameworks and Design Principles Used

The included studies specified the theoretical design framework, the design principles used, both, or neither. Sixteen different theoretical frameworks or design principles were identified across the 29 studies. Co-design approaches were the most commonly used design principles, with 8 studies (28%) reporting a "co-design approach" which included five sub-categories (Table 5). Seven studies (24%) used a Delphi or modified Delphi consensus technique, and six studies (21%) used what they called a "participatory design" approach. Description of the participatory design approach varied, with some studies using consensus generating workshops, expert steering groups, simulation sessions, usability testing, or a combination of these methods (Table 5).

Table 5. Theoretical framework and/or design principle(s) used.

Theoretical frameworks / Design principles used*	Definition / Summary	Number of Studies	Percentage
Co-Design Approach	Co-design: People designing together; designing with users (75)	8	28%
Co-design (general)		1	3%
Experience-based co-design	Enables stakeholders to co-design services or solutions together in partnership by gathering experiences from stakeholders (47,49,62)	3	10%
Agile, collaborative co-design	Co-design with increased pace using agile design (discovering, designing, developing, and testing in a series of sprint cycles) (46)	1	3%
In-the-wild co-design	Shifts focus to participants and their environments where there is less control of external factors (53,56)	2	7%
Co-design based on organizational semiotics	An organization is regarded as an information system and considers an IT system to be designed as part of the organization (55)	1	3%
Delphi consensus / modified Delphi / mini Delphi	Uses the collective opinion of panel members (69)	7	24%
Participatory Design (PD)	Focused on how users are engaged in the research process, priorities, and perspectives (33)	6	21%
User-centered Design / Approach	Ensure the thing being designed meets the needs of the user (75)	4	14%
Action Research (AR)	Collaboration between practitioners and/or researchers to benefit the organization (76)	3	10%
Activity Theory (AT)	Focuses on an entire dynamic "activity system" to facilitate design for meaningful human activity, with longer-term activities comprised of actions (54)	1	3%
Nominal Group Technique (NGT)	Consensus process consisting of 4 elements: silent generation, round robin, clarification, and voting (44)	1	3%
Organizing for Quality (OQ) model	Focuses on 6 challenges that hospitals must meet as part of working on quality and safety in healthcare (44)	1	3%

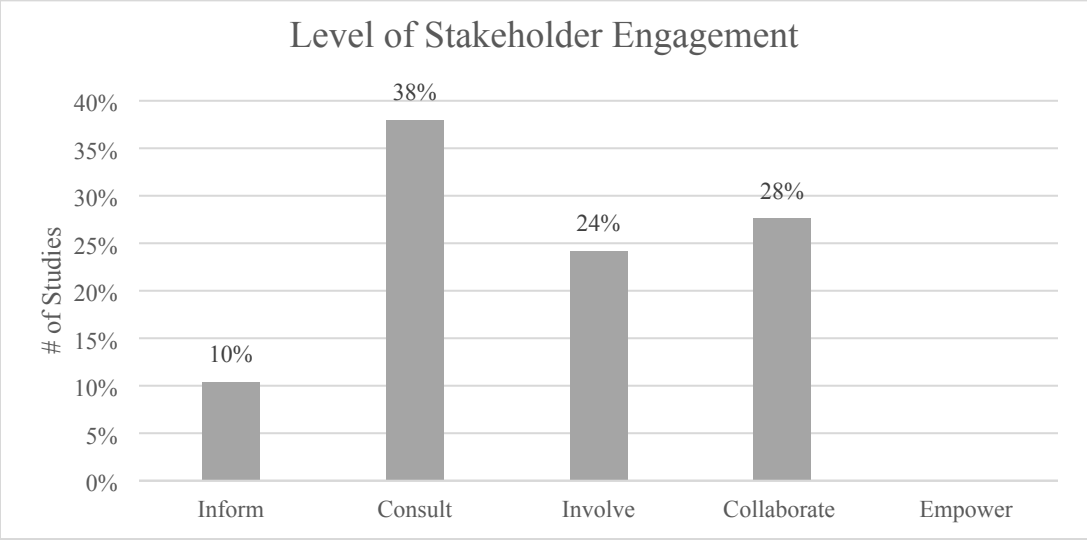
CeHRes Roadmap	Focus on contextual inquiry, value specification, and design and prototyping (43)	1	3%
Minimum Viable Product (MVP) Strategy	Testing the first version of a new product multiple times from an early stage and collecting end-user feedback (43)	1	3%
Theoretical Domains Framework (TDF)	Identifies influences on health professional behavior related to implementation of evidence-based recommendations (77)	1	3%
Work Domain Analysis (WDA)	Identifies information requirements that are event- and time-independent that aims to inform systems design (78)	1	3%
Systems Engineering Initiative for Patient Safety (SEIPS) model	Focuses on system design and its impact on processes, outcomes, and their relationships for patient safety (79)	1	3%
APA's App Evaluation Model	Hierarchical model applied to selecting apps (39)	1	3%
Normalization Process Theory (NPT)	Focuses on the work that individuals and groups do to enable an intervention to become normalized (80)	1	3%
Cross's research through design	Design knowledge resides in the product/artifact (81)	1	3%
None explicitly stated / none described	N/A	2	7%
**some studies may use multiple design principles*			

Level of Stakeholder Engagement

Studies varied in the level of stakeholder engagement according to the IAP2 Spectrum of Participation framework (Fig 2). The greatest percentage of studies were categorized as "Consult," with 11 studies (38%) obtaining feedback from stakeholders on analysis, issues, decisions, etc. None of the 29 studies had an "Empower" level of stakeholder engagement, where final decision-making was placed in the hands of the stakeholders. The majority of studies did not explicitly state what level of participation or engagement the stakeholders had in the study.

The 9 most common design methods used among the studies were also stratified by level of stakeholder engagement using the IAP2 Spectrum of Participation framework (Fig. 3). Results demonstrate studies across the inform, consult, involve, and collaborate levels of stakeholder engagement utilized a variety of design methods, including interviews, literature reviews, simulated scenarios, and surveys. Level of stakeholder engagement was also explored in the context of funding sources available (or not available) for included studies (Fig 4). Of the 29 studies, 76% (n=22) were either nationally, organizationally, or locally/internally funded by a funding source, with 24% (n=7) listing no funding sources. Of the funded

358 studies, the majority (n=13) were nationally funded. Of the 8 studies designated as "collaborate" levels of
359 engagement, n=5 (63%) were nationally funded. Of the 6 studies designated as "inform" levels of
360 engagement, n=4 (67%) did not have funding listed.



