

Online Supplementary Figure 1

Phase 1 – 4 Exemplars from Review Article by Dunn Lopez, Gephart, et al., 2017

Phase 1 Exemplar: Framing the Review Purpose and Aims

To accomplish Phase 1, we identified the overarching purpose was to report the state of the science for clinical decision support for bedside nursing in hospitals. We framed this purpose to focus on three specific aims: 1) To identify methodologic features used in clinical decision support research, 2) To describe the study characteristics of clinical decision support, and 3) To determine study outcomes related to clinical decision support.

Phase 2 Exemplar: Planning and Developing the Extraction Tool

Because our overall objective was to report the state of the science of Clinical Decision Support (CDS) targeting acute care nurses, the first aim was to identify the methodological features of the primary studies. The data categories we extracted for this aim included: 1) sample size, 2) theory used, and 3) study purpose (Step 1). Each of these data categories had a different data type (Step 2) that required a different survey item (Step 3) response option (Step 4) in the survey software.

Example of numerical data type, survey item, and response option. We considered sample size to be numerical data because it is reported as a number. We used a simple survey item: “Report sample size” to extract this data. Although the response option for numerical data may differ among software, in Google Forms, numerical data is reported as a short answer response option with a rule for a whole number between 1 and 1,000.

Example of categorical data type, survey item, and response option. We considered theories to be a categorical data type because we could predict many of the theories used in the primary

studies a priori. Using our expertise in the area, we identified three theories we expected to be used by the primary studies' authors (distributed cognition theory, technology acceptance model, and diffusion of innovation theory). Our survey item was: "Theoretical underpinnings (must be specifically stated by authors, can choose >1)." In Google Forms the multiple options response type that allowed data extractors to choose >1 category because some studies are guided by a combination of theories (in contrast the response type multiple choice only allows the extractor to choose one option). Importantly for categorical data types, we also included options to choose "not included", to allow us to extract studies that did not include a theory and "other," using a short answer format that allowed the data extractors to type in theories that we did not identify a priori.

Example of narrative data type, survey item, and response option. We considered study purpose to be narrative data type because given the wide variety of purposes in our sample of primary studies it would be time prohibitive to develop data categories a priori. Our survey item was "Study Purpose (author's statement of the study purpose)." In the survey software, we chose the short answer response option that allowed the data extractor(s) to type in a short sentence that identified the study's purpose.

Phase 3 Exemplar: Piloting the Extraction Tool

In our development of the extraction tool, we created the data category of "Theory used" and included three theories (e.g., distributed cognition theory, technology acceptance model, and diffusion of innovation theory) as our initial response options. However, during the pilot testing

step, we added sociotechnical theory to the list of response options as we found studies that used this theory in the subsample of articles.

After we felt confident that we identified the majority of theories, we also included a short answer “other theory” as a response option for the “theory used” data category. Adding the “other theory” as a response option proved to be very valuable as we were surprised to find seven additional theories that were used to guide the research in our sample of articles. We also included a “theory not included” response option. This added response option also proved to be very valuable to our extraction process, as the majority of the studies in our sample did not include a theory.

Phase 4 Exemplar: Data Extraction and Synthesis

In our review, four investigators, including one PhD student, extracted data. Initially we used a multiple coder approach with a subset of articles to develop consensus on the definition of each category that could be applied consistently and then assessed interrater agreement in rounds.

Although the kappa statistic provides a stronger assessment of interrater reliability (McHugh, 2012), in the exemplar we measured interrater agreement. We established 89% agreement after 3-4 rounds and then coded independently for the remaining articles.

Once the extraction process was complete, the next step was to synthesize. To begin, we verified that our spreadsheet contained 28 rows of data for each article included in our sample and 23 columns of data that corresponded with the 23 survey items in our tool. Next, we performed simple counts for the categorical data by sorting. In addition to the ease of performing counts, sorting the data in spreadsheet software maintains connections to the other study data categories. We found this very helpful when examining study outcomes. For example, we counted that only

3 studies found positive patient outcomes associated with use of the CDS. On initial examination, this finding seems to indicate that CDS targeting nurse decision making were ineffective in improving patient outcomes. However, because we could sort the findings while maintaining connections to the other data categories, we found that only 5 of the studies included patient outcomes and of these only 4 of the studies had designs and analytic procedures that would allow one to detect changes in patient outcomes. This led to a very different conclusion that 3 of 4 studies with design and analytic procedures to detect change improve patient outcomes, showing the promise of future nurses targeting CDS interventions as well as the need to include patient outcomes in the study designs. Another advantage of sorting using a spreadsheet is that the primary study's author(s), study year, and other data categories stay connected to the data category being analyzed. This allowed us to very efficiently cite studies while viewing the sorted categories as we wrote the results section.

Our final step was to analyze the narrative data using conventional content analysis. For example, because of the disparate ways the study authors phrase the purpose of their studies, we entered the verbatim purpose statement for each article. Then, one investigator on our team grouped study purposes with similar meaning and assigned four descriptive "labels" to the purpose types: 1) predesign needs assessment, 2) redeployment feasibility, 3) comparison to paper tools, and 4) evaluation. Labels were reviewed and modified with the entire investigative team whereby a definition was developed.

Supplemental Materials References

Dunn Lopez, K., Gephart, S. M., Raszewski, R., Sousa, V., Shehorn, L. E., & Abraham, J. (2017). Integrative review of clinical decision support for

registered nurses in acute care settings. *Journal of the American Medical Informatics Association*, 24(2), 441-450.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276-282. doi:10.11613/bm.2012.031