A systematic review of the literature on the evaluation of handoff tools: implications for research and practice

Joanna Abraham, Thomas Kannampallil, Vimla L Patel

ABSTRACT

Objective Given the complexities of the healthcare environment, efforts to develop standardized handoff practices have led to widely varying manifestations of handoff tools. A systematic review of the literature on handoff evaluation studies was performed to investigate the nature, methodological, and theoretical foundations underlying the evaluation of handoff tools and their adequacy and appropriateness in achieving standardization goals.

Method We searched multiple databases for articles evaluating handoff tools published between 1 February 1983 and 15 June 2012. The selected articles were categorized along the following dimensions: handoff tool characteristics, standardization initiatives, methodological framework, and theoretical perspectives underlying the evaluation.

Results Thirty-six articles met our inclusion criteria. Handoff evaluations were conducted primarily on electronic tools (64%), with a more recent focus on electronic medical record-integrated tools (36% since 2008). Most evaluations centered on intra-departmental tools (95%). Evaluation studies were quasi-experimental (42%) or observational (50%), with a major focus on handoff-related outcome measures (94%) using predominantly survey-based tools (70%) with user satisfaction metrics (53%). Most of the studies (81%) based their evaluation on aspects of standardization that included continuity of care and patient safety.

Conclusions The nature, methodological, and theoretical foundations of handoff tool evaluations varied significantly in terms of their quality and rigor, thereby limiting their ability to inform strategic standardization initiatives. Future research should utilize rigorous, multi-method qualitative and quantitative approaches that capture the contextual nuances of handoffs, and evaluate their effect on patient-related outcomes.

BACKGROUND AND SIGNIFICANCE

Handoffs are referred to as a transfer of care involving a transfer of information, responsibility, and authority between clinicians.1 2 Effective handoffs are instrumental for successful patient care management activities such as communication,3 4 coordination,5 and sense making,5 thus contributing to better quality of care.2 6 7 The quality of handoffs, however, is adversely affected by several factors: lack of standardized handoff tools1 3 8; information omissions and inaccuracies3; communication breakdowns related to language, social, and skill issues10–12; lack of training13; and contextual constraints.14 These factors are exacerbated in academic settings by restrictions on resident work hours that increase the frequency of transitions.15 16 To address these challenges, the Joint Commission has suggested the adoption of strategies for standardized and systematic communication.17 18 These standardization efforts have resulted in the widespread development and implementation of handoff strategies, with varying degrees of success.19

Given the complexities of the healthcare environment and the variances in practices across clinical departments, broad guidelines for standardization have led to widely varying manifestations of handoff tools. Regardless of the early successes in handoff implementations,20 broader issues of handoff tool sustainability and consistency, and its ability to maintain seamless transitions across providers and settings, still persist. With these issues unresolved, the success of handoff tools in achieving enhanced patient safety, through standardized information transfer and communication practices, requires further scrutiny.

There are several systematic and integrative reviews on various topical themes related to handoffs including its goals, functions, structure and content, barriers, strategies, and technologies.2 19 21–27 These reviews have advocated for utilizing patient-centered approaches;2 24 formulating best practices,19 20 adopting robust study designs for evaluation,24 designing handoff tools that capitalizes on the content overlap between professions,24 and contextual features.24 While these reviews have recommended robust evaluation studies for the design of contextually-oriented handoff tools, we investigate how the evaluation of handoff tools manifests in actual practice, specifically focusing on the nature, methodological, and theoretical foundations underlying the evaluation of handoff tools in order to explore their adequacy and appropriateness in achieving standardization goals. Such an analysis will provide both practitioners and researchers with an opportunity to reflect on the quality and feasibility, or lack thereof, of current handoff evaluation approaches, as well as on directions for future research.

METHOD

Data sources

We searched the PubMed, Cochrane, and CINAHL databases for research articles published between 1 February 1983 and 15 June 2012. Additionally, we searched for relevant articles on web search engines, from citations of retrieved articles, and from other handoff review articles. The key search
terms used were: handoff(s), hand-off(s), handover(s), shift report(s), shift-report(s), signout(s), sign-out(s), and clinical round(s).

Study selection and inclusion criteria
We included original research articles on the evaluation of handoff tools designed for healthcare practice. Since our focus was on evaluation studies of handoff tools, we excluded articles that examined barriers, the design and development of tools, or the evaluation of process-based strategies related to handoffs. We also excluded reports, perspectives, editorials, and posters.