361

362 **Fig 2. Level of stakeholder engagement.** The twenty-nine studies were categorized by level of
363 stakeholder engagement as Inform (n=3), Consult (n=11), Involve (n=7), Collaborate (n=8), and
364 Empower (n=0).

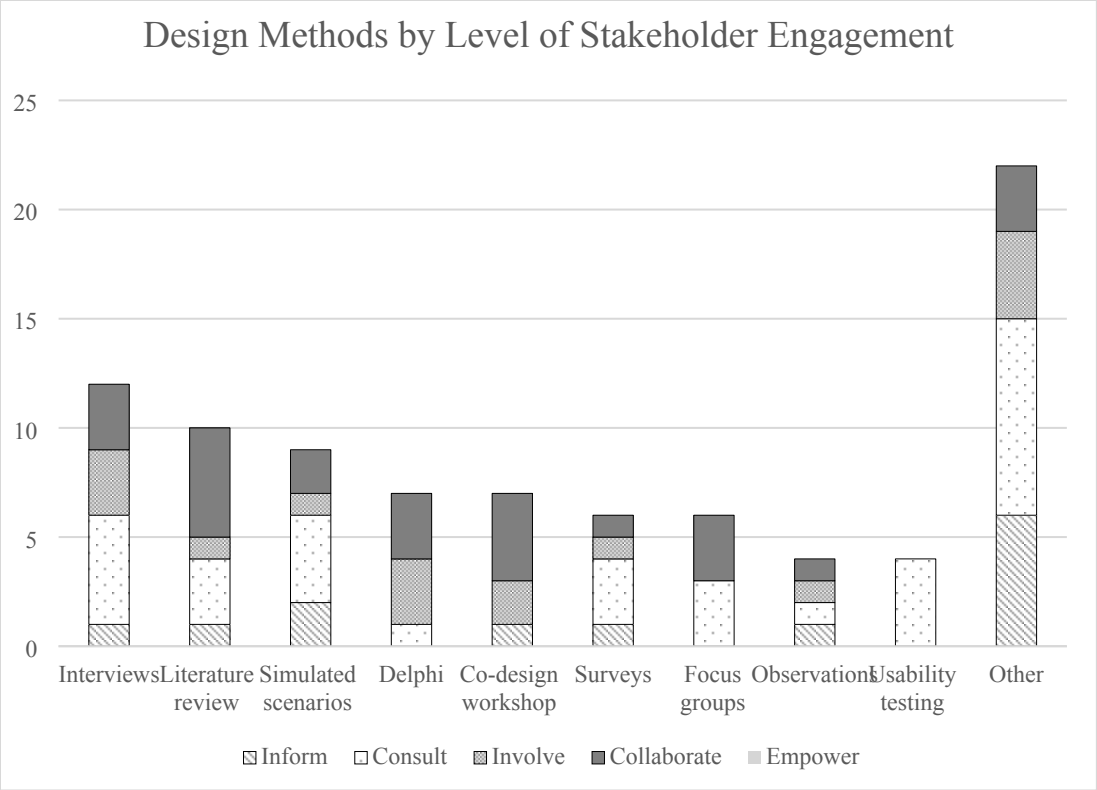


Fig 3. Design methods by level of stakeholder engagement. Most commonly used design methods in relation to level of stakeholder engagement (inform, consult, involve, collaborate, and empower).

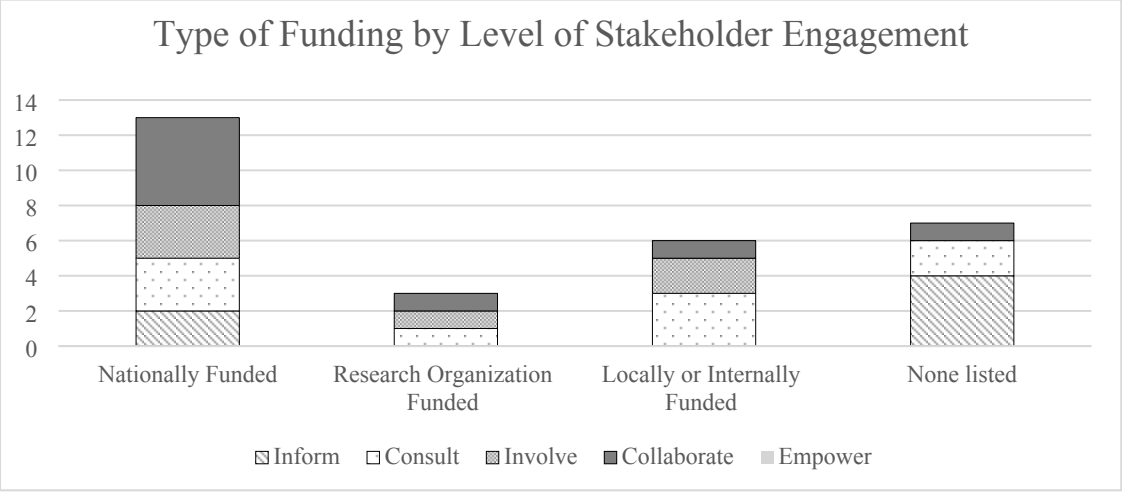


Fig 4. Type of funding in each study by level of stakeholder engagement. Type of funding (whether the study was nationally funded, funded by a research organization, locally/internally funded, or no description of funding at all) in relation to the level of stakeholder engagement (inform, consult, collaborate, and empower).

Stakeholder involvement, or the respective stakeholder roles/groups (e.g., nurses, physicians, patients) that participated in the study, also varied among the studies. Approximately 24% (7/29) had only one primary stakeholder group involved in the study; 76% incorporated two or more different stakeholder roles. Levels of stakeholder engagement, synthesized using the IAP2 Spectrum of Participation framework, were also stratified by one versus multiple stakeholder groups (Fig 5).

