Determining VA Physician Requirements Through Empirically Based Models

Joseph Lipscomb, Kerry E. Kilpatrick, Kerry L. Lee, and Karen S. Pieper

Objective. As part of a project to estimate physician requirements for the Department of Veterans Affairs, the Institute of Medicine (IOM) developed and tested empirically based models of physician staffing, by specialty, that could be applied to each VA facility.

Data Source/Study Setting. These analyses used selected data on all patient encounters and all facilities in VA's management information systems for FY 1989.

Study Design. Production functions (PFs), with patient workload dependent on physicians, other providers, and nonpersonnel factors, were estimated for each of 14 patient care areas in a VA medical center. Inverse production functions (IPFs), with physician staffing levels dependent on workload and other factors, were estimated for each of 11 specialty groupings. These models provide complementary approaches to deriving VA physician requirements for patient care and medical education.

Data Collection/Extraction Methods. All data were assembled by VA and put in analyzable SAS data sets containing FY 1989 workload and staffing variables used in the PFs and IPFs. All statistical analyses reported here were conducted by the IOM.

Principal Findings. Existing VA data can be used to develop statistically strong, clinically plausible, empirically based models for calculating physician requirements, by specialty. These models can (1) compare current physician staffing in a given setting with systemwide norms and (2) yield estimates of future staffing requirements conditional on future workload.

Conclusions. Empirically based models can play an important role in determining VA physician staffing requirements. VA should test, evaluate, and revise these models on an ongoing basis.

Keywords. Physician staffing, production functions, work force planning

This article shows how staffing models estimated from data on current clinical practice in the U.S. Department of Veterans Affairs (VA) can help determine the number of physicians VA *should* have to achieve its missions of patient care, education, and research.[1] These staffing models are central to the strategy developed by the Institute of Medicine (IOM) for estimating VA physician requirements, by specialty and clinical program area, for any VA facility in the country (Institute of Medicine 1991, 1992).[2] Because they are grounded in the current practice of medicine in the VA, these models provide a base against which alternative staffing approaches–such as expert judgment models–can be evaluated. Two alternative variants of the empirically-based models have been developed:

In the *production function* (PF) variant, the volume of workload (e.g., bed-days of care) for a given patient care area (PCA) (e.g., surgery) at a VA medical center (VAMC) was modeled as a function of the number of staff physicians, residents, other personnel, and additional variables possibly associated with physician productivity. Each VAMC was divided into 14 or fewer PCAs within inpatient care: medicine, surgery, psychiatry, neurology, rehabilitation medicine, and spinal cord injury; within ambulatory care: medicine, surgery, psychiatry, neurology, rehabilitation medicine, and other physician services (the latter including emergency care, and admitting and screening); and within long-term care: nursing home and intermediate care. To derive the total physician FTEE (fulltime equivalent employee) in a given specialty (e.g., neurology) or program area (e.g., ambulatory care) required for patient care at a given VAMC, the FTEE required to meet patient workload in each relevant PCA were computed, then summed across PCAs.

In the *inverse production function* (IPF) variant, specialty-specific rather than PCA-specific models were estimated. For a given specialty (e.g., neurology), the quantity of physician FTEE devoted to patient care and resident education across all PCAs at the VAMC was hypothesized to be a function of such factors as total inpatient, ambulatory, and long-term care workload associated with that specialty; the number of residents; and other variables possibly associated with physician time devoted to patient care and resident education. Separate facility-level IPFs were estimated for each of the following 11 specialty groups: internal medicine, surgery, psychiatry, neurology, rehabilitation medicine, anesthesiology, laboratory

---

Address correspondence and requests for reprints to Joseph Lipscomb, Ph.D., Associate Professor of Public Policy Studies and Community and Family Medicine, Sanford Institute of Public Policy, Box 90245, Duke University, Durham, NC 27708-0245. Kerry E. Kilpatrick, Ph.D. is Professor and Chairman, Department of Health Policy and Administration, School of Public Health, University of North Carolina–Chapel Hill; Kerry L. Lee, Ph.D. is Associate Professor of Biostatistics, Department of Community and Family Medicine, Duke University; and Karen S. Pieper, M.S. is Statistician, Department of Medicine, Duke University. This article, submitted to *Health Services Research* on October 5, 1993, was revised and accepted for publication on August 19, 1994.

medicine, diagnostic radiology, nuclear medicine, radiation oncology, and spinal cord injury.

