Where do we stand on the measurement of research utilization?*†

Carole A. Estabrooks
Associate Professor
&
Lars Wallin
Postdoctoral Fellow

Knowledge Utilization Studies Program
Faculty of Nursing, University of Alberta
Edmonton, Alberta, CANADA

July, 2004

* Paper prepared for the 4th Annual Knowledge Utilization Colloquia (KU04), Belfast, Northern Ireland
† With the assistance and feedback of Merry Jo Levers, Shannon Scott-Findlay, Joanne Profetto-McGrath, and Dwight Harley
Where do we stand on the measurement of research utilization?

Nurses of many different titles provide the majority of healthcare globally. In Canada alone nearly 300,000 regulated individuals provided nursing care including 230,957 Registered Nurses\(^1\), 5132 Registered Psychiatric Nurses\(^2\), and 60,123 Licensed Practical Nurses\(^3\) [2002 data]. In addition, many thousands of unregulated workers (e.g., nursing aides, patient care aids, etc.) also deliver nursing care. Increasingly, we read reports documenting how the organization of nursing services influences patient and system outcomes. Lower mortality, decreased length of stay, reduced complications, fewer readmissions in acute care facilities, and decreased restraint use, contractures, and pressure ulcers in long-term care settings are examples of outcomes that we now know are affected by nursing service delivery\(^4\)-\(^9\). To state that improving care improves patient and system outcomes is obvious. What is less obvious are the mechanisms by which this is accomplished. Our team believes that one important approach to improving care is to increase the use of research at the point of care delivery – to systematically improve the care provided by individual delivering nursing services – even incrementally – will have a dramatic effect on patient and system outcomes.

In the past thirty years, we have gained considerable understanding about the concept of research utilization, and in related areas such as knowledge utilization and innovation diffusion\(^8\)-\(^{22}\). Despite these gains in the theoretical base, measuring research utilization validly and reliably has not been adequately addressed, and remains a persistent and unresolved problem in the field\(^{23},^{24},^{25}\). While some existing measures\(^{26}-^{28}\) have been used more than once, the most common approach to measuring research utilization is to develop one’s own measure without much attention to standard psychometric methods or to post-use evaluation of the measure’s performance. Presently, one of the most serious limitations to furthering the study of research utilization is the lack of a sensitive, valid, and reliable tool for its measurement. Developing a robust measure of research utilization would enable: (1) comparisons among groups, settings and/or institutions, (2) an important outcome measure for intervention studies, and (3) accurate assessment of the impact of research use on outcomes.

Review of the literature

Research utilization is the implementation of research-based knowledge (science) in practice. We believe that research utilization is one indicator of an optimum practice environment, an environment that leads to improved patient outcomes. We have evidence to suggest that several additional factors within the health care system also affect outcomes. These include health professionals’ educational level, organizational climate, leadership, nurse-patient relationships, staff well-being and quality improvement strategies\(^{29},^{3,30}-^{32}\). The relative importance of the research utilization indicator remains to be evaluated, but we believe it is a factor whose importance will become more apparent as the body of research examining it grows.

Evidence-based medicine\(^4\) has been described as having five steps: (1) converting information needs into answerable questions, (2) tracking down the best evidence for answering the questions, (3) critically appraising that evidence, (4) implementing the results of this appraisal in clinical practice, and (5) evaluating care performance\(^{33}\). This generic model reflects the

---

\(^1\) There is as Larsen\(^41\) has described an unresolved “terminological tangle” in knowledge utilization and its related and associated fields (e.g., knowledge translation, knowledge mobilization, innovation diffusion, technology transfer, research dissemination, research utilization, evidence-based medicine, evidence-based decision-making, etc.). Throughout this document we use the term research utilization. In the health literature emanating from medicine, evidence-based medicine or evidence-based decision-making is often the term used. In nursing, the term research utilization is commonly used, although since the 1990’s the term evidence-based practice is increasingly seen.
commonly accepted ingredients of research utilization, as it is generally understood among health care providers. Although it can be challenged on such aspects as its assumption of linearity in the change process, the model has evolved to where it is the cornerstone of the evidence-based movement in modern health care. The literature on evidence based care, as well as, research utilization focuses heavily on the critical appraisal phase, with the implementation phase less well understood. Despite the need for expertise and resources in the appraisal process, we and others argue that the major stumbling block to achieving more research based clinical practice lies in the research implementation (use) phase.

