

La fiabilité de la recherche sur la sécurité : le cas de l'évaluation des risques de chute

Janice M. Morse

La plupart des programmes de prévention des chutes comportent deux volets : d'une part, des instruments de prédiction du risque de chute chez les patients et, d'autre part, des stratégies visant à empêcher les chutes ou à prévenir les blessures en cas de chute. Malgré leur rôle essentiel, un grand nombre de ces instruments ont fait l'objet de critiques parce qu'ils ne permettent pas d'identifier avec précision les patients sujets aux chutes. Le présent article examine, à la lumière des essais sur la validité touchant l'échelle de Morse [Morse Fall Scale], les recherches menées au cours des vingt dernières années sur l'évaluation du risque de chute. Certains travaux en la matière s'appuient sur des hypothèses erronées ou des erreurs de conception, tant en ce qui concerne la mise au point des échelles de risque que leur évaluation. Bon nombre de ces instruments ont été élaborés uniquement en fonction de leur validité apparente et n'ont pas bénéficié d'une évaluation adéquate ou, s'ils ont été mis à l'essai en milieu clinique, d'un plan expérimental valide. Enfin, l'usage à mauvais escient des échelles d'évaluation du risque de chute peut accroître le risque de chute chez les patients. L'auteure conclut qu'une grande part de la recherche menée en sciences infirmières sur ce thème ne contribue pas à améliorer la sécurité des patients.

Mots clés : évaluation des risques de chute, prédiction du risque de chute, échelle de Morse, prévention des chutes

The Safety of Safety Research: The Case of Patient Fall Research

Janice M. Morse

Most fall intervention programs consist of 2 components: fall risk prediction instruments to identify the patient who is likely to fall, and fall intervention strategies to prevent the patient from falling or to protect the patient from injury should a fall occur. While critical to the effectiveness of a fall intervention program, many of the fall risk prediction instruments have been criticized for their failure to accurately identify the fall-prone patient. In this article, in the context of the validity assessments conducted on the Morse Fall Scale, the research conducted in the past 2 decades on fall risk assessment is critiqued. Some fall prediction research is based upon invalid assumptions and/or errors in design, both in the development of risk scales and in the evaluation of these instruments. Many of these instruments have been constructed with inappropriate reliance on face validity, have been evaluated inadequately, or have been tested in the clinical setting using an invalid design. Finally, improper use of fall scales in the clinical area may increase the risk of injury to the patient. The author concludes that much nursing research on patient falls does not contribute to improved patient safety.

Keywords: patient fall risk assessment, fall prediction, Morse Fall Scale, fall intervention, Cochrane criteria

All research has consequences. The intent of nursing research is to improve care, and ideally the outcomes of nursing research will lead to improved practice. If the consequences of research are not positive, researchers hope that the unintended outcomes will not cause harm if the recommendations are implemented. But what if our research does not perform as intended and has negative consequences?

In this article, I will review research into patient falls that is targeted towards developing an instrument to screen for risk of falling, using one of the oldest (and therefore most frequently discussed) instruments, the Morse Fall Scale (MFS) (Morse, Morse, & Tylko, 1989). The assumption supporting this research is that if we can predict the patient who is likely to fall, then appropriate fall prevention and protection strategies can be implemented, and either (a) the patient will not fall (i.e., the fall *prevention* strategies were effective), or (b) the patient does fall but is not injured (i.e., the fall *protection* strategies were effective). Using the research published about the MFS, I will review the models of evaluation used to assess this fall risk prediction research. Finally, I will consider the expect-

tations of clinical performance from the perspective of both the scale developer and the clinician, and problems with the clinical utilization of fall prediction instruments.

Fall intervention programs usually consist of two parts: first, identifying the fall-prone patient so that fall interventions can be appropriately targeted, and second, making available and applying appropriate fall intervention strategies. Thus, the key to a fall research program is the accuracy of the instrument used to predict the risk of a patient falling, and therefore enabling the targeting of interventions to those patients actually at risk of a *physiological anticipated fall* (Morse, 1997). The second component, the fall intervention program, is equally essential for patient safety, and the ultimate goal is to prevent injury should a fall occur. Again, both the performance of the fall risk scale and the effectiveness of the preventive or protective interventions subsequently put in place for those patients rated at risk are crucial for preventing the actual fall or, if the fall occurs, protecting the patient from injury. Note that assessing a patient at risk of a fall in itself does not *prevent a patient fall* (the fall prevention strategies are intended to do that) *nor prevent injury* (that is the purpose of fall protection strategies) (Morse, 2002).