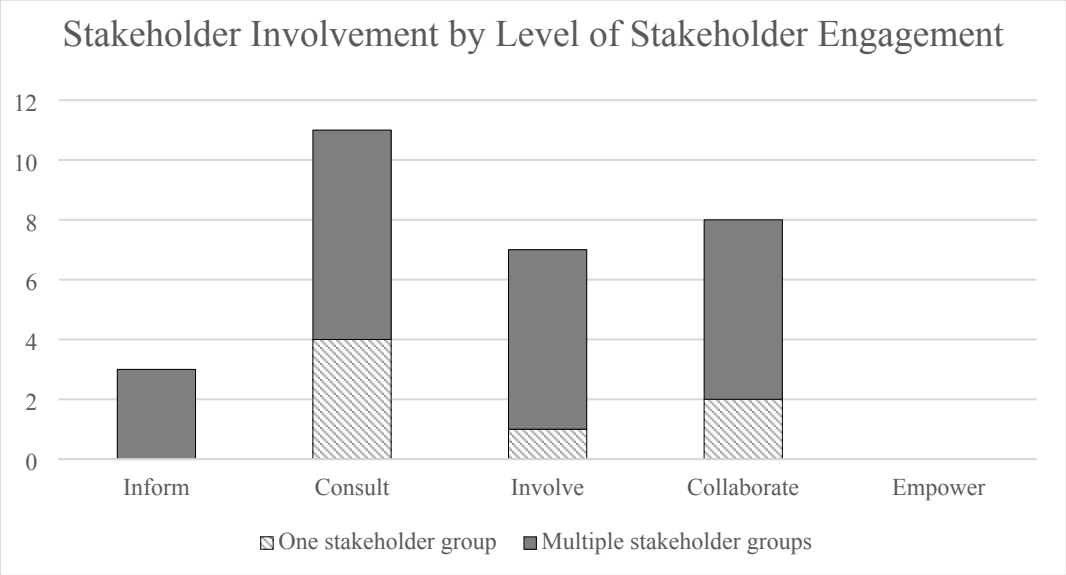


Fig 5. Stakeholder involvement by level of stakeholder engagement. Stakeholder involvement (one vs multiple stakeholder groups involved) in relation to level of stakeholder engagement (inform, consult, involve, collaborate, and empower).

Discussion

This scoping review synthesized evidence from 29 studies by disclosing the current state of design approaches for developing quality checklists in healthcare organizations. We summarized major design approaches and their characteristics, adopted theoretical frameworks and design principles, and levels of stakeholder engagement.

Design Methods

Our analysis suggests that the majority of design methods used to design quality checklists were not collaborative in nature (e.g., involve the cooperation of multiple parties working together), with most methods only involving one researcher and one stakeholder (e.g., interviews, surveys/questionnaires). Studies that used collaborative methods such as Delphi consensus, consensus workshops, and focus groups noted some logistical challenges, such as drop-out observed across consensus meeting phases, virtual-only Delphi meetings hindering productive face-to-face discussion, funding constraints, lack of speaking up in mixed groups, and limitations on participant type and locations (38,44,51,52,65). This may suggest that studies favored more feasible design methods over more collaborative yet logistically-challenging ones.

As anticipated, a wide variety of different design methods were identified among the 29 studies, and several of the studies used multiple design methods in the development of the quality checklists. A third of the studies (n=10/29) started with a form of literature review before developing their quality checklist. This trend likely stems from the diverse existing literature on checklists that can be used to inform the design and development of new quality checklists. However, several studies noted that the quality checklists discussed were part of a quality improvement project in an organization, which may arise from a local need, local evidence base, or existing system process that needs improvement, thereby not prompting a desire for a full literature review to inform the project.

Twelve of the twenty-nine studies (41%) also included a form of interviews in order to gather insights from users and stakeholders prior to checklist development and gather feedback for checklist iterations. Interviews were predominantly described as semi-structured interviews and were done with one to two research team members. All studies used interviews in conjunction with another design method (e.g., observation sessions and interviews, interviews and focus groups). This was anticipated given the

growing use of mixed methods to develop a comprehensive understanding of a topic (triangulation) particularly in qualitative research (82,83). Interview methods varied by channel (video, telephone, in person), length (between 15 minutes to 1 hour), and some described using a prepared interview guide. Notably, several studies described multiple rounds of interviews for different purposes – for example, as part of contextual inquiries for data collection and for value specification, between different checklist prototypes with different focuses, or different interview purposes depending on the type of stakeholder (43,48,49,54). Across the studies using interviews, a variety of primary and occasionally secondary stakeholders were included, in line with checklist development and implementation among high reliability organizations (16). Our findings suggest that while interviews, particularly semi-structured interviews, continue to be the most widely used data collection strategy, organizations adapt interviews as a design method to fit the purposes and needs in the development of quality checklists.

Simulated sessions or scenarios were also frequently used in the design, development, and evaluation of the quality checklist (n=9/29, 31%). These studies predominantly used more than one simulated case and varied by task, environment, situation, and fidelity; for example, one study ran usability testing with multiple, simple scenarios (50), while another used near-live simulation sessions to evaluate features of the checklist (56). Given the healthcare settings of these studies, simulated sessions provide an effective way of performing testing and iterative evaluation of a quality checklist without risking harm to stakeholders such as patients in an emergency care setting (57). Some studies acknowledged that a lack of objective observations in actual settings impacted the validity and limitations of the study (47,54,64). As the focus of this scoping review was on the design and evaluation and not the implementation of quality checklists, many included studies remained in the simulated scenario phase and did not address the transition from simulation to actual (e.g., clinic) settings, although they discussed the intent to apply the checklists in real-world settings in the future.

Other design methods used (Table 4) include 11 different methods that were used in only 1 of the 29 studies. Many of these methods include modeling methods that support conceptual designs or allow the research team to visualize users and their needs, such as personas, storyboarding, and process mapping. These methods were also used in conjunction with other, more commonly used design methods. This may be because the majority of studies had a clear understanding or direct participation of checklist users to qualify data and provide feedback that make these creative modeling methods noncompulsory for research teams. These results indicate opportunities in the design methods of quality checklists that need further exploration. For example, methods such as storyboarding and touchpoint videos have been used with interviews and focus groups to expand the contextual knowledge gathered in the design of electronic patient reported outcomes (84,85). Further research is needed to examine the feasibility and potential added benefit of modeling and creative design methods in conjunction with more established design methods such interviews, focus groups, and simulated sessions/scenarios.

