Inspection of care: Findings from an innovative demonstration

by John N. Morris, Clarence C. Sherwood, and Paul Dreyer

In this article, information is presented concerning the efficacy of a sample-based approach to completing inspection of care reviews of Medicaid-supported nursing home residents. Massachusetts nursing homes were randomly assigned to full (the control group) or sample (the experimental group) review conditions. The primary research focus was to determine whether the proportion of facilities found to be deficient (based on quality of care and level of care criteria) in the experimental sample was comparable to the proportion in the control sample. The findings supported such a hypothesis: Deficient facilities appear to be equally identifiable using the random or full-sampling protocols, and the process can be completed with a considerable savings of surveyor time.

Introduction

In this article, the implementation history and program outcomes are presented of an innovative approach to implementing mandated annual inspection of care (IOC) reviews of Medicaid-supported nursing home residents. Sponsored by the Massachusetts Department of Public Health, the innovative program was approved by the Health Care Financing Administration (HCFA) as a demonstration that waived existing regulations requiring a point-prevalence IOC survey of all Medicaid residents in a nursing home. In its stead, random sampling of residents was permitted. Begun in late 1983 and continuing through 1985, the program was implemented across the Commonwealth of Massachusetts.

The goal of the program was to demonstrate the efficacy of the sampling approach to IOC; that is, to show that the conclusions derived from sample results do not differ significantly from the conclusions derived from the results of full review. Such a demonstration would allow the reallocation of surveyor time—a scarce resource—to other quality assurance activities.

The IOC process focuses on identifying nursing homes with unacceptably high failure rates—"failure" being defined in terms of resident-based, quality of care deficiencies and/or inappropriate level of care placements based on skilled nursing facility (SNF) and intermediate care facility (ICF) criteria. The primary focus in assessing the efficacy of the innovative demonstration program is whether or not failure rates for experimental facilities, which were surveyed using a random-sampling procedure, were comparable to those for controls, which were surveyed using the existing full-sample point-prevalence procedure. The importance of studying this topic is highlighted by the recommendation by the Institute of Medicine (1986) that only a sample of residents be selected for an IOC review.

Reprint requests: John N. Morris, Ph.D., Department of Social Gerontological Research, Hebrew Rehabilitation Center for Aged, Boston, Massachusetts 02131.

Description of the experimental demonstration

In evaluating any change in procedure, a major concern is the difference between what was planned and what actually occurred. Although the system intervention for this demonstration as initially conceived was basically followed, both intended and unintended deviations occurred. The IOC process normally includes the following features, which were followed for the randomly identified control facilities:

- Records are reviewed and bedside visits made for each Medicaid resident.
- The level of care is evaluated.
- The adequacy of that care is assessed.

In altering this process, Massachusetts asserted that the main problems seemed to be excessive cost and consumption of surveyor time. In 1980, surveyors in Massachusetts spent approximately 22,300 hours performing IOC reviews of approximately 30,000 residents. This amounted to 25 percent of all surveyor field time. Many argued that the review of all residents in a nursing home was unnecessary and that the goals of the IOC program could be met by a review of a sample of residents using statistical quality-control techniques.

In August 1983, a HCFA waiver permitted Massachusetts to institute a sampling approach to the IOC survey process. This approach combined elements of sampling projects in Wisconsin and New York (Zimmerman et al., 1984). From the New York plan, the concept of a quick initial walk-through inspection to determine the existence of potential problems of an obvious type was adopted; where such problems were identified (e.g., a large majority of residents were restrained in chairs), all residents were assessed. As in the Wisconsin plan, a more detailed review of a sample of residents was performed. When the walk-through failed to identify an obvious problem, a review of a randomly selected subset of residents was required. If, after review of these cases, the number of out-of-level residents exceeded predetermined cutoff points for the sample (e.g., more than 1 resident per 20 cases sampled) or if

there was an excessive number of adequacy-of-care problems (i.e., 20 percent or more of all areas checked summed across the resident sample), then all residents were reviewed. Surveys were conducted by teams of reviewers, with each team consisting of a nurse and a social worker. For each resident evaluated, the existing three-level assessment was completed, that is, records were reviewed, bedside visits were made, and the level and adequacy of care were evaluated in accordance with the existing criteria.