The searches were limited to English-language articles that appeared in peer-reviewed journals or conferences. We followed the PRISMA guidelines for identifying, screening, selecting, and including papers (see figure 1). Based on the database searches and citation review, the first author (JA) identified a total of 1466 (1438 from databases, 28 from citations) relevant articles. Two authors (JA, TK) then independently assessed the abstracts for inclusion at various stages (see online supplementary table A.4). In cases of disagreement, consensus was reached through discussion. After iterative screening based on title, abstract, and keywords, a total of 88 articles were selected for full-length review. Based on a comprehensive review of these 88 articles, the authors (JA, TK) identified 36 unique articles that met the inclusion criteria.

We evaluated the selected articles using Riesenberg et al’s quality scoring system for comparison across the studies. The scoring system comprised of 12 items, with a maximum score of 16 points. The first item on the study design type (1–3 points) was modified as follows: observational (1 point), non-randomized pre–post (2 points), and randomized controlled trial (3 points). The second item represents the total sample size (0–3 points). All other items are related to reporting and internal validity (0 or 1 point) (see online supplementary table A.3 for itemized scores). Riesenberg’s rating scale was used as it was developed for evaluating the quality of handoff-related studies, accounting for the contextual aspects of handoffs.

Data extraction and synthesis
The first author (JA) extracted data from all the papers (n=36) after a detailed review. The extracted data included study site, study objectives, handoff tool characteristics, study design including the data collection and analysis method(s), outcome measures, results, study implications, theoretical perspectives, funding sources, and study limitations. All abstracted data were then cross-checked and confirmed by the second author (TK) after an independent review. In cases of disagreement, consensus was reached through discussion.

The handoff tool evaluation studies were then categorized along the following dimensions: handoff tool characteristics, standardization initiatives, methodological frameworks, and theoretical perspectives underlying handoff tool evaluations.

Handoff tool characteristics were categorized as follows: type of tool, user, and use. Based on tool type, we classified the handoff tools into: paper-based and electronic. Electronic tools were evaluated to identify whether they were integrated into an electronic medical record (EMR) system or were electronic standalone tools (electronic forms, intranet-based tools). Based on the user dimension, we categorized tools into physician and/or nurse handoff tools. Additionally, we identified the nature of handoff tool use: intra-departmental (eg, nurse shift-report, resident sign-out) or inter-departmental handoff (eg, between an emergency physician and a hospitalist).

Classification based on the standardization initiatives was derived from the National Patient Safety Goal 2E, which called on hospitals to ‘implement a standardized approach to handoff communications, including an opportunity to ask and respond to questions.’ Care continuity and patient safety have been recognized as two primary goals that can be achieved through handoff standardization.

Methodological frameworks were categorized based on study design, data collection methods, and outcome measures. Study designs were categorized as randomized controlled trials (RCT), non-randomized pre–post designs, and observational studies. In RCTs, participants were randomly allocated to intervention or control groups. In non-randomized pre–post designs, outcome measures on the success (or efficacy) of an intervention were measured on participants before and after its implementation. Observational studies involved the evaluation of an existing handoff tool. Articles were also classified based on five commonly used data collection methods: survey/questionnaire, interview, observation, log-file extraction, and audit review of handoff documentation/patient data. Two reviewers (JA, TK) coded the articles with 100% agreement on the handoff characteristics, standardization initiatives, and methodology.

The outcome measures of each study were classified based on the three human factors principles related to the efficacy of tool use: effectiveness, efficiency, and satisfaction. Effectiveness refers to the accuracy and completeness with which users achieve certain goals. Indicators of effectiveness include usefulness and safety in completing a task. Efficiency is the relationship between (a) the accuracy and completeness with which users achieve certain goals using the tool and (b) the resources expended in achieving them. Indicators of efficiency include aspects related to productivity, cost (eg, time), and learnability. Satisfaction refers to the users’ comfort with and attitudes toward handoff tool usage. Indicators of satisfaction include perceived effectiveness and perceived efficiency were used. This scheme did not use mutually exclusive categories (ie, one article could have more than one efficacy category). Two reviewers (JA, TK) employed the scheme to categorize the articles with 98% agreement. Disagreements were resolved through discussion.

