Survey-based Indices for Nursing Home Quality Incentive Reimbursement

by Thomas R. Willemain

Incentive payments are a theoretically appealing complement to nursing home quality assurance systems that rely on regulatory enforcement. However, the practical aspects of incentive program design are not yet well understood. After reviewing the rationale for incentive approaches and recent State and Federal initiatives, the article considers a basic program design issue: creating an index of nursing home quality. It focuses on indices constructed from routine licensure and certification survey results because State initiatives have relied heavily on these readily accessible data. It also suggests a procedure for creating a survey-based index and discusses a sampling of implementation issues.

Nursing Home Quality Incentives

There are about 1.3 million Americans residing in roughly 19,000 nursing homes (Fox and Clauser, 1980). This is a frail, vulnerable population whose lives have been disrupted to the point of requiring institutionalization. For some, the quality of life and quality of care in their institutions are intolerable. For many others, quality is barely tolerable. These problems are chronic and well known (Mendelson, 1974; N.Y. State Moreland Act Commission, 1975; Warner, 1976; Ruchlin, 1979; Kane et al., 1979; Vladek, 1980; Butler, 1980). The recent Health Care Financing Administration (HCFA) "white paper" on long-term care (HCFA Office of Policy Analysis 1981) identified nursing home quality as a major problem and quality assurance as a "major design issue in any long-term care financing program."

Command-and-Control versus Incentives

Most efforts by State and Federal officials to assure quality in nursing homes rely on regulatory policing techniques described by Schultze (1977) as "command-and-control" methods. The limitations of these

methods in the nursing home case are well documented (Beatrice 1980, Butler 1980). The problems derive in large part from a classic regulatory dynamic in which regulators with few resources, ambiguous roles, limited maneuvering room, bureaucratic inertia and restricted clinical knowledge attempt to force changes in the behavior of a much more knowledgeable and maneuverable industry (Schultze 1977).

These difficulties have led some to advocate the development of incentive-based approaches to quality assurance. Usually, the incentives are conceived of as monetary and overlaid on rate-setting procedures (Ruchlin 1979). Some go even further, calling for dismantling much of the regulatory apparatus and substituting a market-like system of outcome-based rewards and penalties (Kane and Kane 1978). Such thinking has stimulated a small but significant number of Federal research efforts and State program initiatives. However, there is certainly no universal agreement on the merits of quality incentive reimbursement in general (Vladek 1980) or of specific forms such as outcome incentives (Willemain 1981). In fact, many of the State initiatives to date have met with limited success, at best.

Federal Research Initiatives

The National Center for Health Services Research (NCHSR) is conducting a controlled experiment in 30 skilled nursing facilities in the San Diego area (Weissert et al., 1980). Three types of incentives are being

tested: admission incentives to facilitate placement of "difficult" residents, discharge incentives to facilitate more appropriate use of beds, and a restricted set of outcome incentives, such as maintenance of good skin condition in comatose residents. Preliminary results are expected in late 1983. The NCHSR has also sponsored a project at the RAND Corporation devoted completely to outcome incentives over several dimensions of resident condition, and is underwriting work by the author on general issues in incentive program design.

State Program Initiatives

Butler (1980) in her excellent work on the quality assurance problem, reviewed the efforts of several States to use reimbursement as a quality assurance device. The State of Michigan uses a two-part reimbursement system that pays costs plus a profit factor. The profit factor is reduced in proportion to the degree of noncompliance with Medicaid-Medicare conditions of participation.

Until 1980, the State of New York used the incentive system described by Butler. In that system, the State used licensure and medical review data to classify homes as "very good," "satisfactory" or "needs improvement." Allowable cost ceilings were set differently for the three groups, with unsatisfactory homes slated to receive reimbursement less than costs. However, the system was abandoned because penalties were never assessed in the belief that poor homes could not improve without adequate funding. A system of fines and public disclosure of quality ratings has apparently replaced the system Butler described.

The Moreland Act Commission (1975) reported some of the history of the system in New York. Beginning in 1970, New York attempted to deny efficiency incentive payments to nursing homes with "significant operating deficiencies." The Commission noted that operational definitions of "significant operating deficiencies" were never developed, leading the state Commissioner of Health to recommend that efficiency incentives not be awarded in order to avoid embarrassing awards to poor quality homes that achieve their efficiencies by reducing the quality of services. Ultimately the State legislature did require the link between quality and reimbursement reported by Butler, but the process was difficult then and, given recent developments, has obviously continued to be difficult.

