Determining the Relationship Between Staffing and Quality

John F. Schnelle, PhD

The article by Zhang and Grabowski (2004) provides important data relevant to two questions: (a) Did the Nursing Home Reform Act (NHRA) of 1987 result in improved quality of care and staffing levels? and (b) How do staffing data generated by nursing homes (NHs) associate with three quality measures?

Zhang and Grabowski (2004) used a sophisticated analytical approach and unique longitudinal secondary data to address these questions. The authors acknowledge that their conclusions may be controversial because of limitations inherent in any design that uses available secondary data to measure staffing and quality. This editorial asserts that a direct intervention paradigm involving prospective data collection is needed to correct these limitations.

Zhang and Grabowski (2004) undertook their study in part to address some of the limitations of prior studies of NH quality and staffing, and their article provides a good discussion of the problems that they could not directly address. In particular, one of their conclusions illustrates the assumptions required whenever risk adjustments are made to quality indicators—a controversial topic in this field. Their analyses showed that the NHRA was associated with a statistically significant reduction in the proportion of residents with physical restraints, catheters, and pressure ulcers, even though the absolute proportion of residents with pressure ulcers increased between 1987 (pre-NHRA) and 1993 (post-NHRA)—the study’s time frame. In fact, the pressure-ulcer rate decreased only after controlling for facility, resident, market, and state factors. Herein lies the dilemma.

In order to accept that the small risk-adjusted decrease in pressure-ulcer prevalence is due to improved quality of care, one must first accept the notion that the control variables reflect resident acuity and, thus, potential risk for pressure-ulcer development. Most notable of the acuity variables used in the risk adjustment is prevalence of bedfastness, which increased dramatically following the NHRA. If this increase in bedfastness reflects increased resident acuity that was not exacerbated by the same poor care processes that could have influenced pressure-ulcer development, then it is reasonable to argue that pressure-ulcer rates improved following the NHRA. The opposing argument, which may be just as reasonable, is that increased bedfast prevalence is itself an outcome and, like increased pressure-ulcer prevalence, associated with poor care. A more complete discussion of the complexities of risk adjustment is provided by Mor and colleagues (2003).

With regard to staffing levels, NHs reported small increases across the pre- and post-NHRA periods, but these staff increases were not related to improved quality except in those NHs that started with the lowest staffing levels. Zhang and Grabowski (2004) list several factors that might explain this finding, including (a) potential error rates in NH reports of staffing and quality, (b) the absence of quality-of-life measures that might be more sensitive to variations in staffing than clinically focused outcomes, and (c) the fact that few NHs may have adequate levels of staffing to affect quality measures. These three points deserve further discussion because they are often raised as points of controversy when arguments are made about the relationship of staffing and quality.

In regard to accuracy, suspicions exist about the accuracy of both staffing and quality measures available in secondary data generated by NH providers. These suspicions would be mitigated if an auditing mechanism were in place to ensure accuracy. However, staffing statistics reported to government agencies are either not audited or only “desk” audited to eliminate outliers (Abt Associates, Inc., 2000). Determining if staffing reports are accurate
using data sources independent of those provided by NHs to government agencies is not easy. One recent study did attempt to document the accuracy of staffing statistics by comparing staff time cards to both Medicaid cost reports and Online Survey, Certification, and Reporting (OSCAR) data. The time-card data were considered the gold standard staffing measure. Medicaid cost reports were more correlated with the time cards than OSCAR, but there were still errors. For example, only 50% to 60% of the facilities that were in the lowest staffed categories according to the time-card reports were consistently identified with either Medicaid cost report or OSCAR data (Abt Associates, Inc., 2000). It is not clear whether this error rate would create enough “noise” to obscure the relationships between quality measures and staffing reports, but these data do not reduce suspicions about the defensibility of research conclusions based on the analysis of NH-generated staffing data.

The problems caused by inaccurate staffing reports are compounded by concerns about the accuracy of quality measures. The U.S. General Accounting Office (GAO) has released two reports that question the accuracy of minimum data set (MDS) data, which are often used to generate care quality measures. The studies indicate that these accuracy problems may exist because most states do not audit MDS data (GAO, 2002a, 2002b). Others, however, have argued that the MDS data are “accurate” without specifying how accuracy was determined (Abt Associates, Inc., Hebrew Rehabilitation Center for Aged Research and Training Institute, & Brown University, 2002). Although it is clear that MDS data can be reliably reproduced by research staff, it is less clear if it can be completed reliably and accurately in the less resource-rich NH environment.