Theoretical perspectives underlying the handoff evaluation were classified based on conceptual themes described in prior research. Although handoffs are traditionally conceptualized as an information transfer activity, researchers have reported on other theoretical foundations underlying handoff communication activity and their functional goals. Cheung et al discussed four conceptual handoff aspects: information processing, supporting the transfer of data through a noisy communication channel; stereotypical narratives, allowing the creation of a narrative and highlighting the deviations in activities; social interaction, supporting co-construction of meaning through shared mental models; and resilience, allowing for cross-checking of assumptions with a fresh perspective. This framework was extended by Patterson et al who introduced three additional categories: accountability, supporting transfer of responsibility.
and authority; distributed cognition, for describing how cognition is distributed across human minds, external cognitive artifacts, and groups of people; and cultural norms, addressing how group values and norms are appropriated, negotiated, and maintained over time. This scheme also did not have mutually exclusive categories and two reviewers used the framework to categorize the articles with 98% agreement. A summary of the analysis dimensions and relevant measures can be found in tables 1 and 2.

RESULTS
We identified 36 (n=36) articles on the evaluation of handoff tools (also see online supplementary tables A.9, A.10, and A.11). A majority of studies were conducted in the USA (n=26), followed by the UK (n=5), Australia (n=3), Canada (n=1), and Ireland (n=1), and often involved multiple clinical departments (47%, n=17) (also see online supplementary tables A.5 and A.6). We have provided a setting-specific analysis (see online supplementary figures A.1 and A.2, and a detailed note on the effect of settings on handoff tools in the online supplementary appendix). Quality scores ranged from 4 to 14 (mean 8.6, SD 2.67) (see online supplementary tables A.1, A.2, A.3, A.9, A.10, and A.11; also refer to the sub-section on the ‘Quality of Handoff Tool Evaluation Studies’ in the online supplementary appendix).

Handoff tool characteristics
Handoff tool characteristics were classified along three dimensions: tool type (paper, electronic), users (physician, nurse), and nature of use (intra-, inter-departmental).

Tool type
Of the handoff tools, 36% (n=13) were paper-based, 27% (n=10) were electronic standalone tools, and 36% (n=13) were EMR-integrated. Approximately 70% of these articles (n=25) were published after 2008 and 36% (n=9)
of these articles described an EMR-integrated handoff tool. Of the articles, 47% (n=17) reported on handoff tools exclusively for physicians,4 57–59 61 63 64 66 34% (n=12) on handoff tools exclusively for nurses,35 36 42 44–46 50 53 55 56 65 68 and 20% (n=7) on tools to support both physicians and nurses.20 40 42 58 60 62 67. A significant percentage of handoff tools used by physicians were either EMR-integrated4 57 59 61 63 64 66 67 (42%, n=7 of 17) or electronic-based48 49 51 54 (24%, n=4 of 17). Similarly for nurses, 50% (n=6 of 12) of handoff tools were either electronic-based (34%, n=4 of 12) or EMR-integrated (16%, n=2 of 12).67 68 Although there was a higher percentage of EMR-integration in physician handoff tools, there was no

Table 1  Categories of data extracted from each article for further synthesis and analysis

