The Effects of Insurance Coverage on the Quality of Prenatal Care

Michael S. Klinkman, MD, MS; Daniel W. Gorenflo, PhD; Tamara S. Ritsema, MPH

Objective: To compare the quality of prenatal care provided to patients with traditional fee-for-service, health maintenance organization, and Medicaid insurance using an evidence-based, community-derived prenatal care guideline.

Design: Retrospective cohort study.

Setting: Seven private and hospital-based prenatal care sites in a suburban county in southeast Michigan.

Patients: A stratified random sample of 267 patients (93 with Medicaid, 92 with health maintenance organization, and 82 with fee-for-service insurance) receiving prenatal care from community physicians (obstetricians-gynecologists and family practitioners) between January 1, 1991, and December 31, 1992.

Main Outcome Measure: Adherence to explicit prenatal care criteria as measured by an evidence-based prenatal care guideline developed by a community panel. "Quality scores" were compared across groups in 4 areas: performance of prenatal screening procedures or tests, visit-based screening, substance use screening, and clinician management of abnormal clinical findings.

Results: Patients with Medicaid insurance presented for prenatal care significantly later in pregnancy (14.5 vs 10.5 weeks, $P<.01$). No significant differences were seen between groups in quality scores for screening tests, clinician management of abnormal clinical findings, visit-based screening, or substance use screening. The overall similarity in quality scores did obscure some significant differences in adherence to individual criteria, particularly in the area of screening tests. Significantly more patients with Medicaid were screened for genital infection ($P<.001$) and fewer for gestational diabetes ($P<.001$) or anemia ($P<.001$) than patients in the other 2 groups.

Conclusions: Although patients with Medicaid presented for prenatal care later in pregnancy and received a different "package" of screening tests than the other 2 groups, there was no overall measurable difference in the quality of prenatal care provided to patients with Medicaid, health maintenance organization, and fee-for-service insurance. Clinicians may have altered screening protocols based on preexisting perceptions of patient risk. Although summary quality measures are a promising tool for comparative research, they provide an incomplete picture of the quality of the prenatal care process and must be interpreted with caution.

Arch Fam Med. 1997;6:557-566

THE BELIEF that managed health care plans reduce the overall cost of medical care has been largely responsible for their rapid expansion in the private and public health care marketplace. This expansion has occurred despite the almost complete absence of useful information about the relative quality of health care services provided in the managed care setting. Efforts are being made to determine whether the quality of care provided under managed care is equivalent to that seen in fee-for-service (FFS) practice.[1-6] These efforts are of great importance to health policy makers and health care consumers: without reliable and systematic ways to measure the quality of care provided in competing health plans, overzealous cost control may lead to serious health consequences, particularly for vulnerable or underserved populations.

Managed health care is viewed as the most promising solution to the problem of escalating cost in the Medicaid program, and states are rapidly moving to implement mandatory enrollment of recipients in either public or contracted private managed care plans. However, there is almost no information available to guide policy in this area. There are only a few published studies that compare the quality of prenatal care received under Medicaid managed care with that received under FFS Medic-

From the Department of Family Practice, University of Michigan, Ann Arbor.

PATIENTS AND METHODS

RATIONALE FOR CHOICE OF PROCESS MEASURES

We elected to use prenatal process rather than outcome measures (such as low birth weight or infant Apgar scores) in this study for several reasons. First, adverse pregnancy outcomes are rare, and power calculations confirmed that a sample size equivalent to several years' worth of deliveries would be required to adequately examine differences in outcomes. Second, outcomes are the result of a complex interaction of genetic predisposition, environmental and social factors, random occurrence, and the quality of prenatal and perinatal care; differences between health care plans in outcome cannot, under many circumstances, be clearly attributed to their relative quality of care. Third, the elements included in process represent the direct efforts of health care providers (physicians, nurses, and physician extenders). Differences in process are arguably more likely to be attributable to insurance type than are outcome differences. Fourth, process measures have been shown to correlate with perinatal outcomes[15,16] and represent the best indirect measure that the health care system has done its part to ensure a good pregnancy outcome.

To examine the potential effect of differences in insurance coverage on health care quality, it was necessary to find or develop a set of quality measures that met the following criteria: (1) based on individual patient-level data, (2) derived from routinely collected clinical data, (3) applicable across all health insurance types, (4) accepted as clinically valid by community physicians, and (5) interpretable by state health policy officials. Several existing methods for estimating the quality of the prenatal care process were reviewed, but none met all 5 criteria. Visit indexes, such as the modified Kesner or Kotelchuk,[17-19] were considered insufficient as a stand-alone quality measure because they reflect patient behavior not under the direct control of health care providers. Clinicians may have relatively little influence over patients' decisions to initiate prenatal care or return for scheduled visits. Simple counts of prenatal procedures or tests[20] were rejected because they do not assess whether the procedures or tests at issue were correctly performed, interpreted, or used by clinicians in the management of the patient. Existing prenatal practice guidelines were reviewed but were found by the panel to contain unnecessary items, omit important elements, or be based on expert opinion rather than evidence.

CRITERIA DEVELOPMENT

We developed explicit criteria for the evaluation of the quality of the process of prenatal care. We reviewed the available scientific literature to identify: (1) specific prenatal screening procedures or tests recommended by expert panels or used in published evaluations of quality of prenatal care; (2) definitions for abnormal screening procedure or test results; and (3) published recommendations for the management of abnormal screening procedure results, test results, or clinical findings. A complete list of all identified screening procedures and tests was prepared for review along with a candidate list of items regarding appropriate identification and management of abnormal procedure results, test results, and clinical findings. A complete list of sources reviewed is available from the authors on request.

