

# Systematic review of applied usability metrics within usability evaluation methods for hospital electronic healthcare record systems

## Metrics and Evaluation Methods for eHealth Systems

Marta Weronika Wronikowska BSc, MSc, PhD<sup>1</sup>  | James Malycha MBBS, FCICM<sup>1,2</sup> |  
Lauren J. Morgan BSc, PhD<sup>3</sup> | Verity Westgate MA (Oxon), MSc, MCLIP, FAUA<sup>1</sup> |  
Tatjana Petrinic MSc<sup>4</sup> | J Duncan Young BM, MD, FRCA, FMedSci<sup>1</sup> |  
Peter J. Watkinson MD, ChB, MRCP, FDICM<sup>1</sup>

<sup>1</sup>Critical Care Research Group, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

<sup>2</sup>Department of Acute Care Medicine, University of Adelaide, Adelaide, Australia

<sup>3</sup>Nuffield Department of Surgical Sciences, University of Oxford, John Radcliffe Hospital, Oxford, UK

<sup>4</sup>Bodleian Health Care Libraries, John Radcliffe Hospital, University of Oxford, Oxford, UK

### Correspondence

Marta Weronika Wronikowska, Nuffield Department of Clinical Neurosciences, University of Oxford, Kadoorie Centre for Critical Care Research and Education, Level 3, John Radcliffe Hospital, Headley Way, OX3 9DU Oxford, UK  
Email: marta.wronikowska@ndph.ox.ac.uk

### Funding information

Wellcome Trust, Grant/Award Number: WT-103703/Z/14/Z; Biomedical Research Centre; National Institute for Health Research; Department of Health; Department of Health & Social Care; Health Innovation Challenge Fund, Grant/Award Number: HICF-R9-524; University of Oxford

### Abstract

**Background and objectives:** Electronic healthcare records have become central to patient care. Evaluation of new systems include a variety of usability evaluation methods or usability metrics (often referred to interchangeably as usability components or usability attributes). This study reviews the breadth of usability evaluation methods, metrics, and associated measurement techniques that have been reported to assess systems designed for hospital staff to assess inpatient clinical condition.

**Methods:** Following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology, we searched Medline, EMBASE, CINAHL, Cochrane Database of Systematic Reviews, and Open Grey from 1986 to 2019. For included studies, we recorded usability evaluation methods or usability metrics as appropriate, and any measurement techniques applied to illustrate these. We classified and described all usability evaluation methods, usability metrics, and measurement techniques. Study quality was evaluated using a modified Downs and Black checklist.

**Results:** The search identified 1336 studies. After abstract screening, 130 full texts were reviewed. In the 51 included studies 11 distinct usability evaluation methods were identified. Within these usability evaluation methods, seven usability metrics were reported. The most common metrics were ISO9241-11 and Nielsen's components. An additional "usefulness" metric was reported in almost 40% of included studies. We identified 70 measurement techniques used to evaluate systems. Overall study quality was reflected in a mean modified Downs and Black checklist score of 6.8/10 (range 1–9) 33% studies classified as "high-quality" (scoring eight or higher), 51% studies "moderate-quality" (scoring 6–7), and the remaining 16% (scoring below five) were "low-quality."

**Conclusion:** There is little consistency within the field of electronic health record systems evaluation. This review highlights the variability within usability methods, metrics,

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Journal of Evaluation in Clinical Practice* published by John Wiley & Sons Ltd.

and reporting. Standardized processes may improve evaluation and comparison electronic health record systems and improve their development and implementation.

**KEYWORDS**

electronic health records, electronic patients record (EPR), systematic review, usability methods, usability metrics

## 1 | INTRODUCTION

Electronic health record (EHR) systems are real-time records of patient-centred clinical and administrative data that provide instant and secure information to authorized users. Well designed and implemented systems should facilitate timely clinical decision-making.<sup>1,2</sup> However<sup>3</sup> the prevalence of poorly performing systems suggest the common violation of usability principles.<sup>4</sup>

There are many methods to evaluate system usability.<sup>5</sup> Usability evaluation methods cited in the literature include user trials, questionnaires, interviews, heuristic evaluation and cognitive walkthrough.<sup>6-9</sup> There are no standard criteria to compare results from these different methods<sup>10</sup> and no single method identifies all (or even most) potential problems.<sup>11</sup>

Previous studies have focused on usability definitions and attributes.<sup>12-17</sup> Systematic reviews in this field often present a list of usability evaluation methods<sup>18</sup> and usability metrics<sup>19</sup> with additional information on the barriers and/or facilitators to system implementation.<sup>20,21</sup> However many of these are restricted to a single geographical region,<sup>22</sup> type of illness, health area, or age group.<sup>23</sup>

The lack of consensus on which methods to use when evaluating usability<sup>24</sup> may explain the inconsistent approaches demonstrated in the literature. Recommendations exist<sup>25-27</sup> but none contain guidance on the use, interpretation and interrelationship of usability evaluation methods, usability metrics and the varied measurement techniques applied to assess EHR systems used by clinical staff. These are a specific group of end-users whose system-based decisions have a direct impact on patient safety and health outcomes.

The objective of this systematic review was to identify and characterize usability metrics (and their measurement techniques) within usability evaluation methods applied to assess medical systems, used exclusively by hospital based clinical staff, for individual patient care. For this study, all components in the included studies have been identified as "metrics" to facilitate comparison of methods when testing and reporting EHR systems development.<sup>28</sup> In such cases, Nielsen's satisfaction attribute is equivalent to the ISO usability component of satisfaction.

## 2 | METHODS

This systematic review was registered with PROSPERO (registration number CRD42016041604).<sup>29</sup> During the literature search and initial analysis phase, we decided to focus on the methods used to assess graphical user interfaces (GUIs) designed to support medical decision-making rather than visual design features. We have changed the title

of the review to reflect this decision. We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines<sup>30</sup> (Appendix Table S1).

### 2.1 | Eligibility criteria

Included studies evaluated electronic systems; medical devices used exclusively by hospital staff (defined as doctors, nurses, allied health professionals, or hospital operational staff) and presented individual patient data for review.

Excluded studies evaluated systems operating in nonmedical environments, systems that presented aggregate data (rather than individual patient data) and those not intended for use by clinical staff. Results from other systematic or narrative reviews were also excluded.

### 2.2 | Search criteria

The literature search was carried out by TP using Medline, EMBASE, CINAHL, Cochrane Database of Systematic Reviews, and Open Grey bibliographic databases for studies published between January 1986 and November 2019. The strategy combined the following search terms and their synonyms: usability assessment, EHR, and user interface. Language restrictions were not applied. The reference lists of all included studies were checked for further relevant studies. Appendix Table S2 presents the full Medline search strategy.

### 2.3 | Study selection and analysis

The systematic review was organized using Covidence systematic review management software (Veritas Health Innovation Ltd, Melbourne).<sup>31</sup> Two authors (MW, VW) independently reviewed all search result titles and abstracts. The full text studies were then screened independently (MW, VW). Any discrepancies between the authors regarding the selection of the articles were reviewed by a third party (JM) and a consensus was reached in a joint session.

### 2.4 | Data extraction

We planned to extract the following data:

1. Demographics (authors, title, journal, publication date, country).
2. Characteristics of the end-users.
3. Type of medical data included in EHR systems.
4. Usability evaluation methods and their types, such as:
  - a. questionnaires or surveys,
  - b. user trials,
  - c. interviews,
  - d. heuristic evaluation.
5. Usability metrics (components variously defined as attributes, criteria,<sup>32</sup> or metrics<sup>33</sup>). For the purpose of this review, we adopted the term “metric” to describe any such component) but we include all metric-similar terms used by authors in included studies:
  - a. satisfaction, efficiency, effectiveness metrics,
  - b. learnability, memorability, errors components,
6. Types and frequency of usability metric analysed within usability evaluation methods.