Under either the PF or IPF variant, total FTEE required at the facility is the sum of the model-derived estimate and separate FTEE estimates for those physician activities not accounted for in these models, including research, continuing education, and miscellaneous assignments. Decisions about the scope of physician activities, and hence FTEE, included in the PFs and IPFs depended on data availability. The PF and IPF variants employ different behavioral assumptions, have different strengths and weaknesses, and provide alternative estimates (and perspectives) on physician staffing requirements; thus, they play complementary roles in the overall methodology.

MODELS

Production Function (PF)

Substantial literature exists on the specification and estimation of production functions, with applications to hospital care (Jensen and Morrisey 1986a and 1986b; Pauly 1980; and Feldstein 1967); to physician services (Gaynor and Pauly 1990; and Reinhardt 1975); and to dentistry (Scheffler 1981; Kushman et al. 1979; Scheffler and Kushman 1977; and Maurizi 1969).

An underemphasized assumption in all such models, including the ones in this article, is that a given output rate can be produced by *any* (technically feasible) combination of inputs with a quality of care that is both acceptable and invariant to input combination. Precisely because this may not be the case, the IOM's proposed staffing strategy recommends that expert clinical judgment be used both to evaluate the empirically based models and to derive alternative estimates of physician staffing consistent with "good" quality care.[3] While the empirically based models presented here benefited significantly from the scrutiny of the study's eight specialty and clinical program panels (Institute of Medicine 1992), there were no explicit adjustments for quality of care variations: all input combinations were implicitly regarded as consistent with "good" quality care.

Notwithstanding, the production functions developed by the IOM differ from earlier work in several respects:

The hospital studies cited are all facility level, and the ambulatory care-oriented studies are all conducted at the practice (or firm) level. In general, the studies were not able to control directly for case-mix variation across production units; one exception was Feldstein's (1967) use of a case-mix proportions vector in his study of British hospitals. Because the PFs below are estimated at the PCA level, and with workload measures that further control for patient acuity *within* the PCA, these PFs should be relatively less affected by patient heterogeneity than earlier models.

The behavioral assumptions undergirding the PFs here differ significantly from those cited earlier. These PFs were set in the private sector and–with the exception of Gaynor and Pauly (1990)[4]–assumed profit-maximizing behavior, implying that inputs are jointly selected with output to minimize the cost of production. These assumptions, together, also imply that ordinary least squares estimates of the parameters of the production function will be biased (Kmenta 1986). In response, additional behavioral assumptions are typically invoked to justify the application of OLS (rather than a simultaneous-equation approach that recognizes the joint determination of inputs and output).

By contrast, a VAMC is a public sector organization charged with multiple missions; profit maximization is not one of them. Rather, a VAMC seeks to meet its patient care mission in a way that balances several concerns: that eligible veterans are treated in a timely manner; quality of care is acceptable; and budget, other resource, and administrative constraints are met.

The *PF analysis* therefore assumes that a VAMC adjusts inputs and workload sequentially: subject to budget and real resource availability, the VAMC sets input levels for each fiscal year in accordance with anticipated workload. Then, during the year, it modulates workload (up or down) to maintain a clinically acceptable relationship between workload and inputs. Statistically, this implies that in a cross-sectional analysis of VAMC production, inputs can be regarded as fixed, and workload as a random variable whose expected value is a function of input levels but whose actual value will also reflect random noise. Under these assumptions, and further ones regarding the model's error term, single-equation estimation of the production function by OLS is justified.

The general form of the production function is as follows:

$$
\begin{aligned}
W_{ij} = f[\{\text{StaffPhys}_{ij}\}, \{\text{ConPhys}_{ij}\}, \{\text{Res}_{ij}\}, \text{C\&A}_{ij}, \text{WOC}_{ij}, \{\text{NPP}ij\}, \\
\text{Nurse}_{ij}, \text{Support}_{ij}, \text{ProdFact}_{ij}, \text{ERROR}_{ij}\}
\end{aligned}
\tag{1}
$$

where

W_{ij} = the annual rate of production of workload in PCA j of VAMC i:

$\{\text{StaffPhy}_{ij}\}$ = a set of variables, each taking the form $\text{StaffPhy}_{ijk} = \text{FTEE}$ for direct patient care in PCA j of VAMC i for physician specialty $k = 1, \ldots 11$;