A community panel was created to review and select items for inclusion as final quality criteria; the panel included clinicians from the University of Michigan departments of family practice and obstetrics-gynecology and community-based obstetrician-gynecologists. No community-based family physicians were available to serve on the panel. All participating clinicians maintained an active obstetrical practice in the community and included patients with Medicaid and private insurance in their practice. Draft versions of criteria were circulated to panel members, who evaluated each based on supporting scientific evidence, accuracy of recording and data retrieval, applicability to the population served in the community, and their own actual practice patterns. Clinicians were also free to propose additional criteria based on their experience. Lists were revised and recirculated until a consensus was reached regarding the inclusion of criteria. This approach was similar to that taken by expert panels, which usually rely on a blend of evidence and expert opinion to arrive at recommendations. However, our intent was to create a true community standard for the process of routine, low-risk prenatal care rather than select an existing standard or guideline, and the final decision on inclusion of criteria rested with panel members.

The final list of criteria (**Table 1**) contained 33 process variables in the following categories: 1 criterion on the initiation of prenatal care; 12 criteria on the performance of screening procedures or tests; 5 criteria measuring the performance of visit-based prenatal examination procedures; 4 criteria on the documentation of diet, medication, and substance use; and 11 criteria concerning clinician identification and management of abnormal procedure results, test results, or clinical findings. Two emerging screening procedures, serum human immunodeficiency virus testing and screening for group B streptococcal colonization of the birth canal, were included as optional procedures but not incorporated into the operational definition of prenatal care quality in this study.

Screening for psychosocial risk factors (eg, presence of adequate social support) was considered by panel members to be an important element of prenatal care quality. However, psychosocial screening items were excluded from the final criteria list because of poor data quality: pilot medical record reviews revealed that psychosocial items were only sporadically recorded.

SAMPLE

A representative sample of women receiving routine, low-risk prenatal care in Washtenaw County, Michigan, during the period between January 1, 1991, and December 31, 1992, was compiled from 7 participating practices using the following procedure. At the university hospital and its affiliated clinics, patients were randomly selected from a master list, stratified by insurance status, of women who had been delivered of neonates during the study period. Patient selection methods varied for the other obstetrical practices included in the study. Four of the participating practices used systematic methods such as computerized databases or index cards for tracking each of their patients; patients from these clinics were stratified and randomly selected using the office's tracking system. One randomly selected using the office's tracking system. One practice did not employ a formal tracking system, and we

relied on office staff and nurses to recall which patients had been delivered of neonates during the study period. All 7 practices provided care for patients with Medicaid and private insurance. High-risk patients were operationally defined as those with preexisting medical conditions that would materially affect the process of prenatal care (such as insulin-dependent diabetes mellitus or systemic lupus erythematosus) or those with known multiple gestation. The medical records for all high-risk patients were excluded from review.

Only those residents whose insurance coverage was unchanged during the entire prenatal and perinatal course were included in the analysis. In the Medicaid sample, patients often became eligible for coverage at the time of initial presentation for prenatal care through the Maternal Support Services program of the Michigan Department of Social Services (for Washtenaw County). In the other 2 groups, insurance coverage predated pregnancy. Records that indicated that transfer of care either into or out of a practice occurred during the prenatal period were excluded from analysis. Finally, because the study was intended to assess routine, low-risk prenatal care, pregnancies that ended in elective or spontaneous abortion and ectopic pregnancies were excluded.

In private FFS and Medicaid samples, clinicians received payment based on the number and level of services provided. All patients in the private HMO group were members of local HMOs that paid clinicians a fixed amount for all prenatal care services for low-risk patients.

An initial power calculation indicated that a sample size of 90 per group would be sufficient to find a small to moderate difference between groups in identification and management of abnormal clinical findings for $\alpha = .05$ and $\beta = .90$.[21] A sample of 100 patients was selected from each group (Medicaid, HMO, and FFS insurance). Some medical records later proved unavailable or incomplete, and a subsequent sample could not be performed, leading to the final sample sizes as reported in the "Results" section.

DATA COLLECTION

All data elements included in the analysis were abstracted from the written prenatal record by 3 trained reviewers during a 1-year period. Standardized search algorithms were developed to operationalize recording of individual data elements, and reviewers were "calibrated" during a full-time 5-month training program by repeated assessment of 3-way interrater concordance using common records. There was essentially complete agreement between reviewers on coding, with the single exception of laboratory tests. An initial 82% concordance across the 3 reviewers for laboratory test coding was found to be due to different search strategies for misfiled data. This problem was solved by a decision rule: laboratory tests not found within a 5-minute search were considered not recorded, as they would likely be unavailable to clinicians as well. Subsequent concordance for coding individual laboratory tests was greater than 90%.

Data were entered directly into a database (FileMaker Pro, Clans Corp, Mountain View, Calif) using a laptop computer (Macintosh, Apple Computers, Cupertino, Calif), then downloaded into a central database (SPSS, SPSS Inc, Chicago, Ill). Validation procedures were employed where possible to minimize data entry error. Details of training program content and hard copies of the data entry screens are available from the authors on request.

ANALYTIC APPROACH: CODING AND SCORING OF CRITERIA

The performance criterion used for the 12 screening test or procedure items was completion as documented in the prenatal record. Patients were considered eligible for every test. Each of the 12 items was assigned 1 of 4 codes: present in the record and normal, present and abnormal, declined by patient, or not recorded. Adherence was determined by the presence of 1 of the first 3 codes. A quality score for the performance of screening tests for each patient was calculated by summing the number of completed tests and dividing by 12.

The performance criterion used for visit-based procedures was documentation of each of the 5 items at each complete prenatal care visit. Visits for non–prenatal care problems such as upper respiratory tract infections, and follow-up visits to address a single prenatal issue occurring within 1 week of another prenatal visit, were identified and excluded from this analysis. The number of prenatal visits at which the blood pressure measurement, urine protein level, and weight were not recorded was tallied for each procedure, then divided by the total number of prenatal visits. The number of visits at which fundal height and presence of fetal heart tones were not recorded was tallied and divided by the number of visits after 20 weeks' gestation. This procedure gave the failure rate for each visit-based procedure. The proportion (1–failure rate) was calculated as the adherence rate for each procedure. A quality score for office-based screening, giving equal weight to each of the procedures, was calculated for each patient by summing the adherence rates, then dividing by 5.