Types of Research Use

Before entering the literature on measuring research use, it is essential to understand the basic typology. To develop a measure of research utilization, it is vital to clarify what concept(s) are being addressed if the instrument is to be valid, reliable and generate data. Research utilization, a subset of knowledge utilization, is a multifaceted concept encompassing different forms of use of research based knowledge. Several conceptualizations have been proposed. Larsen classified knowledge utilization as instrumental and conceptual. Rich and Weiss also discussed these two kinds of knowledge utilization. Instrumental research utilization is a direct use of research knowledge. It is the concrete application of research in clinical practice; either in making specific decisions about patient care, or as the knowledge guiding specific interventions. Conceptual research utilization is the cognitive use of research where the research may be used to change opinion or mind set about a specific practice area, but not necessarily particular actions. This indirect application of research is believed to occur more often in practice than instrumental use, but in a less tangible way. Symbolic (or persuasive) research utilization addresses the use of research knowledge as a political tool in order to influence or legitimate policies and decisions. Stetler introduced instrumental and conceptual use into the nursing literature. Estabrooks empirically verified instrumental, conceptual and symbolic research use by nurses. In addition, she demonstrated that a fourth form of research utilization – overall – could be conceptualized and measured. Subsequent to the original publication, her group has been able to replicate this typology several times.

Although this framework appears in the literature, the tool that we develop will not necessarily reflect these conceptualizations of research use.

Existing research utilization measures in nursing

We conducted a systematic review for instruments/surveys used to measure research utilization in the nursing field. Articles that reported the use or development of an instrument and that included words or phrases such as “use”, “implement”, “utilize” or “change practice” in conjunction with “research findings” in survey items or questions were included. We located two multi-item instruments published in 16 papers. These two most commonly used instruments were the Nurses Practice Questionnaire (NPQ) and the Research Utilization Questionnaire (RUQ). Brett developed the NPQ based on Roger’s innovation diffusion theory. It consists of brief descriptions of 14 specific nursing practice innovations. Seven questions measuring the nurse’s stage of innovation adoption are posed for each of the nursing practice innovations. Champion and Leach developed the RUQ. It has four subscales, of which one is

---

8 Have you read about this nursing practice? Have you heard about this nursing practice? Have you observed this practice in use? Have you learned about this practice from any other source? If appropriate to the practice setting, do you believe a nurse should use this nursing practice? How often do you use this nursing practice? Are you aware of any policies concerning this nursing practice in your workplace?
research utilization. The research utilization subscale has 10 items **measuring the degree to which a nurse believes she/he has incorporated research findings into his/her practice. An additional six multi-item instruments were located^65,66-73. Finally, in 18 published papers, (11 instruments) single-item questions were used to measure research use^74,15,39,75-87.

Common attributes of existing measures in nursing studies

To support the need to develop a sound instrument, we present an appraisal of the following attributes of the measures we located: theoretical framework, conceptualization, operationalization, validity and reliability, and other instrument characteristics.

Theoretical framework (use of theory for instrument development)

While there are many **models** of research utilization in nursing, among them, CURN^88, Goode^89 Ottawa^90, Stetler^91, NCAST(Nursing Child Assessment Satellite Training)^92, Horn^93, Iowa^94, PARIHS^14,95, we were not able to identify instances where any of these models guided instrument development. In only one instrument – the NPQ – was a theoretical framework (Rogers’ theory of innovation diffusion) explicitly used to inform the development of the instrument. Rogers’ model describes five stages an individual undergoes in adopting an innovation. These stages are awareness, persuasion, decision, implementation, and finally confirmation in the adoption process. According to Rogers^96, in the awareness stage the individual comes to know about the innovation. In the persuasion stage, the individual is either favorably or unfavorably disposed to the innovation. Based on this disposition, the individual engages in activities leading to a choice at the decision stage. The implementation stage follows a decision in favor of the innovation. Implementation involves a (usually) deliberate behavior as the innovation is put into practice. In the NPQ, the extent of utilization was indicated by the sum of the all the stage scores. Estabrooks^97 has argued that a drawback to using Roger’s model is the untested assumption that the innovation in classical diffusion theory is equivalent to research findings in the context of clinical nursing practice. We believe that the general under-use of theory to guide instrument development is more a function of a lack of theory than of non-use of existing theory on the part of investigators. While Rogers’ model seems to be one of few sources for this kind of work, we do not believe it should be the automatic choice.