While extremely important and a high priority for patient safety, fall research is difficult and complex. Accordingly, some published research is prone to technical errors in the construction of the scale, design errors in the evaluation of the instruments, confusion about the expectation of performance, and errors in utilization. In practice, these mistakes place the patient at risk, place the staff and the hospital at legal risk, and result in increased health-care costs. Such research provides a false assurance of safety; that is, it provides a façade of care intended to make the patient safe but which is actually not safe. Errors in clinical application further weaken the usefulness of the research. Patient fall research is an excellent example of the importance and significance of nursing research, but the quality of the research needs to be drastically improved.

Quality of the Risk Assessment Instruments to Identify the Fall-Prone Patient

The intent of research-related fall risk prediction scales is to develop an instrument that will quickly triage for those who are at risk of falling, thereby enabling preventive and protective strategies to be immediately put in place to prevent patient injury,¹ and to monitor fall risk throughout their hospital stay. Before continuing with the discussion,

¹ Because the Morse Fall Scale (MFS) was the first of this type of instrument designed to predict which patients are likely to fall, much information is available about it, and I will place this discussion in the context of the MFS.

however, it is important to differentiate fall risk prediction scales (instruments intended to identify the fall prone and to *predict the risk of falling*) from instruments that are used for patient assessment; that is, to assess the individual's *condition* (usually physiologically based factors) that may cause a patient fall, such as gait assessment. *Assessment instruments* are time-consuming to use but provide information about the nature of physiologically based deficits so they can be rectified before a fall occurs (i.e., fall preventive measures, such as exercise or balance training programs to improve gait). By extension they may also assist in identifying the need for fall protective strategies (such as a hip protector to prevent a fractured hip should a fall occur). For example, a fall risk prediction scale might rate gait as normal, weak, or impaired, according to gross indicators based on mobility, while assessment instruments would require actual measurement of strength, balance, and so forth. Note that risk prediction scales provide patient scores that indicate risk of falling but do not tell us why there is a risk or what to do to prevent the fall, just as a thermometer will tell us if the patient has a fever but not what is causing the fever or how it should be treated.

Altman (1997) notes the tension between the purposes of these two types of instrument in trying to “reconcile pragmatism with methodological purity” (p. 1309): clinicians expect risk prediction scales to provide prescriptive information about fall prevention strategies, so they are tempted to add variables that provide diagnostic validity. But adding variables not only invalidates the scale's performance, but also moves the purpose away from fall *risk prediction* towards *fall assessment*. Recently, for instance, McFarlane-Klob (2004) published a “Modified Morse Scale” (without consultation with the developer), and added medication variables. If this researcher had understood how the MFS was developed and how it worked, she would have known that medications were evaluated during the scale construction.² Furthermore, making the scale longer defeats the purpose of efficient rating and does not increase the validity of the scale.

Methods of Scale Development

Fall prediction scales “work” because researchers have developed both the items and the item scores (the weights for those items) in an exploratory process by comparing a large number of variables that may possibly

² Medication is a part of the scale in the secondary diagnosis score. In developing the indices (items), we first included medications that were thought to contribute to falls, then numbers of medications, then combined this item with comorbidity (i.e., secondary diagnosis). Of course, medications contribute to falls, as they relate to the other variables (mainly gait and mental status). These results were also replicated in the Hendrich II (Hendrich, Bender, & Nyhuis, 2003).

contribute to a fall in subjects who *have* fallen, compared to those who have *not* fallen. This comparison of groups enables the identification of items that are statically significant. Computer modelling should be used in an exploratory manner, combining variables to form indices, hence enabling the identification of the minimal number of variables to eventually constitute the scale items, without reducing the ability of the scale to differentiate the fall group from the control. Next, statistical weights of the significant items may be converted to produce item scores, and the scale is subsequently modelled in the data set to assess validity, performance, and cut-off scores³ to determine levels of risk. Of course, these statistical weights as they are first calculated are not likely to be whole numbers, and would not be practical for use in the clinical setting. In the MFS, these numbers were rounded to the next whole number divisible by 5, and then the discriminant function of the scale was re-calculated to ensure that the scale still worked.⁴

This method of scale construction has been used with only two scales — the MFS (Morse et al., 1989) and the recent modification of the STRATIFY tool (Oliver, Britton, Seed, Martin, & Hopper, 1997) in Hamilton, Ontario (Papaiannou et al., 2004). The Hendrich II (Hendrich, Bender, & Nyhuis, 2003) approximates this approach, but it is not clear how the scores were calculated from data presented, why all significant items were not included in the final scale, and if the final scale was subsequently clinically tested.