We extracted data into two stages. Stage 1 relied on the extraction of general data from each of the studies that met our primary criteria based the original data extraction form. Stage 2 extended the extraction to gain more specific information such as the measurement techniques for each identified metric as we observed that these were reported in different ways.

The extracted data was assessed for agreement reaching the goal of >95%. All uncertainties regarding data extraction were resolved by discussion among the authors.

## 2.5 | Quality assessment

We used two checklists to evaluate quality of included studies. First used tool, the Downs & Black (D&B) Checklist for the Assessment of Methodological Quality<sup>34</sup> contains 27 questions, covering the following domains: reporting quality (10 items), external validity (three items), bias (seven items), confounding (six items) and power (one item). It is widely used for clinical systematic reviews because it is validated to assess randomized controlled trials, observational and cohort studies. However, many of the D&B checklist questions have little or no relevance to studies evaluating EHR systems, particularly because EHR systems are not classified as “interventions.” Due to this fact, we modified D&B checklist to have usability-oriented tool. The purpose of our modified D&B checklist, constructed of 10 questions, was quality assessment of the aim of the study (specific to usability evaluation methods) evidence that included methods and metrics were supported by peer reviewed literature. Our modified D&B checklist investigated whether the participants of the study were clearly described and representative of the eventual (intended) end-users, the time period over which the study was undertaken being clearly described and the results reflected the methods and described appropriately. The modified D&B checklist is summarized in the appendix (Appendix Table S3). Using this checklist, we defined “high quality” studies as those which scored well in each of the domains (scores  $\geq$  eight). Those studies, which scored in most but not all domains were

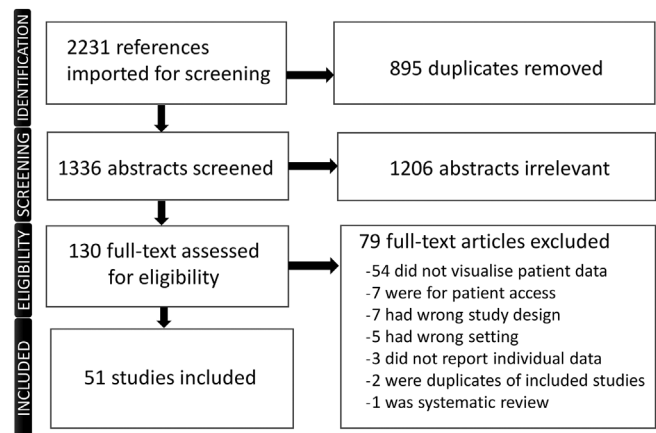


FIGURE 1 Study selection process: PRISMA flow diagram

defined as “moderate quality” (scores of six and seven). The remainder were defined as “low quality” (scores of five and below). We decided to not exclude any paper due to low quality.

## 3 | RESULTS

We followed the PRISMA guidelines for this systematic review (Appendix Table S1). The search generated 2231 candidate studies. After the removal of duplicates, 1336 abstracts remained (Figure 1). From these, 130 full texts were reviewed, with 51 studies eventually being included. All included studies were published between 2001 and 2019. Of the included studies, 86% were tested on clinical staff, 6% on usability experts and 8% on both clinical staff and usability experts. The characteristics of the included studies are summarized in Table 1.

Of the included studies, 16 evaluated generic EHR systems. Eleven evaluated EHR decision support tools (four for all ward patients, one for patients with diabetes, one for patients with chronic pain, one for patients with cirrhosis, one for patients requiring haemodialysis therapy, one for patients with hypertension, one for cardiac rehabilitation and one for management of hypertension, type-2 diabetes and dyslipidaemia). Seven evaluated specific electronic displays (physiological data for patients with heart failure, arrhythmias, also genetic profiles, an electronic outcomes database, longitudinal care management of multimorbid seniors, chromatic pupillometry data, and pulmonary investigation results).

Four studies evaluated medication specific interfaces. Three evaluated electronic displays for patients' clinical notes. Three studies each evaluated mobile EHR systems. Two evaluated EHR systems with clinical reminders. Two evaluated quality improvement tools. Two evaluated systems for use in the operating theatre environment and one study evaluated a sequential organ failure assessment score calculator to quantify the risk of sepsis.

We extracted data on GUIs. All articles provided some description of GUIs, but these were often incomplete, or were a single screenshot. It was not possible to extract further useful information on GUIs.

TABLE 1 Details of included studies

Ref	Author	Year	Country	Participants	Number	System type
35	Aakre et al.	2017	USA	Internal Medicine Residents, Resident, Fellows, Attending Physicians	26	EHR with SOFA <sup>a</sup> score calculator
36	Abdel-Rahman	2016	USA	Physicians, Residents, Nurses, Pharmacologists, Pharmacists, Administrators	28	EHR with the addition of a medication display
37	Al Ghalayini, Antoun, Moacdish	2018	Lebanon	Family Medicine Residents	13	EHR evaluation
38	Allen et al.	2006	USA	“Experts” experienced in usability testing	4	EHR evaluation
39	Belden et al.	2017	USA	Primary Care Physicians	16	Electronic clinical notes
40	Brown et al.	2001	USA	Nurses	10	Electronic clinical notes
41	Brown et al.	2016	UK	Health Information System Evaluators	8	Electronic quality-improvement tool
42	Brown et al.	2018	UK	Primary Care Physicians	7	Electronic quality-improvement tool
43	Chang et al.	2011	USA	Nurses, Home Aides, Physicians, Research Assistants	60	EHR on mobile devices
44	Chang et al.	2017	Taiwan	Medical Students, Physician Assistant Students	132	EHR with the addition of a medication display
45	Devine et al.	2014	USA	Cardiologists, Oncologists	10	EHR with clinical decision support tool
46	Fidler et al.	2015	USA	Critical Care Physicians, Nurses	10	Monitoring – physiology (for patients with arrhythmias)
47	Forsman et al.	2013	Sweden	Specialists Physicians, Resident Physicians, Usability Experts	12	EHR evaluation
48	Fossum et al.	2011	Norway	Registered Nurses	25	EHR with clinical decision support tool
49	Gardner et al.	2017	USA	Staff Physicians, Fellows, Medical Resident, Nurse Practitioners, Physician Assistant	14	Monitoring – physiology (for patients with heart failure)
50	Garvin et al.	2019	USA	Gastroenterology Fellows, Internal Medicine Resident, Interns	20	EHR with clinical decision support tool for patients with cirrhosis
51	Glaser et al.	2013	USA	Undergraduates, Physicians, Registered Nurses	18	EHR with the addition of a medication display
52	Graber et al.	2015	Iran	Physicians	32	EHR with the addition of a medication display
53	Hirsch et al.	2012	Germany	Physicians	29	EHR with clinical decision support tool
54	Hirsch et al.	2015	USA	Internal Medicine Residents, Nephrology Fellows	12	EHR evaluation
55	Hortman, Thompson	2005	USA	Faculty Members, Student Nurse	5	Electronic outcome database display
56	Hultman et al.	2016	USA	Resident Physicians	8	EHR on mobile devices
57	Iadanza et al.	2019	Italy	An evaluator	1	EHR with ophthalmological pupillometry display
58	Jaspers et al.	2008	Netherlands	Clinicians	116	EHR evaluation
59	Kersting, Weltermann	2019	Germany	General Practitioners, Practice Assistants	18	EHR for supporting longitudinal care management of multimorbid seniors
60	Khairat et al.	2019	USA	ICU Physicians (Attending Physicians, Fellows, Residents)	25	EHR evaluation
61	Khajouei et al.	2017	Iran	Nurses	269	Electronic clinical notes
62	King et al.	2015	USA	Intensive Care Physicians	4	EHR evaluation