Four criteria addressed the use of alcohol, the use of other substances, smoking, and the presence of an adequate diet as defined by the clinician. The standardized prenatal record forms used by all practices in this study contained check box items with space for added text for each item. Each item was considered performed if either the appropriate box was completed or any notation was made in the prenatal record of use of cigarettes, alcohol, or other substances or of problems with inadequate diet or referral to either a substance abuse program or a dietitian for a reason other than gestational diabetes (GDM).

The 11 criteria concerning the identification and management of abnormal procedure results, test results, or clinical findings (henceforth labeled "clinician follow-up") were coded and scored as follows. Each was composed as an if-then statement (eg, if an abnormal Papanicolaou smear result was recorded, then was there documentation of follow-up or recommendation?); responses were coded as no (failed criterion), yes (passed criterion), and not applicable (not eligible). There were no individual criteria for which more than 22% of the patients were eligible. Because no patient was eligible for all of these criteria, a quality score for clinician follow-up was calculated for each patient by calculating a percentage equal to the number of items passed divided by the number of items for which the patient was eligible. Sixty-two percent of the patients were eligible for at least 1 of the 11 criteria; the number of criteria for which a patient was eligible ranged from 0 to 4.

Continued on next page

STATISTICAL ANALYSIS

Because all of the participating practices were group practices, patients were extremely unlikely to receive all their prenatal care from a single clinician. Consequently, we did not attempt to measure performance at the individual clinician level. We began by looking for the presence of significant differences in demographic, medical, or obstetrical history variables between groups, using χ^2 tests coupled with the Bonferroni correction for multiple comparisons or a 1-way analysis of variance (ANOVA) coupled with the Scheffe multiple comparison procedure where appropriate. Differences between groups in adherence to individual criteria were assessed by χ^2 tests or ANOVA as appropriate. Mean quality scores for each of the 4 criteria sets (screening tests, visit-based screening, substance use, and identification and management of abnormal procedure results) were calculated for each insurance group; cross-correlations between these 4 dependent variables were found to be minimal using the Pearson product-moment correlation. The mean quality scores for each criteria set were employed as dependent variables for the 3-group comparison using ANOVA.

All analyses were repeated using appropriate techniques to control for the possible confounders of age, educational level, parity, and ethnic status, to assess possible practice effects, and to control for the performance of multiple comparisons. There was evidence of a significant practice effect for only 2 individual criteria: one practice performed significantly fewer screening glucose measurements for patients with Medicaid only ($P<.001$), and another failed to document nutritional screening for any prenatal patients. Reanalyses of mean quality scores and gestational age at the initial visit were performed using a nested ANOVA with practice and insurance type as independent variables, with and without an interaction term for practice by insurance type, and with and without the covariates of age, parity, educational level, and ethnic status.

We found evidence of a significant practice effect for 2 of the 4 quality scores: substance use screening ($P<.001$), fully accounted for by the failure of one practice to document nutritional screening; and laboratory screening tests ($P<.001$), which was only partially attributable to the difference in glucose screening previously mentioned. The practice-insurance interaction term also achieved statistical significance for the substance use quality score ($P=.005$). The inclusion of covariate terms for age, educational level, parity, and ethnic status did not affect the results: none of the terms had a statistically significant effect on the analyses. Consequently, to preserve as much as possible of the statistical power of the analysis, covariate adjustments were not included in the results as reported. The mean quality scores and ANOVA results reported herein are all adjusted for practice identification. All analyses were performed using a statistical software package (SPSS, SPSS Inc) on a microcomputer (Macintosh, Apple Computers).

aid coverage,[7-11] and most of these studies share the notable limitations associated with secondary review of administrative data. Moreover, despite frequently expressed concerns about the quality of health care under the Medicaid program,[12,13] we have almost no information comparing the quality of care received by patients with Medicaid with that received by patients with private health insurance. Only one study could be found that directly addressed this issue, a single-practice comparison of frequency of screening for hypertension, cervical cancer, and hypercholesterolemia.[14] More work in this area is urgently needed, particularly in light of the speed with which states are turning to Medicaid managed care programs.

This article provides the first in a series of reports aimed at addressing the information gap, a comparison of the quality of prenatal care received by patients with Medicaid and privately insured patients in FFS and managed care health maintenance organization (HMO) plans. To our knowledge, this is the first reported comparison of the quality of prenatal care across these 3 insurance groups. The study was conducted as part of a larger project evaluating the implementation of Medicaid managed care in a suburban county in southeast Michigan. Prenatal care was chosen as 1 of 5 target conditions with the largest potential effect on the health of Michigan Medicaid recipients. This report describes the development of a community-derived prenatal care guideline that includes visit indexes, performance of selected prenatal procedures or tests, and clinician management of abnormal clinical findings; the results of the baseline comparison of the quality of prenatal care received by those with Medicaid, FFS, and HMO insurance in the period immediately before the introduction of Medicaid managed care in southeast Michigan are provided.

RESULTS

SAMPLE DEMOGRAPHICS

In general, patients with Medicaid were younger, single, and more likely to be of African American descent, while the HMO and FFS populations were similar (**Table 2**). Household composition was strikingly different between the Medicaid sample and the other 2 groups: significantly more patients with Medicaid were single (71.7% vs ~12% for both other groups) and lived alone, with a partner, or in extended households with their own parents or relatives. Of the patients with Medicaid, 8.7% were younger than 18 years and only 2% were older than 35 years; in the other 2 groups, no patients were younger than 18 years and 13% and 10% were older than 35 years (data not shown).Patients with Medicaid had received significantly less education than those in the other 2 groups in this highly educated sample.