However, most of the fall risk prediction scales available do not follow this design. Some have been developed using a control group to identify statistically significant items, but with the item scores arbitrarily assigned (e.g., Downton Index [Vassallo et al., 2004]; STRATIFY scale [Oliver et al., 1997]). In addition, some scales used retrospective chart reviews as data, rather than patient assessment (e.g., the Scott and White Falls Risk Screener [Yauk, Hopkins, Phillips, & Bennion, 2005]), hence limiting variables that could be identified as significant. Furthermore, some researchers have selected scale items using techniques of *face validity*, which is considered atheoretical, imprecise, and the weakest of all validities (Newfields, 2002). Using their own clinical judgement, these researchers have selected items by surveying other scales for the items most frequently used, or have selected those that they consider, from their own clinical experience, may cause a fall. Some of these instruments are simply checklists⁵ (e.g., Charting tips: Documenting a patient's fall risk,

³ A cut-off score is the lowest high-risk score.

⁴ However, the MFS is often not considered for clinical use, because the scores are still too difficult for nurses to add! Dempsey (2004) writes that the MFS was “considered complicated and time consuming” (p. 481), yet developed her own assessment tool.

⁵ The MFS is not a checklist, because the items are indices (see footnote 2) and weighted.

2000; Haines, Bennell, Osborn, & Hill, 2004); others have arbitrarily assigned scores to the items — scores also based on the researcher's own judgement and convenience. These values are often 1s, 2s, and 3s, selected for the clinicians' ease for totalling the scores, and the resulting scores are used to determine classes of high or low risk of falling (e.g., Browne, Covington, & Davila, 2004). (Note that when easily added numbers were assigned to the MFS, the discriminant function went down to .5 [or to the same ratio that one would obtain by flipping a coin]. It is *both* the combined function of item selection *and* the weight of the score assigned to the item that makes the MFS valid.)

Another criterion of validity of risk prediction scales is that they must work clinically. Scales must be sensitive to patients' conditions by providing a range of scores (the MFS is scored 0 to 120) and also be sensitive to the individual patient's change in condition. Finally, they must have been tested independently by another institution. This criterion was met by McCollam (1995) for the MFS.

Often these poorly constructed scales are used internally by hospitals. Some have been published (e.g., Brown et al., 2003; Dempsey, 2004; Hathaway, Walsh, Lacey, & Saenger, 2000; Undén, Ehnfors, & Sjostrom, 1999), others disseminated via the Internet (Farmer, 2000). These scales are usually "tested" in the clinical area by noting the fall score of the patients who actually fall: if the score of the patient who falls is in the estimated "high risk" range, then the scale is considered to "work" and is declared valid. However, except at a very gross or obvious level, if tested correctly with a control group, these scales probably will not differentiate the fall-prone patients from those who are not fall-prone. Of greatest concern, these instruments do not have the refinement to be able to accurately predict the fall-prone patient, and worse, have not been finely tuned to minimize the false negatives — that is, patients who are actually at risk of falling are not identified. Hence, these scales may have little validity or psychometric standardization (Perell, 2002).

The cost of using poorly constructed scales clinically is that the number of false negatives (or rating a patient *not at risk* when the patient actually *is fall-prone*) is very high, thus risking not identifying patients in need of fall protective and preventive strategies, and placing the patient at risk of injury should a fall occur. *This is the most serious consequence of "homemade" instruments.* The quality of homemade scales is poor and the safety of patients may be jeopardized. Given the availability of scales with diagnostic accuracy, there is no need for facilities to develop their own scales (Perell et al., 2001).

Why do clinical nurse researchers go to all the trouble and expense of developing a homemade scale when scales with reliability and validity data are available? Some nurse researchers have reviewed the MFS and

determined that it was not generalizable for their context. I am puzzled by such comments as “it was developed on Canadians” or “not suitable for our Australian context” (McFarlane-Klob, 2004), because the MFS does not contain contextual variables.