$\{\text{ConPhys}_{ij}\}$ = a set of variables for physicians under contact to VAMC i, such that ConPhys_{ijk} = the contract physician FTEE from specialty k devoted to PCA j;
$\{\text{Res}_{ij}\}$ = a set of variables to account for the net productive contribution of residents, with Res_{ijy} = amount of PGY y resident FTEE allocated to PCA j at VAMC i;
$\text{C\&A}_{ij}$ = for non-VA physicians who perform consulting and attending duties on a fee-for-visit basis, FTEE allocated to PCA j at VAMC i;
WOC_{ij} = for non-VA physicians who perform consulting and attending duties without (monetary) compensation, the amount of FTEE allocated to PCA j at VAMC i;
$\{\text{NPP}_{ij}\}$ = a set of variables, each of the form NPP_{ijm} = FTEE of nonphysician practitioner type m (e.g. a physician assistant) assigned to PCA j at VAMC i;
Nurse_{ij} = the amount of nursing service FTEE allocated to PCA j at VAMC i;
Support_{ij} = for all personnel categories excluding physicians and nurses (in some models also psychologists and social workers), total FTEE allocated to PCA j at VAMC i;
ProdFact_{ij} = variables for factors (e.g., medical school affiliation status) influencing the productivity of physicians and others in PCA j at VAMC i;
ERROR_{ij} = the random-error term for PCA j at VAMC i, assumed to obey the assumptions sufficient for unbiased parameter estimation (Kmenta 1986).

Note that *Res* encompasses *all* postgraduate years of training, including those typically designated for fellows (PGY4 and beyond). For two of these variables, *C&A* and *WOC*, data are not now available on a VA systemwide basis, and they are omitted from all estimated PFs. Note also the omission of capital-related variables from Equation (1). Data on the kinds and amounts of nonlabor inputs at the PCA level of each VAMC were not readily accessible to the IOM committee, and we were not able to derive a satisfactory measure of beds for the PFs. While data do exist on the number of "authorized beds" and "operating beds" for the inpatient and long-term care PCAs, counts of "staffed beds" were not uniformly available in national VA databases. Further, as occupancy levels for major bed sections did not exceed 72 percent on a national basis in FY 1989 (U.S. Department of Veterans Affairs 1990), it is unlikely that bed shortages constrained the way labor inputs were combined to produce output.

Inverse Production Function (IPF)

In the PF variant, the basic question is: What factors account for the production of patient care workload? In the IPF, the basic question is: What factors account for the observed level of staff physician FTEE? The basic behavioral assumption for the IPF is that the VAMC adjusts physician FTEE levels in response to a given projected workload level, controlling for other factors. Thus, the volume of workload per time period cannot be significantly modulated (i.e., it is "exogenously" determined by demand-influencing factors beyond the VAMC's control). Under these assumptions, single-equation OLS estimation is justified. That the PF and IPF have different underlying assumptions does not in any way lead to an *empirical* contradiction. Every model, by design, has its own defining assumptions; the PF and IPF offer alternative perspectives on the same underlying production process.

The general form of the IPF variant is as follows:[5]

$$\mathrm{StaffPhys}_{ik} = g[\{W_{ik}\}, \{\mathrm{Res}_{ik}\}, \{\mathrm{NPP}_{ik}\}, \mathrm{Prodfact}_{ik}, \mathrm{ERROR}_{ik}] \qquad (2)$$

where

$\mathrm{StaffPhys}_{ik}$ = across all PCAs at VAMC*i*, the total amount of specialty *k* staff physician and contract physician FTEE devoted to patient care and resident education;

$\{W_{ik}\}$ = a set of workload variables, each of the form W_{ijk} = the level of workload on PCA *j* of VAMC *i* associated with specialty *k*;

$\{\mathrm{Res}_{ij}\}$ = a set of variables, each of the form Res_{iky} = postgraduate year *y* resident FTEE at VAMC *i* in specialty *k*;

$\{\mathrm{NPP}_{ik}\}$ = a set of variables, each of the form NPP_{ikm} = FTEE of nonphysician practitioner type *m* associated with PCA-related activities of physician specialty *k* at VAMC *i*;

$\mathrm{Prodfact}_{ik}$ = one or more variables for factors influencing the productive efficiency of specialty *k* physicians at VAMC *i*;

ERROR_{ik} = the random-error term for specialty *k* at VAMC *i*.

An additional distinction between the IPF and the PF lies in the scope of what is included in the physician FTEE variable. In the IPF, an attempt is made to account for all physician FTEE devoted to VA mission-related activities across PCAs. Patient care and resident education typically dominate these activities. Hence, the dependent variable in Equation (2) incorporates all patient care–designated FTEE (including physicians under contract) plus FTEE allocated to resident instruction. In the PFs, the focus is on the physician FTEE required for patient care in the PCAs (controlling for other factors of production).