TABLE 1 (Continued)

Ref	Author	Year	Country	Participants	Number	System type
63	Koopman, Kochendorfen, Moore	2011	USA	Primary Care Physicians	10	EHR with clinical decision support tool for diabetes
64	Laursen et al.	2018	Denmark	Human Computer Interaction Experts, Dialysis Nurses and Nephrologist	8	EHR with clinical decision support tool for patients of need of haemodialysis therapy
65	Lee et al.	2017	South Korea	Professors, Fellows, Residents, Head Nurses, Nurses	383	EHR evaluation
66	Lin et al.	2017	Canada	Physicians, Nurses, Respiratory Therapists	22	EHR evaluation
67	Mazur et al.	2019	USA	Residents and Fellows (Internal Medicine, Family Medicine, Paediatrics Specialty, Surgery, Other)	38	EHR evaluation
68	Nabovati et al.	2014	Iran	Evaluators	3	EHR evaluation
69	Nair et al.	2015	Canada	Family Physicians, Nurse Practitioners, Family Medicine Residents	13	EHR with clinical decision support tool for chronic pain
70	Neri et al.	2012	USA	Genetic Counsellors, Nurses, Physicians	7	Electronic genetic profile display
71	Nouei et al.	2015	Iran	Surgeons, Assistants, Other Surgery Students (Residents Or Fellowship)	unknown	EHR evaluation within theatres
72	Pamplin et al.	2019	USA	Physicians, Nurses, Respiratory Therapists	41	EHR evaluation
73	Rodriguez et al.	2002	USA, Puerto Rico	Internal Medicine Resident Physicians	36	EHR evaluation
74	Schall et al.	2015	France	General Practitioners, Pharmacists, NonClinician E-Health Informatics Specialists, Engineers.	12	EHR with clinical decision support tool
75	Seroussi et al.	2017	USA	Nurses, Physicians	7	EHR evaluation
76	Silveira et al.	2019	Brasil	Cardiologists and Primary Care Physicians	15	EHR with clinical decision support tool for patients with hypertension
77	Su et al.	2012	Taiwan	Student Nurses	12	EHR evaluation
78	Tappan et al.	2009	Canada	Anaesthesiologists, Anaesthesia Residents	22	EHR evaluation within theatres
79	Van Engen-Verheul et al	2016	Netherlands	Nurses, Social Worker, Medical Secretary, Physiotherapist	9	EHR evaluation
80	Wachter et al.	2003	USA	Anaesthesiologists, Nurse Anaesthetists, Residents, Medical Students	46	Electronic pulmonary investigation results display
81	Wu et al.	2009	Canada	Family Physicians, Internal Medicine Physician	9	EHR on mobile devices
82	Zhang et al.	2009	USA	Physicians, Health Informatics Professionals	8	EHR evaluation
83	Zhang et al.	2013	USA	Physicians, Health Informatics Professionals	unknown	EHR evaluation
84	Zheng et al.	2007	USA	Active Resident Users, Internal Medicine Residents	30	EHR with clinical reminders
85	Zheng et al.	2009	USA	Residents	30	EHR with clinical reminders

<sup>a</sup>Sequential Organ Failure Assessment.

Appendix Table S4 presents the specification of type of data included in EHR systems.

### 3.1 | Usability evaluation methods

Ten types of methods to evaluate usability were used in the 51 studies that were included in this review. These are summarized in Table 2. We categorized the 10 methods into broader groups: user trials analysis, heuristic evaluations, interviews and questionnaires. Most authors

applied more than one method to evaluate electronic systems. User trials were the most common method reported, used in 44 studies (86%). Questionnaires were used in 40 studies (78%). Heuristic evaluation was used in seven studies (14%) and interviews were used in 10 studies (20%). We categorized thinking aloud, observation, a three-step testing protocol, comparative usability testing, functional analysis and sequential pattern analysis as user trials analysis. Types of usability evaluation methods are described in Table 3.

Three heuristic evaluation methods were used in seven of the included studies. Four studies used the method described by Zhang





TABLE 2 (Continued)

Ref	User trial analysis										
	User trial	Thinking aloud	Observation	Comparative usability testing	A three step testing protocol	Functional analysis	Sequential pattern analysis	Cognitive walkthrough	Heuristic evaluation	Questionnaire / Surveys	Interview
79	*	*							*		
80	*				*					*	
81	*	*	*								
82						*					
83	*									*	
84							*				
85							*				
N	44	23	13	1	1	1	2	3	7	40	10
%	86	45	25	2	2	2	4	6	14	78	20

et al.<sup>75</sup> One study, despite application of this method, also used the seven clinical knowledge heuristics outlined by Devine et al.<sup>37</sup> The three remaining studies used the heuristic checklist introduced by Nielsen.<sup>67,68</sup> The severity rate scale was sometimes used to judge the importance or severity of usability problems.<sup>76</sup> Findings from heuristics analyses are summarized in Appendix Table S5.

Six types of interviews were used in 10 (20%) studies. The interviews were carried out before the user trial, in the middle of user trial or after the user trial.

The purpose of interviews (unstructured,<sup>38</sup> follow-up,<sup>38</sup> and semi-structured<sup>38</sup>) before the user trial was to understand the end-users' needs, their environment, information/communication flow, and identification of possible changes.

The purpose of interviews (contextual<sup>73</sup>) during the user trial was observation by the end-users while using the system to collect information about potential system utility.

The purpose of interviews following the user trial (prestructured,<sup>71</sup> posttest,<sup>38</sup> semi-structured<sup>39,70,72,42,43,74</sup> [one called in-depth debriefing semi-structured interview<sup>69</sup>]) was mainly gathering information about missing data, system's weaknesses, opportunities for improvements, and users' expectations toward further system development.

Findings from interviews are summarized in Appendix Table S6.

Among the questionnaires, the System Usability Scale (SUS) was used in 16 studies, the Post-Study System Usability Questionnaire (PSSUQ) was used in five studies, the Questionnaire of User Interaction Satisfaction (QUIS) was used in four studies, the Computer Usability Satisfaction Questionnaire (CSUQ) was used three times and the NASA-Task Load (NASA-TLX) was used in six studies. The questionnaires used in studies included in this review are summarized in Appendix Table S7.

### 3.2 | Usability metrics

The usability metrics are summarized in Table 4. Satisfaction was measured in 38 studies (75%), efficiency was measured in 32 studies

(63%), effectiveness was measured in 31 studies (61%), learnability was measured in 12 studies (24%), errors was measured in 16 studies (31%), memorability was measured in one study (2%) and usefulness metric that was measured in 20 studies (39%).

### 3.3 | Usability metrics within usability evaluation methods

Table 5 summarizes the variety of usability evaluation methods used to quantify the different metrics. Some authors used more than one method within the same study (e.g., user trial and a questionnaire) to assess the same metric.

Satisfaction and errors: These were assessed using all four categories of usability evaluation methods. Satisfaction (analysed in 38 studies) was measured using questionnaires (in 31 studies), user trials (in 10 studies), interviews (in two studies) and heuristic evaluation (in one study).

