Evaluation of Audit-based Performance Measures for Dental Care Plans

James D. Bader, DDS, MPH; Daniel A. Shugars, DDS, PhD, B. Alex White, DDS, DrPH; D. Brad Rindal, DDS

Abstract

Objectives: *Although a set of clinical performance measures, i.e., a report card for dental plans, has been designed for use with administrative data, most plans do not have administrative data systems containing the data needed to calculate the measures. Therefore, we evaluated the use of a set of proxy clinical performance measures calculated from data obtained through chart audits.* ***Methods:*** *Chart audits were conducted in seven dental programs—three public health clinics, two dental health maintenance organizations (DHMO), and two preferred provider organizations (PPO). In all instances audits were completed by clinical staff who had been trained using telephone consultation and a self-instructional audit manual. The performance measures were calculated for the seven programs, audit reliability was assessed in four programs, and for one program the audit-based proxy measures were compared to the measures calculated using administrative data.* ***Results:*** *The audit-based measures were sensitive to known differences in program performance. The chart audit procedures yielded reasonably reliable data. However, missing data in patient charts rendered the calculation of some measures problematic—namely, caries and periodontal disease assessment and experience. Agreement between administrative and audit-based measures was good for most, but not all, measures in one program.* ***Conclusions:*** *The audit-based proxy measures represent a complex but feasible approach to the calculation of performance measures for those programs lacking robust administrative data systems. However, until charts contain more complete diagnostic information (i.e., periodontal charting and diagnostic codes or reason-for-treatment codes), accurate determination of these aspects of clinical performance will be difficult. [J Public Health Dent 1999;59(3):150-57]*

Key Words: Effectiveness of care; use of services; performance measures; quality of health care; health benefit plans—standards, outcomes, and process assessment; dental records.

Several approaches for assessing the clinical performance of dental care plans could be of use in helping to document the "value" purchasers receive for their dental care expenditures. One approach to measuring plan performance assesses a plan's effect on enrollees' "oral health status," using clinical measures that typically comprise multiple dimensions of oral health (1,2). Another approach assesses a plan's performance through measures of oral health-related quality of life, which are determined either through interview or questionnaire (3). Measures of patient satisfaction with the dental care, provider, and dental care plan represent another approach to assessing plan performance, albeit more indirectly with respect to clinical outcomes (4). Yet another indirect approach is based on the use and the patterns of use of various treatment services (5,6), while a final approach examines specific clinical outcomes such as tooth loss, new caries, and change in periodontal status among plan enrollees (7).

Each of these approaches to assessing the clinical performance of a health care plan offers both theoretical and practical advantages; nevertheless, each also poses specific disadvantages. Measures of oral health status might provide a conceptually simple composite assessment of clinical outcomes associated with plan membership; but oral health status indices typically require costly clinical examinations, do not assess all outcomes (1,7), and may use unvalidated weighting schemes (7). Oral health-related quality of life measures can offer important information from the perspective of the enrollee; however, the measures are not yet fully developed and experience with their longitudinal use is minimal (3,8). Patient satisfaction measures have been used for many years by dental care plans, offering relatively inexpensive means of assessing popularity. However, the relationship between patient satisfaction and most aspects of clinical oral health is unclear (4). Use of services measures, which some plans currently examine in internal quality assessment programs, can provide information about the scope of services provided by a plan and the degree to which specific plan goals are achieved; but this information cannot be used to evaluate appropriateness of care, except in limited situations where the standard of care suggests that a specific service should be provided for all patients with a given condition. Unfortunately, in dentistry, few validated standards of care exist, and still less information is available in patients' charts about their specific clinical conditions (9,10). Finally, specific measures of clinical outcomes offer perhaps the most straightforward information about dental plan performance from a clinical perspective. Yet, their application is by necessity extremely limited

Send correspondence to Dr. Bader, Sheps Center for Health Services Research, University of North Carolina, CB #7590, Chapel Hill, NC 27599-7590. E-mail: jim_bader@unc.edu. Web site: http://www.shepscenter.unc.edu. Dr. Shugars is with the Sheps Center for Health Services Research; Dr. White is with Kaiser Permanente Center for Health Research, Portland, OR; and Dr. Rindal is with HealthPartners Research Foundation, Minneapolis, MN. Reprints will not be available. Supported by the Agency for Health Care Policy and Research, HS09453, and the Centers for Disease Control and Prevention. Manuscript received: 6/22/99; returned to authors for revision: 8/10/99; accepted for publication: 8/13/99.

due to the virtual absence of diagnostic data in dental charts with which to categorize patients' disease status and assess the outcomes of previous preventive and therapeutic treatment (10).