A similar story is emerging in Connecticut. For the past five years, Medicaid payments to nursing homes included efficiency incentive payments which were to be paid only to homes "in full compliance" with State and Federal standards. A one-man grand jury investigation discovered that not one home had been denied the incentive (Condon, 1981; New York Times, 1981). The State legislature responded by preparing legislation authorizing a compliance-based quality index (Foohey, 1981).

Massachusetts instituted in 1979 a set of nursing home incentives relating to willingness to admit Medicaid residents, management efficiency, occupancy rate, and "quality." Quality was defined in terms of a "survey compliance score" formed as a weighted sum of scores on several hundred individual items in the annual licensure and certification surveys. Weights were assigned by State personnel on the basis of their perceptions of the items' salience for quality. Moll (1981) evaluated the early performance of the Massachusetts system, finding no evidence of success:

"So far (through mid-1980), the program has had no significant effect on quality, as measured by the survey compliance score. There appears to be a combination of reasons why the program has failed to produce the anticipated effects on quality. Basically, they are: 1) that the incentive reimbursement system was not well understood, at least in the first year of the program, and 2) that the signals that are provided under the system are not focused clearly on the goal of improving quality."

At the time of this writing, cutbacks in Federal funding of the survey process, rumored changes in the Medicaid conditions of participation and State budget concerns have combined to leave the future of the Massachusetts system unclear.

Summary

Summarizing the state of the art, it is clear that the inherent limitations of command-and-control strategies for quality assurance make incentive strategies look relatively attractive—at least from the somewhat distant vantage point of theory. Ongoing Federal research studies may enrich both our conceptual and empirical understanding of incentive design issues. State initiatives demonstrate that many conceivable designs have not been attempted and that those systems in place merit careful scrutiny. In a word, the art of quality incentive reimbursement is in a formative stage.

Nursing Home Quality Indices

In broad terms, the problem of designing a quality incentive reimbursement system can be decomposed into two parts: "designation" and "reward". Designation identifies providers for reward. For instance, in the Massachusetts system, homes scoring above 80% on the survey compliance score are designated as "acceptable" and given a financial bonus, while homes scoring above 95% are designated as "outstanding" and rewarded somewhat more. The cornerstone of such an incentive system is the quality index used in the designation process.

Quality as an Artifact

States have tended to construct their quality indices from items in routine licensure and certification surveys. This tendency probably arises from the convenience and economy of using readily available data and from the apparent pertinence of the survey items. Since it seems very likely that any further State initiatives will also proceed along these lines, this article focuses on how best to construct such an index.

However, this is not to deny the existence of doubts about the validity of any index based on survey information. Much of the information in such a survey focuses on inputs, such as staffing levels and cleanliness; some focus on elements of process, such as the appropriateness and timeliness of services. Many observers hold that these items are perhaps necessary but not sufficient conditions for quality, and therefore that indices constructed from them cannot adequately distinguish the best facilities. In the words of Donabedian (1981):

"... in order to achieve specificity and a reasonable level of completeness, there is a tendency to attenuate the definition of quality until it is no more than a shadow of its more real, more fully rounded self."

While this point is a powerful one, there are several reasons why it does not necessarily signal the demise of survey-based indices. First, there are the lessons of recent history, in which States have obviously responded to both the face validity and low cost of survey-based indices. Second, there are the obvious practical difficulties of supporting the costs of more subtle and sensitive designations and of establishing that these methods are neither arbitrary nor capricious. Third, there is the fundamental point that in fact there is no prior reality to the notion of quality.

Quality need not be thought of as a tangible but obscure characteristic that is knowable only to the extent that we develop an adequate technology of measurement. Recalling Donabedian's image, our fate need not be that of merely gazing on the shadow of an object hidden from view. Rather than regarding quality as an elusive but nature-given entity, we can regard quality as an artifact, as something we create. In this view, we are more engineers than scientists, making rather than discovering a sense of quality. A quality index embodies our preferences for service attributes. As Kincaid (1981) said,

"... as a practical matter, quality is the degree of conformity to an agreed-upon standard, and without the standard we cannot know whether we have quality. In other words, quality in patient care need not be seen as too fragile or too tenuous to be defined (at least in its technical aspect). Its real working definition is 'quality is what we agree it is ...' "

Policy makers can systematically and fairly arrive at a sense of what kind of service they wish to provide and then carefully embody much of that sense in an index based on the kinds of items found in licensure and certification surveys. While such an index will never embody everything the policy makers might desire, this lack of comprehensiveness is not reason in itself to quit the effort. The indices should be constructed in a novel way and their use tempered by recognition of their deficiencies. The empirical studies reviewed in the next section suggest enough potential value in survey-based indices to justify their consideration.