DATA

While not described in detail here, development of the data set to support the models was itself a substantial effort. All data used were for FY 1989, during which about 97 percent of VA's medical program budget (which totaled $11.6 billion) was allocated to programs and services involving physicians either directly or indirectly. In constructing the models' variables, the IOM committee used data from all 172 Veterans Affairs Medical Centers, pertaining to over 1.1 million inpatient admissions, 22 million outpatient visits, and the time allocations for 200,000 VA employees. Throughout, the committee was closely assisted by VA's Boston Development Center, which provided SAS data sets for all variables investigated in the study. With data of such breadth and depth, we were able to estimate staffing models at an unprecedented level of detail.

Workload

Patient workload variables were defined specific to each of three broad type-of-care areas: inpatient care, ambulatory care, and long-term care.

Inpatient Care. For the PF variant, six PCAs were defined from the VA data systems: medicine, surgery, psychiatry, neurology, rehabilitation medicine, and spinal cord injury. For each PCA, we used the VA Patient Treatment File (PTF) to define three alternative workload variables: *Discharges/Year, Bed-Days of Care/Year,* and an annualized cost-weighted, case mix–adjusted variable called the *Weighted Work Unit* (WWU) score. Each PTF abstract contains sufficient information to assign the patient's discharge to a primary PCA, and to allocate a patient's total bed-days for each stay to the PCA(s) where treatment was rendered.

From the PTF data came the "raw materials" for generating observations on W_{ij} in Equation (1) and W_{ijk} in Equation (2). For the PF equations, workload was defined in terms of the annual sum of *all* WWU scores linked to PCA j. Because the distribution of the WWU scores was typically right-skewed, we applied a natural log transform in each case to create a dependent variable whose distribution more closely approximated the normal. The corresponding production function dependent variable is $W_{ij} = \ln[w_{ij} + 1]$, where w_{ij} is the sum of all medicine, surgery, psychiatry, neurology, and rehabilitation medicine WWUs generated in inpatient PCA j of VAMC i.

In the IPF, the workload variables used in the final equations are not, in fact, PCA-specific, but specific rather to the three type-of-care areas –inpatient, ambulatory, and long-term care (LTC). For specialty k's IPF, the relevant notation for workload variables becomes $W_{i,\text{ Inpatient},\, k}$,

$W_{i,\text{ Ambulatory},\, k}$, and $W_{i,\text{ LTC},\, k}$. For each IPF, one must determine which inpatient PCAs are appropriate for computing $W_{i,\text{ Inpatient},\, k}$, and likewise for the ambulatory care and long-term care workload variables.

Ambulatory Care. The VA Staff Outpatient File contains data on each ambulatory care visit, which can be conceptualized as a sequence of clinic stops encountered by the patient. With the assistance of VA's Boston Development Center, each of 101 clinic stops was mapped into one of six mutually exclusive and exhaustive clinic-stop groups corresponding to the six ambulatory care PCAs. Based on the VA's own analyses of these data, we focused on two workload measures: *Clinic Stops/Year* for ambulatory PCA j at VAMC i, computed as the direct sum of all recorded encounters at all clinic stops in that PCA domain during the year; and the *Capitation Weighted Work Unit* (CAPWWU) score, a cost-weighted workload measure that adjusts partially for visit-mix differences across PCAs by service intensity and patient age (both within and across facilities). For the CAPWWU, the patient is assigned a cost weight, which is then allocated fractionally among the ambulatory PCAs at VAMC i in proportion to the distribution of the patient's clinic-stop encounters across these PCAs.

Long-Term Care. For patients admitted to an intermediate care or nursing home bed, sufficient information is recorded in the VA Patient Treatment File, Extended Patient Treatment File, and Patient Assessment File to calculate several alternative workload measures. The two utilized in this study were *Patient Days/Year* for long-term care PCA j; and the *Resource Utilization Group Weighted Work Unit* (RUGWWU) score, a workload measure that adjusts for LTC treatment cost differences associated with differences in the clinical characteristics and the physical and functional status of LTC patients (Institute of Medicine 1991, 63).

VA Staff Physicians and Nurses

Data for calculating both StaffPhys_{ijk} in the PF variant and the staff physician component of StaffPhys_{ik} in the IPF variant were derived entirely from the VA Cost Distribution Report (CDR). In addition, the PF variables Nurse_{ij} and Support_{ij} were computed from the CDR.