These limitations notwithstanding, there is a need to develop clinical performance measures for dental care plans. As dental PPOs and HMOs continue to increase their market shares (11), it is likely that private purchasers (employers) and public programs (Medicaid, Children's Health Insurance Program) will demand standardized comparisons of plan performance. A similar need in medicine led to the creation of the Health Plan Employer Data and Information Set (HEDIS®), the most widely used set of performance measures for medical plans (12). HEDIS® includes 71 reporting measures and 33 additional testing measures assessing health plan performance across eight domains, two of which—effectiveness of care and use of services—address clinical aspects of care.

In a project funded by the Agency for Health Care Policy and Research's "Q-SPAN" program, we have recently developed a set of dental plan performance measures—a report card—for assessing these clinical domains of effectiveness of care and use of services (13). The measures originally were developed for use in prepaid dental benefit programs, including indemnity and managed care plans. The measures are designed to be used in conjunction with what we anticipate will be the next generation of administrative data systems in dental care plans, namely, systems that include diagnostic information. Because the majority of dental care plans do not have such systems, these performance measures are not immediately universally usable. Therefore, we also have developed a set of proxy performance measures that can be calculated from sample data collected through chart audits. This paper describes the results of the initial testing of these proxy measures in several dental care delivery systems. The purpose of the report is to examine the performance of the measures in terms of their practicality, reliability, sensitivity to expected differences, and, for one program, agreement with administrative-based measures.

TABLE 1
Effectiveness of Care Measures*

1. Current Disease Activity Assessment:

$$\frac{\text{Enrollees in denominator having caries activity and periodontal disease status assessments within 2 years of end of reporting year}}{\text{All enrollees aged 18+ at end of reporting year with continuous enrollment for current and prrevious reporting years}} \times 100$$

2. Preventive Treatment for Caries-active Adults:

$$\frac{\text{Enrollees in denominator receiving fluoride treatment during reporting year}}{\text{All enrollees aged 18+ at end of reporting year with:}} \times 100$$

- continuous enrollment during reporting year
- caries-active classification for entire reporting year

3. New Caries:

$$\frac{\text{Enrollees in denominator receiving restorative, prosthetic, endodontic, or oral surgery treatment due to caries during reporting year}}{\text{All enrollees aged 18+ at end of reporting year with continuous enrollment for current and previous reporting years}} \times 100$$

4. Periodontal Treatment for Perio-present Adults:

$$\frac{\text{Enrollees in denominator receiving periodontal therapy or at least 2 prophylaxes during reporting year}}{\text{All enrollees aged 18+ with:}} \times 100$$

- continuous enrollment for entire reporting year
- perio-present classification for entire reporting year

5. Improvement in Periodontal Status:

$$\frac{\text{Enrollees in denominator where at least one 5+ mm probing depth from previous assessment has resolved to <5 mm, or where at least 1 sextant with PSR score of 4 from previous assessment has resolved to score of 3 or less}}{\text{All enrollees aged 18+ with:}} \times 100$$

- continuous enrollment for entire reporting year
- perio status assessment performed in reporting year
- previous perio-present assessment within previous 2 years

6. Deterioration in Periodontal Status:

$$\frac{\text{Enrollees in denominator where at least one site with previous probing depth of 4 mm or less has increased to 5+ mm, or where at least 1 sextant with previous PSR score <4 is now scored as 4}}{\text{All enrollees aged 18+ with:}} \times 100$$

- continuous enrollment for entire reporting year
- perio status assessment performed in reporting year
- previous perio status assessment within previous 2 years

7. Tooth Loss:

$$\frac{\text{Enrollees in denominator receiving an extraction during reporting year for erupted permanent tooth other than third molar or premolar removed for orthodontic reasons}}{\text{All enrollees aged 18+ at end of reporting year with continuous enrollment for current and previous reporting years}} \times 100$$

*Measures reflect performance for a one-year period. Measures may be calculated separately by caries/and or periodontal disease activity status. Measures 2, 3, and 7 (caries prevention, new caries, tooth loss) also may be calculated separately for children and adults.