Another problem is that in the development of these scales, specialized patient populations are used. For instance, the STRATIFY scale was developed using elderly patients from three hospitals (Oliver et al., 1997). In the development of the MFS, patients were also recruited from three hospitals: acute-care, rehabilitation, and nursing-home hospitals. Although we deliberately tried to make a scale that would be valid for all patients, we did not include outpatients or day surgery, psychiatry, or home-care patients. There is no theoretical rationale, however, why the scale will not perform for these groups, and it would be faster to develop normative scores for those populations than to develop another scale.

Models Used to Evaluate Fall Risk Scales

Unfortunately, researchers have caused harm by inaccurately or improperly evaluating fall risk scales. As a consequence of these errors, excellent research is devalued and even debunked, and research gains are lost. Worse, some of these reviews have been published, so that rather than using completed research, the research effort, of varying quality, has continued in search of a reliable means to predict fall-proneness. The problems of the evaluation research include (1) inappropriate design used for clinical testing, and (2) errors in evaluation.

Inappropriate design used for clinical testing. Once a scale is developed, it is tested for clinical feasibility. Two problems of invalidity have emerged, affecting both homemade scales and those developed more rigorously. These are the Hawthorne effect, and disregarding of interventions that form intervening variables between obtaining the patient’s fall score and the opportunity for a fall to occur.

The Hawthorne effect. Unfortunately, simply implementing a fall intervention program alters the fall rate: (1) staff previously casual about reporting falls may now conscientiously report every fall, causing the fall rate to increase (see, e.g., O’Connell & Myers, 2001); and (2) staff are more aware of fall risk and may adopt fall prevention strategies, causing the fall rate to decrease. Therefore fall rates may be unreliable, and the fall *injury* rate is a more valid statistic for evaluating the efficacy of the fall intervention program. Nurses always file a fall incident report when a patient is injured, but, because injury is a relatively rare event, this may also be unstable for statistical reasons.⁶

⁶ A recent clinical trial randomly assigning matched pairs of clinical units (as control or intervention) tested fall intervention strategies (Healy, Monro, Cockram, & Heseltine,

Problems of design of clinical trials. Researchers often use the number of falls and the fall scores of the patients who fall to assess the efficacy of the risk prediction scale. But the number of falls *evaluates the fall intervention program, not* the efficacy of the scale. Once a patient is rated at risk of falling, staff are obligated to implement fall prevention strategies that actually stop the patient from falling. Therefore these intervention variables interfere with the measurement of the dependent variable and invalidate the trial to the extent that it is unreasonable to use these numbers to ascertain the sensitivity and specificity of a fall risk scale. Implementing such research design is akin to developing a Suicide Prediction Scale and administering it to all pedestrians who walk onto a bridge. Because the bridge is a favourite place from which to leap, barriers have been erected, video surveillance alert guards, and the police prevent anyone from climbing onto the bridge railing in order to leap; hence, no one is able to commit suicide regardless of intent. Does this mean there is anything wrong with the Suicide Predictor Scale? No — the intervening variables interfere with the relationship. Understandably, using a similar research design for determining the validity of a fall risk scale will not provide meaningful information about the validity of the scale. Yet researchers have done this, and published their results in refereed journals, and even wondered why their results obtained using the MFS are at variance with those originally reported (see, for instance, O’Connell & Myers, 2001, 2002).

Errors in evaluation.⁷ Faulty methods of evaluation have also been used. These include the reliance on face validity, failure to use the original publications when assessing performance, and trialing scales against each other and with nurses’ clinical judgement.

Reliance on face validity. Review articles present tables listing all of the scales and comparing the items in each scale (see Evans, Hodgkinson, Lambert, Wood, & Kowanko, 1998, 2001; Joanna Briggs Institute, 1998; Morse, 1993) to determine whether they “fit” some preconceived domain of factors that cause patient falls. Note that the value assigned to

2004) and obtained a statistically significant drop in fall rate, even though the fall interventions may be considered “normal practice.” However, as the fall injury rate actually increased in the intervention units, the difference in the fall rate, as the authors note (p. 391), may have been due to the Hawthorne effect.

⁷ How should the MFS be evaluated? Some researchers investigating falls have reported positive results regarding the validity of the MFS. Camicioli and Licit (2004) noted that the MFS was a predictor of risk of falling in a specialized Alzheimer disease care unit. When investigating the association between medications and falls, Dyer et al. (2004) noted that it was the number of medications, rather than a specific medication or class of medications, that was significant, adding credence to the *secondary diagnosis/comorbidity* item on the MFS.

each item in the respective scales is omitted from these tables, so that the comparisons are meaningless.