Studies of the Validity and Reliability of Survey-based Indices

The literature contains two notable papers that test construct validity by comparing survey-based indices against implicit judgments of quality. Linn (1974) used regression analysis to predict subjective assessments of overall quality of patient care elicited from social workers familiar with 40 facilities. The predictors consisted of yes/no ratings elicited from two researchers on 71 items pertaining to physical plant, meals, administrative policies, safety, services, records, and personnel practices, together with bed total, staffing hours, cost, turnover rate, and staff-to-patient ratio. Linn was able to account for about 25% of the variance in the implicit ratings. Whether the implicit ratings were correct and whether the correspondence between the implicit ratings and the others was good enough for practice remain uncertain.

Similar empirical work was conducted by Gustafson et al. (1980), who reported the results of a pilot study aimed at focusing and streamlining the survey process. The pilot study obtained four data sets on eight facilities. One data set consisted of implicit judgments of overall quality made by six rating teams. A second data set consisted of scores generated by five of the teams using a "screening instrument" specially designed for quality assessment. The third data set contained the ratings of a team using a subset of the usual survey items judged to be most salient for quality. The final data set consisted of a count of deficiencies in the usual survey. Gustafson et al. reported rank-order correlations between the implicit quality ratings and the other data sets as follows: 0.76 with the scores produced by the special screening instrument. 0.32 with the abbreviated survey instrument and 0.11 with the count of deficiencies. Although this pilot study of eight facilities has low statistical power, it does suggest that a survey-type approach can perform even better than Linn reported when the survey is specially designed. Unlike Linn's study, however, the same raters provided both the implicit ratings and the screening scores, so the two data sets do not represent completely independent readings.

Two other empirical studies provide indirect evidence of the link between quality and scores on weighted-sum indices computed from survey data. The Massachusetts Department of Public Health (1980) conducted two analyses of the reliability of its survey compliance score. One involved obtaining three simultaneous assessments of four facilities to study inter-rater reliability among surveyors. The three assessments produced a difference in scores ranging from 5 to 11 points per home on a 100-point scale; the reliability coefficient was 0.9, indicating that inter-rater reliability was fairly high. On the other hand, the other analysis compared the (product-moment) correlation between successive years' scores on the index and found a value of only 0.59. The Department argued that this figure represents a lower bound on the reliability of the index because a facility's quality will fluctuate from year to year, reducing the correlation. However, in practice a survey-based index will in fact be using the last year's data, so both surveyor unreliability and yearly fluctuations in quality are phenomena that must be accounted for when assessing the reliability of an index.

The other empirical study dealing indirectly with the validity of weighted-sum indices was conducted on 54 facilities by Murphy (1975) for the New York State Moreland Act Commission on Nursing Homes and Residential Facilities. Murphy studied the correlation between indices formed from licensure and certification surveys and indices based on the periodic medical review (PMR) program in skilled nursing facilities. PMR is a routine survey that, unlike licensure and certification surveys, focuses on individual residents and requires some actual contact with the residents. Murphy created 10 indices based on PMR results and 10 based on licensure and certification surveys. His 10 PMR indices showed moderately high intercorrelation (typically about 0.30) and his 10 licensure and certification indices showed very high intercorrelation (typically about 0.90). However, the correlations between indices of one type and the other were quite low, never exceeding 0.26 in 100 cases. While this might mean that indices computed from the two data sets provide independent information and could usefully be combined, Murphy was wary of drawing this conclusion. He paid particular attention to areas of "overlap" between the two surveys to see whether they provided consistent ratings of the same attributes and was disappointed to find little consistency between the two data sets. For instance, the correlation between the PMR score for "restorative nursing" and the certification score for "rehabilitative nursing care" was only 0.04.

One must draw mixed conclusions from these studies. If we regard implicit ratings by experienced observers as adequate referents, then survey-based indices provide a moderately effective alternative. Furthermore, it appears that if special efforts are made, both the reliability and validity of survey-based indices can be improved substantially over what they would be in haphazard realizations. Nevertheless, survey-based indices are rather blunt instruments whose appeal has rested primarily on face validity and convenience. They might best be thought of as indicators of conditions that are necessary but not sufficient for sophisticated definitions of quality.