Theoretical Frameworks and Design Principles Used

Studies varied in the way design frameworks and principles were reported. For example, 28% (8/29) of studies described a “co-design approach” which was further described as experience-based, agile/collaborative, in-the-wild, or based in organizational semiotics. The application of co-design methods differed across studies, involving different data gathering techniques (e.g., semi-structured interviews, focus groups, co-design sessions/workshops, use cases, literature review, simulated sessions/scenarios, Delphi consensus group, touchpoint video, and observations). This diversity was somewhat expected given the range of healthcare organization types and specific functions for each quality checklist. Similarly, studies using participatory design (n=6/29) included a variety of design methods, including consensus workshops and meetings, simulated scenarios, mockups, interviews, and questionnaires. Vaughn *et al.* describe participatory research methods and the variety in methods and tools that are participatory by design (33). Rather than emphasizing specific design methods,

participatory design focuses more on how the users and stakeholders are engaged in the research process (33). Despite often nuanced differences in practice, we postulate that co-design (28%), participatory design (21%), and user-centered design (14%) approaches were among the more prevalent design approaches due to study inclusion criteria specifically looking at the design, development, and evaluation of quality checklists (Table 1).

Although not the primary focus of our review, there was a small but meaningful amount of overlap between theoretical frameworks and design principles used and implementation methods or frameworks adopted. While 55% of studies did not involve implementation of the checklist and 31% of studies involved implementation but did not describe an implementation framework (Table 3), 4 studies integrated implementation frameworks with the design and development principles (46,48,49,55,58). These studies used frameworks such as Action Research / the Action Research Cycle, Theoretical Domains Framework (TDF), a co-design methodology for an integrated pathway management approach, and the Normalization Process Theory (NPT). Eleven of the sixteen distinct theoretical frameworks or design principles were identified in only 1 of the 29 studies (Table 5). These theoretical frameworks and design principles appear to be specific to the type of quality checklist, environment, or checklist purpose within the study. For example, one study used the American Psychiatric Association's (APA) App Evaluation Model to support decision-making for a digital health checklist because of its application to selecting apps (40,86). Another study used the Systems Engineering Initiative for Patient Safety (SEIPS) work system model which focuses on system design and its impact on processes, outcomes, and their relationships for patient safety, to identify safety hazards for a preliminary safety checklist (45). Further research is needed to identify and evaluate the optimal use of theoretical frameworks for the development and implementation of quality checklists.

Overall, the scoping review revealed considerable variation in design quality and definition of design principles and frameworks, with no one definitive framework emerging as preferable or exceptional.

Some studies used participatory design, co-design, and/or user-centered design interchangeably or in combination when describing the study methods or design strategies (45,46,49,60,62). Other studies stated a particular approach was used but provided limited information on why the approach was used, definitions for the principle or framework mentioned, and how these principles and frameworks impacted the methods or work accomplished (37,51). Notably, the majority of studies using a participatory design, co-design, or user-centered design approach did *not* include nor clearly specify a validated theoretical framework (47,50,51,56,57,59,60,63,66,67).

Level of Stakeholder Engagement

Level of stakeholder engagement did not appear to be associated with the design approaches or methods used (Fig. 3). For example, the 8 studies that used co-design approaches had different levels of stakeholder participation and engagement, including consult (obtain feedback on analysis, issues, decisions, etc. from stakeholders), involve (work with stakeholders to ensure concerns and aspirations are considered and understood), and collaborate (partner with stakeholders in each aspect of the decision-making) levels. However, of the design methods used in the studies categorized as "consult" levels, the majority were non-collaborative methods in nature (e.g., individualistic methods that only required one stakeholder at a given time, i.e., surveys, interviews). In comparison, methods used in "collaborate" levels of engagement studies often involved design methods that brought stakeholders altogether (i.e., co-design or participatory design sessions, Delphi consensus groups, and focus groups). This suggests a shift towards partnership with multiple stakeholders at once with increased levels of engagement. However, none of the studies were categorized as having "empower" levels of engagement (placing the final decision-making in the hands of stakeholders). This may be because studies developing a quality checklist may not have or require the level of resources needed to empower their stakeholders, or the organizations may not have the capacity to turn over full or final decision-making to stakeholders (87). Nevertheless,

515 this absence suggests a gap in the literature, indicating empowerment of stakeholders to make final
516 decisions in the design of quality checklists is a potential area warranting further exploration.
517
518 The level of stakeholder engagement was also synthesized in the context of funding source availability.
519 We found a potential association of increased levels of stakeholder engagement (i.e., "collaborate" levels
520 of stakeholder engagement) and the presence of national funding, compared to less engagement by
521 stakeholder (i.e., "inform" levels of stakeholder engagement) with no disclosed funding (Fig 4).
522 Specifically, 63% of studies that used "collaborate" levels of engagement (partnered with stakeholders in
523 each aspect of the decision-making) having some type of national funding source, while 67% of studies
524 using "inform" levels of engagement (provide stakeholders with objective information) did not cite
525 funding of any type. This suggests studies with funding likely have more resources for stakeholder
526 engagement and therefore are more capable of engaging stakeholders. Literature has demonstrated the
527 importance of research funding towards meaningful stakeholder engagement and increased financial
528 efforts and support to engage stakeholders in the design of interventions (88–90). Further research is
529 warranted to understand how funding sources may impact design methods and stakeholder engagement in
530 healthcare organizations.
531
532 The type of stakeholder involved in the study did not appear to be associated with stakeholder
533 engagement (Fig 5). Studies ranged from one population type (24% of the studies; e.g., just physicians in
534 a particular department) to a variety of stakeholder groups (76% of the studies; e.g., nurses, pharmacists,
535 speech language therapist, physicians, respiratory therapists, dieticians, social workers, survivors, and
536 family members) involved in one study. Studies with only one stakeholder group were spread across
537 inform, consult, involve, and collaborate levels, as were those involving multiple stakeholder groups .
538 One study that identified physicians as the primary user group of the quality checklist and nurses as a
539 secondary user group conducted usability evaluations with only physicians but acknowledged that
540 requirements of nurses (that did not participate in usability evaluations) may differ from those that were

identified using only physicians in the evaluation process (64). Other studies acknowledged that further validation of the quality checklist was needed utilizing other stakeholders or a wider range of users (39,45,63,65). Some studies also used design methods for different purposes depending on the stakeholder involved. For example, Chudleigh *et al.* held semi-structured interviews with healthcare professionals to explore intervention acceptability, feasibility, and usability and perceptions of influential factors, while they conducted semi-structured interviews with parents of newborns to ascertain experiences and perceptions of the co-designed interventions (48,49). The quantity or variety of stakeholder roles participating in the study did not appear to impact the level of engagement, suggesting the potential feasibility for an increased level of engagement while maintaining a large variety of stakeholder roles involved in the study.