Despite these demographic differences, no significant differences were seen between groups in either general health or prior obstetrical history (Table 2). The Medicaid sample had a slightly lower mean number of medical problems and a slightly higher proportion of gravida 1 patients, but these differences were not statistically significant. There were no differences in the prevalences of individual health problems between groups (data not

Table 1. Final Criteria Used in Quality-of-Care Assessment*

Variable	Quality of Evidence†	Groups Making the Recommendation‡
Initiation of prenatal care		
The first prenatal visit should occur during the first trimester	II-1	1
Screening laboratory tests		
Every woman should be screened for Rh factor and antibody titer at the prenatal intake visit	II-2	1-4 and 6
Every woman should undergo a serological test for rubella immunity before delivery of a neonate or have prior documentation of rubella immunity	II-2	1-3 and 6
Every woman should be screened for anemia at the prenatal intake visit	III	1-4
All women should undergo a screening glucose test at the prenatal intake visit	III	None
A 50-g, 1-hour glucose challenge test should be performed on all women at 24 to 28 weeks of gestation	I	1-3
Every woman should have either Papanicolaou smear results available from within 12 months of the first prenatal visit or a Papanicolaou smear performed at the prenatal intake visit	III	None
Every woman should be screened for cervical gonorrhea infection at the prenatal intake visit	III	1-5
Every woman should be screened for cervical chlamydia infection at the prenatal intake visit	III	None
All women should be screened for hepatitis B surface antigen before delivery of a neonate	II-2	1-6
A nontreponemal screening test should be performed on all women at the prenatal intake visit	II-3	1-5
A urine specimen for culture should be obtained from all women at the prenatal intake visit	I	1-4
All women should be offered an MSAFP determination at between 15 and 20 weeks of gestation	III-2	None
HIV screening tests should be offered to all pregnant women at the prenatal intake visit (optional)	III	None
Every woman should be screened for group B streptococcus genital infection at the prenatal intake visit (optional)	III	6
Office visit–based screening		
Blood pressure measurements should be taken at each prenatal visit	II-2	1-4
Weight should be recorded at each prenatal visit	III	None
Urine protein should be assessed at each prenatal visit	III	None
Fundal height measurement should be recorded at each visit after 20 weeks of gestation	III	1 and 2
Fetal heart tones should be auscultated at each visit after 12 weeks of gestation	III	None
A smoking history should be obtained from every woman at the prenatal intake visit	I	1 and 2
A history of alcohol use should be obtained from every woman at the prenatal intake visit	II-2	1 and 2
A substance abuse history should be obtained from every woman at the prenatal intake visit	III	1 and 2
The clinician should assess the adequacy of the patient's diet at the prenatal intake visit	III	None
Clinician follow-up of abnormal laboratory or screening test results		
The clinician should provide follow-up to abnormal Papanicolaou smear results	III	None
Every woman with an HCT <30% at screening should receive supplemental iron therapy	III	None
Women with positive genital cultures should have a posttreatment follow-up culture	III	None
Women with positive urine cultures should have a posttreatment follow-up culture	I	None
Women with abnormal 50-g glucose challenge results (≥7.8 mmol/L [≥140 mg/dL]) should undergo a 3-hour plasma glucose tolerance test	I	3
Subsequent screening, ultrasonography, or amniocentesis should be performed on all women with an abnormal MSAFP test result	II-2	None
The clinician should investigate possible cause if there is more than a 3-cm difference between fundal height and gestational age	III	None
Women in whom preterm labor is diagnosed should be given tocolytic medications until at least 36 weeks of gestation unless contraindicated	III	None
Twice weekly fetal monitoring should begin at 41 weeks and continue until labor (spontaneous or induced) begins	III	None
If the woman is a smoker, the clinician should document that the patient was warned about the potential harms of smoking, was advised to quit smoking, or both	III	None
If the woman drinks alcohol, the clinician should document that the patient was warned about the potential harms of drinking, was advised to stop drinking, or both	III	None

*MSAFP indicates maternal serum alpha-fetoprotein; HIV, human immunodeficiency virus; and HCT, hematocrit.

†The quality of evidence was assessed using the US Preventive Services Task Force methodology and criteria for supporting evidence.[30] Level I indicates randomized controlled trial; level II-1, controlled trials without randomization; level II-2, well-designed cohort or case-control studies; level II-3, multiple time-series studies; and level III, descriptive studies and opinion-based publications.

‡Groups making the recommendation are coded as follows: 1, US Public Health Service Expert Panel on the Content of Prenatal Care; 2, US Preventive Services Task Force; 3, American College of Obstetricians and Gynecologists; 4, Canadian Task Force on Periodic Health Examination; 5, Centers for Disease Control and Prevention; and 6, American Academy of Pediatrics.

shown). The average gravidity, parity, and number of elective and spontaneous abortions in each group were similar.

INITIATION AND TIMING OF PRENATAL CARE

There were significant differences between groups in gestational age at initial presentation but not in the mean number of prenatal visits (**Table 3**). Patients with Medicaid presented for prenatal care significantly later in pregnancy than either HMO or FFS patients, and a significantly higher proportion of patients with Medicaid did not begin care until the second trimester of pregnancy. All patients were receiving prenatal care by the beginning of the third trimester (data not shown). Despite their late presentation for prenatal care, patients with Medicaid had the same number of prenatal visits as the other 2 groups.

Table 2. Baseline Comparisons Between Groups (N=267)*

Variable	Type of Insurance: Medicaid (n=93)	HMO (n=92)	FFS (n=82)	P
Ethnic status				
White	77.4	80.4	93.8	.01†
African American	19.4	9.8	2.5	
Asian or other	3.3	9.9	3.7	
Household				
Single, living alone	26.1	4.3	4.9	<.001†
Single, living with a partner	17.4	4.3	3.7	
Married	28.3	87.0	87.8	
Separated or divorced	4.3	2.2	0	
Extended family household	22.8	2.2	3.7	
Mean amount of schooling, y	11.5	14.5	14.9	<.001‡
Highest grade completed				
Less than high school	41.9	3.3	3.7	<.001†
High school diploma	34.4	31.5	28.0	
Some college	15.1	15.2	17.1	
College graduate or beyond	8.6	50.0	51.2	
Mean age at delivery of the neonate, y	23.2	29.9	28.5	<.001‡
Prior obstetrical history				
Average gravidity§	2.52	2.48	2.55	.95
Average parity§	1.00	0.84	0.78	.30
Mean No. of medical problems	0.33	0.47	0.52	.15
Gravida I	31	21	23	.29
Para I	45	40	40	.74
With prior elective abortion	22	20	23	.84
With prior spontaneous abortion	15	23	28	.11

*All data given as the percentage of patients in each group unless otherwise specified. HMO indicates health maintenance organization; FFS, fee-for-service.