TABLE 2
Use of Services Measures

1. Receipt of Prophylaxes	
All enrollees in denominator who receive prophylaxis procedure during reporting year / All enrollees aged 18+ with continuous enrollment for entire reporting year	× 100
2. Preventive Treatment : Restorative Treatment Ratio	
Total number of preventive procedures provided during reporting year to enrollees aged 18+ / Total number of direct restorative procedures provided during reporting year to enrollees aged 18+	× 100
3. Casting : Large Direct Filling Ratio	
Total number of casting procedures provided during reporting year to enrollees aged 18+ / Total number of large direct filling procedures provided during reporting year to enrollees aged 18+	× 100
4. Endodontic Treatment: Extraction Ratio	
Total number of teeth treated endodontically during reporting year for enrollees aged 18+ / Total number of nonthird molar teeth extracted during reporting year for enrollees aged 18+	× 100
5. Receipt of Third Molar Extractions	
Total number of enrollees in denominator who receive at least 1 third molar extraction during reporting year / Enrollees aged 18–24 (inclusive) at any time during reporting year	× 100
6. Mean Number of Third Molars Extracted	
Total number of teeth extracted among enrollees in denominator / Enrollees aged 18–24 (inclusive) at any time during reporting year who receive at least 1 third molar extraction	× 100

Methods

The seven effectiveness of care measures and six use of services measures that we developed for use with administrative data are summarized for adult patients in Table 1 and Table 2. The effectiveness of care measures are based on the assumption that plans should assess the disease activity or disease risk of all enrollees and target preventive treatment to those with higher levels of disease or risk. The measures address risk assessment, the prevention and management of dental caries and periodontal diseases, and tooth loss. Diagnostic information identifying both caries requiring restorative treatment and the presence of periodontal pocketing is necessary to calculate the effectiveness of care measure. The use of services measures address services for which there are alternative approaches to preventive and reparative treatment, and are intended to provide insight into a plan's philosophy of care.

The details of the development and testing of these measures are described elsewhere (13). Briefly, we worked with two panels of stake holders in refining a set of preliminary measures we had specified earlier (7). We were guided by HEDIS® criteria (12) in specifying the effectiveness of care measures—namely, that they must address outcomes of importance, or processes strongly linked to outcomes through controlled studies, that they be population based, patient centered, and risk adjustable or stratifiable. We specified existing use of services measures that we had identified as being employed by one or more plans, and that provided information useful in identifying plans' philosophies of care. The stake-holder panels—one composed of senior managers of 11 dental managed care plans and the other composed of employer benefits managers, directors of a state dental program and a state dental Medicaid program, and practicing dentists—worked with us to refine the preliminary measures using a modified Delphi process. The panels considered the potential validity, accuracy, sensitivity to change, and interpretability of each measure, as well as the likelihood that plans would begin to collect the data necessary to calculate the measure within the next five years.

The measures to be calculated from chart audit data were designed to parallel the administrative-based measures as closely as possible. All use of services measures could be specified using criteria identical to those employed in the administrative measures. For the effectiveness of care measures, tooth loss, periodontal disease improvement and deterioration, and receipt of appropriate caries-preventive and periodontal maintenance services could be specified identically. However, criteria for the remaining effectiveness of care measures required departures from the criteria specified for the administrative-based measures.

The first effectiveness of care measure concerns the assessment of disease activity, and can be satisfied by use of a two-category classification for caries activity (high, not high) and periodontal disease (present, not present). Because few dental charts contain formal notations of caries activity assessment, diet counseling, caries susceptibility testing, and notations of fluoride prescriptions were considered in their absence to be indicators that a patient's caries activity level had been assessed. Similarly, full mouth periodontal charting, charting of pockets over a minimum depth, and referral to a periodontist were considered indicators of periodontal disease activity assessment. The substitute criteria for caries do not yield a classification, however, so "presumptive high caries activity classification" was assigned to patients experiencing two or more restorations of any type in the preceding year. This classification facilitated cal-

culation of the caries preventive treatment and caries experience measures. Finally, in the absence of any standard approach to recording caries diagnoses or caries-related reasons for treatment in dental charts, we used the receipt of one or more restorations within the audit period as an indicator of "new presumptive caries." Because previous research has suggested that caries is associated with slightly less than half of the filled tooth increment in adults (14), we multiplied the proportion of patients receiving at least one restoration by 0.5 as a very crude adjustment for this overestimation.