Concerns also have been expressed that the MDS and other available NH-generated data do not capture quality-of-life measures that may be more sensitive to staffing levels than clinical measures. For example, there is good evidence that NH consumers prefer consistent assistance in activities of daily living (ADLs) delivered in a respectful manner (Grau, Chandler, & Saunders, 1993; Gustafson & Gustafson, 1996). Intuitively, one would expect that consistent, timely ADL assistance would be more directly related to staffing levels than clinical outcomes, which are known to be influenced by immutable resident-acuity factors. Unfortunately, there are no data available in secondary data sets that capture these quality-of-life measures.

Finally, even if one assumes accuracy in NH-generated staffing and quality reports, there are concerns that many of the nation’s NHs may not be adequately staffed to affect quality measures. If this is the case, efforts to identify relationships between staffing and quality by analyzing data from large groups of suboptimally staffed NHs are problematic. Zhang and Grabowski (2004) provide data relevant to this issue. They argue that the relationship between staffing and quality may not be linear, as evidenced by an association between quality improvement and staffing increases only in the lowest staffed homes. A recent report by the Centers for Medicare and Medicaid Services (CMS) underscores the potential severity of this problem. This report indicated that 92% of the nation’s NHs were not adequately staffed to provide care that could prevent poor clinical outcomes or which expert consensus data indicate as necessary to provide good care (Abt Associates, Inc., Hebrew Rehabilitation Center for Aged Research and Training Institute, & Brown University, 2002). If this 92% estimate is correct, then efforts to identify the relationship between quality-of-care measures and staffing levels will be difficult to detect by analyzing small staffing variations in numerous low-staffed NHs.

In short, conclusions in either direction about the relationship of staffing to quality that are derived from secondary data will continue to be challenged by the problems described in this editorial. One partial solution to this problem is to wait until the accuracy of data generated by NHs is improved with auditing technologies and the addition of quality-of-life measures. CMS has acknowledged the importance of these issues, and there are planning efforts under way to improve the accuracy and comprehensiveness of staffing and MDS data generated by NH providers. Unfortunately, even if these projects are operationalized into practice, they will not address concerns that there may be few NHs that are adequately staffed to provide care that could significantly improve quality.

Alternatively, these issues could be addressed by a more direct approach. This approach would actively provide care processes to residents that have been identified in regulatory and practice guidelines as associated with good outcomes and quality of life. The purpose would be to determine the staffing requirements to provide this care under conditions of high productivity. There will be no need for controversial risk adjustments because clinical measures will not be the primary criteria used to identify relationships between staffing and quality. Instead, the focus will be on care processes based on the premise that all NH residents deserve to receive good care even if they suffer from conditions that may preclude good clinical outcomes. For example, residents who prefer and need walking assistance should receive that assistance even if it may preclude the ability to walk independently. In addition, research staff could implement these care processes and utilize scientifically defensible measurement protocols to identify the time required to provide that care. Thus, it would not be necessary to use NH-generated data to identify staffing requirements nor to prove that the care process implementation actually occurred.

This alternative approach described in this editorial is a departure from traditional health-service
Methodologies that attempt to identify variations in measures between large groups of NHs in the hope that important variations exist and are reported accurately. However, this direct approach to determine the labor requirements to provide a high-quality service or product is not unique in manufacturing and some service settings. Analytical technologies to accomplish accurate labor requirement projections have been developed and widely used in other settings (Edmonds & O’Connor, 1999; Hershey, Pierskalla, & Wander, 1981). It may be time to apply these methodologies to the NH setting to answer fundamental questions about staffing and quality that should have been answered long ago.

It should be noted that a scientifically defensible answer to most of the staffing questions being posed today may not be politically popular (e.g., Should there be mandated minimum staffing requirements?). There is some preliminary evidence that the excellent care described in practice and regulatory guidelines may be very labor intensive (Schnelle et al., 2002; Schnelle, Simmons, & Cretin, 2001). There is a possibility that even NHs that have motivated their staff to work at high productivity levels, despite limited staffing and low salaries, still do not have the capacity to provide care consistent with regulatory requirements. The policy implication of this finding is that increased reimbursement will be required to provide NHs with the staffing resources necessary for excellent care.

References

Received August 21, 2003
Accepted September 12, 2003
Decision Editor: Linda S. Noelker, PhD