Strengths and Limitations

Strengths

The literature review was performed by searching 7 different databases: PubMed, APA PsycInfo, CINAHL, Embase, Scopus, ACM Digital Library, and IEEE Xplore. These databases were selected by recommendation from an expert committee as they span a diverse range of healthcare, biomedical, life sciences, behavioral and social sciences, psychology, physical sciences, health sciences, nursing and allied health, computing, technology, and engineering fields for identification of manuscripts related to healthcare organizations.

Title and abstract review and full text review were performed with two reviewers (EK and AC), with a third reviewer (LM) resolving any conflicts at each step. Inter-rater reliability statistics were calculated for the title and abstract review stage and the full text review stage, generating a Cohen's Kappa of 0.632 and 0.745, respectively. Kappa value interpretations by Landis and Koch (1977) indicate the strength of

agreement for both stages was "substantial" (91). This suggests there was a substantial degree of agreement among the independent reviewers, lending validity to literature review results.

Limitations

Limitations to the literature review include studies with small participant sizes and observational study designs. Thirteen of the 29 studies (45%) had sample sizes of less than twenty participants, which likely influenced the design methods and approaches that were reported. Participant size may have been influenced by the number of stakeholder groups or funding sources. Studies may have also been able to achieve data saturation through the use of various design principles, with research suggesting a range of 9-17 participants were adequate in qualitative research and usability studies suggesting approximately 15-16 participants to test the usability of a design (92,93).

There was moderate strength of evidence identified by the quality assessment as interpreted using the AHRQ EPIC approach (34). Due to the nature of the literature review, which aimed to answer the primary question of, "what is the current state of utilizing various design approaches for developing quality checklists in healthcare organizations," studies were observational studies and not conventional randomized controlled trials. As a result, this was expected to impact the study limitations, directness, consistency, precision, and reporting bias domains when assessing for quality.

Conclusion

This literature review demonstrated the variety of design methods, theoretical frameworks, and design principles used in the design and development of quality checklists across healthcare organizations. The studies used predominantly non-collaborative design methods (e.g., interviews, surveys/questionnaires), suggesting a preference for more feasible design methods over more collaborative methods that present logistical challenges. The review also revealed design terminology discrepancies and instances where

terms were used interchangeably. The analysis of stakeholder engagement indicated a gap in studies that empowered their stakeholders in the quality checklist design process. Future research needs to clarify theoretical frameworks or design principles and justify the selection of specific design methods, and examine how they impact the outcomes. Further research is also needed into optimal ways to empower stakeholders in the design process, and how different levels of stakeholder engagement might impact design and implementation outcomes. These propositions address the need for a highly effective, standardized methodology for the design of quality checklists that may improve the use, adoption impediments, and implementation barriers that exist in healthcare.

Author Contributions

Elizabeth Kwong and Lukasz Mazur designed and conceptualized the study. Elizabeth Kwong, Amy Cole, and Lukasz Mazur performed data curation, collection, and extraction of the data. Elizabeth Kwong conducted the formal analysis of the data. Elizabeth Kwong, Amy Cole, Dorothy Sippo, Fei Yu, Karthik Adapa, Christopher M. Shea, Carlton Moore, Shiva Das, and Lukasz Mazur interpreted the data. Elizabeth Kwong prepared the original draft of the manuscript. All authors reviewed, edited, and approved the final manuscript.

Registration

The scoping review protocol was not registered.

Competing Interests

The authors have declared that no competing interests exist.

Data Availability Statement

Data underlying the findings are fully available without restriction from the time of publication. All relevant data are within the manuscript.

Acknowledgements and Support

This work would not have been made possible without the financial support of The National Library of Medicine (NLM) Biomedical Informatics and Data Science T-15 Fellowship (T15-LM012500). We would like to give special thanks to Jamie Lynn Conklin and Jennifer Bissram for their assistance in the scoping review and their valuable time and contributions.

References

1. Hales B, Terblanche M, Fowler R, Sibbald W. Development of medical checklists for improved quality of patient care. Int J Qual Health Care. 2008 Feb;20(1):22–30.

2. Checklists | PSNet [Internet]. [cited 2023 Aug 21]. Available from: https://psnet.ahrq.gov/primer/checklists

3. Quality Checklist VS Quality Checksheet - PM Vidya [Internet]. [cited 2023 Aug 21]. Available from: https://pmvidya.com/blog/quality-checklists-vs-quality-checksheets/

4. Higgins WY, Boorman DJ. An Analysis of the Effectiveness of Checklists when combined with Other Processes, Methods and Tools to Reduce Risk in High Hazard Activities [Internet]. 2016 [cited 2023 Aug 18]. Available from: http://www.boeing.com/resources/boeingdotcom/specialty/innovation-quarterly/December-2016/BTJ_Checklist_full.pdf

5. Chaparro A, Keebler JR, Lazzara EH, Diamond A. Checklists: A review of their origins, benefits, and current uses as a cognitive aid in medicine. Ergonomics in Design: The Quarterly of Human Factors Applications. 2019 Apr;27(2):21–6.

6. Pronovost P, Needham D, Berenholtz S, Sinopoli D, Chu H, Cosgrove S, et al. An intervention to decrease catheter-related bloodstream infections in the ICU. N Engl J Med. 2006 Dec 28;355(26):2725–32.

7. Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat A-HS, Dellinger EP, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. N Engl J Med. 2009 Jan 29;360(5):491–9.

8. Wolff AM, Taylor SA, McCabe JF. Using checklists and reminders in clinical pathways to improve hospital inpatient care. Med J Aust. 2004 Oct 18;181(8):428–31.

9. Molina RL, Benski A-C, Bobanski L, Tuller DE, Semrau KEA. Adaptation and implementation of the WHO Safe Childbirth Checklist around the world. Implement Sci Commun. 2021 Jul 8;2(1):76.