Several alternative ways of accounting for the productivity influence of nurses (primarily RNs and LPNs, but not including nurse practitioners) were examined in the PF models. The expression involving nurse FTEE that proved most satisfactory, both conceptually and statistically, was Nurse_{ij} / StaffPhys_{ij}, where the denominator is the total FTEE for physicians involved in hands-on delivery of care in PCA j of VAMC i. Such a variable allows examination of the effect of nursing staff *intensity* on physician productivity–which seems the appropriate focus here. Moreover, when Nurse_{ij} was used in versions of Equation (1), implausible estimates often emerged for the coefficients of the VA staff physician variables.[6]

Support Staff

For almost all VA providers who are not physicians or nurses, the allocation of time to activities across patient care areas cannot be tracked through the CDR (two important exceptions are psychologists and social workers). Consequently, it is not possible to allocate to PCAs the FTEE of physician assistants, nurse practitioners, and others. In response, we proceeded as follows: for each PCA *j*, total physician FTEE was added to total nurse FTEE, and the result subtracted from total (from all sources) PCA FTEE. The result is the residual FTEE quantity labeled Support_{ij}. It is a composite measure of all nonphysician, nonnurse FTEE in that PCA. (Whenever psychologists or social workers were included in an equation, their FTEE was subtracted, as well, in computing Support_{ij}.)

Contract Physicians

Observations on the variable ConPhys_{ijk} were not derived from the CDR, but from the FY 1989 version of a survey conducted annually by VA Central Office.

Residents and Fellows

Observations on Res_{ijy}–the total FTEE of postgraduate year (PGY) *y* residents in PCA *j* at VAMC *i*–were derived from two sources. The VA Office of Academic Affairs provided the number of VA-supported residency positions actually filled, by specialty and PGY, at each VAMC. From each facility's CDR, a rough allocation of resident FTEE to PCAs, by specialty, was inferred for all VA-supported residents. These comprised well over 90 percent of the authorized residency positions.

Nonpersonnel Factors Influencing Physician Productivity

For the PF and the IPF equations, two basic types of control variables were tested as proxies for productivity-influencing factors (Prodfact). Each facility was classified into one of six resource allocation model (RAM) groups derived from a cluster analysis of facility attributes (Stefos, LaVallee, and Holden 1992): small affiliated VAMCs (HGROUP1); small general unaffiliated (HGROUP2); midsize affiliated (HGROUP3); midsize general unaffiliated (HGROUP4); metro affiliated (HGROUP5); and psychiatric (HGROUP6). One rationale for these variables is that a physician's style of practice and overall productivity may vary by type of facility, all else equal.[7] Dummies for the VA geographic location of the facility did not perform satisfactorily in most PFs and IPFs, and were dropped from the final models.

ESTIMATED EQUATIONS

For each PF and IPF the aim was to generate the best-fitting clinically plausible model, given the available data. Each model evolved during the course of the study through an interactive process involving the study committee, its data and methodology panel, the appropriate specialty or clinical program panel, the IOM study staff, and the project's statistical consultants. The important advisory role played by the committee's eight specialty and clinical program panels and the iterative regression strategy used to select the final PF and IPF candidates are described in Institute of Medicine (1992, 478–480).

In addition, Monte Carlo simulation analyses based on "bootstrap sampling" (Efron and Gong 1983) were conducted on selected PFs and IPFs to examine the models' stability and reproducibility and to correct goodness-of-fit measures for over-fitting. For each PF or IPF, this involves sampling with replacement from the available cross-section of data to generate any number of additional "samples" from which the model can be reestimated repeatedly. This, in turn, provides the basis for studying the robustness of the model's original specification and goodness of fit. Our bootstrap analyses (apparently the first reported in health work force analysis) provided encouraging evidence on the robustness of the particular models studied: the PFs for the inpatient PCAs of medicine, surgery, and psychiatry, and the IPFs for the specialties of internal medicine, surgery, and psychiatry (Institute of Medicine 1992, 489–505).

In a more conventional check on model quality, we plotted the normalized residuals of each PF and IPF against the corresponding predicted values of the dependent variable. If the OLS assumptions hold, the plotted residuals should appear random; that this was generally the case (some outliers notwithstanding) is evidenced by the plots reported in Institute of Medicine (1992, 506–30).