The most frequently reported metrics of user trials were efficiency and effectiveness (both used in 29 studies). For heuristic evaluation it was errors, for interviews' it was usefulness (in four studies) and for questionnaires it was satisfaction (in 31 studies) and usefulness (in 11 studies).

Results were reported in different ways regardless of types of usability evaluation methods or types of usability metric applied, so we created a list of measurement techniques.

### 3.4 | Usability metrics' measurement techniques

We found that different measurement techniques (MT) were used to report the metrics. The number of measurement techniques used to report the identified usability metrics differed from 1 to 25 per single metric. Appendix Table S8 presents all types of measurement techniques applied for all identified metrics and how the

**TABLE 3** Description of the methods included as User Trials Analysis

Method	Description	References
User trial	A process through which end-users (or potential end-users) complete tasks using the system under evaluation. Every participant should be aware of the purpose of the system and analysis. According to Neville et al. <sup>35</sup> participants should be “walked through” through the task under analysis. One of the main objectives for a user trial is to collect observation data but sometimes information comes from the post-test interviews or questionnaires.	35 Studies using user trials are indicated in Table 2
Thinking aloud	Verbal reporting method that generates information on the cognitive processes of the user during task performance. The user must verbalize their thoughts as they interact with the interface	36,37,38,39,40,41,42,43,44,45,46,47,48-51
Observation	Direct and remote observation of users interacting with the system	52
Comparative Usability Testing	Examines the time to acquire information and accuracy of information	53
Three Step Testing Protocol	Tests for intuitiveness within the system. Step one asks users to identify relevant features within the interface. Step two requires users to connect the clinical variables of interest. Step three asks users to diagnose clinical events based on the emergent features of the display	54
Functional Analysis	Measures “functions” within the EHR and classifies them into either Operations or Objects. Operations are then sub classified into Domains or Overheads.	55,56
Sequential Pattern Analysis	Searches for recurring patterns within a large number of event sequences. Designed to show “combinations of events” appearing consistently, in chronological order and then in a recurring fashion.	57,58
Cognitive Walkthrough	Walkthrough of a scenario with execution of actions that could take place during completion of the task completion with expression of comments about use of the interface. It measures ease of learning for new users.	8,36,39,59-62
Heuristic evaluation	Method that helps to identify usability problems using a checklist related to heuristics. Types of HE methods are reported in Appendix Table S5.	7,37,49,59,63-66,67, 68
Questionnaire/ Survey	Research instrument used for collecting data from selected group of respondents. The questionnaires used in studies included in this review are summarized in Appendix Table S7.	Appendix Table S7
Interview	Structured research method, which may be applied before the user-trial, in the middle of user trials or after the user trial. We identified six types of interviews (follow-up, unstructured, prestructured, semi-structured, contextual and post-test interviews), described in Appendix Table S6. The purpose of interviews (unstructured, follow-up and semi-structured), applied before the user trial, was understanding the end-users' needs, their environment, information/communication flow and identification of possible changes, which could improve the process/workflow. The goal of interviews (contextual), applied during user trial, was end-users observation while they work to collect information about potential utility of systems. The purpose of interviews (prestructured, posttest, semi-structured), applied after the user trial, was mainly gathering information about missing data, system's weaknesses, opportunities for improvements and users' expectations toward further system's development.	38,39,42,43,69-74

measurement technique was used (e.g., within a user trial, survey/questionnaire, interview or heuristic evaluation). The greatest variety in usability metric reporting was found in the case of Nielsen's errors quality component (23 measurement techniques were used), ISO 9001 effectiveness (15 measurements techniques used) and our

newly identified usefulness metric (12 measurement techniques used).

User errors, reported using 23 different measurement techniques, were most often reported as the number of errors ( $n = 4$ ) or percentage of errors made ( $n = 6$ ). Authors sometimes provided contextual





TABLE 4 Usability metrics

Ref	Satisfaction	Efficiency	Effective-ness	Learn-ability	Memor-ability	Errors	Useful-ness	Total
77		*						1
36	*	*	*	*			*	5
78	*	*	*			*		4
63					*			1
53	*	*	*					3
79	*	*	*					3
59	*		*			*		3
69	*	*	*			*	*	5
80	*	*	*				*	4
81	*	*		*			*	4
37	*	*	*			*	*	5
82	*	*		*		*		4
38	*	*	*	*				4
39	*	*	*				*	4
83	*			*			*	3
70	*	*	*	*			*	5
84	*	*	*					3
85	*	*	*			*	*	5
86		*	*					2
40	*	*	*					3
87	*		*	*		*		4
41	*	*	*					3
64						*		1
60	*	*		*				3
71								0
72	*	*					*	3
88	*							1
42	*							1
43	*	*	*			*		4
65								0
89	*		*				*	3
44	*	*	*			*	*	5
90		*	*			*		3
66			*			*		2
45	*			*				2
46	*	*	*			*	*	5
73	*						*	2
74	*	*	*					3
91	*	*	*	*				4
48	*	*	*	*		*	*	6
92	*						*	2
93						*	*	2
47	*	*	*	*				4
94	*	*	*				*	3
49		*	*			*		3
54	*							1

(Continues)

TABLE 4 (Continued)

Ref	Satisfaction	Efficiency	Effective-ness	Learn-ability	Memor-ability	Errors	Useful-ness	Total
50	*	*	*				*	4
55	*	*	*				*	4
95								0
57								0
58								0
Total	38	32	31	12	1	16	20	
%	75	63	61	24	2	31	39	

	User Trials	Heuristic Evaluation	Interviews	Questionnaires
Satisfaction	10	1	2	31
Efficiency	29	0	0	2
Effectiveness	29	0	0	2
Learnability	4	0	0	10
Memorability	0	1	0	0
Errors	11	5	1	1
Usefulness	5	0	4	11

TABLE 5 Usability metrics and the usability methods used to measure them. Values are the number of studies

information about the type of errors ( $n = 5$ ), or reason for errors ( $n = 1$ ). These measurement techniques were investigated within user trials.

The effectiveness metric was reported with 15 measurement techniques. The most frequent ones used were: number of successfully completed tasks (in eight studies), percentage of correct responses (in four studies) and the percentage of participants able to complete tasks (in three studies).

Efficiency was mostly reported as time to complete tasks ( $n = 27$ ). Sometimes this was reported as a comparator against an alternative system ( $n = 13$ ). Task completion was also measured by number of clicks ( $n = 11$ ). Five studies measured the number of clicks compared to a predetermined optimal path. In two cases the time of assessing the patient's state was also measured.

Satisfaction was reported by eight measurement techniques. This was most frequently by questionnaire results ( $n = 31$ ), by general user comments related to the system satisfaction ( $n = 10$ ), by recording the number of positive comments ( $n = 4$ ) or the number of negative comments ( $n = 4$ ) or users preferences across two tested system versions ( $n = 1$ ).

The usefulness metric was reported using 12 different measurement techniques. These included users' comments regarding the utility of the system in clinical practice ( $n = 5$ ), comments about usefulness of layout ( $n = 1$ ), average score of system usefulness ( $n = 5$ ), and total mean scores for work system-useful-related dimensions ( $n = 1$ ).

### 3.5 | Quality assessment

Results for the quality assessment are summarized in the appendix (Appendix Table S9). We did not exclude articles due to poor quality. For the D&B quality assessment, the mean score (out of a possible

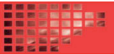
32 points) was 9.9 and the median and mode score were 10. The included studies scored best in the reporting domain, with seven out of the 10 questions generating points. Studies scored inconsistently (and generally poorly) in the bias and confounding domains and no study scored points in the power domain (Appendix Table S10).