<table>
<thead>
<tr>
<th>Data category</th>
<th>Description of the category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study site</td>
<td>Three aspects of the study site were recorded</td>
</tr>
<tr>
<td>Location: country in which the study was conducted</td>
<td></td>
</tr>
<tr>
<td>Type of hospital: teaching, non-teaching</td>
<td></td>
</tr>
<tr>
<td>Type of unit: critical care, non-critical care</td>
<td></td>
</tr>
<tr>
<td>Study objectives</td>
<td>Specific study objectives, including the purpose of the study</td>
</tr>
<tr>
<td>Handoff tool characteristics</td>
<td>Type: paper, electronic (also, if EMR-integrated)</td>
</tr>
<tr>
<td>User: physician (attending, resident, fellow), nurses, support personnel</td>
<td></td>
</tr>
<tr>
<td>Nature of handoff supported: intra-departmental, inter-departmental</td>
<td></td>
</tr>
<tr>
<td>Study design</td>
<td>Observational: studying the effect of an existing tool</td>
</tr>
<tr>
<td>Non-randomized pre–post: studying the effect of an intervention using a pre (prior to the intervention)–post evaluation</td>
<td></td>
</tr>
<tr>
<td>Randomized control trials: studying the effect of an intervention through random allocation of participants to control and intervention groups based on selected outcome measures</td>
<td></td>
</tr>
<tr>
<td>Data collection method(s)*</td>
<td>One (or more) of the following data collection methods were used for classification: survey/questionnaire, interview, observation, log-file analysis, and audit review of handoff/patient data</td>
</tr>
<tr>
<td>Outcome measures*</td>
<td>Effectiveness: aspects related to usefulness and safety in completing a task</td>
</tr>
<tr>
<td>Efficiency: aspects related to productivity, cost (in terms of time), and learnability (in terms of use)</td>
<td></td>
</tr>
<tr>
<td>User satisfaction: aspects related to perceived effectiveness and perceived efficiency</td>
<td></td>
</tr>
<tr>
<td>Standardization initiatives</td>
<td>Continuity of care: support for interactive communications, up-to-date and accurate information, an opportunity to review relevant historical data and limit interruptions</td>
</tr>
<tr>
<td>Patient safety: process for verification, patient harm, adverse patient events, and hospital-acquired conditions</td>
<td></td>
</tr>
<tr>
<td>Theoretical underpinning(s)*</td>
<td>Information processing: transfer of data through a noisy communication channel</td>
</tr>
<tr>
<td>Stereotypical narratives: creation of a narrative and highlighting the deviations in activities</td>
<td></td>
</tr>
<tr>
<td>Social interaction: co-construction of meaning through shared mental models</td>
<td></td>
</tr>
<tr>
<td>Resilience: cross-checking of assumptions with a fresh perspective</td>
<td></td>
</tr>
<tr>
<td>Accountability: transfer of responsibility and authority</td>
<td></td>
</tr>
<tr>
<td>Distributed cognition: how cognition is distributed across human minds, external cognitive artifacts, and groups of people</td>
<td></td>
</tr>
<tr>
<td>Cultural norms: how group values and norms are appropriated, negotiated, and maintained over time</td>
<td></td>
</tr>
<tr>
<td>Funding sources</td>
<td>Specific funding sources that were mentioned in the paper</td>
</tr>
</tbody>
</table>

*Subcategories were not mutually exclusive.
†Presence or absence of either one of these standardization criteria.
EMR, electronic medical record.

Key aspects of paper-based tools for supporting handoffs included single-page organization and tabular or checklist-based templates with basic patient information such as demographic data, reason for admission, medications, to-do lists, IV fluids, oxygen levels, tube feeds, and monitor settings. EMR-integrated tools were characterized by features for automated download of handoff information with minimal manual entry, interface into other ancillary clinical information systems, automatic population information, alerting capabilities, and support for clinical handoff workflow including pre-turnover, handoff, and post-turnover phases.

Table 2  Measures used for categorizing outcome measures, standardization initiatives, and theoretical underpinnings

<table>
<thead>
<tr>
<th>Data category</th>
<th>Measures used for each category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome measures</td>
<td>Effectiveness: information gaps and errors</td>
</tr>
<tr>
<td>Efficiency: handoff duration, improvement in learnability (in terms of use), and usability</td>
<td></td>
</tr>
<tr>
<td>User satisfaction: perceived effectiveness and perceived efficiency</td>
<td></td>
</tr>
<tr>
<td>Standardization initiatives</td>
<td>Continuity of care: up-to-date and accurate information, missed information, tasks, and interruptions</td>
</tr>
<tr>
<td>Patient safety: medical errors, adverse patient events, hospital-acquired conditions</td>
<td></td>
</tr>
<tr>
<td>Theoretical underpinning(s)</td>
<td>Information processing: accuracy and completeness of information transferred</td>
</tr>
<tr>
<td>Stereotypical narratives: consistent synopsis of patient care plan and narrative</td>
<td></td>
</tr>
<tr>
<td>Social interaction: shared mental models, conversational grounding, team climate</td>
<td></td>
</tr>
<tr>
<td>Resilience: support for detection of inaccurate information and inconsistencies in information, support for error recovery</td>
<td></td>
</tr>
<tr>
<td>Accountability: task completion, inappropriate tasks transferred, dropped/missed patients</td>
<td></td>
</tr>
<tr>
<td>Distributed cognition: effective coordination of care, technical errors</td>
<td></td>
</tr>
<tr>
<td>Cultural norms: educational interventions, handoff policies and procedures, changes in priorities, values, acceptable behaviors</td>
<td></td>
</tr>
</tbody>
</table>

statistically significant association between the type of user (physician, nurse) and type of tool (paper, electronic) ($\chi^2(1) = 0.0423$, $p = 0.837$) (figure 2).