†Significant differences between groups, χ^2 testing using the Bonferroni correction.

‡Significantly lower for patients with Medicaid than patients with HMO and FFS insurance, 1-way analysis of variance, Scheffe test, P<.05.

§Data not given as a percentage.

Table 3. Initiation and Timing of Prenatal Care (N=267)

Variable	Type of Insurance: Medicaid (n=93)	HMO (n=92)	FFS (n=82)	P
Gestational age at initial visit, wk	14.5	10.4	10.7	<.001†
Percentage of patients with no visit in the first trimester	33	2	9	<.001‡
Mean No. of prenatal visits	11.9	11.8	10.9	.09

*HMO indicates health maintenance organization; FFS, fee-for-service.

†Patients with Medicaid initiated prenatal care significantly later than patients with HMO and FFS insurance, Scheffe test, P<.05.

‡Significant differences between groups, χ^2 testing using the Bonferroni correction.

ADHERENCE TO QUALITY-OF-CARE CRITERIA FOR SCREENING TESTS AND PROCEDURES

Table 4 displays adherence rates for patients in each insurance group for each of the 12 selected screening procedures, as well as the summary quality score for screening tests. The quality score was not significantly different across groups, with adjusted scores of 82% for patients with Medicaid, 83% for patients with HMO insurance, and 81% for patients with FFS insurance ($P=.77$). However, this aggregate score obscured significant differences in rates of performance of individual tests. Significantly more patients with Medicaid were screened for genital infection with gonorrhea and chlamydia fluorescent antibody studies, and significantly fewer patients with Medicaid were screened for anemia or GDM at 28 weeks' gestation.

Documentation of urinary tract infection screening and the offer of maternal serum α-fetoprotein screening was rather low across all groups. Group B streptococcal infection screening was infrequently performed in all 3 groups. Human immunodeficiency virus screening (either the performance of a test or the statement that the patient declined the test) was documented more frequently for patients with FFS insurance than for the other 2 groups.

The differences in the performance of individual tests do not correspond with the differences between groups in the proportion of abnormal test results, with a single exception: patients with HMO and FFS insurance were significantly more likely than patients with Medicaid to have abnormal 28-week GDM screening test results (20% vs 8%, $P<.001$) (data not shown).

ADHERENCE RATES FOR VISIT-BASED SCREENING CRITERIA

Adjusted quality scores for visit-based screening showed no significant differences between groups (**Table 5**), with rates of 98% for patients with Medicaid and 96% for patients with HMO and FFS insurance ($P=.34$). No significant differences were seen between groups for any of the 5 individual criteria.

ADHERENCE RATES FOR SUBSTANCE USE SCREENING CRITERIA

A significant ($P<.001$) difference between groups was seen in raw quality scores for substance use screening, but this was found to be due to a combination of practice effect and practice-insurance interaction (one practice, heavily weighted with patients with FFS insurance, failing to document nutritional assessment for any patients). Adjusted quality scores for the 3 groups did not significantly differ: 87% for patients with Medicaid, 88% for patients with HMO insurance, and 81% for patients with FFS insurance ($P=.25$). The proportion of abnormal or positive responses to screening questions did vary across groups, with patients with Medicaid significantly more likely to report use of cigarettes (41% vs 11% [HMO] and 17% [FFS], $P<.001$) and other drug use (12% vs 0% [HMO] and 2% [FFS], $P=.01$) and more likely to have a diet rated as inadequate (74% vs 14% [HMO] and 21%

Table 4. Adherence to Screening Laboratory Test Criteria (N=267)*

Laboratory Test	Type of Insurance: Medicaid (n=93)	HMO (n=92)	FFS (n=82)	P
Required				
Blood type or antibody screen	97	100	100	.09
Rubella antibody titer	97	97	100	.29
CBC or hematocrit	83	96	96	.001†
Screening glucose	24	58	44	<.001†
24-28 Week glucose	82	93	100	<.001†
Papanicolaou smear	100	98	95	.05‡
Gonorrhea	96	72	65	<.001§
Chlamydia	98	78	63	<.001§
Hepatitis B	88	97	96	.03‡
VDRL (syphilis)	89	90	92	.88
Screening urine culture	51	55	55	.77
MSAFP‖	65	72	83	.03‡
Optional				
HIV‖	69	63	90	<.001¶
Group B streptococcus	3	14	11	.13
Adjusted quality score	*82*	*83*	*81*	*.77*

**All data given as the percentage of patients in each group who adhered to the criteria. χ^2 analysis was used to compare the mean adherence rates across groups for individual tests. Analysis of variance was used to compare mean adjusted quality scores across groups (see text for details). Optional tests were not used in the calculation of the quality score. HMO indicates health maintenance organization; FFS, fee-for-service; CBC, complete blood cell count; MSAFP, maternal serum alpha-fetoprotein; and HIV, human immunodeficiency virus.*

†Patients with Medicaid had significantly lower adherence than patients with HMO and FFS insurance.

‡No significant differences were noted between groups after application of Bonferroni correction for multiple comparisons or on χ^2 analysis of 2 by 2 tables.

§Patients with Medicaid had significantly greater adherence than patients with HMO and FFS insurance.

‖For these criteria, adherence was defined by either presence of the test in the medical record or a written statement in the medical record that the patient declined the test.

¶Patients with FFS insurance had significantly greater adherence than patients with Medicaid and HMO insurance.