Several alternative methods of model validation, such as split-sample analysis using the (cross-sectional) observations from FY 1989 or reestimating the PFs and IPFs with subsequent years of data, were not feasible during the IOM study because of time and resource constraints.

Following Jensen and Morrisey (1986a,b), the functional form we specified for each PF and IPF was the flexible quadratic, which allows for the testing of all direct-effort, squared, and first-order interaction terms derivable from a set of possible determinants. With few exceptions, the committee's final PFs and IPFs specify patient workload in terms of weighted work units (either WWUs, CAPWWUs, or RUGWWUs as the context requires). These workload measures generally led to models with a slightly better goodness of fit than the available alternative measures, which do not control for case mix or patient acuity.[8]

For compactness, we now present and discuss one representative PF (for the Inpatient Medicine PCA) and one IPF (for the specialty of Internal Medicine). The remaining 13 PFs and 10 IPFs are reported analogously in Institute of Medicine (1991, 74–83 and 86–94).

Production Functions

Variable names introduced in Equation (1) and in the Data section are adopted below except where greater specificity requires additions. For simplicity, subscripts are suppressed; for example, W_{ij} in Equation (1) is written simply as W.

For the *Inpatient Medicine* PCA we have

$$\begin{aligned} W = 6.144 &+ \underset{(9.431)}{0.213}\ \text{MED\_MD} - \underset{(-4.276)}{0.007}(\text{MED\_MD})^2 + \underset{(2.700)}{0.138}\ \text{SUR\_MD} \\ &+ \underset{(2.373)}{0.163}\ \text{PSY\_MD} + \underset{(1.703)}{0.106}\ \text{NEU\_MD} + \underset{(3.681)}{0.015}\ \text{RESIDENTS} + \underset{(1.909)}{0.015}\ \text{FELLOWS} \\ &+ \underset{(10.001)}{0.048}\ \text{SUPPORT/MD} + \underset{(2.629)}{0.048}\ \text{SOCW} - \underset{(-4.455)}{0.237}\ \text{HGROUP6} \\ &- \underset{(-2.843)}{0.003}\ (\text{MED\_MD} \times \text{FELLOWS}) \end{aligned} \tag{3}$$

with $\bar{R}^2 = 0.822$ and $N = 159$

where

$W = \ln[w_{ij} + 1] =$ the natural logarithm of total WWUs, plus 1, produced in the inpatient medicine PCA during the fiscal year;
MED_MD = VA staff physician FTEE from internal medicine allocated to direct care in the inpatient medicine PCA;
$(\text{MED\_MD})^2 =$ variable testing for diminishing marginal returns to internal medicine physicians in the production of workload;
SUR_MD = VA staff surgeon FTEE allocated to direct care in this PCA;
PSY_MD = VA staff psychiatrist FTEE allocated to direct care in this PCA;

NEU_MD = VA staff neurologist FTEE allocated to direct care in this PCA;
SUPPORT/MD = support staff FTEE divided by total FTEE for physicians involved in hands-on delivery of care in the inpatient medicine PCA;
RESIDENTS = PGY 2 and PGY 3 FTEE allocated to this PCA (PGY 1 not significant);
FELLOWS = FTEE of residents PGY 4 and above allocated to this PCA;
SOCW = social worker FTEE allocated to this PCA;
HGROUP6 = categorical variable = 1 if facility is in RAM Group 6 (psychiatric hospital);
(MED_MD × FELLOWS) = interaction term for the joint influence of VA staff physicians in internal medicine and fellows; and
N = number of inpatient medicine PCAs in the sample.

Beneath each estimated coefficient is its t-statistic. All variables had estimated coefficients statistically significant at $p < .05$ (two-tail t-test), except NEU_MD and FELLOWS ($p < .10$).

From Equation (3), it can be inferred that in their consulting roles, surgeons, psychiatrists, and neurologists contribute significantly to workload production. Residents, fellows, and social workers are clearly important, although there is a negative interaction involving fellows and staff physicians in internal medicine (i.e., their total contribution to workload is less than the sum of their individual contributions, all else equal). The intensity of support staff positively influences productivity in the inpatient medicine PCA (although NURSE/MD was not significant). Medicine PCAs in VA psychiatric hospitals produce significantly fewer WWUs/year than at other VAMCs, controlling for other factors. Interaction terms involving each of the other hospital group variables with MED_MD–which might detect whether this physician FTEE is either accounted for differently (within the CDR) or else used differently in the PCA at affiliated versus unaffiliated hospitals–were not significant in this PF.