We developed an instruction manual and a recording form for the chart audit that were designed to be self-instructional. The manual guided the auditor through the process of selecting a random sample of charts for audit, and provided detailed instructions and definitions for all items to be audited. The recording form summarized the criteria for each of 36 specific audit items to be recorded on the form. With the exception of identifiers and dates, all items were recorded by circling a preprinted dichotomous response or entering a number.

We pilot tested the audit data recording form in two local dental practices, observing dental assistants and receptionists as they read the instructions, selected a small sample of charts, and audited the selected charts. On the basis of these pilot tests, several additions were made in the wording of the instructions and the criteria to clarify appropriate action under unanticipated circumstances. We then conducted a larger pilot test involving a two-hour, onsite training session for three auditors who subsequently each audited 50 charts in three separate clinics affiliated with a state dental public health program. Each auditor also photocopied 10 audited charts that were subsequently audited by the investigators. Auditors were urged to call with any questions that arose in the course of the audits, and two calls were received from one auditor requesting clarification of definitions.

Following these pilot tests, audits were conducted at seven programs to evaluate the process of collecting data and calculating measures under a variety of conditions. The audits all took place in 1998, and covered care provided in calendar year 1997. Some of the effectiveness of care measures also required information from calendar years 1995 and 1996. The programs included three public health clinics, two preferred provider organizations (PPO), and two dental health maintenance organizations (DHMO). In the DHMOs the programs' enrollment data were used to generate true random samples of individuals enrolled during 1996 and 1997. In the public health clinics, charts were sampled systematically following randomly selected starting points until the targeted number of eligible charts that reflected visits in 1996 and 1997 was identified. Patients with 1997 claims in 15 private dental offices for one PPO, and in five offices for the other, were identified by the respective PPOs, and these charts were included in the audit if visits had occurred in 1996 as well as 1997.

In all but two programs, the audits focused on adults (18 years and older) to maximize sample sizes for subgroup estimates. In one DHMO and one public health clinic, both children and adults were included in the sample; however, analyses reported here involve only adults. In all but one DHMO, the intended sample size was 150 charts, which yielded a maximum confidence interval of ±8 percent around estimate proportions. This sample size was a compromise between the approximately 400 charts needed to ensure 5 percent confidence intervals for proportions and the limited resources available to test the measures in these environments. Sample sizes less than 150 reflect the elimination of children from the samples in two programs, and the termination of one PPO's audit cycle prior to completion of all office visits. In one DHMO the sample size was increased to 400 to

TABLE 3
Effectiveness of Care Results for Seven Dental Programs

	Public Clinics (%)			PPOs (%)		DHMOs (%)	
Measure	Prog. A (n=82)	Prog. B (n=150)	Prog. C (n=150)	Prog. D (n=150)	Prog. E (n=102)	Prog. F (n=402)	Prog. G (n=102)
Current disease activity assessment:							
Caries activity assessment	0	19	8	2	1	64	72
Periodontal disease activity assessment	0	27	10	70	35	71	87
Both assessments	0	12	2	1	1	60	66
(Presumptive high caries activity)	12	39	51	15	18	34	18
Preventive treatment for caries-active adults	0	18	27	13	5	18	28
New presumptive caries:							
Low-caries-activity adults	16	20	36	25	21	9	18
High-caries-activity adults	35	26	37	36	29	30	42
Preventive treatment for perio-present adults	*	60	100	89	96	82	94
Improvement in periodontal status	*	33	0	22	35	41	25
Deterioration in periodontal status	*	17	14	16	29	50	29
Tooth loss	24	10	11	16	9	5	6

*Data unavailable.

permit direct comparison of the audit-based measures with measures based on administrative data.

The auditors for the DHMOs and the public health clinics were program staff, either dentists or dental hygienists. One DHMO utilized two auditors; the remaining DHMO and the public health clinics used single auditors. The same dentist audited charts for both PPOs. Auditors read the self-instruction manual and then participated in a brief question-and-answer session prior to beginning the audits. Auditors were urged to call the investigators for any necessary clarifications during the course of the audits; four of the seven auditors did so.

