8-26-2004

Studying Effects of Primary Care Physicians and Patients on the Trade-Off Between Charges for Primary Care and Specialty Care Using a Hierarchical Multivariate Two-Part Model

John W. Robinson
Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics, jrobinso@jhsph.edu

Scott L. Zeger
The Johns Hopkins Bloomberg School of Public Health, szeger@jhsph.edu

Christopher B. Forrest
Johns Hopkins Bloomberg School of Public Health, Department of Health Policy & Management, cforrest@jhsph.edu

Studying Effects of Primary Care Physicians and Patients on the
Trade-Off Between Charges for Primary Care and Specialty Care
Using a Hierarchical Multivariate Two-Part Model

John W. Robinson, M.D., Ph.D., Scott L. Zeger, Ph.D., and
Christopher B. Forrest, M.D., Ph.D.

August 26, 2004

ABSTRACT

Objective. To examine effects of primary care physicians (PCPs) and patients on the association between charges for primary care and specialty care in a point-of-service (POS) health plan.

Data Source. Claims from 1996 for 3,308 adult male POS plan members, each of whom was assigned to one of the 50 family practitioner-PCPs with the largest POS plan member-loads.

Study Design. A hierarchical multivariate two-part model was fitted using a Gibbs sampler to estimate PCPs' effects on patients' annual charges for two types of services, primary care and specialty care, the associations among PCPs' effects, and within-patient associations between charges for the two services. Adjusted Clinical Groups (ACGs) were used to adjust for case-mix.

Principal Findings. PCPs with higher case-mix adjusted rates of specialist use were less likely to see their patients at least once during the year (estimated correlation: –.40; 95% CI: –.71, –.008) and provided fewer services to patients that they saw (estimated correlation: –.53; 95% CI: –.77, –.21). Ten of 11 PCPs whose case-mix adjusted effects on primary care charges were significantly less than or greater than zero ($p < .05$) had estimated, case-mix adjusted effects on specialty care charges that were of opposite sign (but not significantly different than zero). After adjustment for ACG and PCP effects, the within-patient, estimated odds ratio for any use of primary care given any use of specialty care was .57 (95% CI: .45, .73).

Conclusions. PCPs and patients contributed independently to a trade-off between utilization of primary care and specialty care. The trade-off appeared to partially offset significant differences in the amount of care provided by PCPs. These findings were possible because we employed a hierarchical multivariate model rather than separate univariate models.

Key Words. Provider profiling; Referral to specialists; Point-of-service health plan; Gibbs sampling

INTRODUCTION

Previous studies have found substantial variation among primary care physicians (PCPs) in the proportion of their patients that visit specialists, even after case-mix adjustment (Salem-Schatz et al. 1994; Franks et al. 2000). Two characteristics of PCPs' practices that have been found to contribute to this variation are range of diagnoses treated and number of visits per hour. PCPs that treat a narrower range of diagnoses have been found to have a higher proportion of their patients visit specialists (Franks et al. 2000), possibly because their narrower scope of practice makes it necessary for them to refer some clinical presentations to specialists that other PCPs would evaluate and manage themselves; and PCPs that see a greater number of patients per hour have also been found to have higher proportion of their patients visit specialists (Salem-Schatz et al. 1994), possibly because time constraints cause them to refer patients to specialists that they would otherwise treat themselves or induce patients to self-refer to specialists because of difficulty obtaining timely appointments. Hence, these two practice characteristics that are associated with an increase in the proportion of a PCP's patients that visit specialists, might also be associated with a decrease in the amount of care provided by the PCP on a per-patient basis.

Multivariate Provider Profiling

A relatively simple approach to looking for evidence of such a trade-off between PCPs' effects on patients' use of specialty care and primary care would be to estimate PCPs' effects on the use of each type of service separately, and then to compare results, PCP by PCP. Such an analysis amounts to the performance of two separate, univariate provider profiling procedures. The essential limitation of such an approach is that it does not allow for estimation of measures of association between the use of primary care and specialty care and the uncertainty of those estimates, at the level of either PCP or patient.

On the other hand, if PCPs' effects on utilization of primary care and specialty care are estimated simultaneously, using a hierarchical multivariate model, it is possible to estimate measures of association between utilization of primary care and specialty care and the uncertainty of those estimates, at the levels of both PCP and patient. Here, we describe a study in which this was done, using as measures of utilization, annual per-patient charges in a point-of-service (POS) health plan, for outpatient evaluation and management services provided by PCPs and medical and non-ophthalmologic surgical specialists. This study is an example of *multivariate* provider profiling, which has the capability of estimating not only providers' effects on multiple dependent variables but also, the associations among these effects (Bronskill, Normand, Landrum, and Rosenheck 2002; Burgess, Lourdes, and West 2000; Landrum, Bronskill, and Normand 2000; Landrum, Normand, and Rosenheck 2003).

Referral to Specialists Under Point-of-Service Plans

When patients belong to a health maintenance organization (HMO) which requires that PCPs, functioning as "gatekeepers", approve all specialist referrals, it is reasonable to assume that PCPs have played a role in essentially all specialist visits for which claims have been paid (Salem-Schatz et al. 1994; Franks et al. 2000). However, when patients belong to a managed care plan that includes a PCP gatekeeping function but also allows patients to self-refer to specialists at higher out-of-pocket cost, claims include specialist visits that were not approved by PCPs. One such type of plan is a "point-of-service" (POS) plan, that blends HMO, preferred-provider, and traditional indemnity benefits in a manner that allows members to obtain referrals from their PCPs or alternatively, to self-refer to specialists, but at higher out-of-pocket cost (Bodenheimer 1996).