Using the Modified D&B checklist the mean score was 6.8 and the median was 7.0 out of a possible 10 points. Seventeen studies (33%) were classified as "high-quality" (scoring eight or higher), 26 studies (51%) were "moderate-quality" (scoring six or seven), and the remaining eight studies (16%) were "low-quality" (scoring five or below). The relationship between the two versions of the D&B scores is shown in the appendix (Appendix Figure S1).

## 4 | DISCUSSION

### 4.1 | Main findings

This review demonstrates wide variability in both methodological approaches and study quality in the considerable amount of research undertaken to evaluate EHR systems. EHR systems, despite being expensive and complex to implement, are becoming increasingly important in patient care.<sup>96</sup> Given the pragmatic, rather than experimental nature of EHR systems, it is not surprising that EHR systems evaluation requires an observational or case-controlled study. Common methodological failings were unreferenced and incorrectly named usability evaluation methods, discrepancies between study aims, methods and results (e.g., authors did not indicate their intention to measure certain metrics and then subsequently reported these metrics in the results or described the usability evaluation methods in method section but did not present the results).



In the future, well-conducted EHR system evaluation requires established human-factor engineering driven evaluation methods. These need to include clear descriptions of study aims, methods, users and time-frames. The Medicines and Healthcare Regulation Authority (MHRA) requires this process for medical devices and it is logical that a comparable level of uniform evaluation may benefit EHRs.<sup>97</sup>

## 4.2 | Strengths

We have summarized the usability evaluation methods, metrics, and measurement techniques used in studies evaluating EHR systems. To our knowledge this has not been done before. Our results' tables may therefore be used as a goal-oriented matrix, which may guide those requiring a usability evaluation method, usability metric, or combination of each, when attempting to study a newly implemented electronic system in the healthcare environment. We identified usefulness as a novel metric, which we believe has the potential to enhance healthcare system testing. Our modified D&B quality assessment checklist was not validated but has the potential to be developed into a tool better suited to assessing studies that evaluate medical systems. By highlighting the methodological inconsistencies presented by researchers in this field we hope to improve the quality of research in the field, which may in turn lead to better systems being implemented in clinical practice.

## 4.3 | Limitations

The limitations of the included studies were reflected in the quality assessment: none of the included studies scored >41% in the original D&B checklist, which is indicative of poor overall methodological quality. Results from the modified D&B quality assessment scale, offered by our team, were better but still showed over half the studies were of low or medium quality. A significant proportion of the current research into EHR systems usability has been conducted by commercial, nonacademic entities. These groups have little financial incentive to publish their work unless the results are favourable, so although this review may reflect publication bias, it is unlikely to reflect all current practices. It was sometimes difficult to extract data on the methods used in studies included in this review. This may reflect a lack of consensus on how to conduct studies of this nature, or a systematic lack of rigour in this field of research.

## 5 | CONCLUSION

To our knowledge, this systematic review is the first to consolidate applied usability metrics (with their specifications) within usability evaluation methods to assess the usability of electronic health systems used exclusively by clinical staff. This review highlights the lack of consensus on methods to evaluate EHR systems' usability. It is possible that healthcare work efficiencies are hindered by the resultant inconsistencies.

The use of multiple metrics and the variation in the ways they are measured, may lead to flawed evaluation of systems. This in turn may

lead to the development and implementation of less safe and effective digital platforms.

We suggest that the main usability metrics as defined by ISO 9241-1 (efficiency, effectiveness, and satisfaction) used in combination with usefulness, may form part of an optimized method for the evaluation of electronic health systems used by clinical staff. Assessing satisfaction via reporting the users positive and negative comments; assessing efficiency via time to task completion and time taken to assess the patient state; assessing effectiveness via number/percentage of completed tasks and quantifying user errors; and assessing usefulness via user trial with think-aloud methods, may also form part of an optimized approach to usability evaluation.

Our review supports the concept that high performing electronic health systems for clinical use should allow successful (effective) and quick (efficient) task completion with high satisfaction levels and they should be evaluated against these expectations using established and consistent methods. Usefulness may also form part of this methodology in the future.

## ACKNOWLEDGEMENTS

We would like to acknowledge Nazli Farajidavar and Tingting Zhu who translated screenshots of medical systems in Farsi and Mandarin languages and Julie Darbyshire for extensive editing and writing advice. JM would like to acknowledge Professor Guy Ludbrook and the University of Adelaide, Department of Acute Medicine, who are administering his doctoral studies. This publication presents independent research supported by the Health Innovation Challenge Fund (HICF-R9-524; WT-103703/Z/14/Z), a parallel funding partnership between the Department of Health & Social Care and Wellcome Trust. The views expressed in this publication are those of the author(s) and not necessarily those of the Department of Health or Wellcome Trust. The funders had no input into the design of the study, the collection, analysis, and interpretation of data nor in writing the manuscript. PW is supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC).

## CONFLICT OF INTEREST

The authors declare that they have no competing interests.

## ETHICS STATEMENT

Ethical approval is not required for the study.

## AUTHORS' CONTRIBUTIONS

MW, LM and PW designed the study, undertook the methodological planning and led the writing. TP advised on search strategy and enabled exporting of results. JM, VW and DY assisted in study design, contributed to data interpretation, and commented on successive drafts of the manuscript. All authors read and approved the final manuscript.

## DATA AVAILABILITY STATEMENT

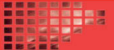
The data that supports the findings of this study are available in the supplementary material of this article