Ten charts from each of the public health clinics were photocopied and reaudited by the investigators. No reliability analyses for the PPO audit data were attempted due to perceived provider concern over patient confidentiality. Reliability in one DHMO was determined through re-audits of 43 charts by a second onsite auditor. Reliability was assessed as percent agreement and kappa score at the level required by the audit form, usually the presence or absence of a particular notation or the receipt or nonreceipt of a specific procedure and, occasionally, the number of a specific group of procedures.

Statistical testing of differences among programs for the performance measures was not done for three reasons. First, the PPO samples are not random samples of the PPO programs' enrollees. Second, the denominators for the PPOs and public health clinics were determined differently than those for the DHMOs. Third, the purpose of this report is to examine the performance of the performance measures, rather than the programs. Thus, while general differences among plan types are of interest in interpreting how the measures performed, identification of specific significant differences among plans is not appropriate. Finally, the performance measures for one DHMO were compared to similar measures calculated form the DHMO's administrative data (13). This comparison also was purely descriptive.

Results

Table 3 shows values of the seven effectiveness of care measures for the seven programs, as well as the proportion of the audit sample determined to exhibit presumptive high caries activity based on receipt of restorations in the previous year. Disease assessment rates range from zero in one public health clinic to about two-thirds of charts in both DHMOs, which had clinical protocols recommending this assessment for all patients every two years. In all programs the rate for recording periodontal disease assessments tended to be higher than the rate for recording caries activity assessments. The rate of receipt of appropriate preventive treatment among high caries activity adults was quite low, with little more than one-quarter of such patients receiving preventive treatment in any program. New presumptive caries rates ranged from 9 percent to 36 percent among patients classified as having low caries activity, and from 26 percent to 42 percent among patients with high caries activity. Periodontal therapy for patients classified as having periodontal disease was uniformly high in the six programs where the measure could be calculated, while no distinct pattern was found between periodontal improvement and deterioration rates. Tooth loss rates ranged from 5 percent to 24 percent.

Table 4 shows values of the use of services measures for the seven programs. Receipt of prophylaxes varied across programs, ranging from 27 percent in one public program to 99 percent in another. The ratio of preventive to restorative treatment displayed about a fivefold range, with the DHMOs exhibiting the highest rates. The ratio of castings to large direct restorations shows the greatest variation across programs, due in part to smaller sample sizes in the two programs with the highest rates, and the relatively infrequent provision of either castings or large direct restorations in these programs. The ratio of endodontic to extraction procedures was clearly lower in the public health clinics than in the PPOs and DHMOs. Finally, the two third molar measures reflect somewhat different situations, in that plans with lower proportions of patients experiencing an extraction tend to remove more third molars per patient.

The reliability results shown in Table 5 indicate that for most audit items, agreement between the auditor and the investigators was greater than 90 percent, with kappa scores in the good and excellent ranges (15). In the program where two auditors' assessments were compared, reliability tended to be only fair for some counts of services received, and for notations

TABLE 4
Use of Services Results for Seven Dental Programs

	Public Clinics			PPOs		DHMOs	
Measure	Prog. A (n=82)	Prog. B (n=150)	Prog. C (n=150)	Prog. D (n=150)	Prog. E (n=102)	Prog. F (n=402)	Prog. G (n=102)
Receipt of prophylaxis	27%	61%	99%	87%	71%	61%	87%
Preventive treatment : restorative treatment ratio	0.4	1.0	0.9	1.2	0.9	1.3	1.9
Casting : large direct filling ratio	2.5	0.3	0.1	1.0	1.2	0.1	3.4
Endodontic treatment : extraction ratio	0.1	0.1	0.3	1.5	1.6	0.9	1.5
Receipt of third molar extractions	0%	3%	1%	9%	3%	1%	<1%
Mean number of third molars extracted	0	1.0	1.0	1.4	1.3	2.2	4