It does appear that under POS plans, PCPs still play a role in most decisions to use specialists. To begin with, in a study of three geographically diverse POS plans, 70% to 83% of members who visited specialists were referred by their PCPs and the remaining minority referred themselves, exercising their preferred-provider or indemnity benefits (Forrest et al. 2001). Additionally, PCPs appeared to play a role in the decision-making of many self-referring members, since in response to a survey of members who had visited specialists under one of the plans, those who self-referred reported lower satisfaction and less continuity with their PCPs compared with those who were PCP-referred (Braun et al. 2003). Thus, it appears that in the POS setting, PCPs affect the great majority of specialist referrals either directly by approving or disapproving referrals or indirectly through the quality of their work with patients.

Case-Mix Adjustment

Because the assignment of patients to profiled providers is not random, it is important in any profiling study to adjust providers' estimated effects on dependent variables for patient morbidity and demographic variables of known importance. In this study, that adjustment was accomplished using Adjusted Clinical Groups (ACGs), a morbidity taxonomy of 93 mutually exclusive categories based on age, gender, and 12 months of diagnoses, developed by researchers and practicing physicians to sort patients, solely based on information from healthcare claims, into face-valid categories predictive of current and future healthcare utilization (Health Services Research and Development Center 2001). The ability of ACGs to predict utilization can be roughly measured by an analysis of variance of annual, per patient charges for ambulatory health services, using ACGs as a one-way classification. Investigators have done so, using data from numerous public- and private-sector health plans, and have obtained values of R^2 ranging from

.34 to .47, indicating that 34% to 47% of the variability in charges can be explained by ACG classification (Reid et al. 2001; Weiner et al. 1991).

Since ACGs have been shown to explain just one-third to one-half of the variation in ambulatory charges, for this study we had to assume that this diagnosis-based risk adjustment approach would not completely control for differences in case-mix. Further, we anticipated that incomplete risk-adjustment would be most likely to bias inferences regarding PCPs with small case-loads, for whom the effect of a few incompletely-adjusted charges would be greatest. This was one motivation for fitting a hierarchical model, which would induce an adjustment, or "shrinkage" of individual PCP effects toward the overall mean of PCP effects, thereby partially offsetting any bias due to incomplete risk-adjustment. Further, such shrinkage would be greatest for PCPs with the smallest case-loads and most extreme risk-adjusted mean charges, precisely those PCPs for whom the likelihood of incomplete risk-adjustment was greatest (Burgess et al. 2000; DeLong et al. 1997; Goldstein and Spiegelhalter 1996, Shaihan et al. 2001).

A Hierarchical Multivariate Two-Part Model

In studies such as this one, that include users and non-users of a healthcare service, charges for that service are a mixture of zeros and highly-skewed, positive values that cannot be approximated by a normal distribution or any other simple parametric form. Such a mixture can be thought of as resulting from a two-part process, the first part determining whether any use of the service occurred and the second part determining the amount of charges if use occurred. This is the conceptual basis for the two-part model (Duan et al. 1983), under which charges for a single service are represented by a binary variable that indicates whether any use occurred and if use occurred, also by a continuous variable, the natural logarithm of charges, here referred to as the "log of positive charges". The log of positive charges is used because its distribution is generally, more nearly normal than the distribution of positive charges. The effects of covariates (such as indicator variables representing PCP assignment) on the binary and continuous dependent variables are estimated separately, using regression procedures; and then results from these regressions are combined to estimate effects of the covariates on charges for the service.

Although the above model involves two dependent variables, we will refer to it as a *univariate* two-part model; because the relationship between the two dependent variables is deterministic, not stochastic, and each of the two parts of the model involves a single dependent variable. For this study, because we were interested in the associations between PCPs' effects on charges for two services, primary care and specialty care, we employed a hierarchical *multivariate* two-part model, which has been described in detail elsewhere (same authors as this manuscript 2004). As will be described in Methods, this model does involve stochastic relationships between dependent variables.

METHODS

The POS Health Plan Study

The study presented here involves a sample of PCPs who participated in a POS health plan offered by a not-for-profit insurer in the Northeast. The plan was one of three that contributed administrative data to a study of referral patterns in POS plans (Forrest et al. 2001). Each member of the northeastern POS plan selected a PCP from among those participating; female members could optionally select an obstetrician-gynecologist as a second PCP (an "ObGyn-PCP"); and all members could change PCPs and/or ObGyn-PCPs as often as they wished. When a member exercised the HMO benefit, the member's PCP or ObGyn-PCP functioned as a gatekeeper, deciding whether to authorize specialist referrals. Alternatively, a member could exercise the preferred-provider or indemnity benefit and self-refer to a specialist, at higher out-of-pocket cost.

Profiling Sample Selection

We elected to profile PCPs from just one primary care specialty, family practice so that we could reasonably assume that PCPs' effects on utilization, which would be represented by random parameters, were identically distributed. From a total of 266 family physicians participating as PCPs, we selected the 50 with the largest POS plan member-loads in order to assure that each PCP's patient sample included an adequate number of users of specialty care.