Table 5. Adherence to Visit-Based Screening, Substance Use Screening, and Clinician Follow-up Criteria (N=267)*

Procedure	Type of Insurance: Medicaid (n=93)	HMO (n=92)	FFS (n=82)	P
Visit-based screening criteria				
Blood pressure	99	99	99	.76
Maternal weight	99	99	99	.72
Urine protein or glucose	96	97	98	.41
Fundal height	96	90	93	.09
Fetal heart rate	98	95	95	.26
Adjusted quality score	*98*	*96*	*96*	*.34*
Substance use screening criteria				
Screening for smoking	100	100	100	.99
Screening for alcohol use	100	100	99	.85
Screening for other drug use	99	97	96	.35
Screening for diet adequacy	63	63	42	<.001†
Adjusted quality score	*87*	*88*	*81*	*.25*
Clinician follow-up criteria				
Adjusted quality score	*83* (n=65)	*81* (n=53)	*88* (n=47)	*.40*

**All data given as the percentage of patients in each group who adhered to the criteria. HMO indicates health maintenance organization; FFS, fee-for-service.*

†Patients with FFS insurance had significantly lower adherence than patients with Medicaid and HMO insurance.

[FFS], $P<.001$). The high proportion of inadequate diet recorded in the Medicaid sample is likely the result of the need to provide documentation to support enrollment in the Michigan Department of Social Services Maternal Support Services program.

ADHERENCE RATES FOR CLINICIAN FOLLOW-UP CRITERIA

No significant differences were seen between groups in this quality score ($P=.40$) (Table 5): the highest adjusted quality score was observed for patients with FFS insurance, followed by patients with Medicaid and HMO insurance. Because of the small number of patients eligible for each individual criterion, statistical analysis was not performed for individual items.

COMMENT

In this study, we systematically reviewed existing recommendations and guidelines regarding the provision of prenatal care to create evidence-based community standards for the process of routine, low-risk prenatal care in 4 areas: the performance of routine prenatal screening tests; the performance of visit-based screening procedures; substance use screening; and clinician identification and management of abnormal procedure results, test results, or clinical findings. We then developed a systematic method for collecting, coding, and aggregating data present in the prenatal record to allow comparison against those community standards. Our intent was to rigorously search for differences in the quality of prenatal care provided to patients with Medicaid, HMO, and traditional FFS insurance. To our knowledge, this is the first reported comparison of the quality of prenatal care across these 3 insurance groups.

Overall, we found no significant differences in the 4 summary quality scores (screening tests, visit-based screening, substance use screening, and clinician follow-up of abnormal findings) across the 3 groups. This null finding reflects favorably on the relative quality of care provided to Medicaid recipients in this area. However, the overall similarity between groups in quality scores did obscure some significant differences in adherence to individual quality-of-care criteria in the area of screening tests.

First, patients with Medicaid were significantly more likely to be screened for genital infection than were pa-

tients with HMO or FFS insurance. This may reflect the clinical perception that prenatal patients with Medicaid are at higher risk for infection and ought to be carefully screened. However, there is no published evidence to support this perception. We found cases of both types of genital infection (chlamydia and gonorrhea) in the HMO and FFS samples and no significant differences between the 3 groups in the proportion of abnormal screening test results. The recommendation for universal screening made by our panel was based in large part on the high rate of asymptomatic chlamydial infection across all socioeconomic strata and its potential consequences.[22,23] Although the adherence rate for patients with HMO and FFS insurance is at or above the rates seen for gonorrheal or chlamydial screening reported from other settings,[24,25] we believe the lower rate of screening seen in patients with HMO and FFS insurance represents a true difference in quality of care and provides an opportunity for quality improvement for patients in the HMO and FFS groups.

SECOND, PATIENTS with Medicaid were significantly less likely to undergo screening tests for GDM. Late screening for prenatal care in the Medicaid population may have affected the initial serum glucose screening, but all patients were receiving prenatal care by the 28th week of gestation. Again, the difference in rate may reflect the clinical perception of less risk of GDM in the younger Medicaid population. In our sample, patients with Medicaid did have a significantly lower rate of abnormal 28-week screening test results than either the HMO or FFS samples; however, at 8% this still represents considerable risk.

Third, patients with Medicaid had a significantly lower rate of documented screening for anemia than the other 2 groups. Although it is possible that anemia screening for Medicaid recipients might occur at a facility other than the clinician's office (such as the county health department), at the time of this study there were no sites known to provide this service on a routine basis. Moreover, regardless of where screening might have occurred, clinicians at each practice were responsible for documenting the results of screening to manage anemia if present. The 82% adherence rate for anemia screening observed in this study is lower than the greater than 90% rates reported elsewhere.[24-26] We believe this represents a real difference in the quality of care between groups, particularly in light of the documented concern regarding dietary adequacy in the Medicaid population, and an opportunity for quality improvement for patients with Medicaid.

This pattern of results may reflect clinician overvigilance for sexually transmitted disease and undervigilance for more common problems such as anemia and GDM in the Medicaid population. Although this is only a tentative interpretation based on a small sample of patients, it raises an interesting question about how clinicians perform risk assessment in practice. Do clinicians attribute risk profiles to a group of patients (eg, the Medicaid population) or do they make individualized screening decisions? This issue warrants further exploration.

The other major difference between groups was seen in the initiation and timing of prenatal care. In this study, patients with Medicaid presented for care at a significantly later gestational age than patients with HMO or FFS insurance. This finding is consistent with findings reported elsewhere[8,9]; however, as previously noted, it represents a more indirect measure of clinician or office performance than the medical record–derived quality scores. Although the early initiation of prenatal care is unquestionably an important factor in the overall quality of prenatal care, late presentation is also dependent on factors beyond the control of office or clinician. Although this finding clearly represents an opportunity for quality improvement for patients with Medicaid, changes in clinician practice or procedure will likely have minimal effect on the problem of late presentation for prenatal care unless coupled with other community outreach efforts.[27,28]

Several potential limitations in this study warrant discussion. First, the study population is somewhat atypical, and findings from this well-served, well-educated Midwestern university community should be generalized to other communities with caution. Second, to preserve statistical power in this relatively small sample size, reported results were adjusted only for practice identification. As other potential confounders were examined and found to have no significant effect, it is unlikely that this limitation affected our major findings. The third and most important potential limitation is the choice of summary quality measures to compare prenatal care quality across the 3 groups.