TABLE 5
Reliability for Individual Audit Items in Four Programs

	Programs							
	A (*n*=10)		B (*n*=13)		C (*n*=9)		F (*n*=43)	
Audit Item	% Agree	Kappa	% Agree	Kappa	% Agree	Kappa	% Agree	Kappa
Exam	90	.80	92	.76	100	*	89	.75
Prophylaxis	80	.62	100	1.0	100	*	86	.70
Periodontal scaling	90	.74	100	1.0	100	*	95	.65
Periodontal surgery	100	*	100	*	100	*	100	1.0
Periodontal referral	100	*	100	*	100	*	95	.64
Fluoride	90	.74	100	*	100	1.0	89	.46
Sealant							98	.66
# preventive services	60	.29	100	1.0	100	1.0	58	.37
# direct restorations	90		92	.76	78	.53	88	.77
# previous restorations	100	1.0	100	1.0	89	.77	81	.39
# large direct restorations	100	*	100	1.0	78	.57	88	.70
# cast restorations	100	*	100	1.0	100	*	95	.48
# endodontic procedures	100	*	100	*	100	*	100	*
# extractions	100	1.0	100	*	100	1.0	98	.79
# third molar extractions	100	*	100	*	100	1.0	100	*
Diet counseling	100	*	100	*	100	*	100	*
Caries susceptibility testing	100	*	100	*	100	*	100	*
Fluoride prescriptions	100	*	100	*	100	*	100	*
Caries risk note	100	*	100	*	100	*	95	.89
Caries activity note	100	*	100	*	100	*	98	*
Other caries note	100	*	100	1.0	100	*	95	*
Probing depths	100	*	100	*	100	*	88	.73
No pockets/WNL	100	*	100	1.0	100	*	84	.37
Periodontist referral	100	*	100	1.0	100	*	91	*

NOTES: For programs A, B and C, auditors' determinations were compared to those of an investigator. In program F, the determinations of two auditors were compared. For numerical values, simple kappa was used (match vs no match). When there are no discordant pairs, kappa=1.0. When data fall into only two cells (one matching pairs, one discordant pairs) kappa=*. Sealants were recorded only in program F.

TABLE 6
Comparison of Administrative Data-based and Chart Audit-based Performance Measures

Measure	Administrative (*N*=94,251)	Chart Audit (*N*=402)
Effectiveness of care		
Current disease activity assessment	64%	64%
Preventive treatment for caries active adults	13%	18%
Preventive treatment for perio-present adults	81%	82%
Tooth loss	5%	5%
Use of services		
Receipt of prophylaxis	57%	61%
Preventive treatment : restorative treatment ratio	1.8	1.3
Casting : large direct filling ratio	0.1	0.1
Endodontic treatment : extraction ratio	0.5	0.9
Receipt of third molar extractions	2%	1%
Mean number of third molars extracted	2.1	2.2

of no periodontal disease. Follow-up questioning revealed that misunderstanding about specific criteria led to most disagreements. Auditors in four programs responded to informal questions about time required for the audit procedures. They reported that audits required from three to 20 minutes per chart, depending on the clarity of the entries and the number of visits and procedures recorded in the previous two years. Three of the auditors indicated the "usual" time required to audit a chart was five minutes.

Table 6 presents a comparison of the chart audit-based performance measures for program F with the measures calculated using administrative data. New caries experience is not included because these data were not available

from the program's administrative data (13). Periodontal improvement and deterioration are not included because the same audit-based data were used to calculate the measure for both the administrative and audit-based reports. The paired values for the rate measures are in generally good agreement, all within five percentage points. Values for two of the three ratio measures, however, reflect substantial discrepancies.

Discussion

The initial testing of the audit-based performance measures was undertaken to examine their sensitivity to differences among programs, reliability, practicality, and agreement with similar measures calculated from a programs' administrative data. At a general level, the measures do seem to reflect differences among the dental delivery systems in which they were calculated, although these differences could not be tested statistically. Both DHMOs routinely perform caries and periodontal disease "risk" assessments, and this situation is reflected in the data. The 60 percent and 66 percent rates probably reflect some slippage in application of the assessments as well as the proportion of enrollees who had an examination in the two-year period.

In these analyses, the DHMO samples were drawn from all enrolled individuals, regardless of past utilization, while for all other dental programs, audited charts were required to reflect at least one visit in each of the two calendar years prior to the audit. For purposes of testing only, the rate for assessment of disease was disaggregated into separate rates for caries and periodontal assessments. For the PPOs and clinics, the periodontal disease assessment rate reflects variation among providers known to exist in the rate with which periodontal probing information is recorded in the chart (16). The low proportion of patients for whom information describing periodontal status is available means that subsequent effectiveness of care measures examining appropriate periodontal preventive treatment and improvement and deterioration in periodontal conditions will be based on only a few observations, effectively lowering the reliability of these estimates.