Our patient sample consisted of all male POS plan members, ages 18 - 64, who were enrolled in the plan for 12 months of 1996, were assigned solely to one of the 50 selected PCPs, were not diagnosed with a psychiatric disorder, and received at least one claimed health service of any type. Reasons for the exclusions were as follows: females and members that changed PCPs during 1996 because our model could not estimate the combined effects of two or more PCPs on a patient's utilization; members younger than 18 because we planned to later compare family practice PCPs with internal medicine PCPs, who do not generally treat that age group; members 65 and older because the POS plan did not cover seniors; members enrolled fewer than 12 months because we required 12 months of claims data; members with psychiatric diagnoses because psychiatric claims data was not available; and members without claims because at least one claim (for any type of health service) was necessary for risk adjustment. Applying these criteria, 3,308 patients were included, resulting in within-PCP sample sizes ranging from 30 to 152. (A total of 102,943 persons were enrolled in the POS plan for all or part of 1996. Of these, 38,878 were adults, enrolled for all 12 months and of these, 18,375 were men.)

To measure utilization, we used annual allowed, per-patient charges for two services, primary care and specialty care, where primary and specialty care refer to outpatient evaluation and management services provided by PCPs and medical and non-ophthalmologic surgical specialists, respectively. We chose allowed rather than billed charges, because allowed charges were set by the managed care plan and therefore generally uniform across providers.

Case-Mix Adjustment Using ACGs

For the POS plan study, each patient had been assigned to an Adjusted Clinical Group (ACG). Of the 93 ACGs, only 40 were represented in the profiled sample because many applied only to women or children. Some of the 40 ACGs were represented so infrequently in the profiled sample that we could not validly estimate their effects on utilization solely on the basis of information contained in that sample. For example, six ACGs had frequencies of less than 10 among the 3,308 sampled patients, and five had frequencies of one or two among the 1,004 patients with positive charges for specialty care. However, all of the ACGs were well represented in the overall POS plan membership from which our sample had been drawn. So, using the 38,878 adults enrolled in the plan for 12 months of 1996, we ranked the 40 ACGs on each of the two services, primary care and specialty care, on the basis of proportion of members using the service and mean charges among users of the service, resulting in four sets of "prior" ranks. Thus, for each ACG that was infrequent in our profiled sample, we had identified other ACGs of similar prior rank that were better represented, allowing us through statistical modeling, to use observations on the better-represented ACGs to increase the power of estimation of the effects of infrequent ACGs.

Statistical Model and Estimation

For statistical analysis we employed a hierarchical multivariate two-part model (same authors as this manuscript 2004), salient features of which are depicted in Figure 1. The model involves four dependent variables, two for each service: Binary variables, U_1 and U_2 equal 1 or 0, indicating whether or not there was any use of primary care and specialty care, respectively. Continuous variables, Y_1 and Y_2 equal the log of positive annual charges for primary care and specialty care, if the respective services were used.

The first level of the hierarchy is *within-patient*: Figure 1 depicts the six within-patient associations between the four dependent variables. The association between any use of the two services is represented by an odds ratio. The association between the log of positive charges for the two services is represented by a correlation. The association between any use of one service and the log of positive charges for the *other* service is represented by a regression of Y| U=1 for one service on U for the other. The association between any use of a service and the log of positive charges for the *same* service is deterministic, not stochastic. (For example, U_1 determines whether Y_1 is observed.)

The second level is *within-PCP*: The effects of PCPs on the four dependent variables are represented using parameters from the regressions of the log odds of any use of each service and the log of positive charges for each service on 50 indicator ("dummy") covariates representing patients' PCP assignments and additional covariates (described below) representing patients' ACG assignments. Since there are four regressions, each PCP has four regression effects, labeled b_{U1}, b_{U2}, b_{Y1}, and b_{Y2} in Figure 1, which can be correlated within-PCP. Also, because model-based predictions of per-patient charges will depend upon estimates of both the mean and variance of the log of positive charges, we allow the variance of log of positive charges for each service to differ by PCP, or be "PCP-specific" (not depicted in Figure 1). Thus, each PCP is represented by four regression effects and two variances; all regarded as random variables under the hierarchical model.

The third level is *between-PCP* (and not fully depicted in Figure 1): The set of four regression effects for each of the 50 PCPs is assumed to be an independent random sample from a multivariate normal distribution with mean zero, implying that correlations between the four regression effects are the same for all PCPs. Separately for each of the two services, the 50 PCP-specific variances are assumed to be independent random samples from an inverse chi-square distribution, parameterized in such a way that its mean approximately equals the harmonic mean of the 50 PCP-specific variances.

For case-mix adjustment, the effects of ACGs on the four dependent variables are represented as parameters of the regressions of the log odds of any use of each service and the log of positive charges for each service on piecewise polynomial expansions of the prior ACG ranks and 40 random indicator ("dummy") covariates representing the 40 ACGs. Details of this aspect of the model are not included in Figure 1, but are presented elsewhere (same authors as this manuscript 2004).

The model was fitted using a Gibbs sampler, a method of simulating the joint posterior distribution of model parameters given observed data (Gelfand 2000). As a Markov chain Monte Carlo method, Gibbs sampling is typically used to fit Bayesian models such as this one, for which parameter estimates and measures of their uncertainty cannot be expressed as formulaic functions of observed data values (Gilks, Richardson, and Spiegelhalter 1996). The Gibbs sampler output, in the form of the simulated joint posterior distribution of parameter values, can be used to accurately compute estimates of model parameters and measures of uncertainty for those estimates. Here, we use the mean and 95% credible interval of the simulated parameter value as the parameter estimate and measure of its uncertainty, respectively. (The 95% credible interval is the range within which the parameter can be found with 95% probability, given the model assumptions and observed data values.)

A program for the Gibbs sampler was written and implemented in SAS Interactive Matrix Language (SAS Institute 1999). A detailed description of the statistical composition of the sampler and its implementation is presented elsewhere (same authors as this manuscript 2004).