Failure to use the original source when assessing performance. A review is valid only if it is complete. Yet in the review of fall risk scales reported by the Joanna Briggs Institute in Australia (Evans et al., 1998) this was not the case. Instead of using the publication reporting the MFS development (i.e., Morse et al., 1989), they used a publication describing the characteristics of types of fall (Morse, Tylko, & Dixon, 1987). This is a surprising error,⁸ for the original source is cited in many earlier publications, and the research program is even summarized in a book (Morse, 1997). Given their omission of key publications, one must challenge Evans et al.'s strong conclusion that "Falls risk assessment tools are very inaccurate...no evidence to suggest that the generic risk tools...offer any additional benefits over tools that are used within a single institution and have been developed based on that population's characteristics...no particular risk assessment tool can be assessed" (p. 30).

Trialing of scales against each other and with nurses' clinical judgement. Some researchers have trialed risk assessment scales against nurses' clinical judgement and, when finding neither excellent, have recommended the use of a combined approach (both the scale and clinical judgement) (Moore, Martin, & Stonehouse, 1996). However, these trials are inadequately designed: researchers did not control for nurses' prior knowledge about falls or knowledge about fall assessment. Of greater concern, the study by Eagle et al. (1999) testing three methods of assessment — nurses' clinical judgement, the Functional Reach Test to measure balance (Duncan, Weiner, Chandler, & Studenski, 1990), and the MFS — the researchers used the MFS incorrectly, scoring the patients using retrospective chart review rather than assessing them. The MFS cannot be validly completed by using chart data — patients *must* be examined — but these evaluators did not do this. Further, while the raters and the nurses were blind to the patients' MFS scores, it was not known if raters (who were using their clinical judgement) had used the MFS and/or other methods to rate patient risk of falling previously. In other words, there was no control over the nurses' knowledge about fall risk assessment. Was their clinical judgement blind to research knowledge? This threat to validity would be very difficult to control.

Invalid Design of Clinical Testing

The most problematic design of fall intervention program research is the simultaneous testing of the fall risk prediction scale and the fall

⁸ The Joanna Briggs Institute is responsible for evaluating research for a number of institutions internationally. These reviews are highly specialized and accuracy is an imperative.

interventions. The O'Connell and Myers (2001) study used this design, but it was further confounded by a second fall intervention study conducted simultaneously, but unknowingly, by the occupational therapy staff. Despite these problems (which included the intervention program interfering with their dependent variable, the fall rate), O'Connell and Myers (2001, 2002) were still critical of the predictive validity of the MFS. Their false positive rate (i.e., 79% of patients rated at risk of falling and who did not fall) perhaps meant that their interventions were working, not that the scale was problematic, with limited generalizability, as they concluded.

How did we therefore obtain sensitivity and specificity statistics for the MFS that apparently cannot be replicated? First, we studied patients who fell *at the time of the fall (confirmed fallers)* and controls — those who had *not* fallen — and this provided sensitivity of 78% and specificity of 83% (Morse et al., 1989). These results were satisfactory, but were still not without problems, for there were a number of errors — false positives (patients who had not fallen and who were considered by the computer to be at risk) and false negatives (patients who had fallen and were rated as not at risk). We investigated these errors by examining the charts of these patients 10 weeks after the initial analysis. We found that the false positive group had a high rate of falls (5 of the 17 patients had fallen; one patient fell 3 times) and concluded the computer was correct — these patients were at risk but had not had the opportunity to fall before the time of the original data collection, and they increase the sensitivity to the scale to 91%. The falls that were experienced by patients who rated at risk of falling by the MFS we labelled *physiological anticipated falls*. Next, by examining the circumstances of falls that occurred in patients in the false negative group, we identified two additional types of fall: the accidental fall (true accidents, slips and trips in those who are rated at not risk of falling), and the unanticipated physiological fall (falls due to a seizure or fainting in patients who also scored not at risk) (Morse et al., 1987). Recalculating the ability of the scale to discriminate after making these corrections, the sensitivity and specificity of the scale increases to 84% sensitivity. But the importance of recognizing the three types of fall is that the scale will never identify 100% of falls in a hospital, and staff should always try to determine what type of fall occurred, record statistics accordingly, and be aware that the preventive and protective strategies for each type of fall differ (Morse, 1997). The site of accidental falls must be investigated to prevent recurrence, and strategies implemented to protect those with unanticipated falls from injury should a second fall occur.