Use
Evaluation of tools that supported inter-departmental handoffs was reported in only two articles (5%, n=2).38 44 The remaining articles were on tools for intra-departmental handoffs (see online supplementary table A.7). Additionally, we analyzed hospital policy on handoff tool use as being either voluntary or mandatory (see online supplementary table A.8).

Standardization initiatives
Most articles (81%, n=29) utilized at least one of the considered standardization measures.4 20 35–39 41–47 50–54 56–58 61–66 68 Figure 3 shows the distribution of the articles across paper-based, electronic (standalone), and EMR-integrated tools. Additionally, approximately 66% of the standardization initiatives were on electronic or EMR-integrated tools.

Study design
Of the articles, 8% (n=3) utilized RCT designs,39 63 64 42% (n=15) used a non-randomized pre-post design,35 37 38 40 41 45 47 49 51 52 54 57 58 62 66 and 50% (n=18) used observational approaches.4 20 36 42–44 46 48 50 53 55 56 59–61 65 67 68

We also compared the type of handoff tools (ie, paper vs electronic) that were used in the control (pre-) and experimental (post-) conditions in pre-post and RCT studies. In several studies the tool type used as the control (pre-) was unspecified or was a combination of approaches (eg, verbal and partially written) and was classified as ‘other’ (n=10). We found that the most common pre-post pair was paper (pre)—electronic (post): 27% (n=5 of 19).

Figure 4 shows the frequency of tools that were used in the pre- and post-conditions (pre-post and RCT-based studies only). The x axis shows the pre- and post-intervention tools (first line) and the experimental design (pre-post or RCT). The y axis shows the number of articles in each category. The most preferred control was paper (42%, n=8 of 19) and the most preferred intervention was electronic (both standalone and EMR-integrated) (58%, n=11 of 19).

Data collection approaches
Of the articles, 50% (n=18) adopted a multi-method approach for data collection.4 20 38 39 43 44 47–49 53 56 37 59 60 62–64 66 For example, Palma et al62 utilized surveys and analyses of EMR log data to systematically track handoff tool usage, its effects on workflow, and provider satisfaction. The most common method for data collection was surveys/questionnaires (70%, n=25).4 20 35 36 38 39 41 44 47–49 53 56 37 59 60 62–64 66 followed by audits and review of handoff documents (42%, n=15).4 20 37 39 40 42–44 47–49 53 60 62–64 66 67 interviews (20%, n=7),4 38 48 49 53 56 63 64 log-file analysis (12%, n=4),20 38 47 52 57 66 and observations (14%, n=5).

Outcome measures
We found that a majority of the studies (94%, n=34) used handoff activity-related outcome measures: information gaps,65 39 49 handover duration, number of patients handed off, interruptions,64 care quality, frequency of tool use,20 handoff efficiency, and length of shift-report.50 Only two articles utilized
patient-related outcome measures.40 53 Ryan et al40 evaluated the efficacy of an electronic handover system to support communication between surgical teams using patient length of stay as the outcome measure. Similarly, Roberts et al53 used patient fall rate as a measure to determine the impact of an interdepartmental handoff tool on quality of nurse communication.

As previously described, we also categorized the outcome measures in terms of three efficacy categories: *effectiveness*, *efficiency*, and *satisfaction*32 34 69 (see figure 5).

**Effectiveness**
A majority of the articles (59%, n=21) evaluated the effectiveness of the handoff tool. Measures related to effectiveness included analysis of information gaps,4 37 42 49 52 errors,38 64 information recall accuracy, and fall rates.53 These articles were spread across 10 (48%) studies on paper-based tools,35 37–39 42–47 five (24%) on electronic standalone tools,68 52–54 56 and six (29%) on EMR-integrated tools.4 57 62–64 67

**Efficiency**
In terms of efficiency, 34% (n=12) of the articles reported on a variety of measures, including handover duration, number of patients handed off, interruptions, pre-rounding time and duty hours,66 patterns of tool usage,60 61 length of stay for patients,40 monthly rate of use of the sign-out tool, and frequency of use.59 Of these 12 articles, four (33%) were on electronic-standalone tools,20 49 55 56 three (25%) on paper-based,35 40 44 and five (42%) on EMR-integrated tools.59–63