How would adding one more VA staff physician in internal medicine affect workload production in the medicine PCA? The "marginal product" of MED_MD can be derived in several steps: (1) take the partial derivative of W with respect to MED_MD ($\partial W/\partial \mathrm{MED\_MD}$), obtaining an expression that is linear in MED_MD, FELLOWS, and the derived variable SUPPORT/MD^2; (2) following Jensen and Morrisey (1986a,b), substitute the VA systemwide FY 1989 mean value of each of these variables into this expression, obtaining in this instance $\partial W/\partial \mathrm{MED\_MD} = 0.017$;

(3) since W is a nonlinear function of WWUs, the marginal product will depend on the assumed baseline WWU level, so adopt the FY 1989 national mean workload level for inpatient medicine PCAs of 2,484 WWUs; and (4) calculate the marginal product of MED_MD at the "statistically typical" VAMC to be 42 WWUs/Year–which is not large relative to the baseline workload level. (Of course, marginal products will be higher or lower at individual VAMCs.)

By comparison, similar analyses show that the marginal product of the VA staff surgeon in the statistically "typical" ***inpatient surgery PCA*** was 141 WWUs/year, and the marginal product of the staff psychiatrist in the "typical" ***inpatient psychiatry PCA*** was 187 WWUs/year in FY 1989 (Institute of Medicine 1992, 496–98).

Inverse Production Functions

When possible, the variable names introduced in Equation (2) and in the Data section are adopted in Equation (4), but, as with the PF equation, more specific definitions are sometimes required. Subscripts again are suppressed, so that W_{ik} in Equation (4) becomes simply W.

For the specialty area of ***Internal Medicine*** we have[9]

$$
\begin{aligned}
\text{MED\_MD}' = {} & 1.234 + \underset{(3.846)}{3.982}\ \text{MEDWWU} + \underset{(3.428)}{0.00078}\ \text{MEDCAPWWU} \\
& + \underset{(1.990)}{0.014}\ \text{MEDRUGWWU} + \underset{(1.869)}{0.012}\ \text{FELLOWS} + \underset{(2.416)}{0.608}\ \text{HGROUP2} \\
& + \underset{(2.370)}{0.767}\ \text{HGROUP3} + \underset{(2.551)}{0.667}\ \text{HGROUP4} + \underset{(2.079)}{0.822}\ \text{HGROUP5} \\
& + \underset{(2.367)}{0.332}\ \text{HGROUP6} - \underset{(-1.986)}{2.707}\ (\text{MEDWWU} \times \text{HGROUP2}) \\
& - \underset{(-2.705)}{3.311}\ (\text{MEDWWU} \times \text{HGROUP3}) - \underset{(-1.971)}{2.380}\ (\text{MEDWWU} \times \text{HGROUP4}) \\
& - \underset{(-2.801)}{3.271}\ (\text{MEDWWU} \times \text{HGROUP5})
\end{aligned}
\tag{4}
$$

with $\bar{R}^2 = 0.595$ and $N = 164$

where

MED_MD′ = log of the sum of all VA internal medicine physician FTEE devoted to direct care and resident training across all PCAs, plus 1;

MEDWWU = total medicine WWUs produced in the inpatient PCAs of medicine, surgery, psychiatry, neurology, and rehabilitation medicine (÷10,000);

MEDCAPWWU = total CAPWWUs produced in the ambulatory PCAs of medicine and other physician services ($\div$ 10,000);
MEDRUGWWU = total RUGWWUs produced in the long-term care PCAs of nursing home and intermediate care ($\div$ 10,000);
FELLOWS = total FTEE of medicine residents PGY4 and above at the VAMC; and
N = number of VAMCs reporting FTEE for internal medicine physicians in FY 1989.

CAPWWU encompasses not only ambulatory medicine above, but also the "other physician services" PCA because the latter includes the emergency unit, and admitting and screening, which are important clinic stops with heavy internal medicine involvement. All coefficient estimates above are statistically significant at the .05 level (two-tail-test), except for FELLOWS (for which $p < .10$).

From Equation (4) it can be inferred that inpatient, ambulatory care, and long-term care WWUs all influence internal medicine physician FTEE required for direct care and resident education; required physician FTEE is positively related to the number of fellows; and the relationship between RAM group and physician FTEE is complex and depends, in particular, on the absolute level of MEDWWU.