The sampling process was instituted on an experimental basis; that is, facilities were randomly assigned to sampling or full-review conditions. The decision was made to implement the demonstration on a statewide basis, with only 50 (of more than 500) facilities randomly designated for full review (the controls). This distribution ensured an adequate control group, while practically permitting the maximum number of facilities possible to be exposed to the new sample survey process.

Program monitoring findings

For the IOC demonstration, we looked at the reliability of the walk-through process; the number of facilities for which a sample review was completed; and whether the demonstration reduced the time required to complete the survey.

For a walk-through to make a functional contribution to quality assurance procedures, the judgments of staff must be reliable. That is, there must be reason to believe that surveyors make similar judgments when confronted by similar conditions. To measure the reliability of the judgments of Department of Public Health (DPH) staff, a substudy was conducted within the overall quality assurance project. A sample of 43 demonstration facilities was selected, and two-person review teams consisting of a nurse and a social worker were formed. Each member of the team performed the walk-through separately in each assigned facility, with a subsequent statistical comparison of these assessments.

Two reliability measurement procedures were utilized in this analysis of the reliability of walk-through judgments: an essentially descriptive procedure, using percent of agreement between judges as the criterion; and a more statistically oriented analysis-of-variance-based reliability measurement procedure that produces the Spearman-Brown coefficient of reliability.

Agreement among evaluators

Thirteen teams, involving eight nurses and nine social workers, were assigned to the study of walk-through judgment reliability. The walk-through evaluation checklist consisted of 22 items. Because 41 facilities were evaluated, a total of 924 judgments were made by the nurses and another 924 by social workers. In Table 1, the percentage of agreement between nurses and social workers is shown.

Table 1

Percent of agreement between nurses and social workers performing nursing home inspection of care reviews for 41 facilities: Massachusetts, 1984 and 1985

Item	Percent
Staff attention to residents	100.0
Staff appearance	100.0
Staffing shortage in resident care	100.0
Behavioral problems ignored	100.0
Patient mood	97.6
Patient grooming	93.0
Bedsores	95.3
Contractures	95.3
Patient positioning	93.0
Lack of privacy	93.0
Physical abuses	100.0
Resident's general physical condition	100.0
Catheters	97.7
Restraints	88.4
Level of alertness	100.0
Meaningful activities	97.7
Dirty facility	95.3
Odorous facility	90.7
Safety hazards	88.4
Patient complaints	100.0
Level II (Medicaid SNF) residents that are ambulatory	97.7
Level III (ICF) residents that are ambulatory	86.0

NOTES: SNF is skilled nursing facility. ICF is intermediate care facility.

SOURCE: Morris, J. N., Dreyer, P., Sherwood, C. C., and Atkins, C.: *Nursing Home Quality Assurance by Random Sampling*. Final Report. Contract No. 11-P-98260. Prepared for Health Care Financing Administration. Boston, Massachusetts Department of Public Welfare, Medical Assistance Division, Sept. 1987.

The average percent agreement for the 22 items was 95.9. This would seem to be a quite satisfactory percentage, given the possibility of some degree of clerical and data processing errors. It would seem unreasonable to expect a much higher degree of agreement.

The percent agreement for each of the 13 teams is shown in Table 2. Although some variation occurred, agreement appears to be satisfactorily consistent across the evaluation teams.

Spearman-Brown coefficient of reliability

The findings that resulted from application of the Spearman-Brown analysis-of-variance-based formula for reliability of paired judgments are presented in Table 3. They tended to be somewhat lower than the percentages of agreement. The Spearman-Brown coefficients would have been higher had there been greater variability among the judgments, given the same level of disagreement among the judges. Reliabilities above .70 are generally regarded as satisfactory; the findings here would thus appear to be generally quite satisfactory. Clarification of the criteria and perhaps some additional team training are indicated for the item "Restraints." This low level of reliability resulted from a relative lack of team consensus in those few instances in which excessive levels of restraint might have been present. A

Table 2

Agreement of individual teams performing nursing home inspection of care reviews for 41 facilities: Massachusetts, 1984 and 1985