**User satisfaction**
User satisfaction aspects were evaluated in 53% (n=19) of the articles. Measures related to tool satisfaction that were commonly examined included perceptions of care quality, perceived efficiency, perceived information omissions,20 35 50 34 provider-reported sign-out accuracy, staff satisfaction and workflow,53 38 62 65 handoff-related patient safety, quality, and efficiency.57 User satisfaction was primarily measured in EMR-based tools (48%, n=9 of 19),4 57 58 62 64–66 68 followed by electronic standalone tools (27%, n=6 of 19),20 48 50 51 54 55 and paper-based tools (22%, n=4 of 19).36 38 41 46

**Theoretical perspectives**
One or more theoretical perspectives were used to evaluate handoff tool functionalities in 34% of the articles (n=12).20 35 38 41 45 50 54 59 65–68 Similar to prior research,5 we found that *information processing* was the primary and most used theoretical framing in a majority of the articles (89%, n=32).4 20 35–39 41–46 48–60 62 63 65 67 68 For example, the underlying feature supported by the information processing perspective provided by Flanagan et al4 included the effectiveness of retrieving and transferring patient-related information using a standardized structured tool.

Other theoretical perspectives included distributed cognition (25%, n=9),4 38 40 45 50 59 61 67 68 accountability (12%, n=4),34 54 63 66 cultural norms (5%, n=2),35 66 and social interaction (2%, n=1).65 For example, Frank et al51 described a sign-out system that supported easy retrieval of patient information from a clinical system with better integration of physician work activities within the clinical workflow. Studies utilizing the underpinnings of accountability described tracking of missed patients during rounds, and dropped and inappropriate handing-off of tasks.41 34 66 Cultural norms and social interaction framings were used to describe the supporting preparatory tasks for avoiding delays in shift starting times and compliance with duty hours,35 and the ability to ask and respond to questions,4 respectively.

**Research support**
We found that only 14% (n=5) of articles reported on any form of research/funding support. Two were internally supported research conducted at Veterans Affairs (VA) hospitals,4 37 and one each were supported by a McNeil award (American Association of Family Practitioners),51 a National Research Service Award,63 and an NLM training grant.67

**DISCUSSION**
We found that a majority of the evaluation studies were conducted in the USA (73%) on physician-based (67%, included physician users), electronic handoff tools (64%, both electronic and EMR-integrated). The focus on EMR-integrated tools has been prominent since 2008 (36%), possibly driven by the federally mandated push toward standardized, integrated electronic healthcare applications.70 A preponderance of articles used observational (50%) or simple non-randomized pre-post comparison-based (42%) quasi-experimental study designs with survey-based data collection methods (70%). The outcome measures were primarily handoff-related (94%) with a focus on user-satisfaction-based (53%) or effectiveness-based (59%) measures. Standardization efforts related to care continuity or patient safety were reflected in 81% of the articles. Unfortunately, there seems to be limited evidence of extramural funding (14% of articles) for handoff studies. We discuss the implications of our findings on standardization efforts.

**Handoff tool type**
A socio-technically driven change was the decisive shift toward EMR-integrated handoff tools. This transformation is not only driven by federal mandates, but also by improvements in health information technology through improvements in data persistence and consistency. Integrated environments accelerate standardization initiatives by seamlessly retrieving data from multiple sources and presenting it in a shareable electronic format. Nevertheless, based on our review, two aspects of EMR-integrated tools require further research attention. First,
most EMR-integrated handoff tools were built for specific departments. Given the contextual nature of handoffs, the interoperability of these tools in other similar settings would require further evaluation. Second, the theoretical focus of evaluation studies was on the handoff tools’ information processing capabilities, with limited emphasis on its resilience capabilities, and ability to effectively highlight deviations from expected practice through the use of stereotypical narratives. There was a significant variation between tools evaluated across study settings (also see online supplementary table A.6). Nevertheless, we found that several studies were conducted in settings that had significant patient turnover (eg, surgery), and on handoff tools that utilized a patient-problem oriented content framework (see section on the ‘Role of Context in Handoff Tool Evaluation Studies’ in the online supplementary appendix).