PCPs' Predicted Annual Charges and Deviations

Using parameter values from the fitted model, we can compute for each patient, an expected annual charge for each service (primary care and specialty care) given the regression effects of that patient's ACG and PCP assignments and the PCP-specific variance. We refer to the mean of these expectations for a service, for a PCP's patient sample, as the "predicted" mean annual charge for that service, for that PCP. It is the mean annual charge that the model predicts would result were that PCP to manage a second sample of patients with the same ACG-mix as that PCP's profiled sample. (Note that the model does *not* predict that the mean annual charge for such an ACG-matched patient sample would be the same as the mean of observed charges for the profiled sample, or observed mean.)

If we replace the PCP regression effects by zeros (the mean of the prior distribution of these effects) and replace the PCP-specific variances by the harmonic mean of PCP-specific variances, and then perform the same computation as was used to derive the predicted mean annual charge, the resultant quantity is a "standardized" mean annual charge. The standardized mean annual charge for a service represents the mean annual charge that would be expected were that PCP's patient sample managed by a hypothetical, or "reference" PCP whose effects on charges are set equal to the mean of all profiled PCPs' effects.

Because the predicted mean annual charge incorporates assumptions contained in the hierarchical model, regarding the distribution of PCP regression effects and PCP-specific variances, it is adjusted, or shrunken away from the observed mean and toward the standardized mean. As a result, it is less vulnerable than the observed mean to bias stemming from limited sample size and incomplete risk-adjustment and thus, a better predictor of a PCP's future performance.

The difference between a PCP's predicted and standardized mean annual charges for a service will be referred to as a "predicted deviation", and the difference between a PCP's observed and standardized mean annual charges will be referred to as an "observed deviation". A deviation is essentially a case-mix adjusted, estimate of a PCP's effect on per-patient, annual charges for a given service. Since it employs the predicted mean rather than the observed mean annual charge, the predicted deviation is a more reliable measure of PCP effect than is the observed deviation. The observed deviation is however, useful for demonstrating the shrinkage that results from hierarchical model assumptions and thus, will be included in the results solely for that purpose. Because they are functions of the model parameters, the predicted and standardized mean annual charges and observed and predicted deviations and their 95% credible intervals can be accurately estimated using the Gibbs sampler output (Gelfand 2000; Gilks, Richardson, and Spiegelhalter 1996).

RESULTS

Patient Characteristics

Of the 3,308 adult male patients meeting inclusion criteria, 17%, 25%, 32%, and 26% were ages 18-30, 31-40, 41-50, and 51-64, respectively. For comparison, among the 21,715 adult males who were members of the POS plan for any portion of 1996 and used a health service of any kind, the respective percentages by age group were 20%, 24%, 29%, and 27%.

Primary and specialty care services were used by 81% (2,666) and 30% (1,004) of patients, respectively. Both services were used by 25% (834) and neither service was used by 14% (472) of patients. Mean per-patient annual charges for primary and specialty care were $122 and $52, respectively; and ranges of per-patient annual charges were $0 - $1,827 and $0 - $3,405, respectively.

PCPs' Predicted and Standardized Mean Annual Charges

Figure 2 shows observed mean annual charges and estimates of standardized and predicted mean annual charges for (a) primary care and (b) specialty care for each of the 50 PCPs, ordered by standardized mean charge, where "estimates" are estimated Bayesian posterior means. The standardized mean is the mean annual charge that would have been expected had a given PCP's patient sample been treated by a hypothetical, "reference" PCP whose effects on charges equaled the mean of all profiled PCPs' effects. The ranges of estimated standardized means, $95 - $148 for primary care and $33 - $83 for specialty care, represent the extent to which ACG-mix explains the ranges of PCPs' predicted means.

The predicted mean is the most reliable predictor of a PCP's mean charge when treating a second patient sample with the same ACG-mix as that PCP's profiled sample. Because it incorporates the hierarchical model assumptions, the predicted mean tends to be shrunken away from the observed mean and toward the standardized mean, a tendency evident in Figure 2. The range of observed mean charges is $61 - $172 for primary care and $12 - $98 for specialty care. The estimated predicted means have narrower ranges: $82 - $167 and $25 - $85, respectively.

Correlation Among PCP Regression Effects

The Table represents the estimated variance-covariance matrix of PCP risk-adjusted regression effects in terms of standard deviations and correlations. The right-hand side of the table displays estimates (estimated Bayesian posterior means) and 95% credible intervals (CIs) for the within-PCP correlations among the (random) PCP regression effects. The estimated correlations demonstrate three important findings: First, PCPs with a higher rate of specialist use were less likely to see their patients at least once during the year (estimated correlation: –.40; 95% CI: –.71, –.008). Second, PCPs with a higher rate of specialist use also provided fewer services to patients that they saw (estimated correlation: –.53; 95% CI: –.77, –.21). And third, PCPs that were more likely to see their patients at least once during the year provided more services to patients that they saw (estimated correlation: .45; 95% CI: .086,.72).

PCPs' Predicted Deviations

Figure 3 displays estimates and 95% CIs for predicted deviations of charges for (a) primary care and (b) specialty care for each of the 50 PCPs, ordered by estimated predicted deviation. Estimates of observed deviations are included for comparison. The deviation is essentially a case-mix adjusted, estimate of a PCP's effect on per-patient, annual charges for a given service. Since the predicted deviation employs the predicted rather than the observed mean annual charge, it is a more reliable measure of a PCP's effect on annual charges than the observed deviation. In Figure 3(a), four credible intervals exclude zero from below and seven from above; thus, four PCPs provided significantly less and seven provided significantly more primary care than expected after risk-adjustment. In Figure 3(b), none of the credible intervals excludes zero thus, none of the PCPs had an effect on charges for specialty care that was significantly different than that of any other.