## ORCID

Marta Weronika Wronikowska  <https://orcid.org/0000-0002-5416-7872>

## REFERENCES

- Cresswell KM, Sheikh A. Inpatient clinical information systems. *Key Advances in Clinical Informatics: Transforming Health Care through Health Information Technology*; UK: Academic Press; 2017. <https://doi.org/10.1016/B978-0-12-809523-2.00002-9>.
- Sittig DF. Category definitions. *Clinical Informatics Literacy*; UK: Academic Press; 2017. <https://doi.org/10.1016/b978-0-12-803206-0.00001-8>.
- Kruse CS, Mileski M, Alaytsev V, Carol E, Williams A. Adoption factors associated with electronic health record among longterm care facilities: a systematic review. *BMJ Open*. 2015;5:1-9. <https://doi.org/10.1136/bmjopen-2014-006615>.
- Zahabi M, Kaber DB, Swangnetr M. Usability and safety in electronic medical records Interface design: a review of recent literature and guideline formulation. *Hum Factors*. 2015;57:805-834. <https://doi.org/10.1177/0018720815576827>.
- Fernandez A, Insfran E, Abrahão S. Usability evaluation methods for the web: A systematic mapping study. *Inform Software Technol*. 2011; 53(8):789-817. <https://doi.org/10.1016/j.infsof.2011.02.007>.
- Bowman DA, Gabbard JL, Hix D. A survey of usability evaluation in virtual environments: classification and comparison of methods. *Presence Teleoperators Virtual Environ*. 2002;11(4):404-424. <https://doi.org/10.1162/105474602760204309>.
- Nielsen J, Molich R. Heuristic Evaluation of User Interfaces. In: ACM CHI'90 Conference (Seattle, WA, 1-5 April), 1990:249-256.
- Lewis C, Polson PG, Wharton C, Rieman J. Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Empowering People - CHI '90, 1990:235-242. doi: <https://doi.org/10.1145/97243.97279>
- Savoy A, Patel H, Flanagan ME, Weiner M, Russ AL. Systematic heuristic evaluation of computerized consultation order templates: Clinicians' and human factors Engineers' perspectives. *J Med Syst*. 2017;41(8):129. <https://doi.org/10.1007/s10916-017-0775-7>.
- Hartson HR, Andre TS, Williges RC. Criteria for evaluating usability evaluation methods. *Int J Human Comput Interact*. 2003;15(1):145-181. [https://doi.org/10.1207/S15327590IJHC1501\\_13](https://doi.org/10.1207/S15327590IJHC1501_13).
- Horsky J, McColgan K, Pang JE, et al. Complementary methods of system usability evaluation: surveys and observations during software design and development cycles. *J Biomed Inform*. 2010;43(5):782-790. <https://doi.org/10.1016/j.jbi.2010.05.010>.
- Weichbroth P. Usability attributes revisited: a time-framed knowledge map. In: Proceedings of the 2018 Federated Conference on Computer Science and Information Systems, 2018. doi:<https://doi.org/10.15439/2018f137>
- Seffah A, Donyaee M, Kline RB, Padda HK. Usability measurement and metrics: a consolidated model. *Softw Qual J*. 2006;14(2):159-178. <https://doi.org/10.1007/s11219-006-7600-8>.
- Abran A, Khelifi A, Suryan W, Seffah A. Usability meanings and interpretations in ISO standards. *Software Quality J*. 2003;11:325-338. <https://doi.org/10.1023/A:1025869312943>.
- Schneiderman B. Designing the user Interface: strategies for effective. *Human Comput Interact*. Pearson; 1998;624 ISBN-10: 0321537351. ISBN-13: 978-0321537355.
- International Organization for Standardization. ISO 9241-11: ergonomic requirements for office work with visual display terminals (VDTs) - part 11: guidance on usability. *Int Organ Stand*. 1998; 1998(2):28. <https://doi.org/10.1038/sj.mp.4001776>.
- Nielsen J. Usability. *Engineering*. Morgan Kaufmann; 1993;362. ISBN-10: 0125184069. ISBN-13: 978-0125184069.
- Ellsworth MA, Dziadzko M, O'Horo JC, Farrell AM, Zhang J, Herasevich V. An appraisal of published usability evaluations of electronic health records via systematic review. *J Am Med Informatics Assoc*. 2017;24(1):218-226. <https://doi.org/10.1093/jamia/ocw046>.
- Dukhanin V, Topazian R, DeCamp M. Metrics and evaluation tools for patient engagement in healthcare organization- and system-level DecisionMaking: a systematic review. *Int J Heal Policy Manag*. 2018;7 (10):889-903. <https://doi.org/10.15171/ijhpm.2018.43>.
- Gesulga JM, Berjame A, Moquiala KS, Galido A. Barriers to electronic health record system implementation and information systems resources: a structured review. *Proc Comput Sci*. 2017;124:544-551. <https://doi.org/10.1016/j.procs.2017.12.188>.
- Kruse CS, Stein A, Thomas H, Kaur H. The use of electronic health records to support population health: a systematic review of the literature. *J Med Syst*. 2018;42:1-16. <https://doi.org/10.1007/s10916-018-1075-6>.
- Kavuma M. The usability of electronic medical record systems implemented in sub-Saharan Africa: a literature review of the evidence. *J Med Internet Res*. 2019;6(1):1-11. <https://doi.org/10.2196/humanfactors.9317>.
- Arnhold M, Quade M, Kirch W. Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50years or older. *J Med Internet Res*. 2014;16(4):1-20. <https://doi.org/10.2196/jmir.2968>.
- Sagar K, Saha A. A systematic review of software usability studies. *Int J Inf Technol*. 2017:1-24. <https://doi.org/10.1007/s41870-017-0048-1>.
- Carayon P. Human factors and ergonomics in health care and patient safety. Salvendy G, *Handb Hum Factors Ergon Heal Care Patient Saf*. John Wiley & Sons, Inc.; 2012.
- Greiner L. Usability 101. *netWorker*. 2007;11(2):11. <https://doi.org/10.1145/1268577.1268585>. Print ISBN:9780470528389. Online ISBN:9781118131350.
- Nguyen LTL, Bellucci E. Electronic health records implementation: an evaluation of information system impact and contingency factors. *Int J Med Inform*. 2014;83(11):779-796. <https://doi.org/10.1016/j.ijmedinf.2014.06.011>.
- Tullis T, Albert B. *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*: 2nd ed. USA: Morgan Kaufmann; 2013.
- Burt A, Morgan L, Petrinic T, Young D, Watkinson P. Usability evaluation methods employed to assess information visualisations of electronically stored patient data for clinical use: a protocol for a systematic review. *Syst Rev*. 2017;6(1):4-7. <https://doi.org/10.1186/s13643-017-0544-1>.
- Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):1-6. <https://doi.org/10.1371/journal.pmed.1000097>.
- Veritas Health Innovation. Covidence. [www.covidence.org](http://www.covidence.org).
- Jeng J. Usability assessment of academic digital libraries: effectiveness, efficiency, satisfaction, and learnability. *Libri*. 2005;55(2-3):96-121. <https://doi.org/10.1515/LIBR.2005.96>.
- Kopanitsa G, Zhanna T, Tsvetkova Z. Analysis of metrics for usability evaluation of EHR management systems. *Stud Health Technol Inform*. 2012;180:358-362. <https://doi.org/10.3233/978-1-61499-101-4-358>.
- Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health*. 1998;52(6):377-384. <https://doi.org/10.1136/jech.52.6.377>.
- Stanton NA, Salmon PM, Rafferty LA, Walker GH, Baber C, Jenkins DP. *Human Factors Methods: A Practical Guide for Engineering and Design*. 2nd ed.; UK: Ashgate Publishing Ltd; 2005 <https://www.scopus.com/inward/record.url?eid=2-s2.0-84900228923&partnerID=tZOTx3y1>.
- Abdel-Rahman SM, Breitkreutz ML, Bi C, et al. Design and testing of an EHR-integrated, busulfan pharmacokinetic decision support tool