Clinical Errors When Using the Scale

Essential to the clinical performance of a scale is its correct use in patient assessment. This assessment is reasonably quick for the MFS (it takes 1–3 minutes), but users need to have received instruction.⁹

Not using the MFS according to directions. As noted above, the lack of correct assessment and inaccurate scores results in errors. Despite the availability of instructional tools for the MFS, some clinicians do not realize that scoring the patient requires patient examination. As with all forms of assessment, if the scale is not used correctly, regardless of its reliability and validity, it will not perform as expected clinically. Patient safety will be jeopardized.

Failure to acknowledge the sensitivity of the scale. The second problem occurs when the staff record the patient's score as *high risk* or *low risk* and do not record the total score (Perell et al., 2001). This is akin to recording a patient's temperature as high or low without recording the actual figure, so that staff do not know if the temperature is increasing or decreasing or the severity of the fever. Similarly, if the actual fall score is not recorded, then staff will not know how high the fall risk is, and whether it changes throughout the 24 hours. As the goal of care is to reduce the score, if the actual score is not recorded, then it will not be possible to gauge improvement (and decrease of fall risk) or an increasing score (and therefore increased risk of falling).

"All patients scored at high risk." A frequent complaint is that all of the patients scored high risk of falling — that is, the scale does not discriminate adequately. It is possible that all of the patients are, for instance, at high risk. Raising the level of risk will not change this fact, and will place those who are at risk in the "not at risk" category (i.e., a false negative). But if each patient's actual score is recorded, then the staff will recognize that there are discernable degrees of high risk.

Infrequent scoring. The final problem is not scoring the patient frequently enough. An emerging standard is that patients should be scored upon admission, and thereafter if a patient's condition changes. This is not frequent enough for patients in acute care, who should be scored at least once per shift. In long-term care, where patient fall risk is more stable, the patients should be scored frequently over several 24-hour periods, until their pattern is recognized, and then scored less often — as infrequently as once a week — if the resident's condition remains stable.

⁹When the scale was first developed, an instructional videotape was available to teach the use of the scale. In 2003, this was replaced with an instructional DVD, provided without charge by Hill Rom Industries (safetyprograms@hill-rom.com).

Discussion: So What? What Is at Risk?

Given the poor quality of this clinical evaluation research, and unrealistic expectations of the scale's performance, it is not surprising that the quest for a perfect — or at least improved — scale has continued since the development of the MFS. Patient falls is probably one of the most researched clinical problems in nursing. The responsibility for patient falls has been placed squarely on the shoulders of nursing. We feel guilty if a patient falls, blaming ourselves for not remaining vigilant and perhaps even for neglecting basic care (“I should have asked this patient if she needed toileting”). Because of this firm link to basic nursing skills, many nurses have attempted to examine the problem of falls in various ways. Researchers are motivated by clinical *problems* — they hope interesting problems — those that will improve nursing care and change patient outcomes. Thus patient falls has been researched and researched by nurses, and this research continues to the present time.

However, the research is extraordinary. Each project is conducted in relative isolation from other projects, so that the research is not cumulative overall. Project after project is conducted with the aim of developing yet another fall risk prediction instrument. The failure to utilize the work of others has “levelled the playing field” and often results in mighty steps backwards.¹⁰ The problem is compounded by invalid methods of evaluating and testing the available instruments and a lack of rigorous, funded inquiry by experienced researchers.

Review articles, including the Cochrane reviews, do link fall research, but these are not without error and omissions — which are then perpetuated by means of meta reviews (see, for instance, Burrows, 1999).

Astonishingly, this research is by and large being conducted using “opinion,” albeit under the guise of clinical judgement. The Cochrane criteria are correct: opinion (be it “clinical judgment,” “intuition,” or “expert committee decisions”) results in poor research involving measurement and in a low level of evidence. In the case of nursing fall research, this over-reliance on soft data results in the paradox of applying “qualitative” data¹¹ to a quantitative problem. It is the poorest of qualitative work, inappropriately applied, with the results masquerading as a quantitative tool that jeopardizes patient safety. Patients risk injury and even death.

¹⁰ But this independence has extended to the developers of the instruments, who are rarely consulted about proposals for evaluating their tools or asked to comment on the accuracy of articles evaluating their work prior to publication.