10. Halasyamani L, Kripalani S, Coleman E, Schnipper J, van Walraven C, Nagamine J, et al. Transition of care for hospitalized elderly patients--development of a discharge checklist for hospitalists. J Hosp Med. 2006 Nov;1(6):354–60.

11. Soong C, Daub S, Lee J, Majewski C, Musing E, Nord P, et al. Development of a checklist of safe discharge practices for hospital patients. J Hosp Med. 2013 Aug;8(8):444–9.

12. Crandell BC, Bates JS, Grgic T. Start using a checklist, PRONTO: Recommendation for a standard review process for chemotherapy orders. J Oncol Pharm Pract. 2018 Dec;24(8):609–16.

13. Grelat M, Pommier B, Portet S, Amelot A, Barrey C, Leroy H-A, et al. Patients with Coronavirus 2019 (COVID-19) and Surgery: Guidelines and Checklist Proposal. World Neurosurg. 2020 Jul;139:e769–73.

14. Cheng L, Zhu H, Pan H, Huang S, Zhuang Y, Zhu C, et al. Development and application of quality checklist for the prevention and control of COVID-19 in fever clinic and isolation ward of the general hospital. Zhejiang Da Xue Xue Bao Yi Xue Ban. 2021 Feb 25;50(1):74–80.

15. Lingard L, Regehr G, Orser B, Reznick R, Baker GR, Doran D, et al. Evaluation of a preoperative checklist and team briefing among surgeons, nurses, and anesthesiologists to reduce failures in communication. Arch Surg. 2008 Jan;143(1):12–7; discussion 18.

16. Thomassen Ø, Espeland A, Søfteland E, Lossius HM, Heltne JK, Brattebø G. Implementation of checklists in health care; learning from high-reliability organisations. Scand J Trauma Resusc Emerg Med. 2011 Oct 3;19:53.

17. Thomassen Ø, Brattebø G, Søfteland E, Lossius HM, Heltne JK. The effect of a simple checklist on frequent pre-induction deficiencies. Acta Anaesthesiol Scand. 2010 Nov;54(10):1179–84.

18. de Vries EN, Prins HA, Crolla RMPH, den Outer AJ, van Andel G, van Helden SH, et al. Effect of a comprehensive surgical safety system on patient outcomes. N Engl J Med. 2010 Nov 11;363(20):1928–37.

19. Winters BD, Gurses AP, Lehmann H, Sexton JB, Rampersad CJ, Pronovost PJ. Clinical review: checklists - translating evidence into practice. Crit Care. 2009 Dec 31;13(6):210.

20. Cross N. A history of design methodology. In: Vries MJ, Cross N, Grant DP, editors. Design Methodology and Relationships with Science. Dordrecht: Springer Netherlands; 1993. p. 15–27.

21. van der Linden J, Pedroso de Lacerda A, Pedro Ornaghi de Aguiar J. The evolution of design methods. ResearchGate; 2011.

22. Friis Dam R, Siang TY. The History of Design Thinking | IxDF [Internet]. [cited 2023 Sep 12]. Available from: https://www.interaction-design.org/literature/article/design-thinking-get-a-quick-overview-of-the-history

23. Burian BK, Clebone A, Dismukes K, Ruskin KJ. More than a tick box: medical checklist development, design, and use. Anesth Analg. 2018 Jan;126(1):223–32.

24. What Makes a Good Checklist | PSNet [Internet]. [cited 2023 Sep 11]. Available from: https://psnet.ahrq.gov/perspective/what-makes-good-checklist

25. Verdaasdonk EGG, Stassen LPS, Widhiasmara PP, Dankelman J. Requirements for the design and implementation of checklists for surgical processes. Surg Endosc. 2009 Apr;23(4):715–26.

26. Fletcher KA, Bedwell WL. Cognitive Aids. Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care. 2014 Jun;3(1):148–52.

27. Petkovic J, Magwood O, Lytvyn L, Khabsa J, Concannon TW, Welch V, et al. Key issues for stakeholder engagement in the development of health and healthcare guidelines. Res Involv Engagem. 2023 Apr 28;9(1):27.

28. Boaz A, Hanney S, Borst R, O'Shea A, Kok M. How to engage stakeholders in research: design principles to support improvement. Health Res Policy Syst. 2018 Jul 11;16(1):60.

29. Hollmann S, Regierer B, Bechis J, Tobin L, D'Elia D. Ten simple rules on how to develop a stakeholder engagement plan. PLoS Comput Biol. 2022 Oct 13;18(10):e1010520.

30. Concannon TW, Fuster M, Saunders T, Patel K, Wong JB, Leslie LK, et al. A systematic review of stakeholder engagement in comparative effectiveness and patient-centered outcomes research. J Gen Intern Med. 2014 Dec;29(12):1692–701.

31. Riva JJ, Malik KMP, Burnie SJ, Endicott AR, Busse JW. What is your research question? An introduction to the PICOT format for clinicians. J Can Chiropr Assoc. 2012 Sep;56(3):167–71.

32. Bammer G. Key issues in co-creation with stakeholders when research problems are complex. Evid Policy. 2019 Aug 1;15(3):423–35.

33. Vaughn LM, Jacquez F. Participatory research methods – choice points in the research process. J Particip Res Methods. 2020 Jul 21;

34. Berkman ND, Lohr KN, Ansari MT, Balk EM, Kane R, McDonagh M, et al. Grading the strength of a body of evidence when assessing health care interventions: an EPC update. J Clin Epidemiol. 2015 Nov;68(11):1312–24.

35. Page M, McKenzie J, Bossuyt P, Boutron I, Hoffmann T, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.

36. Carayon P, Hoonakker P, Hundt AS, Salwei M, Wiegmann D, Brown RL, et al. Application of human factors to improve usability of clinical decision support for diagnostic decision-making: a scenario-based simulation study. BMJ Qual Saf. 2020 Apr;29(4):329–40.

37. Higby C, Pye K. Improving discharge from the paediatric oncology unit. Paediatr Nurs. 2009 May;21(4):30–2.

38. Ashiru-Oredope D, Garraghan F, Olaoye O, Krockow EM, Matuluko A, Nambatya W, et al. Development and Implementation of an Antimicrobial Stewardship Checklist in Sub-Saharan Africa: A Co-Creation Consensus Approach. Healthcare (Basel). 2022 Sep 6;10(9).