Team number	Number of facilities	Number of judgments	Number of agreements	Percent agreement
Total	41	924	884	95.7
1	4	88	88	100.0
2	5	132	126	95.5
3	2	44	41	93.2
4	5	110	102	92.7
5	1	22	21	95.5
6	1	22	19	86.4
7	1	22	21	95.5
8	1	22	22	100.0
9	1	22	22	100.0
10	1	22	21	95.5
11	6	132	127	96.2
12	9	198	191	96.5
13	4	88	83	94.3

SOURCE: Morris, J. N., Dreyer, P., Sherwood, C. C., and Atkins, C.: *Nursing Home Quality Assurance by Random Sampling.* Final Report. Contract No. 11-P-98260. Prepared for Health Care Financing Administration. Boston, Massachusetts Department of Public Welfare, Medical Assistance Division, Sept. 1987.

Table 3

Spearman-Brown coefficient of reliability for items in nursing home inspection of care reviews: Massachusetts, 1984 and 1985

Item	Coefficient of reliability
Staff attention to residents	1.00
Staff appearance	1.00
Staffing shortage in resident care	1.00
Behavioral problems ignored	1.00
Patient mood	.80
Patient grooming	.88
Bedsores	.85
Contractures	.85
Patient positioning	.72
Lack of privacy	.72
Physical abuses	1.00
Resident's general physical condition	1.00
Catheters	.88
Restraints	.39
Level of alertness	1.00
Meaningful activities	.80
Dirty facility	.87
Odorous facility	.82
Safety hazards	.75
Patient complaints	1.00
Level II (Medicaid SNF) residents that are ambulatory	.88
Level III (ICF) residents that are ambulatory	.71
Average reliability for 22 items	.86

NOTES: SNF is skilled nursing facility. ICF is intermediate care facility.

SOURCE: Morris, J. N., Dreyer, P., Sherwood, C. C., and Atkins, C.: *Nursing Home Quality Assurance by Random Sampling.* Final Report. Contract No. 11-P-98260. Prepared for Health Care Financing Administration. Boston, Massachusetts Department of Public Welfare, Medical Assistance Division, Sept. 1987.

sufficiently high degree of agreement among reviewers provides a basis for confidence in the data; measurement error is not a serious threat to the validity of the experiment.

Application of experimental design

According to DPH policy, experimental facilities should receive a full review if they failed the walk-through (had what appeared to be obvious deficiencies) or if too many of the randomly sampled residents were found to have deficiencies. Full reviews were also mandated when there was a change in ownership, if the facility was decertified, or if the facility was small (having 15 or fewer Medicaid residents). In 1984, 354 out of a possible 463 experimental facilities (76 percent) received only a sample review. This compares with 375 out of 460 (81 percent) in 1985. Thus, only about 20 percent of the experimental facilities received a full review.

At the same time, the number of facilities failing the walk-through procedures dropped precipitously from 6.9 percent in 1984 to 0.7 percent in 1985. This calls into question the appropriateness of the decrease in the number of facilities receiving a full review in 1985. Nevertheless, this change in rate reflects a purposeful alteration in the criteria for determining whether the walk-through was failed—a better focus of our concern may be on the usefulness of the walk-through procedures themselves, at least as applied in this demonstration.

Time saving

A major goal of the demonstration was to reduce time spent on IOC without harming the integrity of the IOC decisionmaking process. In this section, the inspector time-reporting (ITR) data by IOC survey teams in 1985 are described. The survey teams consisted of a registered nurse (RN) and a graduate social worker (MSW). They recorded the amount of time spent on a variety of tasks, including travel time and document time (time spent doing paper work at the office).

Table 4

Means and ranges of time required to conduct inspection of care reviews, by type of inspector and type of facility: Massachusetts, 1984 and 1985

Review status of facility	Number of facilities	All inspections		Type of inspector				Type of facility			
				RN		MSW		Skilled nursing		Intermediate care	
		Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
		Inspection time in hours									
Total	384	35	7-201	17	2-98	18	2-124	24	3-130	24	3-127
Control	39	63	7-201	30	3-98	33	4-124	47	4-130	36	5-113
Experimental	345	32	7-149	16	2-70	16	2-89	21	3-109	22	3-127
Sample review	278	28	12-56	14	3-30	14	2-34	17	3-49	20	3-45
Full review	67	48	7-149	23	2-70	26	4-89	44	7-109	34	8-127

NOTES: RN is registered nurse. MSW is graduate social worker.