Figure 4 demonstrates that PCPs' predicted deviations of charges for primary and specialty care have an inverse relationship, consistent with the negative correlations between b_{U1} and b_{U2} and between b_{Y1} and b_{U2} in the Table. Further, careful comparison of Figures 3(a) and 4 reveals that estimates of predicted deviations of charges for specialty care are positive for all four PCPs whose predicted deviations of charges for primary care were significantly less than zero and negative for six of seven PCPs whose predicted deviations of charges for primary care were significantly greater than zero.

Figure 4 includes estimates of observed deviations and demonstrates the shrinkage from observed to predicted deviations with solid lines. Because only 1,004 patients used specialty care, compared with 2,666 that used primary care, effective sample sizes for estimating deviations of charges for specialty care were generally much smaller, explaining the overall greater shrinkage from observed to predicted deviation of charges for specialty care (the vertical dimension in Figure 4), as compared with primary care.

Within-Patient Associations

Of the four non-deterministic, within-patient associations depicted in Figure 1, only the association between any use of primary care and any use of specialty care was statistically significant. The odds ratio was estimated to be .57 (95% CI: .45, .73), indicating that after adjustment for ACG and PCP effects, the probability that a patient had visited a specialist was significantly reduced if that patient had visited his or her PCP at least once during the year,. Estimates of the other three within-patient associations, adjusted for ACG and PCP effects, were as follows: the estimated correlation between log of positive charges for primary care and log of positive charges for specialty care was .033 (95% CI: –.035, .10); the estimated change in log of positive charges for primary care, given any use of specialty care was –.033 (95% CI: –.092, .025); and, the estimated change in log of positive charges for specialty care, given any use of primary care was –.036 (95% CI: –.18, .11).

DISCUSSION

Both PCPs and patients contributed significantly to a trade-off between utilization of primary care and specialty care. After adjustment for ACG-mix, PCPs with higher rates of per-patient specialist use were significantly less likely to see their patients at least once during the year and provided significantly fewer services to patients that they saw; and after adjustment for the effects ACG and PCP assignments, any patient who visited a specialist during the year had a reduced probability of visiting his or her PCP. These findings were possible because we employed a hierarchical multivariate model rather than separate univariate models. Had we fitted a pair of univariate two-part models, one for primary care and one for specialty care, we could not have estimated measures of association between utilization of the two types of service at the level of either PCP or patient.

We found statistically significant differences in the amount of primary care delivered by PCPs and a strong suggestion that these differences were partially offset by the trade-off between provision of primary and specialty care. Specifically, of 11 PCPs whose predicted deviations of annual charges for primary care were significantly less than or greater than zero, 10 had estimated predicted deviations of annual charges for specialty care that were of opposite sign, although not significantly different than zero.

The quality implications of the trade-off between utilization of primary and specialty care depend upon the underlying causes. For example, two possible causes noted in the introduction, PCP time constraints and scope of practice limitations have very different implications for quality of care. Constraints on a PCP's time might result in PCP- or patient-initiated referral to a specialist for evaluation and management of a condition that is within the PCP's scope of practice, but which the PCP is too busy to adequately address, with an adverse effect on continuity and coordination of care for that patient. On the other hand, referral of a patient whose condition is outside of a PCP's scope of practice clearly improves quality of care for that patient; although if the condition is within the scope of practice of the *typical* but not the treating PCP, this may be an indication for widening the treating PCP's practice scope through additional education and training.

It is somewhat surprising that only one of the four non-deterministic within-patient associations was statistically significant, namely, the association between any use of the two services. For instance, we might have anticipated that among patients who visited a specialist, those who also used primary care, and more of it, would have been less likely to see the specialist for ongoing care following an initial consultation. If this had been the case, we could have found that either the regression effect of any use of primary care on the log of positive charges for specialty care or the correlation between the logs of positive charges for the two services was significantly less than zero; but we found neither.

An important limitation of this study is its limited size and scope; PCPs from only one primary care specialty and only male patients were included. Also, by focusing on PCPs with the largest POS plan member-loads, we may have inadvertently selected PCPs with busier practices, who were more likely to refer patients to specialists or induce patients to self-refer to specialists as a consequence of time constraints. Another limitation is that we did not profile PCPs with respect to their effects on use of laboratory and imaging services or pharmaceuticals, important components of ambulatory care costs that are largely controlled by PCPs and specialists.

REFERENCES

Bodenheimer, T. 1996. "The HMO Backlash -- Righteous or Reactionary." *Journal of the American Medical Association* 335: 1601-1604.

Braun, B. L., J. B. Fowles, C. B. Forrest, E. A. Kind, S. S. Foldes, and J. P. Weiner. 2003. "Which Enrollees Bypass Their Gatekeepers in a Point-of-Service Plan?" *Medical Care* 41: 836-841.

Bronskill, S. E., S-L. T. Normand, M. B. Landrum, and R. A. Rosenheck. 2002. "Longitudinal Profiles of Health Care Providers." *Statistics in Medicine* 21: 1067-1088.

Burgess, J. F., C. L. Christiansen, S. E. Michalak, and C. N. Morris. 2000. "Medical Profiling: Improving Standards and Risk Adjustments Using Hierarchical Models." *Journal of Health Economics* 19: 291-309.