The estimates of "presumptive" caries activity are problematic because not all restorations in adults are placed for reasons due to caries, and because the attempt to minimize this source of error by assuming caries was involved in only one-half of all instances where one or more restorations was placed in a given year cannot be evaluated without additional data. Only for program G is information available in the administrative data with which to identify enrollees receiving treatment related to caries. These administrative data indicate that the proportion of enrollees with new caries was 7 percent and 28 percent in 1997 for low and high caries activity adults, respectively (13). The comparable rates calculated from the chart audits were 18 percent and 42 percent. These differences suggest that despite the attempted correction, the audit-based presumptive caries rates overestimate actual caries experience. It should be noted, however, that the sample size of 102 in program G permits considerable sampling error. These putative caries estimates seem to be sensitive to caries activity classification, consistently showing a greater proportion of adults classified as having high caries activity experiencing "putative new caries." However, it must be remembered that because most of the high caries activity classifications were made on the basis of restorations received in the previous year, this relationship may be due to factors unrelated to propensity to caries.

The uniformly low rates for receipt of appropriate preventive treatment for caries-active adults, a criterion that can be satisfied by any fluoride treatment or prescription, received independent verification at both DHMOs. The rates surprised clinical directors, and led to internal studies to verify the situation and institute corrective action. Several of the use of services measures, as well as the tooth loss measure, show differences that may reflect assumed differences in practice philosophy by type of practice. The DHMOs had the highest preventive treatment : restorative treatment ratios and the lowest extraction rates, reflecting a strong preventive emphasis. The PPOs had the highest endodontic treatment : extraction ratios, which may reflect differences in reimbursement arrangements. Given these differences, it might be expected that the PPOs also would have the highest casting : large direct restoration ratios; however, one DHMO and one clinic each had ratios at least twice as high. Further examination indicated that the numbers of casting and large restorations delivered in these programs were small; coupled with their reduced sample sizes, the high rates may be artifactual.

The chart audits proved to be reasonably reliable despite the limited training available to the auditors. Attention to the elimination of any unclear criteria or explanatory material during the development of the audit materials was probably the major reason for this outcome, aided by the principally objective determinations required of auditors. Unfortunately, with respect to the overall quality of dental charts, for audit items recording the presence or absence of specific notations in some programs, 100 percent agreement was achieved because the chart contained few notations other than procedural information. Again, the absence of periodontal information exemplifies the situation and its implications for evaluating both change over time for individual patients and effectiveness of care for the program.

The reliability of the measures, in terms of agreement with similar performance measures calculated from administrative data, was reasonably good in the one program where agreement was tested. Only for two use of services ratio measures did the two sets of measures show substantial disagreement. The preventive treatment: restorative treatment ratio was lower for the chart audit partly because preventive procedures that were listed separately in the administrative record were "bundled" in the chart notation and not counted separately. The discrepancy in the endodontic : extraction ratio is probably the result of low frequencies of service provision, around 5 percent for extractions and 0.5 percent for endodontic treatment. At these frequencies, repeated samples of 400 can yield strikingly different estimates. The caries experience and periodontal improvement and deterioration measures were not included in the comparison. The DHMO had designed but not yet fully implemented a limited set of diagnostic codes to be associated with all treatment procedures at the time the administrative data-based calculations were made. The DHMO also did not

include periodontal charting data in the administrative data set; thus, the information from the chart audits was used to calculate these values for both the administrative and audit-based measures.

It is clear that the current state of the typical dental chart seriously impairs the practicality of calculating clinical performance measures using data collected via chart audits. The principal impedimenta are the difficulty in determining when caries is the reason for treatment, the frequent absence of any information describing periodontal status, and, to a lesser extent, the lack of any written assessment of disease activity or disease risk status. This latter omission is less serious because crude assessments can be made if the caries diagnostic and periodontal status information is available. In addition to these obstacles, obtaining a representative sample may be resource-intensive, depending on the program for which the performance measures are to be calculated. If charts for all patients in the program are available in a central location, then the audits may be accomplished fairly easily, although a minimum of 40 hours' total auditor time will be required to learn the materials and perform 400 audits. If the program operates multiple treatment centers with separate charts, sampling may become more complex, and logistics of auditing increase with the increasing number of sites at which audits must be conducted. Also, based on the results of this initial test, it may be necessary to increase the sample size to obtain stable estimates for some of the measures. Finally, some equitable means of determining appropriate denominators must be devised for those dental programs that do not have a defined, enumerated population such as "enrollees" or "eligibles."