SOURCE: Morris, J. N., Dreyer, P., Sherwood, C. C., and Atkins, C.: *Nursing Home Quality Assurance by Random Sampling.* Final Report. Contract No. 11-P-98260. Prepared for Health Care Financing Administration. Boston. Massachusetts Department of Public Welfare, Medical Assistance Division, Sept. 1987.

As can be seen from Table 4, on average, approximately twice as much time (63 hours) was spent inspecting a control facility as an experimental facility (32 hours), and the pattern was virtually the same for RNs and MSWs (30 hours versus 16 for RNs and 33 hours versus 16 for MSWs). One interesting finding is that considerably more time was spent inspecting the control SNFs (a mean of 47 hours) than in inspecting control ICFs (36 hours). The means were virtually identical for experimental SNFs and ICFs; the SNF mean was actually slightly lower. The means for the sampled experimental facilities were, as would be expected, even lower.

Impact of the intervention

Applying level of care standards

Our inspection of available data appeared to reveal instances in which survey staff did not apply DPH criteria for determining whether or not a facility failed the IOC review process based on level of care standards. Many facilities that appeared to require a full review based on IOC deficiencies were not so reviewed. Discussion with staff revealed a complex set of operational exceptions. For one, if a reviewer identified a resident who had been placed in an inappropriate level of care, and the resident's record at the facility indicated that the facility knew the resident was "out of level" and an alternative placement was being sought, the resident was not included in the calculation of failure status.

Other exceptions occurred in facilities in the more rural western part of the State, where there were difficulties in securing necessary transfers, because of the unavailability of alternative beds. In this instance, residents were excluded from the calculation of failure status if two conditions were met. First, it was necessary for surveyors to indicate that the resident had been placed in an inappropriate level of care; second, the surveyors had to have determined that, given the absence of alternative beds in the area, keeping the resident in his or her current bed status was appropriate. In addition, if a facility had assessed a resident as out of level for more than 1 year, the resident was not used in the calculation of failure.

Of these exclusionary criteria, most were used in this analysis, although only partial information was available in DPH files on residents for whom alternative placements were being actively sought by the facilities. These data were available on handwritten documents in the central office files and, when available, the residents identified were not used in calculating whether or not the facility had an unacceptable number of residents in the inappropriate level of care category.

Using the preceding adjustment criteria, the findings come as close as possible to reflecting accurately the pass or fail status of facilities, based on DPH criteria.

The primary impact question of this experiment is whether the demonstration protocol resulted in facility failure rates for experimental nursing homes that are comparable to those of control nursing homes; that is, were the deficient facilities identified?

The annual IOC process seeks to verify whether SNF and ICF residents are appropriately placed and whether they are receiving adequate care. The unit of analysis is the facility, and failure status is based on aggregated data for Medicaid residents surveyed in the nursing home. A facility fails if too large a proportion of surveyed residents are found to be inappropriately placed or if too many quality deficiencies are found for these residents.

The IOC algorithm used by the Department of Public Health determines facility failure based on one or both of two criteria:

- The total number of standard quality of care variables, when summed across all residents in a facility, is greater than 25 percent of all assessments (a DPH-established standard that indicates that the home is excessively deficient in responding to the care needs of residents).
- The number of residents in an inappropriate level of care is greater than 20 percent of the number of Medicaid residents in the facility.

Using these criteria for ICFs, the experimental and control, pass or fail distributions were not significantly different in 1984 or 1985 (Table 5). The estimates of the proportion of facilities that failed in the two samples were within the range of what might have been expected by chance alone.

For SNFs, only 1985 data were available, and, once again, there were no significant differences in the experimental and control distributions of failed facilities. The two values were within the range of what might have been expected by chance alone (Table 6).