Evaluation studies
Quasi-experimental and observational studies were predominant, even with their purported limitations regarding scientific rigor. While the methodological rigor of some studies may be questioned, the value of quasi-experimental/observational studies of handoff tools cannot be discounted, as they provide a platform to identify and establish appropriate metrics for handoff tool evaluation, insights on the socio-cultural factors that affect handoff, and potentially relevant process-related measures. Additionally, based on the analysis of the quality rating scores, we found that RCT studies had higher overall scores than observational studies. The most significant difference in terms of the quality scoring was related to the sample size (mean_{observational} 1.1, mean_{pre-post} 2.1, mean_{RCT} 2.3), with observational studies having significantly smaller sample sizes (see the online supplementary appendix section on the ‘Quality of Handoff Tool Evaluation Studies,’ and tables A.1, A.2, and A.3).

While RCTs are preferred due to their scientific rigor, it is often challenging to run such studies in clinical settings due to a variety of factors: challenges of incorporating a new handoff intervention without disrupting the existing clinical workflow; recruiting a random sample of clinicians, especially among trainees, who transition on a monthly basis; lack of universally accepted outcome and evaluation measures for handoffs; difficulty in linking to patient-related data for evaluation; and a high rate of susceptibility to patient or treatment selection biases. Alternative approaches such as practice-based evidence designs or regression discontinuity designs are scientifically rigorous approaches. Two studies published in The New England Journal of Medicine argued that observational studies and RCTs produce similar results. Given that RCTs may not be feasible in most settings for evaluating handoff outcomes, alternatives are certainly worth exploring.

However, the limitations of current handoff evaluation studies are worth mentioning. For example, there was a widespread use of surveys or questionnaires for evaluation studies—these surveys were often not validated for reliability, had a very small sample of respondents (n<20), relied on users’ recall (eg, ‘how many patients did you miss’), and included no contextual information. Nature of outcome measures
Lack of attention to patient-related outcomes has been raised as a source of concern by a number of researchers. Most studies (94%) used handoff-related outcome measures for evaluating tool efficacy and usability. For example, information gap (eg, missing or incorrect information) was commonly used to evaluate handoff quality. However, this measure is rarely tied to patient outcome data (eg, procedural or treatment delays, adverse patient events, or re-hospitalization rate), and as a result, it is difficult to ascertain the effect of such handoff-related variables on care continuity and patient safety. In other words, such measures only provide localized metrics for ascertaining the efficacy of handoffs and are often unable to provide holistic perspectives regarding their impact on overall quality of care. The narrow emphasis of localized metrics can be mitigated through the utilization of broader process-oriented outcome measures. Such measures would help in developing trace-based metrics predicated on situating handoffs within the greater context of overall clinical workflow, and in identifying the impact of interdependencies in the overall handoff process.

Limitations
We report on several limitations in our review. First, we considered only English language articles. However, our search strategy, progressing from general handoff content to specific handoff evaluation studies, helped in streamlining this process. Second, recent research has highlighted the importance of the contextual aspects of handoffs. In our analysis, we aggregated results across multiple settings due to the small number of articles evaluating handoff tools. Nevertheless, even with this smaller sample, we have provided a detailed analysis on the effect of context on handoff tools (see the online supplementary section on the ‘Role of Context in Handoff Tool Evaluation Studies’ in the appendix).

CONCLUSION
Using a systematic review of the literature on handoff tool evaluation, we found that most studies were predicated on developing preliminary perspectives on handoffs. While most handoff tools had certain general characteristics, they varied widely in their structural organization (eg, free-text or problem-oriented) and implementation in practice (eg, single to multiple units). The importance of electronic handoff tools has been reflected in the significant number of handoff tools that were evaluated after 2008. Additionally, given the known patient safety issues with handoffs, it is remarkable that very little research in this area has been funded. Our evaluation provides insights into the challenges of handoff tool evaluation and an ample backdrop for rethinking strategies for future empirical studies.

Contributors IA and TK conceived the study, collected the data, and performed the analysis. All authors participated in the interpretation of data, helped to draft the article or revise it critically for important intellectual content, and gave final approval of the version to be published.

Funding The study was supported by a grant from the James S McDonnell Foundation to VLP to investigate cognitive complexity and errors in critical care (No. 220020152).

Competing interests All authors have completed the Unified Competing Interests form at http://www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: all authors were supported by funding from the James S McDonnell Foundation; no financial relationships with any other organizations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES


Jt Comm J Qual Transform 2007;91:91. 00.


70 Patient Protection and Affordable Care Act, 2010.