39. Nebeker C, Bartlett Ellis RJ, Torous J. Development of a decision-making checklist tool to support technology selection in digital health research. Transl Behav Med. 2020 Oct 8;10(4):1004–15.

40. Bartlett Ellis R, Wright J, Miller LS, Jake-Schoffman D, Hekler EB, Goldstein CM, et al. Lessons Learned: Beta-Testing the Digital Health Checklist for Researchers Prompts a Call to Action by Behavioral Scientists. J Med Internet Res. 2021 Dec 22;23(12):e25414.

41. Bearsley-Smith C, Sellick K, Chesters J, Francis K, Gippsland Adolescent Depression Research Group. Treatment content in child and adolescent mental health services: development of the treatment recording sheet. Adm Policy Ment Health. 2008 Sep;35(5):423–35.

42. Benton SE, Hueckel RM, Taicher B, Muckler VC. Usability assessment of an electronic handoff tool to facilitate and improve postoperative communication between anesthesia and intensive care unit staff. Comput Inform Nurs. 2020 Oct;38(10):500–7.

43. Bentvelsen RG, van der Vaart R, Veldkamp KE, Chavannes NH. Systematic development of an mHealth app to prevent healthcare-associated infections by involving patients: 'Participatient.' Clinical eHealth. 2021;4:37–44.

44. Bergerød IJ, Braut GS, Fagerdal B, Gilje B, Wiig S. Developing a Next-of-Kin Involvement Guide in Cancer Care-Results From a Consensus Process. Cancer Nurs. 2021 Dec 1;44(6):E447–57.

45. Bowie P, Ferguson J, MacLeod M, Kennedy S, de Wet C, McNab D, et al. Participatory design of a preliminary safety checklist for general practice. Br J Gen Pract. 2015 May;65(634):e330-43.

46. Campbell-Yeo M, Dol J, Richardson B, McCulloch H, Hundert A, Foye S, et al. A co-design of clinical virtual care pathways to engage and support families requiring neonatal intensive care in response to the COVID-19 pandemic (COVES study). J Neonatal Nurs. 2021 Dec;27(6):463–70.

47. Ceulemans M, Brughmans M, Poortmans L-L, Spreuwers E, Willekens J, Roose N, et al. Development and pilot testing of a dispensing protocol on emergency contraceptive pills for community pharmacists in belgium. Pharmacy (Basel). 2022 Jun 1;10(3).

48. Chudleigh J, Holder P, Moody L, Simpson A, Southern K, Morris S, et al. Process evaluation of co-designed interventions to improve communication of positive newborn bloodspot screening results. BMJ Open. 2021 Aug 27;11(8):e050773.

743 49. Chudleigh J, Shakespeare L, Holder P, Chinnery H, Hack G, Gill T, et al. Co-designing Improved
744 Communication of Newborn Bloodspot Screening Results to Parents: Mixed Methods Study. J
745 Particip Med. 2022 Jul 27;14(1):e33485.

746 50. Flohr L, Beaudry S, Johnson KT, West N, Burns CM, Ansermino JM, et al. Clinician-Driven
747 Design of VitalPAD-An Intelligent Monitoring and Communication Device to Improve Patient
748 Safety in the Intensive Care Unit. IEEE J Transl Eng Health Med. 2018 Mar 5;6:3000114.

749 51. Kerwin T, Hittle B, Stredney D, De Boeck P, Wiet G. Multi-Institutional Development of a
750 Mastoidectomy Performance Evaluation Instrument. J Surg Educ. 2017 May 20;74(6):1081–7.

751 52. Kirkham EJ, Crompton CJ, Iveson MH, Beange I, McIntosh AM, Fletcher-Watson S. Co-
752 development of a Best Practice Checklist for Mental Health Data Science: A Delphi Study. Front
753 Psychiatry. 2021 Jun 10;12:643914.

754 53. Kulp L, Sarcevic A. Design in the wild: lessons from researcher participation in design of
755 emerging technology. Ext Abstr Hum Factors Computing Syst. 2017 May;2017:1802–8.

756 54. Kuo P-Y, Saran R, Argentina M, Heung M, Bragg-Gresham JL, Chatoth D, et al. Development of
757 a checklist for the prevention of intradialytic hypotension in hemodialysis care: design
758 considerations based on activity theory. Proceedings of the 2019 CHI Conference on Human
759 Factors in Computing Systems. New York, NY, USA: ACM; 2019. p. 1–14.

760 55. Li W, Liu K, Yang H, Yu C. Integrated clinical pathway management for medical quality
761 improvement – based on a semiotically inspired systems architecture. Eur J Inf Syst. 2014
762 Jul;23(4):400–17.

763 56. Mastrianni A, Sarcevic A, Hu A, Almengor L, Tempel P, Gao S, et al. Transitioning Cognitive
764 Aids into Decision Support Platforms: Requirements and Design Guidelines. ACM Trans Comput-
765 Hum Interact. 2023 Jan 30;

766 57. McIntosh CA, Donnelly D, Marr R. Using simulation to iteratively test and re-design a cognitive
767 aid for use in the management of severe local anaesthetic toxicity. BMJ STEL. 2017 Aug
768 20;4(1):bmjstel-2017-000221.

769 58. Moody G, Choong YY, Greenwood J. An action research approach to the development of a
770 clinical pathway for women requiring Caesarean sections. Contemp Nurse. 2001 Dec;11(2–
771 3):195–205.

772 59. Østergaard KL, Simonsen J, Hertzum M. The Handover from Intensive Care Unit to General
773 Ward: Baseline Performance and Participatory Design of an Electronic Follow-Up Plan. Stud
774 Health Technol Inform. 2019 Aug 21;264:1303–7.

775 60. Østergaard K. A New Model for Intensive Care Unit Follow-up. CIN: Computers, Informatics,
776 Nursing. 2023 Apr;41(4):195–204.

777 61. Rebic N, Munro S, Norman WV, Soon JA. Pharmacist checklist and resource guide for
778 mifepristone medical abortion: User-centred development and testing. Can Pharm J (Ott). 2021
779 May 28;154(3):166–74.