Finally, we estimated the total number of facilities that failed the IOC review process (Table 7). Once

Table 5

Number of facilities and rate of failure of inspection of care reviews, for intermediate care facilities, by review status of facility: Massachusetts, 1984 and 1985

Review status of facility	1984			1985		
	Percent failed	Number	Significance level	Percent failed	Number	Significance level
Experimental	6.0	466	.74	2.6	422	.42
Control	7.3	41		4.8	42	

SOURCE: Morris, J. N., Dreyer, P., Sherwood, C. C., and Atkins, C.: *Nursing Home Quality Assurance by Random Sampling*. Final Report. Contract No. 11-P-98260. Prepared for Health Care Financing Administration. Boston. Massachusetts Department of Public Welfare, Medical Assistance Division, Sept. 1987.

Table 6

Number of facilities and rate of failure of inspection of care reviews for skilled nursing facilities, by review status of facility: Massachusetts, 1985

Review status of facility	Percent failed	Number	Significance level
Experimental	3.9	281	.83
Control	3.1	32	

SOURCE: Morris, J. N., Dreyer, P., Sherwood, C. C., and Atkins, C.: *Nursing Home Quality Assurance by Random Sampling*. Final Report. Contract No. 11-P-98260. Prepared for Health Care Financing Administration. Boston. Massachusetts Department of Public Welfare, Medical Assistance Division, Sept. 1987.

Table 7

Number of facilities and rate of failure of inspection of care reviews for all facilities, by review status of facility: Massachusetts, 1985

Review status of facility	Percent failed	Number	Significance level
Experimental	3.6	470	.35
Control	6.4	47	

SOURCE: Morris, J. N., Dreyer, P., Sherwood, C. C., and Atkins, C.: *Nursing Home Quality Assurance by Random Sampling*. Final Report. Contract No. 11-P-98260. Prepared for Health Care Financing Administration. Boston. Massachusetts Department of Public Welfare, Medical Assistance Division, Sept. 1987.

again, there were no statistically significant differences, and the experimental and control estimates of deficient facilities were within the range of what would have been expected by chance alone.

It is important to remember that the estimate for the experimental facilities is based only on those nursing homes in which surveyors ultimately assessed all Medicaid residents in the facility. We again note that this did not occur in a number of instances in which distributions for the sample of Medicaid residents assessed in the facility suggested facility failure and therefore should have required a full review of all such residents. If we include these sites in our analysis for the full complement of nursing homes, in 1985, more of the experimentals than controls would have failed (8.5 percent versus 6.4 percent). Thus, had the surveyors fully carried out the specified protocol, there is reason to believe that not only would there have been an absence of a statistically significant difference (as previously indicated), but that the discrepancy in the raw proportions of failed facilities would no longer apply—that is, the number of failed facilities identified for the experimental and control protocols would not differ. On an absolute basis, the survey process would not have missed homes with deficiency problems.

Impact on quality

A second impact question addressed in this research relates to the quality of care provided in control and experimental homes. The annual IOC process seeks to verify whether SNF and ICF residents are appropriately placed and whether they are receiving adequate care. It is this process that was systematically modified in the demonstration. Independent of the modified IOC review activity, a second annual survey was completed at the facility level by another survey staff for purposes of Medicaid and Medicare certification. In this second certification survey, limited quality of care information was gathered relative to the continued licensure and certification status of the facilities. Certification survey results for each facility were used to determine the effectiveness of the separate IOC sampling methodology in detecting quality problems as found by the certification process.

Distributions of the number of deficiencies for these Massachusetts facilities were virtually identical in each of the two study years. More than 40 percent of facilities had no deficiencies; approximately two-thirds of the facilities had either no deficiency or only 1; and three-quarters of the facilities had 2 or fewer deficiencies in the 56 items surveyed. At the other extreme, depending on the year, either 3 or 5 percent of facilities exhibited seven or more deficiencies.