- for the point-of-care clinician. *Front Pharmacol.* 2016;7(Mar):1-12. <https://doi.org/10.3389/fphar.2016.00065>.
37. Devine EB, Lee C-JJ, Overby CL, et al. Usability evaluation of pharmacogenomics clinical decision support aids and clinical knowledge resources in a computerized provider order entry system: a mixed methods approach. *Int J Med Inform.* 2014;83(7):473-483. <https://doi.org/10.1016/j.ijmedinf.2014.04.008>.
  38. Forsman J, Anani N, Eghdam A, Falkenhav M, Koch S. Integrated information visualization to support decision making for use of antibiotics in intensive care: design and usability evaluation. *Informatics Heal Soc Care.* 2013;38(4):330-353. <https://doi.org/10.3109/17538157.2013.812649>.
  39. Fossum M, Ehnfors M, Fruhling A, Ehrenberg A. An evaluation of the usability of a computerized decision support system for nursing homes. *Appl Clin Inform.* 2011;2(4):420-436. <https://doi.org/10.4338/ACI-2011-07-RA-0043>.
  40. Hirsch JS, Tanenbaum JS, Gorman SL, et al. HARVEST, a longitudinal patient record summarizer. *J Am Med Informatics Assoc.* 2015;22(2):263-274. <https://doi.org/10.1136/amiajnl-2014-002945>.
  41. Hultman G, Marquard J, Arsoniadis E, et al. Usability testing of two ambulatory EHR navigators. *Appl Clin Inform.* 2016;7(2):502-515. <https://doi.org/10.4338/ACI-2015-10-RA-0129>.
  42. King AJ, Cooper GF, Hochheiser H, Clermont G, Visweswaran S. Development and preliminary evaluation of a prototype of a learning electronic medical record system. *AMIA Annu Symp Proc.* 2015;1967-1975.
  43. Koopman RJ, Kochendorfer KM, Moore JL, et al. A diabetes dashboard and physician efficiency and accuracy in accessing data needed for high-quality diabetes care. *Ann Fam Med.* 2011;9(5):398-405. <https://doi.org/10.1370/afm.1286>.
  44. Lin YL, Guerguerian AM, Tomasi J, Laussen P, Trbovich P. Usability of data integration and visualization software for multidisciplinary pediatric intensive care: a human factors approach to assessing technology. *BMC Med Inform Decis Mak.* 2017;17(1):1-19. <https://doi.org/10.1186/s12911-017-0520-7>.
  45. Nair KM, Malaeekeh R, Schabort I, Taenzer P, Radhakrishnan A, Guenter D. A clinical decision support system for chronic pain Management in Primary Care: usability testing and its relevance. *J Innov Heal Informatics.* 2015;22(3):329-332. <https://doi.org/10.14236/jhi.v22i3.149>.
  46. Neri PM, Pollard SE, Volk LA, et al. Usability of a novel clinician interface for genetic results. *J Biomed Inform.* 2012;45(5):950-957. <https://doi.org/10.1016/j.jbi.2012.03.007>.
  47. Su KW, Liu CL. A mobile nursing information system based on human-computer interaction design for improving quality of nursing. *J Med Syst.* 2012;36(3):1139-1153. <https://doi.org/10.1007/s10916-010-9576-y>.
  48. Schall MC Jr, Cullen L, Pennathur P, Chen H, Burrell K, Matthews G. Usability evaluation and implementation of a health information technology dashboard of evidence-based quality indicators. *Comput Inf Nurs.* 2016;35:281-288. <https://doi.org/10.1097/cin.0000000000000325>.
  49. Van Engen-Verheul MM, Peute LWP, de Keizer NF, Peek N, Jaspers MWM. Optimizing the user interface of a data entry module for an electronic patient record for cardiac rehabilitation: a mixed method usability approach. *Int J Med Inform.* 2016;87:15-26. <https://doi.org/10.1016/j.ijmedinf.2015.12.007>.
  50. Wu RC, Orr MS, Chignell M, Straus SE. Usability of a mobile electronic medical record prototype: a verbal protocol analysis. *Informatics Heal Soc Care.* 2008;33(2):139-149. <https://doi.org/10.1080/17538150802127223>.
  51. Rose AF, Schnipper JL, Park ER, Poon EG, Li Q, Middleton B. Using qualitative studies to improve the usability of an EMR. *J Biomed Inform.* 2005;38(1):51-60. <https://doi.org/10.1016/j.jbi.2004.11.006>.
  52. Thompson SM. Remote observation strategies for usability testing. *Inf Technol Libr.* 2003;22(1):22-31.
  53. Belden JL, Koopman RJ, Patil SJ, Lowrance NJ, Petroski GF, Smith JB. Dynamic electronic health record note prototype: seeing more by showing less. *J Am Board Fam Med.* 2017;30(6):691-700. <https://doi.org/10.3122/jabfm.2017.06.170028>.
  54. Wachter SB, Agutter J, Syroid N, Drews F, Weinger MB, Westenskow D. The employment of an interactive design process to develop a pulmonary graphical display. *J Am Med Inf Assoc.* 2003;10(4):363-372. <https://doi.org/10.1197/jamia.M1207.Despite>.
  55. Zhang Z, Walji MF, Patel VL, et al. Functional analysis of interfaces in U.S. military electronic health record system using UFuRT framework. *AMIA Annu Symp Proc.* 2009;2009:730-734. <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed9&NEWS=N&AN=20351949>.
  56. Zhang J, Butler KA. UFuRT: a work-centered framework and process for design and evaluation of information systems. *Proc HCI Int.* 2007;2-6.
  57. Zheng K, Padman R, Johnson MP. User interface optimization for an electronic medical record system. *Stud Health Technol Inform.* 2007;129(Pt 2):1058-1062.
  58. Zheng K, Padman R, Johnson MP, Diamond HS. An interface-driven analysis of user interactions with an electronic health records system. *J Am Med Informatics Assoc.* 2009;16(2):228-237. <https://doi.org/10.1197/jamia.M2852>.
  59. Brown B, Balatsoukas P, Williams R, Sperrin M, Buchan I. Interface design recommendations for computerised clinical audit and feedback: hybrid usability evidence from a research-led system. *Int J Med Inform.* 2016;94:191-206. <https://doi.org/10.1016/j.ijmedinf.2016.07.010>.
  60. Jaspers MWM. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. *Int J Med Inform.* 2009;78(5):340-353. <https://doi.org/10.1016/j.ijmedinf.2008.10.002>.
  61. Lewis C, Wharton C. Cognitive walkthroughs. *Handb Human-Computer Interact.* 1997;717-732. <https://doi.org/10.1016/B978-044481862-1.50096-0>.
  62. Wharton C, Bradford J, Jeffries R, Franzke M. Applying cognitive walkthroughs to more complex user interfaces: experiences, issues, and recommendations. *J Chem Inf Model.* 1992;53:1689-1699. <https://doi.org/10.1017/CBO9781107415324.004>.
  63. Allen M, Currie LM, Bakken S, Patel VL, Cimino JJ. Heuristic evaluation of paper-based web pages: a simplified inspection usability methodology. *J Biomed Inform.* 2006;39(4):412-423. <https://doi.org/10.1016/j.jbi.2005.10.004>.
  64. Iadanza E, Fabbri R, Luschi A, Melillo P, Simonelli F. A collaborative RESTful cloud-based tool for management of chromatic pupillometry in a clinical trial. *Health Technol (Berl).* 2019;10:25-38. <https://doi.org/10.1007/s12553-019-00362-z>.
  65. Laursen SH, Buus AA, Brandt L, Vestergaard P, Hejlesen OK. A decision support tool for healthcare professionals in the management of hyperphosphatemia in hemodialysis. *Stud Health Technol Inform.* 2018;247:810-814. <https://doi.org/10.3233/978-1-61499-852-5-810>.
  66. Nabovati E, Vakili-Arki H, Eslami S, Khajouei R. Usability evaluation of laboratory and radiology information systems integrated into a hospital information system. *J Med Syst.* 2014;38(4):1-7. <https://doi.org/10.1007/s10916-014-0035-z>.
  67. Nielsen J. 10 heuristics for user interface design: article by Jakob Nielsen. *Jakob Nielsen's Alertbox.* 1995:1-2. doi:<https://doi.org/10.1016/B978-0-12-385241-0.00003-8>
  68. Nielsen J. Enhancing the explanatory power of usability heuristics. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Celebrating Interdependence, Boston, Massachusetts, United States.* 1994:152-158. doi:<https://doi.org/10.1145/191666.191729>
  69. Brown B, Balatsoukas P, Williams R, Sperrin M, Buchan I. Multi-method laboratory user evaluation of an actionable clinical performance information system: implications for usability and patient safety. *J Biomed Inform.* 2018;77(May 2017):62-80. <https://doi.org/10.1016/j.jbi.2017.11.008>.