Conceptualization (Definition of research use)

Over half of the nursing studies we reviewed had no explicit definition of research utilization. Only two of the studies reporting single-item instruments defined research utilization. In those studies that defined research utilization, the definition was not always operationalized in the instrument used, i.e., we found studies that defined research utilization as a process, but measured it as an event. This unclear conceptualization of research use results in several measurement problems. Primarily – when the instrument is not clear as to which type of research utilization is measured, the interpretation of outcomes is consequently unclear. One way to handle this is to give respondents guiding descriptions of various kinds of research use to clarify what is actually in question.

Operationalization

**Type of use:** Several kinds of research use were assessed, including overall, instrumental and conceptual. One study also measured symbolic (or persuasive) use. Most often, however, the type of research use had to be inferred because it was not stated explicitly. **Process or event:** All

---

**Sample items include:** I base my practice on research, I apply research results to my own practice, I use research to supply my nursing practice.
studies except those using the NPQ measured research use as an event. The NPQ measures research use as a process and does this focusing on specific innovations. In most of the nursing research utilization models, research use is seen as a process that includes steps such as awareness, critical appraisal and implementation. The event approach, which focuses on the implementation stage, may tell us more about what finally benefits the patient.

Validity and reliability
The psychometric properties of the instruments in the reviewed studies are not well addressed. More than half of the 41 studies we reviewed did not report any reliability assessment of the instrument. Of those studies actually addressing reliability issues, 13 reported Cronbach α scores (solely). Approximately 60% of the studies reported some sort of validity assessment; most of these assessments related to content validity by expert panels. The procedure for content validity assessment was not well described in these studies, leaving us skeptical as to the quality of content validity. Only two studies assessed construct validity. Pain et al.65 examined construct validity by testing the relationship between the construct of interest (research orientation) and similar constructs. Estabrooks40, demonstrated construct validity by testing the fit of the construct to other variables using structural equation modeling.

Other instrument characteristics
The focus across all reviewed studies was on measuring individual nurses’ use of research. Organizational dimensions of research utilization were not commonly studied. If organizational variables were examined the unit of analysis was the individual, suggesting the existence of errors in the unit of analysis. We believe this focus on individual nurse’s use of research is questionable, or at best insufficient. Some reports did emphasize research use as more of a system (organizational) characteristic rather than an individual attribute 35,93–101. In addition to the assumption that evidence-based practice is an individual responsibility, many of the weaknesses in the instruments found in these studies are related to other unacknowledged assumptions. Some examples are: research utilization is always good (all research is applicable to practice), research utilization is linear (a sequential process of dissemination → critical appraisal → implementation), research utilization is a valuable outcome in and of itself, and finally, decisions to use research are the decisions of rational actors in rational environments (which was evident in more or less all instruments)102–104. For example, the use of the NPQ illustrates the result of holding an assumption of linearity. The scoring of the NPQ presumes completion of prior stages and hence a linear progression through stages of research use. This assumption runs counter to the complex and often messy conditions that prevail in most health care facilities today.

Existing research utilization measures outside of nursing
We reviewed the non-nursing literature on knowledge utilization and identified 10 measures of knowledge utilization 105–114. These non-nursing studies investigated either the extent or the determinants of knowledge utilization. The broader term ‘knowledge utilization’ included social science research findings, program evaluation, and educational innovations in fields such as public policy, business, and education. One instrument, the “stage of knowledge utilization scale” (SKUS), appeared in multiple studies reviewed27,111,115,116. The instrument in Johnson’s 1995 paper110 was a single item measure. All the other instruments have multiple items. As in the nursing literature on measuring research use, this field is underdeveloped.
Theoretical framework (use of theory for instrument development)

Only one instrument, the “stage of concern questionnaire” (SoCQ) was developed on a previously specified theory of knowledge utilization, the Concern-Based Adoption Model\textsuperscript{117}. None of the other non-nursing studies had an explicit theoretical framework informing the development of the knowledge utilization measures.

Conceptualization (Definition of research use)

The conceptualization of knowledge utilization was poor, and very often there was no conceptualization (in terms of an explicit definition of knowledge or research utilization) prior to its operationalization (e.g., Green & Kvidahl\textsuperscript{108}, Johnson\textsuperscript{109,113}). When conceptualizations were present, there was no widely accepted specification of the term utilization. Various definitions of knowledge utilization were proposed in different studies\textsuperscript{105,111,116,118,110} and the conceptualizations did not correspond to the operationalizations, that is, we found instances where knowledge use was measured as a process, but the conceptualization implied that it was an event (e.g., Cousins & Leithwood\textsuperscript{106}).