In Table 8, one can see that experimental facilities had, on average, about one and one-half deficiencies, whereas control facilities had approximately two deficiencies. Neither of the 2 annual mean comparisons was significant at either the .05 or .10 levels. This suggests that there is no reason to suspect that the experimental IOC intervention resulted in either improved or deteriorated quality of care in Massachusetts nursing homes—at least when assessed by the admittedly crude measure of facility compliance with licensure and certification standards. On a raw-score basis, however, we note that these slight differences in mean rates may be attributed to two factors. First, there were more experimental facilities with zero deficiencies than control facilities

Table 8

Mean number of certification of licensure deficiencies in surveyed facilities, by review status of facility: Massachusetts, 1984 and 1985

Review status of facility	1984		1985	
	Mean	Number	Mean	Number
Experimental	1.55	497	1.50	490
Control	2.02	43	1.98	48
Significance level	.15		.10	

SOURCE: Morris, J. N., Dreyer, P., Sherwood, C. C., and Atkins, C.: *Nursing Home Quality Assurance by Random Sampling*. Final Report. Contract No. 11-P-98260. Prepared for Health Care Financing Administration. Boston. Massachusetts Department of Public Welfare, Medical Assistance Division, Sept. 1987.

(44 percent versus 35 percent in 1984; 46 percent versus 29 percent in 1985). In fact, if we based the analysis on the dichotomy of whether deficiencies were or were not present, the chi-square statistic for the 1985 data would approach statistical significance. Second, a greater proportion of control facilities had six or more deficiencies (14 percent versus 8 percent in 1984; 11 percent versus 6 percent in 1985).

Summary

This study describes an innovative demonstration program for implementing IOC reviews of Medicaid-eligible nursing home residents in Massachusetts nursing homes. The demonstration substituted the review of a random sample of residents for the review of all residents required by the Social Security Act. The new sample-review procedures were implemented in all but 50 of the State's 550 nursing homes; these 50 facilities served as a randomly selected control group, in which full-review procedures were retained. The new procedures included a facility walk-through to determine quickly whether a full review was necessary; otherwise, a sample review was conducted and the results compared with preestablished criteria to determine the acceptability of the sample findings. Unacceptable findings led to a full review.

The experimental IOC intervention did not result in a lower quality of care in Massachusetts nursing homes; deficient facilities appear to be equally identifiable using the random or full-sampling protocols; and the process can be completed with a considerable savings of surveyor time. More specifically:

- Walk-through items were reliably scored by surveyors. The mean level of agreement across 13 independent teams was 95.8 percent, and the average reliability coefficient across the 22 walk-through items was .86. More importantly, however, few homes failed the walk-through (less than 1 percent in 1985), and the usefulness of the procedure (from a cost-benefit perspective) can be questioned.
- The sample procedures took considerably less time than the full-review procedures. The average number of hours to complete an IOC was 63 in the full-review control facilities versus 32 in the experimental facilities.
- Decision rules for performing full reviews were not always applied as required by the demonstration protocol.
- Differences in the ability of experimental and control procedures to detect poor facility performance were not statistically significant; however, extrapolation of actual detection rates suggests that more poorly performing experimental homes would escape detection than would poorly performing control homes.
- The aforementioned difference in rates disappears when sample failure is recalculated according to the appropriate demonstration protocol.
- Quality of care (as measured by performance on annual licensure certification surveys) was no worse in the sample-review facilities than in the full-review facilities.

These findings taken together suggest that a sampling approach to quality assurance can be as effective as full review in detecting poor facility performance, if sampling procedures are followed according to protocol. Additional tests of sampling procedures should include stronger controls to assure that pass or fail decisions are made according to specified procedures. In addition, when implemented fully, there would appear to be little reason to replicate the walk-through process in the future.

Finally, it is important to note that this research does not specifically address a number of changes in the survey process that have been recommended by others (Institute of Medicine, 1986; Zimmerman et al., 1984), including the assessment of non-Medicaid residents, the joining of the IOC and certification reviews, the expansion of outcome areas, and the need for better followup mechanisms. These are all important issues but are largely beyond the scope of our research. We have focused on the issue of sampling residents for the IOC review and conclude that quality of care was not harmed and that deficient facilities appear to be equally identifiable using the random or full-sampling protocols.

References

Institute of Medicine: *Improving the Quality of Care in Nursing Homes*. Washington, D.C. National Academy Press, 1986.

Zimmerman, D., Egan, R. J., Gustafson, D. et al.: Evaluation of the State Demonstrations in Nursing Home Quality Assurance Processes. A Final Report to MPR A Mathematica Company. Madison, Wis. University of Wisconsin, 1984.