Burgess, J. F., V. Lourdes, and M. West. 2000. "Profiling Substance Abuse Provider Trends in Health Care Delivery Systems." *Health Services and Outcomes Research Methodology* 1: 253-276.

DeLong, E. R., E. D. Peterson, D. M. DeLong, L. H. Muhlbaier, S. Hackett, and D. B. Mark. 1997. "Comparing Risk-Adjustment Methods for Provider Profiling." *Statistics in Medicine* 16: 2645-2664.

Duan, N., W. G. Manning, C. N. Morris, and J. P. Newhouse. 1983. "A Comparison of Alternative Models for the Demand for Medical Care." *Journal of Business and Economic Statistics* 1: 115-126.

Forrest, C. B., J. P. Weiner, J. Fowles, C. Vogeli, K. D. Frick, K. W. Lemke, B. Starfield. 2001. "Self-Referral in Point-of-Service Health Plans." *Journal of the American Medical Association* 285: 2223-2231.

Franks, P., G. C. Williams, J. Zwanziger, C. Mooney, and M. Sorbero. 2000. "Why Do Physicians Vary So Widely in Their Referral Rates?" *Journal of General Internal Medicine* 15: 163-168.

Gelfand, A. E. 2000. "Gibbs Sampling." *Journal of the American Statistical Association* 95: 1300-1304.

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter (eds.). 1996. Markov Chain Monte Carlo in Practice. London: Chapman & Hall.

Goldstein, H., and D. Spiegelhalter. 1996. "League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance." *Journal of the Royal Statistical Society* (Series A) 159: 385-444.

Health Services Research and Development Center at The Johns Hopkins University Bloomberg School of Public Health. 2001. The Johns Hopkins ACG Case-Mix System, Version 5.0. Baltimore: The Johns Hopkins University.

Landrum, M. B., S. E. Bronskill, and S-L. T. Normand. 2000), "Analytical Methods for Constructing Cross-Sectional Profiles of Health Care Providers," *Health Services and Outcomes Research Methodology* 1: 23-47.

Landrum, M. B., S-L. T. Normand, and R. A. Rosenheck. 2003. "Selection of Related Multivariate Means: Monitoring Psychiatric Care in the Department of Veterans Affairs." *Journal of the American Statistical Association* 98: 7-16.

Reid, R. J., L. MacWilliam, L. Verhulst, N. Roos, and M. Atkinson. 2001. "Performance of the ACG Case-Mix System in Two Canadian Provinces." *Medical Care* 39: 86-99.

Salem-Schatz, S., G. Moore, M. Rucker, and S. D. Pearson. 1994. "The Case for Case-Mix Adjustment in Practice Profiling: When Good Apples Look Bad." *Journal of the American Medical Association* 272: 871-874.

Same authors (as this manuscript). 2004. "A Hierarchical Multivariate Two-Part Model for Profiling Providers' Effects on Healthcare Charges." Currently under review by the *Journal of the American Statistical Association*.

SAS Institute Inc. 1999. SAS/IML User's Guide, Version 8. Cary, NC: SAS Institute, Inc.

Shahian, D. M., S. T. Normand, D. F. Torchiana, S. M. Lewis, J. O. Pastore, R. E. Kuntz, and P. I. Dreyer. 2001. "Cardiac Surgery Report Cards: Comprehensive Review and Statistical Critique." *Annals of Thoracic Surgery* 72: 2155-2168.

Weiner, J. P., B. H. Starfield, D. M. Steinwachs, and L. M. Mumford. 1991. "Development and Application of a Population-Oriented Measure of Ambulatory Care Case-Mix." *Medical Care* 29: 452-472.

TABLES

Table. Standard Deviations and Correlations of PCP Regression Effects: Estimates* and 95% Credible Intervals (CIs)

PCP Regression Effect	**Std Dev: Estimate (CI)**	**Correlation: Estimate (95% CI)**				
			$\mathbf{b_{U1}}$	$\mathbf{b_{U2}}$	$\mathbf{b_{Y1}}$	$\mathbf{b_{Y2}}$
$\mathbf{b_{U1}}$: log odds of any use of PC†	.40 (.29,.53)	$\mathbf{b_{U1}}$	1			
$\mathbf{b_{U2}}$: log odds of any use of Spc†	.39 (.28,.52)	$\mathbf{b_{U2}}$	–.40 (–.71, –.008)	1		
$\mathbf{b_{Y1}}$: log of positive charges for PC†	.17 (.13,.21)	$\mathbf{b_{Y1}}$	.45 (.086, .72)	–.53 (–.77, –.21)	1	
$\mathbf{b_{Y2}}$: log of positive charges for Spc†	.10 (.059,.16)	$\mathbf{b_{Y2}}$	–.078 (–.60, .48)	–.002 (–.50, .53)	.19 (–.33, .64)	1