70. Garvin JH, Ducom J, Matheny M, et al. Descriptive usability study of cirrords: clinical decision and workflow support tool for management of patients with cirrhosis. *J Med Internet Res.* 2019;21(7):1-12. <https://doi.org/10.2196/13627>.
71. Kersting C, Weltermann B. Evaluating the feasibility of a software prototype supporting the management of multimorbid seniors: mixed methods study in general practices. *J Med Internet Res.* 2019;21(7):1-10. <https://doi.org/10.2196/12695>.
72. Khairat S, Coleman C, Newlin T, et al. A mixed-methods evaluation framework for electronic health records usability studies. *J Biomed Inform.* 2019;94(March):103175. <https://doi.org/10.1016/j.jbi.2019.103175>.
73. Nouei MT, Kamyad AV, Soroush AR, Ghazalbash S. A comprehensive operating room information system using the Kinect sensors and RFID. *J Clin Monit Comput.* 2015;29(2):251-261. <https://doi.org/10.1007/s10877-014-9591-5>.
74. Pamplin J, Nemeth CP, Serio-Melvin ML, et al. Improving clinician decisions and communication in critical care using novel information technology. *Mil Med.* 2019;00:1-8. <https://doi.org/10.1093/milmed/usz151>.
75. Zhang J, Johnson TR, Patel VL, Paige DL, Kubose T. Using usability heuristics to evaluate patient safety of medical devices. *J Biomed Inform.* 2003;36(1-2):23-30. [https://doi.org/10.1016/S1532-0464\(03\)00060-1](https://doi.org/10.1016/S1532-0464(03)00060-1).
76. Nielsen J. *Usability Inspection Methods: Heuristic Evaluation Chapter 2*. USA: John Wiley & Sons, Inc.; 1994. doi:<http://www.sccc.premiumpdw.com/readings/heuristic-evaluation-nielsen.pdf>
77. Aakre CA, Kitson JE, Li M, Herasevich V. Iterative user interface design for automated sequential organ failure assessment score calculator in sepsis detection. *JMIR Hum Factors.* 2017;4(2):e14. <https://doi.org/10.2196/humanfactors.7567>.
78. Al Ghalayini M, Antoun J, Moacdieh NM. Too much or too little? Investigating the usability of high and low data displays of the same electronic medical record. *Health Informatics J.* 2018;26:88-103. <https://doi.org/10.1177/1460458218813725>.
79. Brown SH, Hardenbrook S, Herrick L, St Onge J, Bailey K, Elkin PL. Usability evaluation of the progress note construction set. *Proc AMIA Symp.* 2001;76-80.
80. Chang P, Hsu CL, Mei Liou Y, Kuo YY, Lan CF. Design and development of interface design principles for complex documentation using PDAs. *CIN - Comput Informatics Nurs.* 2011;29(3):174-183. <https://doi.org/10.1097/NCN.0b013e3181f9db8c>.
81. Chang B, Kanagaraj M, Neely B, Segall N, Huang E. Triangulating methodologies from software, medicine and human factors industries to measure usability and clinical efficacy of medication data visualization in an electronic health record system. *AMIA Jt Summits Transl Sci Proceedings AMIA Jt Summits Transl Sci.* 2017;2017:473-482. <https://doi.org/10.1016/j.foodres.2017.03.015>.
82. Fidler R, Bond R, Finlay D, et al. Human factors approach to evaluate the user interface of physiologic monitoring. *J Electrocardiol.* 2015;48(6):982-987. <https://doi.org/10.1016/j.jelectrocard.2015.08.032>.
83. Gardner CL, Liu F, Fontelo P, Flanagan MC, Hoang A, Burke HB. Assessing the usability by clinicians of VISION: a hierarchical display of patient-collected physiological information to clinicians. *BMC Med Inform Decis Mak.* 2017;17(1):1-9. <https://doi.org/10.1186/s12911-017-0435-3>.
84. Glaser D, Jain S, Kortum P. Benefits of a physician-facing tablet presentation of patient symptom data: comparing paper and electronic formats. *BMC Med Inform Decis Mak.* 2013;13(1):1-8. <https://doi.org/10.1186/1472-6947-13-99>.
85. Graber CJ, Jones MM, Glassman PA, et al. Taking an antibiotic time-out: utilization and usability of a self-stewardship time-out program for renewal of Vancomycin and Piperacillin-Tazobactam. *Hosp Pharm.* 2015;50(11):1011-1024. <https://doi.org/10.1310/hpj5011-1011>.
86. Hirsch O, Szabo E, Keller H, Kramer L, Krones T, Donner-Banzhoff N. Arriba-lib: analyses of user interactions with an electronic library of decision aids on the basis of log data. *Informatics Heal Soc Care.* 2012;37(4):264-276. <https://doi.org/10.3109/17538157.2012.654841>.
87. Hortman PA, Thompson CB. Evaluation of user interface satisfaction of a clinical outcomes database. *CIN - Comput Informatics Nurs.* 2005;23(6):301-307. <https://doi.org/10.1097/00024665-200511000-00004>.
88. Khajouei R, Abbasi R. Evaluating nurses' satisfaction with two nursing information systems. *CIN - Comput Informatics Nurs.* 2017;35(6):307-314. <https://doi.org/10.1097/CIN.0000000000000319>.
89. Lee K, Jung SY, Hwang H, et al. A novel concept for integrating and delivering health information using a comprehensive digital dashboard: an analysis of healthcare professionals' intention to adopt a new system and the trend of its real usage. *Int J Med Inform.* 2017;97:98-108. <https://doi.org/10.1016/j.ijmedinf.2016.10.001>.
90. Mazur LM, Mosaly PR, Moore C, Marks L. Association of the Usability of electronic health records with cognitive workload and performance levels among physicians. *JAMA Netw Open.* 2019;2(4):e191709. <https://doi.org/10.1001/jamanetworkopen.2019.1709>.
91. Rodriguez NJ, Murillo V, Borges JA, Ortiz J, Sands DZ. A usability study of physicians interaction with a paper-based patient record system and a graphical-based electronic patient record system. *Proc AMIA Symp.* 2002;667-671.
92. Séroussi B, Galopin A, Gaouar M, Pereira S, Bouaud J. Using therapeutic circles to visualize guideline-based therapeutic recommendations for patients with multiple chronic conditions: a case study with go-dss on hypertension, type 2 diabetes, and dyslipidemia. *Stud Health Technol Inform.* 2017;245:1148-1152. <https://doi.org/10.3233/978-1-61499-830-3-1148>.
93. Silveira DV, Marcolino MS, Machado EL, et al. Development and evaluation of a mobile decision support system for hypertension management in the primary care setting in Brazil: mixed-methods field study on usability, feasibility, and utility. *J Med Internet Res.* 2019;21(3):1-12. <https://doi.org/10.2196/mhealth.9869>.
94. Tappan JM, Daniels J, Slavin B, Lim J, Brant R, Ansermino JM. Visual cueing with context relevant information for reducing change blindness. *J Clin Monit Comput.* 2009;23(4):223-232. <https://doi.org/10.1007/s10877-009-9186-8>.
95. Zhang Z, Wang B, Ahmed F, et al. The five Ws for information visualization with application to healthcare informatics. *IEEE Trans Vis Comput Graph.* 2013;19(11):1895-1910. <https://doi.org/10.1109/TVCG.2013.89>.
96. Boonstra A, Versluis A, Vos JFJ. Implementing electronic health records in hospitals: a systematic literature review. *BMC Health Serv Res.* 2014;14(1):370.
97. MHRA. Human factors and usability engineering - guidance for medical devices including drug-device combination products. *Med Healthc Prod Regul Agency.* 2017;212(September):1-30. <https://doi.org/10.1016/j.molliq.2015.09.041>.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Wronikowska MW, Malycha J, Morgan LJ, et al. Systematic review of applied usability metrics within usability evaluation methods for hospital electronic healthcare record systems. *J Eval Clin Pract.* 2021;27(6):1403-1416. <https://doi.org/10.1111/jep.13582>