Recommended Procedure for Creating a Survey-Based Index

Exigencies of practice can be expected to tempt the program designer to construct a survey-based index in an *ad hoc* way. It is easy to imagine that the design of the index may not be systematic after much energy has been expended persuading providers that quality incentive reimbursement is not a danger, budget officials that it is not a boondoggle, and client advocates that it is not a sham. Faced with literally hundreds of survey items, the designer may find great appeal in forming a simple weighted sum of item scores, arriving at the weights in some expedient manner.

Such a simple weighted sum was used in the original design of the Massachusetts system. The annual survey consisted of several hundred items, with scores of 0, 1 or 2 assigned to items for which the nursing home was considered to be in varying degrees of compliance and scores of 3, 4 or 5 assigned for various degrees of noncompliance. The survey compliance score was formed by multiplying these item scores by item weights. The weights ranged from 1 to 10 according to the opinions of a group of experienced long-term care inspectors about the items' importance for quality.

A better index can be devised by following the procedure described next. This procedure avoids several problems inherent in an index like that devised in Massachusetts, although it is somewhat more complicated.

The proposed method proceeds in three steps. The first two steps produce quality salience weights in the form of a pool of points associated with each survey item, with more important items receiving more points. The third step competitively divides the points associated with each item among facilities: those that do relatively well receive a disproportionate share of the points, subsidized by facilities that do relatively poorly. The method will be described in terms of the original Massachusetts system, but the general approach is useful for any survey data set with many items and two or more "passing grades".

Weighting Survey Items

At least three principles should guide the weightings that begin the process. First, the weights assigned to each of the 18 Medicaid conditions of participation (such as nursing, dietary) should be consciously controlled. This means the varying number of items under each condition of participation cannot

influence the relative weight of that condition. Second, distinguish between "policy weights" and "clinical weights." A group with the appropriate standing should assign weights to each condition of participation in a way that defines a balance of aspiration for "quality of care" and "quality of life". It is these weights that define "quality". Then clinical experts should subdivide the weights in each condition of participation to reflect the technical importance of each item in that condition. Third, promote simplicity and clarity by focusing attention on only the most salient items in each condition of participation.

The process of assigning policy weights might proceed as follows. First, an appropriate group must be established. This group should be representative of constituencies whose expertise or vantage point entitles them to a voice in defining quality. In Massachusetts, this group has been a task force with membership from State agencies, nursing home associations, academia and a consumer advocacy group. To an important extent, the legitimacy of the index will derive from the perception that policy weights were assigned by a balanced, broad-based process. Second, those doing the weighting might rank-order the 18 conditions of participation and some aggregate ranking be determined from the individuals' rankings. Carter and Kosinski (1981) documented the rankings produced by a task force in Michigan that reached consensus on the most important among 182 items. Lastly, using the rankings as a guide, a fixed pool of points should be apportioned among the conditions of participation. There is no single, compelling way to execute any of the steps of constituting a group, ranking the conditions of participation, then assigning points, but even a more or less arbitrary process can proceed fairly, thoroughly and wisely.

Dividing Points Among Facilities

The two-stage process of policy and clinical weighting will produce a pool of points for each item. This pool should then be divided among facilities, again according to three principles. First, since every facility should be in compliance, no points should be given to a facility that is not in compliance on an item. In effect, this treats the surveyor ratings as 0, 1, 2, 3-5, which we might think of as three passing and one failing grades: A, B, C and F.

Second, the significance of a given item score should depend on the distribution of scores across all nursing homes on that item. If all facilities score equally on an item, all should receive an equal share of the pool of points associated with that item. Facilities that do poorly on an item on which most others do well should receive a smaller share of the points. Conversely, facilities that do well on an item when most do poorly should receive a larger share. This competition adds an important dimension missing in a system that does not use the distribution of item scores because it takes account of the discriminating

power of each survey item. With the proposed change, the greatest number of points will be awarded for doing better than most facilities on a clinically important item in a high-priority condition of participation. Note that there is "ratcheting" built into the competitive point-sharing formula. Improvements beyond the average level of compliance are rewarded, and when most homes are solidly in compliance, backsliding is costly. An additional advantage here is that the index automatically adjusts priorities to the current state of compliance across survey items by always offering extra rewards for doing well on an item on which most facilities have trouble.