We believe that the criteria chosen for inclusion in quality scores accurately reflect the quality of much of the prenatal care process. They were derived from existing evidence-based guidelines or recommendations as amended by a local expert panel, and each item was approved by each panel member as an important and measurable element in the quality of prenatal care. The use of a community panel proved quite valuable in that it created clinician trust in the quality assessment method and enhanced community "buy in" to the implementation of managed care. Although the final criteria closely resemble those published by the American College of Obstetricians and Gynecologists,[29] US Preventive Services Task Force,[30] and the Expert Panel on the Content of Prenatal Care of the US Public Health Service,[31] there were a few important differences. Our screening laboratory criteria included maternal serum alpha-fetoprotein screening, a recent Papanicolaou smear, and a screening serum glucose determination. We included more detailed measures of office-based and substance use screening. Most importantly, we created criteria to assess whether clinicians responded to abnormal prenatal laboratory or screening test results, an area almost entirely missing from existing criteria. We also believe that our results represent differences in the actual care provided to patients, not simply differences in documentation or data retrieval. All practices in this study used standardized obstetrical prenatal records and flow sheets, and reviewers followed structured protocols to obtain data. Criteria that could not be reliably found in the record were dropped. Finally, reviewers did not report any differences in the

degree of difficulty of data retrieval between individual practices.

However, the criteria used in this study do not completely reflect the process of prenatal care. Some important aspects of the process of prenatal care could not be assessed or measured. We were unable to measure the interpersonal aspects of care across all 3 insurance groups. The retrospective nature of the study precluded a survey of patient satisfaction or clinician-patient communication. Although a subset of the patients with Medicaid included in this study completed a semistructured interview regarding their perceptions of the Medicaid program, their relationships with health care providers, and perceived barriers to health care services as part of this study, survey responses could not be directly linked to individual patients. As previously mentioned, poor data quality prevented our measurement of any psychosocial aspects of prenatal care. We were also unable to determine the effectiveness or appropriateness of interventions made in response to abnormal screening test results, only that recognition or intervention occurred. Future work in this area would greatly benefit from the use of multimethod research designs,[27,32-34] which could couple qualitative data on the interpersonal and psychosocial aspects of care[27] or observational data on clinician-patient interaction[35] or both with the quantitative data available from summary process measures.

This is the third study published in recent months to assess the quality of prenatal care using explicit process criteria. Murata et al[24] assessed the quality of prenatal care in 6 HMO settings across the United States. An expert panel consisting of obstetricians and perinatologists working across geographic regions used an evidence-based methodology to identify prenatal care process criteria similar to those used in this study. Baldwin et al[25] compared the performance of different provider groups in Washington state using existing American College of Obstetricians and Gynecologists' prenatal care process guidelines. Although the studies ask slightly different questions and employ different analytic methods, results are provided in similar formats and are in general quite comparable. Many of the criteria are sufficiently similar to allow basic comparisons or "benchmarking" across studies: for example, screening for hepatitis B was documented for 77.9% of the sample in the study by Murata et al, 79.0% of the sample in the study by Baldwin et al, and 94.0% of our sample. Murata et al also reported adherence rates for condition-specific care, operationally defined as "diagnostic and treatment interventions following abnormal screening tests and condition specific care designed to mitigate the effects of pregnancy complications"; their criteria are similar to our clinician follow-up criteria, and their overall adherence rate of 69.8% is comparable with our 82.0% rate.

As discussed in the article by Murata et al and its accompanying editorial,[36] the derivation and use of summary quality measures offers a promising approach to evaluation of the quality of various health care plans. A carefully derived composite score can provide a meaningful basis for comparing health plan performance. However, our results also highlight an important caveat to the use of summary quality measures: they can obscure notable differences in the individual items that compose the summary score. At their current level of development, summary measures should be interpreted with caution and probably should only be reported in combination with their individual components.

CONCLUSION

In this study, the quality of care provided to prenatal patients covered by 3 different types of health insurance was compared using community-derived explicit process criteria. We found no significant differences in the overall quality of prenatal care provided to patients with Medicaid, HMO, and FFS insurance as measured by 4 summary quality scores; the adherence rates for specific criteria were similar to those reported in recent studies examining the prenatal care process. However, patients with Medicaid presented for prenatal care significantly later in pregnancy and received a different package of screening tests (increased screening for genital infection and decreased screening for GDM and anemia) than the other 2 groups; clinicians may have altered screening protocols based on preexisting perceptions of patient risk.

Although summary quality measures show promise as a tool for comparing the quality of care provided in diverse health care settings, at present they provide an incomplete picture of the quality of the prenatal care process and must be interpreted with caution. Future work would benefit from the use of multimethod research, which could couple qualitative or observational data on the psychosocial aspects of prenatal care with the quantitative data available from summary process measures.

Accepted for publication November 14, 1996.

This study was supported by grant HDP 92-006 from the Michigan Department of Social Services, Lansing.

Corresponding author: Michael S. Klinkman, MD, MS, Department of Family Practice, University of Michigan, 1018 Fuller Rd, Ann Arbor, MI 48109-0708 (e-mail: mklinkma@umich.edu).

REFERENCES

1. Ware JE, Brook RH, Rogers WH, et al. Comparison of health outcomes at a health maintenance organization with those of fee-for-service care. *Lancet.* 1986;1:1017-1022.
2. Yelin EH, Shearn MA, Epstein WV. Health outcomes for a chronic disease in prepaid group practice and fee for service settings: the case of rheumatoid arthritis. *Med Care.* 1986;24:236-247.
3. Udvarhelyi S, Jennison K, Phillips RS, et al. Comparison of the quality of ambulatory care for fee-for-service and prepaid patients. *Ann Intern Med.* 1991;115:394-400.
4. Francis AM, Polissar L, Lorenz AB. Care of patients with colorectal cancer: a comparison of a health maintenance organization and fee-for-service practices. *Med Care.* 1984;22:418-429.
5. Greenfield S, Rogers W, Mangotich M, Carney MF, Tarlov AR. Outcomes of patients with hypertension and non–insulin-dependent diabetes mellitus treated by different systems and specialties: results from the Medical Outcomes Study. *JAMA.* 1995;274:1436-1444.
6. Ware JE, Bayliss MS, Rogers WH, Kosinski M, Tarlov AR. Differences in 4-year health outcomes for elderly and poor, chronically ill patients treated in HMO and fee-for-service systems: results from the Medical Outcomes Study. *JAMA.* 1996;276:1039-1047.