These limitations notwithstanding, it is important to begin the process of evaluating the clinical performance of dental delivery organizations. This generation of performance measures, both the parent administrative data-based set and the proxy audit-based set described here, is envisioned as being transitional, available for use until better data permit the formulation of more sophisticated measures. Because most dental programs' administrative data systems are generally not sufficient to support calculation of the measures, the audit-based measures were developed to help programs begin to gain experience with performance measurement. From the initial testing reported here, these measures appear sufficiently sensitive and reliable to warrant their use in internal programs of quality improvement. Due to the problems of practicality, however, interprogram public comparisons of effectiveness of care and use of services calculated from chart audits through the circulation of report cards are not yet warranted. More experience with the caries and periodontal disease outcome measures, as well as with the ratio measures, are needed to ensure that they provide reasonably stable, valid estimates of program performance. With improvements in the information available in the dental chart, however, these measures may offer a near-term approach to obtaining performance information with which to compare delivery programs. Of course, it is likely that programs will begin to develop more powerful administrative data systems as their role in increasing the efficiency of provision of appropriate care becomes apparent. As these systems become more prevalent, so too should use of the administrative data-based version of these and similar measures. The introduction of the American Dental Association's diagnostic code set and the association's support for the use of the codes in evaluating the appropriateness of care (17) should help facilitate their inclusion in the claims information provided by practitioners. In time, similar improvements should occur in how periodontal status is documented. With the availability of these data elements, the first generation of performance measures should be fully operational.

Acknowledgments

The authors thank the administrators and staff of the participating dental programs for their assistance with the project.

References

1. Coulter I, Marcus M, Atchison K. Measuring oral health status: theoretical and methodological challenges. Soc Sci Med 1994;38:1531-41.
2. Singer L. DWI: Dental Wellness Index helps practices predictably measure treatment outcomes and assure that quality levels remain high. Dent Econ 1998;88:117-21.
3. Slade G, Strauss R, Atchison K, Kressin N, Locker D, Reisine S. Conference summary: assessing oral health outcomes—measuring health status and quality of life. Community Dent Health 1998;15:3-7.
4. Kress G, Shulman J. Consumer satisfaction with dental care: where have we been, where are we going? J Am Coll Dent 1997;64:9-15.
5. Hayden W, Marcus M, Lewis C. Developing population profiles from claims data for conceptualizing effectiveness of care. J Dent Educ 1989;53:619-28.
6. Anonymous. Minnesota Delta ends capitation, changes pay rules. Managed Dent Care 1999:4:3.
7. Bader J, Shugars D, Hayden W, White B. A health plan report card for dentistry. J Am Coll Dent 1996;63:29-38.
8. White BA. Use of oral health-related quality of life measures in managed dental care organizations. Community Dent Health 1998;15:27-31.
9. Bader J, Shugars D. Variation, treatment outcomes, and practice guidelines in dental practice. J Dent Educ 1995;59:61-96.
10. Bader J, Shugars D. A case for diagnoses. J Am Coll Dent 1997;64:44-6.
11. National Association of Dental Plans. 1998 dental HMO/PPO industry profile. Dallas, TX: NADP, 1988.
12. National Committee for Quality Assurance. HEDIS® 3.0. Vol 1. Narrative. Washington, DC: NCQA, Jan 1997.
13. Bader J, Shugars D, White B, Rindal D. Development of effectiveness of care and use of services measures for dental care plans. J Public Health Dent 1999;59:142-9.
14. Bader J, Shugars D, Rozier R. Relationship between epidemiologic coronal caries assessments and practitioners' treatment recommendations in adults. Community Dent Oral Epidemiol 1993;21:96-101.
15. Landis J, Koch G. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-79.
16. Bader J, Rozier R, McFall W, et al. Evaluating and influencing periodontal diagnostic and treatment behaviors in general practice. J Am Dent Assoc 1990;121:720-4.
17. McKee L. SNODENT to provide inclusive means of transmitting dental information. ADA News 1999;30:1, 9.