¹¹ I am using qualitative as a descriptor for non-numerical data, not to indicate a legitimate qualitative method.

Safety research is important, but it must be safe. It must be given adequate funding, conducted by researchers with appropriate qualifications, implemented wisely, and evaluated appropriately. Fall risk prediction is not easy to research: the outcome variable is intercepted; fall risk changes rapidly, and — particularly in the acute-care setting — is unstable, so that frequent assessments are essential. Fall intervention programs are not a low-cost add-on in the clinical area; they are expensive in time and dollars, but are essential to safe care. Fall risk assessment is a task that can be achieved only through the education of nurses, some time commitment in their workload, some attention by the quality assurance department to the recording of scores and fall statistics, and some investment on the part of administration for program costs. Fall intervention programs require all of these commitments, plus funding of a position for a clinical specialist to organize the program, funding for fall prevention and protection devices, funding to ensure that the building and equipment are as safe as possible, and vigilance and responsiveness to the program as a whole. Without the complete package, fall injuries in hospitals will not be reduced.

References

- Altman, D. G. (1997). Study to predict which elderly patients will fall shows difficulties in deriving and validating a model. *British Medical Journal*, *315*, 1309.
- Browne, J. A., Covington, B. G., & Davila, Y. (2004). Using information technology to assist in redesign of a fall prevention program. *Journal of Nursing Care Quality*, *19*(3), 218–225.
- Burrows, E. (1999). *What strategies exist for the effectiveness of fall prevention strategies of older patients in institutional settings?* Clayton, Victoria, Australia: Centre for Clinical Effectiveness, Monash University. Retrieved October 10, 2005, from <http://www.med.monash.edu.au/healthservices/cce/evidence/pdf/b/234.PDF>
- Camicioli, R., & Licitis, L. (2004). Motor impairment predicts falls in specialized Alzheimer care units. *Alzheimer Disease and Associated Disorders*, *18*(4), 214–218.
- Charting tips: Documenting a patient's fall risk. (2000). *Nursing*, *30*(11), 14.
- Dempsey, J. (2004). Falls prevention revisited: A call for a new approach. *Journal of Clinical Nursing*, *13*, 479–485.
- Duncan, P. W., Weiner, D. K., Chandler, J., & Studenski, S. (1990). Functional reach: A new clinical measure of balance. *Journal of Gerontology*, *45*, M192–M197.
- Dyer, A. E., Taylor, G. J., Reed, M., Dyer, C. A., Robertson, D. R., & Harrington, R. (2004). Falls prevention in residential care homes: A randomized controlled trial. *Age and Ageing*, *33*, 596–602.
- Eagle, D. J., Salama, S., Whitman, D., Evans, L. A., Ho, E., & Olde, J. (1999).