Third, the discrimination principle must be moderated by concern for surveyor unreliability and bias, so the share differentials among A, B and C grades must not be too extreme. Taken together, these principles should produce an appropriate allocation of points across nursing homes. We can implement these three principles of point sharing if we are willing to pay the price of a little complexity in the system of point allocation.

Four examples of how this method works are given in Exhibit A. Example 1 shows that, if all facilities score the same, they all receive the same share of the points associated with that item. Example 2 shows that if nearly all facilities score A on an item but a few score lower, all the A's receive a bit (3%) more than an equal share, while the B's and C's subsidize the A's by receiving less than equal shares (40% and 25% of equal shares, respectively). Example 3 shows an item for which it is tougher to do well. In this case, a score of A merits a larger (10%) increase over an equal share, while scores of B or C are not penalized so heavily. Finally, example 4 shows an item which is very difficult to pass. Here an A score merits 2.52 times an equal share, but since even passing this item is an accomplishment and there are so few facilities that do very well, even a grade of C merits most (90%) of the points that would have been provided if all those passing shared equally. These four examples show how the proposed method takes proper account of the distribution of scores on an item to establish incentives and disincentives.

The method of share allocation works as follows. First the distribution of scores is used to convert the "grades" of A, B, C or F into numbers. These numbers are calculated by the method suggested by Mosteller and Tukey (1977), which satisfies the second principle by taking account of the distribution of scores. The numbers are then linearly scaled so that F corresponds to 0.0 and A corresponds to 1.0. This rescaling excludes points for homes not in compliance, in accordance with the first principle, and limits the range of the difference, in accordance with the third principle. The rescaled values are used as weights that, along with the number of facilities receiving each grade, determine the share of an item's point pool allocated to all the facilities receiving a particular grade on this item. Technical details are provided in Exhibit B.

Implementation Issues

A number of implementation issues have arisen in Massachusetts that are of general interest. For the most part they deal with the wider process of incentive reimbursement that rests on the quality index.

Piggybacking

The process of constructing a quality index can become an occasion to raise other issues, with the intention of piggybacking "fixes" in the overall reimbursement system onto the quality initiative. A case in point was the desire of some Massachusetts providers to build case-mix adjustments into the quality index. These providers argued that certain facilities were performing well given their case-mix and reimbursement rates, even though they received poor survey reports. The providers claimed that the index should make allowance for special difficulties. However, to grant that reimbursement rates may be too low and that case-mix matters to quality is not to accede to the distraction of adding case-mix factors onto the quality index. In the first place, the survey process should already be sensitive to case-mix and hold providers accountable for offering appropriate care. Secondly, adding "fudge factors" will destroy whatever usefulness the index has as a means of documenting the quality level in a home, whatever its determinants. A more appropriate response is to make the facility's base rate sensitive to case-mix (Willemain 1980a and 1980b), not to weaken the validity of the quality index.

A more subtle but nevertheless potentially damaging form of piggybacking is the dominance of the incentive program in Massachusetts by efficiency and access incentives. Moll (1981) pointed out that few homes in Massachusetts were denied incentive money because of a low survey compliance score; efficiency standards were most often the binding criterion. If the quality standards are set too low, the index might be thought of as an instrument for "laundering" money intended for purposes besides quality improvement. Though these other purposes may be laudable in themselves, they crowd quality issues off the agenda.

Clarity

Even a "simple" weighted-sum index can be difficult to understand, especially when embedded in a complicated set of incentive rules and regulations. Moll's (1981) analysis of a survey of providers showed that the index and the system built upon it were largely opaque to the providers. Under such circumstances, it is meaningless to speak of sending "signals" to the industry. Quality incentive systems and the indices on which they are based must be explained. In particular, it is desirable and feasible to use computer analysis of survey results to indicate to individual facilities where they lost points and how to improve.

Reactions

It is possible to foresee some of the likely reactions to the use of a quality index of the type recommended. One reaction that was triggered by the Massachusetts prototype was scrutiny of the survey process itself, as in the special study of surveyor reliability undertaken by the Massachusetts Department of Public Health (1980). While good professional practice dictates that the survey process be monitored and standardized, it seems that States have not been very aggressive in this regard. The special study in Massachusetts uncovered several surveyors whose judgments were aberrant. Providers are especially sensitive to the possibility of encountering a "tough" inspector, whereas client advocates fear the opposite. Documentation and standardization of surveyor judgments became important issues.