* Estimated Bayesian posterior means

† PC = primary care, Spc = specialty care

FIGURE LEGENDS

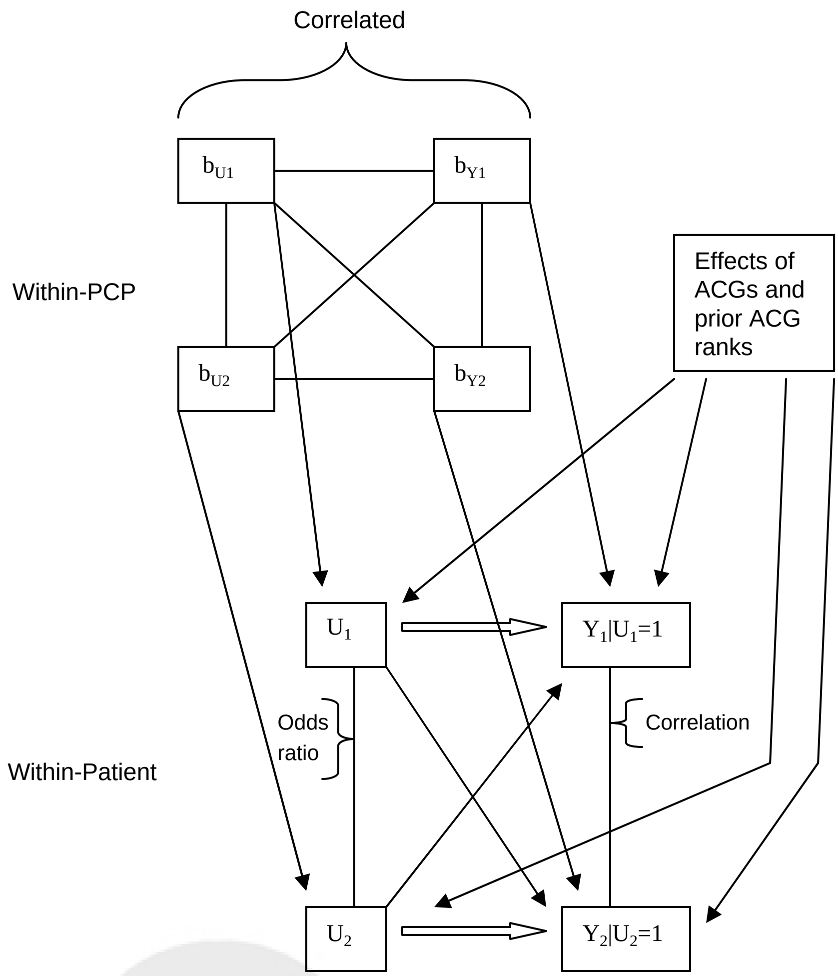
Figure 1. Important features of the hierarchical multivariate two-part model. Dependent variables, U_1 and U_2 represent any use of primary care and specialty care, respectively. Dependent variables, $Y_1| U_1=1$ and $Y_2 | U_2=1$ represent logs of positive charges for primary care and specialty care, respectively. Corresponding to these 4 dependent variables are random PCP regression effects, b_{U1}, b_{U2}, b_{Y1}, and b_{Y2}. Lines represent correlations or odds ratio. Solid arrows represent regression relationships with arrow pointing toward dependent variable. Block arrows represent deterministic relationships.

Figure 2. Observed (dot) and estimates of standardized (star) and predicted (square) mean annual charges in dollars for (a) primary care and (b) specialty care for 50 PCPs ordered by standardized mean charge. (Estimates are estimated Bayesian posterior means.)

Figure 3. Deviations of annual charges in dollars, for (a) primary care and (b) specialty care for 50 PCPs ordered by estimated predicted deviation. Scored bars represent estimates and 95% credible intervals of predicted deviations. Dots represent estimates of observed deviations. (Estimates are estimated Bayesian posterior means.)

Figure 4. Estimates of predicted (square) and observed (dot) deviations of annual charges in dollars, for primary care v. specialty care for 50 PCPs. Connecting lines represent shrinkage. (Estimates are estimated Bayesian posterior means.)

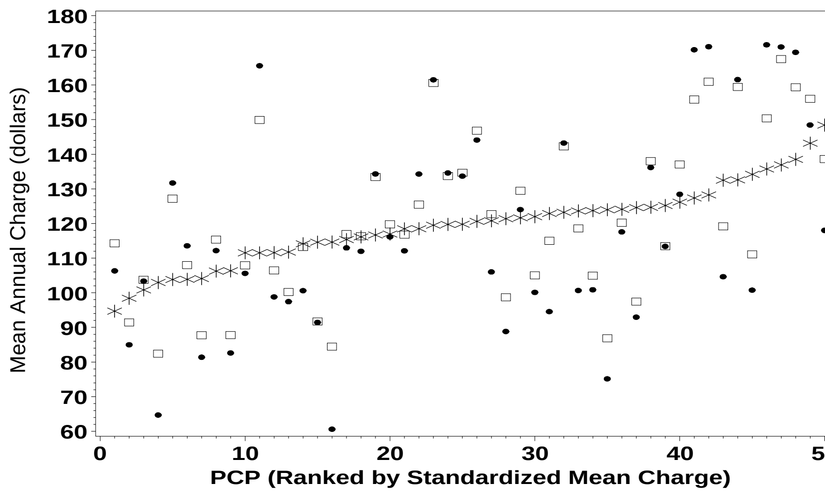
Figure 1



Correlated
Within-PCP
b_{U1}
b_{Y1}
b_{U2}
b_{Y2}
Effects of ACGs and prior ACG ranks
U_1
$Y_1|U_1=1$
Odds ratio
Correlation
Within-Patient
U_2
$Y_2|U_2=1$