7. Hurley RE, Freund DA, Paul JE. *Managed Care in Medicaid: Lessons for Policy and Program Design.* Ann Arbor, Mich: Health Administration Press; 1993.
8. Carey TS, Weis K, Charles H. Prepaid versus traditional Medicaid plans: lack of effect on pregnancy outcomes and prenatal care. *Health Serv Res.* 1991;26:165-181.
9. Krieger JW, Connell FA, LoGerfo JP. Medicaid prenatal care: a comparison of use and outcomes in fee-for-service and managed care. *Am J Public Health.* 1992; 82:185-190.
10. Carey T, Weis K, Homer C. Prepaid versus traditional Medicaid plans: effect on preventive health care. *J Clin Epidemiol.* 1990;43:1213-1220.
11. Goldfarb NI, Hillman AL, Eisenberg JM, Kelley MA, Cohen AV, Dellheim M. Impact of a mandatory Medicaid case management program on prenatal care and birth outcomes: a retrospective analysis. *Med Care.* 1991;29:64-71.
12. Rosenbaum S, Hughes D, Butler E, Howard D. Incantations in the dark: Medicaid, managed care, and maternity care. *Milbank Q.* 1988;66:661-693.
13. Jencks SF, Benedict MB. Accessibility and effectiveness of care under Medicaid. *Health Care Financing Rev.* 1990;11(suppl):47-56.
14. Hueston WJ, Spencer E, Kuehn R. Differences in the frequency of cholesterol screening in patients with Medicaid compared with private insurance. *Arch Fam Med.* 1995;4:331-334.
15. Kogan MD, Alexander GR, Kotelchuk M, Nagey DA. Relation of the content of prenatal care to the risk of low birthweight: maternal reports of health behavior advice and initial prenatal care procedures. *JAMA.* 1994;271:1340-1345.
16. Murata PJ, McGlynn EA, Siu AL, Brook RH. *Prenatal Care: Literature Review and Quality Assessment Criteria.* Santa Monica, Calif: Rand Corp; 1992.
17. Kesner DM, Kalk CE, Singer J. Assessing health quality: the case for tracers. *N Engl J Med.* 1973;288:189-194.
18. Poland ML, Ager JW, Olson KL, Sokol RJ. Quality of prenatal care; selected social, behavioral, and biomedical factors; and birth weight. *Obstet Gynecol.* 1990; 75:607-611.
19. Kotelchuk M. *Overview of Prenatal Care Utilization Index.* Chapel Hill: School of Public Health, University of North Carolina; 1987.
20. Hansell MJ. Sociodemographic factors and the quality of prenatal care. *Am J Public Health.* 1991;81:1023-1028.
21. Cohen J. *Statistical Power Analysis for the Social Sciences.* 2nd ed. Hillsdale, NJ: Lawrence A Erblaum Associates; 1988.
22. Faro S. *Chlamydia trachomatis*: female pelvic infection. *Am J Obstet Gynecol.* 1991;164:1767-1770.
23. Cohen I, Veille JC, Calkins B. Improved pregnancy outcome following successful treatment of chlamydial infection. *JAMA.* 1990;263:3160-3163.
24. Murata PJ, McGlynn EA, Siu AL, et al. Quality measures for prenatal care: a comparison of care in six health care plans. *Arch Fam Med.* 1994;3:41-50.
25. Baldwin LM, Raine T, Jenkins LD, Hart LG, Rosenblatt R. Do providers adhere to ACOG standards? the case of prenatal care. *Obstet Gynecol.* 1994;84:549-556.
26. Melnikow J, Alemagno S. Adequacy of prenatal care among inner-city women. *J Fam Pract.* 1993;37:575-582.
27. Bosch SJ, Merino R, Daniels MS, Fischer EP, Rosenthal M. A proposed network to improve access to high-quality health care for Medicaid-eligible families. *J Community Health.* 1979;4:302-311.
28. Brown SS, ed. *Prenatal Care: Reaching Mothers, Reaching Infants.* Washington, DC: National Academy Press; 1988.
29. Frigoletto FD, Little GA. *Guidelines for Perinatal Care.* Elk Grove Village, Ill: American Academy of Pediatrics; and Washington, DC: American College of Obstetricians and Gynecologists; 1988.
30. US Preventive Services Task Force, US Public Health Service. *Guide to Clinical Preventive Services: An Assessment of the Effectiveness of 169 Interventions.* Baltimore, Md: Williams & Wilkins; 1989.
31. Public Health Service Expert Panel on the Content of Prenatal Care. *Caring for Our Future: The Content of Prenatal Care.* Washington, DC: Public Health Service, US Dept of Health & Human Services; 1989.
32. Brewer J, Hunter A. *Multimethod Research: A Synthesis of Styles.* Newbury Park, Calif: Sage Publications; 1989.
33. Stange KC, Zyzanski SJ. Integrating qualitative and quantitative research methods. *Fam Med.* 1989;21:448-451.
34. Stange KC, Miller WL, Crabtree BF, O'Connor PJ, Zyzanski SJ. Multimethod research: approaches for integrating qualitative and quantitative methods. *J Gen Intern Med.* 1994;9:278-282.
35. Callahan EJ, Bertakis KD. Development and validation of the Davis Observation Code. *Fam Med.* 1991;23:19-24.
36. Longo DR, LeFevre ML. The quality of prenatal care: issues for consumer reports and health care reform. *Arch Fam Med.* 1994;3:37-39.