Operationalization

**Type of use:** Some of the non-nursing studies limited the utilization to instrumental use (e.g., Johnson\textsuperscript{110}). Others also investigated conceptual use (e.g., Cousins & Leithwood\textsuperscript{106}, Johnson\textsuperscript{109}). The types of utilization were not differentiated in many other studies (e.g., Oh & Rich\textsuperscript{113}).

**Process or event:** Most of the non-nursing studies measured utilization as an event. The two instruments that apparently employed a process approach to utilization were the SKUS and the SoCQ. However, the scale systems used in the two instruments were different from that in the NPQ found in nursing. The stage (SKUS) or the peak stage (SoCQ) at which the respondents locate themselves indicates the extent of utilization. A high stage suggests a high degree of knowledge utilization.

Validity and reliability

As in the nursing studies, the measures were short of reliability and validity evidence. About half of the reviewed non-nursing studies discussed reliability. Similar to the nursing studies, internal consistency (Cronbach alpha) was most frequently addressed. Only one study presented validity evidence\textsuperscript{107}. The validity of the scores on the stage of concern questionnaire have been demonstrated with the inter-stage correlation matrix, the relationship with the judgments of concerns based on interview data, and confirmation of expected group differences and changes over time. Such evidence can only tell us that this questionnaire may be a valid measure of ‘concern’. When used as an instrument of knowledge utilization, the inference from the stage of concern to the stage of utilization has not been demonstrated.

Other instrument characteristics

Various indicators of research utilization were used in the non-nursing studies. The most common indicators were “use”, “apply”, “change”, “modify”, and other verbs that describe the action of utilization. Knowledge utilization was also inferred by policy impact (e.g., the identifiable effects of social research in decisions at the stage of formulating the problem\textsuperscript{114} and concerns about innovations (e.g., concern about the impact of the innovation on students\textsuperscript{107}). A variety of data collection methods were used. These non-nursing studies employed qualitative approaches, such as interview with content analysis (e.g., Caplan, Morrison & Stambaugh\textsuperscript{105}), whereas in the nursing studies, investigators commonly used self-report surveys. Various response anchors (scales) were used to determine the extent of research utilization. Approaches
such as measuring frequency of use\textsuperscript{108}, degree of agreement with statements on utilization\textsuperscript{106}, number of occasions of use\textsuperscript{105} and ways of using research findings\textsuperscript{112} were reported.

**Summary of literature review**

Developing a high quality instrument on research utilization is important from a theoretical, as well as, a practical perspective. Use of research-based practice and being able to evaluate this practice is essential if nurses are to provide high quality patient care. Our review of the literature suggests that the measurement of research utilization is underdeveloped. In summary, the existing instruments on research utilization have two major problems. **First**, the early efforts to measure research utilization were, for the most part, not based on prior theory or conceptualization\textsuperscript{23}. Twenty years ago Mandell and Sauter\textsuperscript{119} identified the specification of the construct “use” as one of four conceptual and methodological problems. Our review indicates that this conceptual problem still exists; almost all existing instruments clearly lacked clarification of the construct of research utilization. Most measures were not based on a theoretical framework with the exception of the NPQ and SoCQ. Various definitions of knowledge utilization were proposed in different studies, but most often, there was no explicit definition of the term utilization prior to its operationalization. The construct, research utilization, can be clarified with a theoretical framework or an explicit definition. The lack of progress in this issue is one of the major obstacles to establishing a sound measure of research use.

**Second**, existing instruments lacked psychometric assessments based on measurement theory. More than half of the instruments were not evaluated concerning reliability. Those that reported reliability evidence most often limited it to internal consistency. Internal consistency is the degree to which items tap the same construct. This is an important aspect of reliability but depending on the purpose of the instruments, other types of reliability need to also be considered. For example, when tracking the trend of research utilization across time, test-retest reliability should be demonstrated. Further, most of the instruments reported no validity evidence. Content validity was reported for some instruments, but we rarely saw an assessment of construct validity. This can be tied directly to the lack of a sound conceptualization of research utilization, i.e., construct clarity. This returns us to our first conclusion and to Dunn’s\textsuperscript{23} claim nearly two decades ago that construct validity is a serious and unresolved problem in the field.
References


76. Logsdon C, Winders D, Hawkins B, Parker B, Peden A. Factors related to research


121. Estabrooks CA, Floyd JA, Scott-Findlay S, O'Leary KA, Gushta M. Individual


126. Morgan D. Focus groups as qualitative research. *Qualitative Research Methods Series.* 1997;16.