A second possible reaction is "distortion to the measure." If the index focuses on only the most salient survey items, there may be a tendency to ignore the items which do not figure prominently in the incentive system. The incentives should not be seen as entirely superseding the usual regulatory apparatus.

Finally, an interesting feature of the dynamic behavior of the index is that if the incentive system is successful it will gradually diminish the discriminatory power of the index, since the variation in item scores will decrease. This is as it should be, since if there is little variation in survey compliance there is correspondingly little reason for using a survey-based index to sort the facilities.

The Danger of Diversion

The greatest danger arising from using a survey-based quality index is that the long term care system may be permanently diverted from implementing a superior index, if one exists. The survey compliance score in Massachusetts was a way to begin the evolution of an effective quality incentive system, but there was conflicting evidence about the State's willingness to let the system evolve. The changes advocated above were discussed by the State task force, but on the whole the Department of Public Health was tentative about its responsibility to evaluate and modify the original system. The Department restricted the scope of the task force's work to minor modifications and ultimately dissolved the task force before changes were made. While it is true that quality rating is a sensitive process, and too many changes in the system will defeat the purpose of sending a clear signal on quality, it is also true that quality incentives are an important and promising approach and should be improved when possible.

Survey-based quality indices have proven attractive because they are inexpensive and accessible. This paper has reviewed their properties and recommended a way to construct them, but the basic question of whether they work well enough or should be replaced by other approaches remains unanswered. Survey-based quality indices may focus attention on marginal rather than comprehensive reform, and they may provide false assurance that quality will be improved through their use in incentive schemes. To bend an old aphorism, we do not yet know whether the easy is the enemy of the good.

EXHIBIT A: EXAMPLES OF POINT SHARING

Example	Score	% of Homes	Relative Share of Points[1]
1	A	100	1.00
	B	0	0
	C	0	0
	F	0	0
2	A	95	1.03
	B	3	.40
	C	1	.25
	F	1	0
3	A	80	1.10
	B	15	.63
	C	4	.36
	F	1	0
4	A	1	2.52
	B	4	1.79
	C	45	0.90
	F	50	0

[1]Ratio of share based on grade to equal share for all homes scoring C or better.

EXHIBIT B: METHOD FOR ALLOCATING POINT SHARES

Step 1: Convert ordinal grades to metric scores using Score (grade) = $(B(D) - B(d)/(D - d)$

where d = fraction of homes with better grade

D = fraction of homes with better or same grade

and $B(x) = x \log(x) + (1 - x)\log(1 - x)$

where $B(0) = B(1) = 0$.

Step 2: Rescale scores to range from 0 to 1 using Rescale (grade) = (Score(grade) − Score(F))/(Score(A) − Score(F)).

Step 3: Convert the rescaled scores into point share using Share (grade) =

$$\frac{\text{Number (grade)} \times \text{Rescale (grade)}}{\sum_{\text{grades}} \text{Number (grade)} \times \text{Rescale (grade)}}$$

Step 4: Each home then receives Points (item, grade) = (Share(grade)/Number(grade)) × Pool of points (item).

The details of example 4 in Exhibit A will illustrate this method.

	Grade			
	A	B	C	F
Fraction of homes	.01	.04	.45	.5
Fraction > Grade	0	.01	.05	.5
Fraction ≥ Grade	.01	.05	.5	1
Score	− 5.60	− 3.56	− 1.10	1.39
Rescale	1.00	.71	.36	0
Share	.05	.14	.81	0
Fraction of Passing Homes	.02	.08	.9	—
Ratio of Shares to Fraction Passing	2.52	1.79	.90	—

(Note: all figures rounded to 2 decimals)

For instance, the 1% of nursing homes scoring A on the item will divide among themselves 5% of the item's pool of points. If all homes with a passing grade (C or better) shared equally, the homes scoring A would share only 2% of the pool of points since they represent 2% of all passing homes. Thus the weighting scheme gives each home scoring A a total of 2.52 times as many points as a scheme that does not take account of the rarity of an A grade on this item.

References

Beatrice, D. F. "Licensure and Certification in Nursing Homes: Assuring Quality of Care?" In *Federal Health Programs: Problems and Prospects*, S. H. Altman, and H. M. Sapolsky, Lexington, MA: Lexington Books, 1981.