- Comparison of three instruments in predicting accidental falls in selected inpatients in a general teaching hospital. *Journal of Gerontological Nursing*, 25(7), 40–45.
- Evans, D., Hodgkinson, B., Lambert, L., Wood, J., & Kowanko, I. (1998). *Falls in acute hospitals: A systematic review*. Adelaide, Australia: Joanna Briggs Institute for Evidence Based Nursing and Midwifery.
- Evans, D., Hodgkinson, B., Lambert, L., Wood, J., & Kowanko, I. (2001). Falls risk factors in hospital settings: A systematic review. *International Journal of Nursing practice*, 7, 38–45.
- Farmer, B. C. (2000, May). Fall risk assessment. *Try this: Best practices in nursing care to older adults, Hartford Institute for Geriatric Nursing*, 8. Retrieved October 16, 2005, from <http://www.hartfordign.org/publications/trythis/issue08.pdf>
- Haines, T. P., Bennell, K. L., Osborne, R. H., & Hill, K. D. (2004). Effectiveness of targeted falls prevention programme in a subacute hospital setting: Randomised control trial. *British Medical Journal*, 328(7441), 676.
- Hathaway, J., Walsh, J., Lacey, C., & Saenger, H. (2000). Insights obtained from an evaluation of a falls prevention program set in a rural hospital. *Australian Journal of Rural Health*, 9, 172–177.
- Healey, F., Monro, A., Cockram, A., Adams, V., & Heseltine, D. (2004). Using targeted risk factor reduction to prevent falls in older inpatients: A randomized controlled trial. *Age and Ageing*, 33, 390–395.
- Hendrich, A. L., Bender, P. S., & Nyhuis, A. (2003). Validation of the Hendrich II Fall Risk Scale: A large concurrent case/control study of hospitalized patients. *Applied Nursing Research*, 16(1), 9–21.
- Joanna Briggs Institute. (1998). Falls in hospitals. *Best Practice: Evidence Based Practice Information Sheets for Health Professionals*, 2(2), 1–6.
- McCollam, M. E. (1995). Evaluation and implementation of a research-based falls assessment innovation. *Nursing Clinics of North America*, 30(3), 507–514.
- McFarlane-Klob, H. (2004). Falls risk assessment: Multi-targeted interventions and the impact on hospital falls. *International Journal of Nursing Practice*, 10, 199–206.
- Moore, T., Martin, J., & Stonehouse, J. (1996). Predicting falls: Risk assessment tool versus clinical judgment. *Perspectives*, 20(1), 8–11.
- Morse, J. M. (1993). Nursing research on patient falls in health care institutions. *Annual Review of Nursing Research*, 11, 299–326.
- Morse, J. M. (1997). *Preventing patient falls*. Thousand Oaks, CA: Sage.
- Morse, J. M. (2002). Enhancing the safety of hospitalization by reducing patient falls. *American Journal of Infection Control*, 30(6), 376–380.
- Morse, J. M., Morse, R. M., & Tylko, S. J. (1989). Development of a scale to identify the fall-prone patient. *Canadian Journal on Ageing*, 8(4), 366–377.
- Morse, J. M., Tylko, S. J., & Dixon, H. A. (1987). Characteristics of the fall-prone patient. *Gerontologist*, 27(4), 516–522.
- Newfields, T. (2002). Challenging the notion of face validity. *JALT Testing and Evaluation*, 6(3), 19–20.
- O'Connell, B., & Myers, H. (2001). A failed fall prevention study in an acute care setting: Lessons from the swamp. *International Journal of Nursing practice*, 7, 126–130.

- O'Connell, B., & Myers, H. (2002). The sensitivity and specificity of the Morse Fall Scale in acute care settings. *Journal of Clinical Nursing, 11*, 134–136.
- Oliver, D., Britton, M., Seed, P., Martin, F. C., & Hopper, A. H. (1997). Development and evaluation of an evidence-based risk assessment tool (STRATIFY) to predict which elderly patients will fall: Case control and cohort studies. *British Medical Journal, 31*, 1049–1053.
- Papaioannou, A., Parkinson, W., Cook, R., Ferko, N., Coker E., & Adachi, J. D. (2004). Prediction of falls using a risk assessment tool in acute care. *BMC Medicine*. Retrieved November 1, 2005, from <http://www.biomedcentral.com/17417015/2/1>
- Perell, K. L. (2002). Assessing the risk of falls: Guidelines for clinical practice settings. *Generations, 3*, 56–59.
- Perell, K. L., Nelson, A., Goldman, R. L., Luther, S. L., Prieto-Lewis, N., & Rubenstein, L. Z. (2001). Fall risk assessment measures: An analytic review. *Journal of Gerontology. Series A: Biological Sciences and Medical Sciences, 56A*(12), M761–766.
- Undén, G., Ehnfors, M., & Sjostrom, K. (1999). Use of an initial risk assessment and recording as the main nursing intervention in identifying risk of falls. *Journal of Advanced Nursing, 29*(91), 145–152.
- Vassallo, M., Vignaraja, R., Sharma, J. C., Hallam, H., Binns, K., Briggs, R., Ross, I., & Allen, S. (2004). The effect of changing practice on fall prevention in a rehabilitative hospital: The Hospital Injury Prevention Study. *Journal of the American Geriatrics Society, 52*(3), 335–339.
- Yauk, S., Hopkins, B. A., Phillips, C. D., & Bennion, J. (2005). Predicting in-hospital falls: Development of the Scott and White Falls Risk Screener. *Journal of Nursing Care Quality, 20*(2), 128–133.

Author's Note

Comments or queries may be directed to Dr. J. Morse, 6–10 University Extension Centre, University of Alberta, Edmonton, Alberta T6G 2T4 Canada. E-mail: Janice.morse@ualberta.ca

Janice M. Morse, RN, PhD, D. Nurs (Hon), FAAN, is Professor, Faculty of Nursing, and Scientific Director, International Institute for Qualitative Methodology, University of Alberta, Edmonton, Canada.