Figure 2

(a) Primary Care



(b) Specialty Care

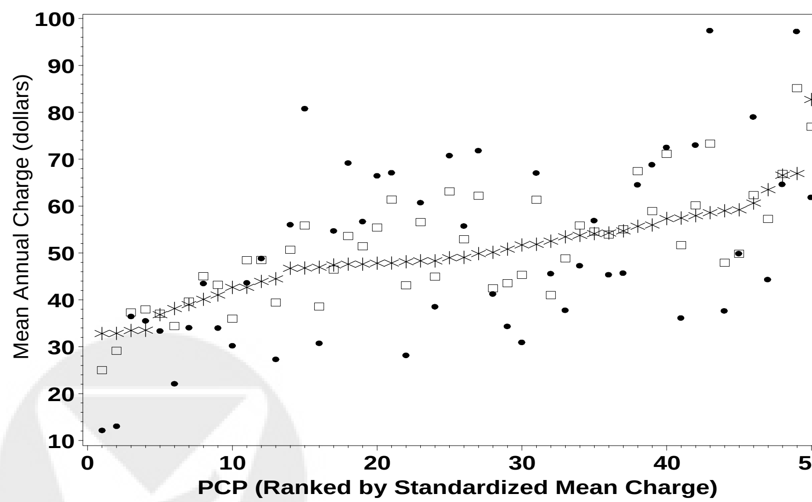
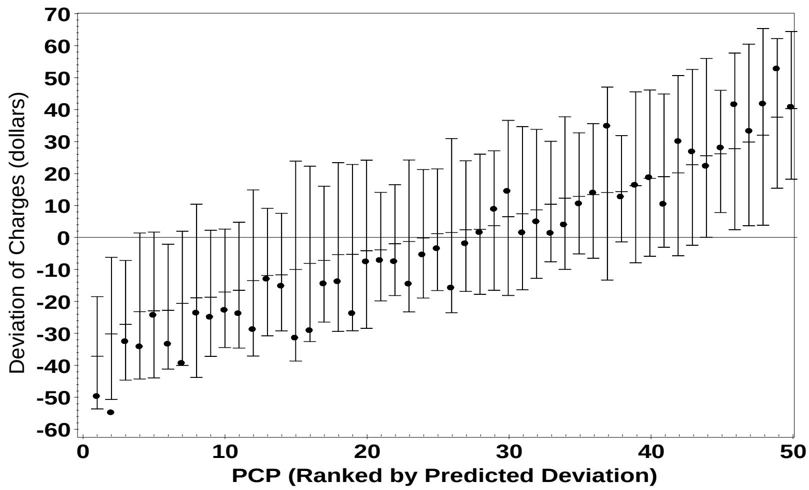


Figure 3

(a) Primary Care



(b) Specialty Care

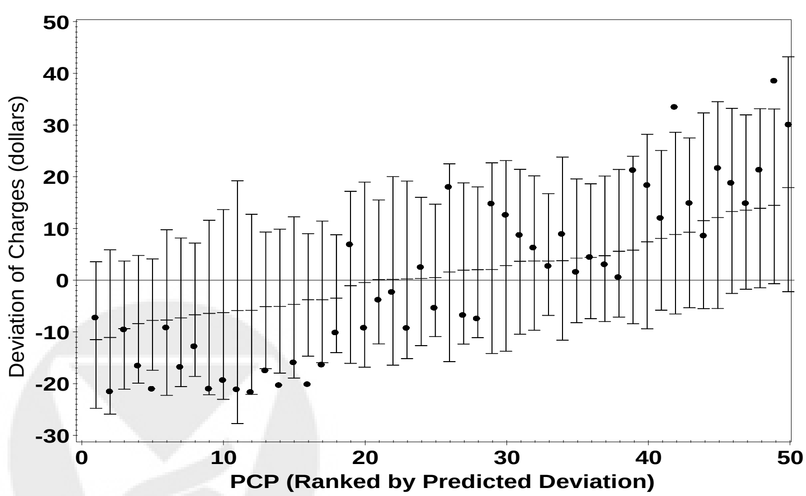
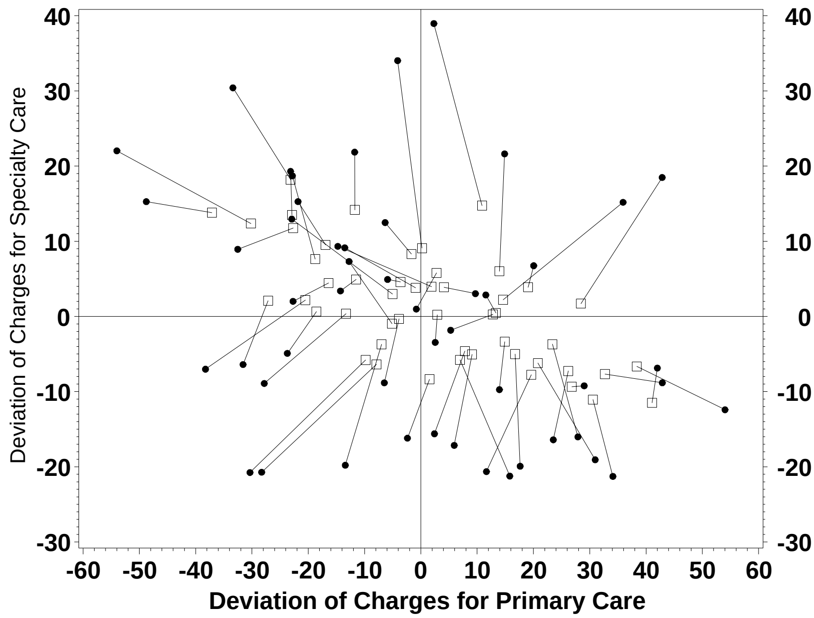


Figure 4


Deviation of Charges for Specialty Care
Deviation of Charges for Primary Care
40
30
20
10
0
-10
-20
-30
-60 -50 -40 -30 -20 -10 0 10 20 30 40 50 60