Butler, P. "Assuring the Quality of Care and Life in Nursing Homes: The Dilemma of Enforcement." *North Carolina Law Review* 57 (1979): 1317.

Carter, R. K. and R. D. Kosinski. "Doing Research in a Politically Charged Environment: The Michigan Study of Nursing Home Rules." *New England Journal of Human Services*, Spring 1981, p. 33.

Condon, T. "Nursing Home Payment Criticized." *Hartford Courant*, 12 March 1981.

Foohey, B. J. "Quality in Long-term Care: An Analysis of Factors that Affect the Inspection Process." unpublished Master of Public Health essay, Yale University, 1981.

Fox, P. D. and S. B. Clauser. "Trends in Nursing Home Expenditures: Implications for Aging Policy." *Health Care Financing Review*, Volume 2, No. 2 (1980): 75.

Gustafson, D. H., C. J. Fiss Jr., J. C. Fryback, P. A. Smelser, and M. E. Hiles. "The Wisconsin System: Quality of Care in Nursing Homes." *The American Health Care Association Journal*, September 1980, pp. 3-12.

HCFA Office of Policy Analysis. "Long-term Care: Background and Future Directions." Discussion paper, 1981.

Kane, R. and R. Kane. "Care of the Aged: Old Problems in Need of New Solutions." *Science* 200: 913, 17 May 1978.

Kane, R. *et al.* "The PSRO and the Nursing Home: Volume 1, An Assessment of PSRO Long-term Care Review." Rand Corporation report R-2459/1-HCFA, 1979.

Kincaid, W. H. "The Future of Quality Assurance in Patient Care." Proceedings of the Fifth Annual Symposium on Computer Applications in Medical Care, Washington, D.C., November 1981, p. 635.

Linn, M. "Predicting Quality of Care in Nursing Homes." *The Gerontologist*: 225, June 1974.

Massachusetts Department of Public Health. "The Reliability of Long-term Care Surveyor Judgments: Data, Analysis and Recommendations." unpublished, 1980.

Mendelson, M. A. *Tender Loving Greed.* New York: Knopf, 1974.

Moll, A. "Towards Improving Quality of Care in Nursing Homes: Can Incentive Reimbursement Work?" Policy Analysis Exercise, Kennedy School of Government, Harvard University, April 1981.

Mosteller, F. and J. W. Tukey. *Data Analysis and Regression.* Reading, MA: Addison Wesley, 1977, pp. 103-107.

Murphy, S. E. "Comparison of Periodic Medical Reviews with Surveys of Skilled Nursing Facilities in the New York City Health Region." in New York State Moreland Act Commission, *Regulating Nursing Home Care: The Paper Tigers*, Report 1, pp. 86-115, October 1975.

New York State Moreland Act Commission on Nursing Homes and Residential Facilities. *Regulating Nursing Home Care: The Paper Tigers.* 1975.

New York Times. "A Double Take on Nursing Homes." 15 March 1981.

Ruchlin, H. S. "An Analysis of Regulatory Issues and Options in Long-term Care." In *Reform and Regulation in Long-term Care*, V. LaPorte and J. Rubin, editors. New York: Praeger, 1979.

Schultze, C. L. *The Public Use of Private Interest.* Washington, D.C.: Brookings, 1979.

Vladek, B. *Unloving Care: The Nursing Home Tragedy.* New York: Basic Books, 1980.

Warner, G. M. "Nursing Home Investigations and Reform: Their Effects on the Quality of Care in New York State." Paper presented at the annual meeting of the American Public Health Association, Miami Beach, FL. 1976.

Weissert, W., W. Scanlon, T. Wan and D. Skinner. "Encouraging Appropriate Care for the Chronically Ill: Design of the NCHSR Experiment in Nursing Home Incentive Payments." Paper presented at the annual meeting of the American Public Health Association, Detroit, MI., 1980.

Willemain, T. R. "Nursing Homes Levels of Care: Reimbursement of Resident-Specific Costs." *Health Care Financing Review*, Volume 2, No. 2 (1980a): 47.

Willemain, T. R. "A Comparison of Patient-centered and Case-Mix Reimbursement for Nursing Home Care." *Health Services Research.* Winter 1980b, pp. 365-377.

Willemain, T. "Second Thoughts about Outcome Incentives for Nursing Homes." In *Research in Public Policy Analysis and Management.* J. P. Crecine, editor. Greenwich, CT: JAI Press, 1981.