780 62. Rose L, Istanboulian L, Amaral ACK-B, Burry L, Cox CE, Cuthbertson BH, et al. Co-designed
781 and consensus based development of a quality improvement checklist of patient and family-
782 centered actionable processes of care for adults with persistent critical illness. J Crit Care. 2022
783 Dec;72:154153.

784 63. Sarcevic A, Rosen BJ, Kulp LJ, Marsic I, Burd RS. Design challenges in converting a paper
785 checklist to digital format for dynamic medical settings. Int Conf Pervasive Comput Technol
786 Healthc. 2016 May;2016:33–40.

787 64. Schild S, Sedlmayr B, Schumacher A-K, Sedlmayr M, Prokosch H-U, St Pierre M, et al. A Digital
788 Cognitive Aid for Anesthesia to Support Intraoperative Crisis Management: Results of the User-
789 Centered Design Process. JMIR Mhealth Uhealth. 2019 Apr 29;7(4):e13226.

790 65. Tarola CL, Hirji S, Yule SJ, Gabany JM, Zenati A, Dias RD, et al. Cognitive Support to Promote
791 Shared Mental Models during Safety-Critical Situations in Cardiac Surgery (Late Breaking
792 Report). IEEE Conf Cogn Comput Asp Situat Manag. 2018 Jun;2018:165–7.

793 66. Tyler N, Wright N, Gregoriou K, Waring J. Improving mental health care transitions through
794 information capture during admission to inpatient mental health services: a quality improvement
795 study. BMC Health Serv Res. 2021 Oct 21;21(1):1132.

796 67. Wu L, Cirimele J, Leach K, Card S, Chu L, Harrison TK, et al. Supporting crisis response with
797 dynamic procedure aids. Proceedings of the 2014 conference on Designing interactive systems -
798 DIS '14. New York, New York, USA: ACM Press; 2014. p. 315–24.

799 68. Lateef F. Simulation-based learning: Just like the real thing. J Emerg Trauma Shock. 2010
800 Oct;3(4):348–52.

801 69. Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: How to decide its
802 appropriateness. World J Methodol. 2021 Jul 20;11(4):116–29.

803 70. What Is Codesign? — updated 2024 | IxDF [Internet]. Interaction Design Foundation. 2023 [cited
804 2024 Aug 27]. Available from: https://www.interaction-design.org/literature/topics/codesign

71. A guide for Focus Group, Steering Group and Advisory Panel members. Cambridge Biomedical Research Centre [Internet]. [cited 2024 Aug 27]; Available from: https://cambridgebrc.nihr.ac.uk/wp-content/uploads/2017/03/PPI-panel-focus-groups.pdf

72. What is Usability Testing? — updated 2024 | IxDF [Internet]. Interaction Design Foundation. 2016 [cited 2024 Aug 27]. Available from: https://www.interaction-design.org/literature/topics/usability-testing

73. Beyer H. Contextual Design: Defining Customer-Centered Systems (Interactive Technologies). 1st ed. Morgan Kaufmann; 1997.

74. Dam RF, Siang TY. Personas – A Simple Introduction | IxDF [Internet]. Interaction Design Foundation. 2024 [cited 2024 Aug 26]. Available from: https://www.interaction-design.org/literature/article/personas-why-and-how-you-should-use-them

75. Sanders EB-N. From user-centered to participatory design approaches. CRC Press. 2002 Apr 25;18–25.

76. Williamson K, Johanson G. Research Methods: Information, Systems, and Contexts. 2nd ed. Cambridge, MA, United States: Chandos Publishing; 2017.

77. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. Implement Sci. 2017 Jun 21;12(1):77.

78. Hajdukiewicz JR, Vicente * KJ. A theoretical note on the relationship between work domain analysis and task analysis. Theoretical Issues in Ergonomics Science. 2004 Nov;5(6):527–38.

825 79. Carayon P, Schoofs Hundt A, Karsh BT, Gurses AP, Alvarado CJ, Smith M, et al. Work system
826 design for patient safety: the SEIPS model. Qual Saf Health Care. 2006 Dec;15 Suppl 1:i50-8.

827 80. Murray E, Treweek S, Pope C, MacFarlane A, Ballini L, Dowrick C, et al. Normalisation process
828 theory: a framework for developing, evaluating and implementing complex interventions. BMC
829 Med. 2010 Oct 20;8:63.

830 81. Cross N. Designerly ways of knowing. Design Studies. 1982 Oct;3(4):221–7.

831 82. Lambert SD, Loiselle CG. Combining individual interviews and focus groups to enhance data
832 richness. J Adv Nurs. 2008 Apr;62(2):228–37.

833 83. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville .Alan J. The use of triangulation in
834 qualitative research. Oncol Nurs Forum. 2014 Sep;41(5):545–7.

835 84. Dunlop E, Ferguson A, Mueller T, Baillie K, Laskey J, Clarke J, et al. Involving patients and
836 clinicians in the design of wireframes for cancer medicines electronic patient reported outcome
837 measures in clinical care: mixed methods study. JMIR Formativ Res. 2023 Dec 21;7:e48296.

838 85. Rodríguez-Fuertes A, Reinares-Lara P, Garcia-Henche B. Incorporation of the emotional
839 indicators of the patient journey into healthcare organization management. Health Expect. 2023
840 Feb;26(1):297–306.

841 86. Nebeker C, Gholami M, Kareem D, Kim E. Applying a digital health checklist and readability
842 tools to improve informed consent for digital health research. Front Digit Health. 2021 Jul
843 15;3:690901.

844 87. Spectrum of Public Participation – Organizing Engagement [Internet]. [cited 2024 Jan 29].
845 Available from: https://organizingengagement.org/models/spectrum-of-public-participation/

88. Concannon TW, Meissner P, Grunbaum JA, McElwee N, Guise J-M, Santa J, et al. A new taxonomy for stakeholder engagement in patient-centered outcomes research. J Gen Intern Med. 2012 Aug;27(8):985–91.

89. Lee DJ, Avulova S, Conwill R, Barocas DA. Patient engagement in the design and execution of urologic oncology research. Urol Oncol. 2017 Sep;35(9):552–8.

90. Hacker KE, Smith AB. Engaging stakeholders and patient partners. Surg Oncol Clin N Am. 2018 Oct;27(4):665–73.

91. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977 Mar;33(1):159–74.

92. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '93. New York, New York, USA: ACM Press; 1993. p. 206–13.

93. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med. 2022 Jan;292:114523.