APPLICATIONS: COMPARING MODEL-DERIVED WITH ACTUAL PHYSICIAN STAFFING AT A VAMC

To illustrate how these empirically based models can be put to use, we focus on one actual VAMC, a large medical school–affiliated hospital in a metropolitan area. In the first application, all 11 estimated IPF models are used to compare the predicted physician FTEE for direct care and resident education in FY 1989 with the actual FTEE. In the second application, all 14 estimated PF models are used to derive physician requirements for direct patient care, by specialty, at this facility; these "projections" are compared with the actual FY 1989 FTEE allocation for patient care.

IPF Analysis of Physician FTEE

Given specified workload rates (and values for other variables in the equation), the IPF will predict the specialty k FTEE required for patient care and resident education at VAMC i. If predicted FTEE exceeds actual FTEE, this may be interpreted as a *signal* that VAMC i is "understaffed" in specialty k.

(But, alone, this is not sufficient for a policy conclusion that *i* is understaffed in *k*–unless the VA decision maker has elected to build the physician staffing methodology exclusively around the IPF.) Likewise, if actual FTEE exceeds predicted FTEE, there is an indication that the facility may be overstaffed.

In Table 1, actual and predicted physician FTEE for direct patient care and resident education are compared for each of 11 specialties at the selected VAMC;[10] the "prediction intervals" yield information about the statistical precision of each IPF prediction. Frequently, there is a substantial divergence (in percentage terms) between actual and IPF-predicted physician FTEE. As suggested, such divergence constitutes *prima facie* evidence that staffing in that specialty departs substantially from VA system norms. Subsequent discussions with the facility could reveal any of several explanations for the divergence: an error in data from the VA cost distribution report; significant assistance from non-VA consulting physicians that reduced staff physician requirements (assuming predicted FTEE exceeds actual); or an especially severe case mix within the DRGs assigned, so that WWUs understate the demands on physician time (assuming actual exceeds predicted). On the other hand, these discussions might indicate that the facility is indeed understaffed or overstaffed, relative to the VA system "norms" embedded implicitly in the IPF equations.

PF Analysis of Physician FTEE

The estimated PF equations provide an alternative approach for deriving physician FTEE requirements for direct patient care. Compared with the IPF approach, a PF-based strategy has the advantage of allowing physician patient care requirements to be derived, by specialty, at the PCA level while taking into account the productivity contributions of other physician and nonphysician personnel.

However, the PF approach does present some complications. If the variable representing a certain specialty does not merit inclusion in a given PF on statistical grounds, the required staffing for that specialty in the associated PCA will always be computed as zero, whatever the specialty's actual involvement in patient care. Also, because the dependent variable is workload and not physician FTEE, prediction intervals on FTEE requirements cannot be computed directly.[11] Moreover, for any given workload level, say W^*, projected for a PCA, there is typically not one, but an unlimited number of provider combinations which, when substituted into the PCA's production function, yield a predicted value of W equal to W^*. In traditional production theory in economics, the issue is resolved because the assumption of profit maximization implies that the firm will choose the one combination of inputs that minimizes cost, given the desired rate

Table 1: CDR-Based Actual Physician FTEE and IPF-Predicted FTEE for Direct Patient Care and Resident Education by Specialty at a Selected VAMC, FY 1989

	Physician FTEE		
Specialties	*Actual**	*Predicted**	*95% Prediction Interval†*
Internal medicine	44.52	43.98	(19.99, 96.13)
Surgery	14.51	15.08	(8.52, 26.30)
Psychiatry	19.33	22.85	(12.16, 42.07)
Neurology	4.32	5.06	(2.47, 9.54)
Rehabilitation medicine	2.51	4.08	(2.03, 7.59)
Spinal cord injury	0.76	1.61	(0.55, 3.39)
Anesthesiology	5.72	7.92	(3.59, 16.39)
Laboratory medicine	7.16	5.88	(3.06, 10.68)
Diagnostic radiology	13.70	9.93	(4.92, 19.12)
Nuclear medicine	1.92	1.68	(0.78, 3.07)
Radiation oncology	1.21	2.86	(1.03, 6.32)

*Includes all physician FTEE for direct care and resident education associated with the specialty's cost distribution report cost center, across all patient care areas (and thus encompassing the emergency, and admitting and screening areas of the "other physician services" patient care area); excludes physicians in that specialty who are assigned to a CDR cost center other than the one normally associated with the specialty.

†This corresponds to the 95 percent confidence interval around the predicted value for physician FTEE in specialty k at VAMC i. It is important to distinguish these prediction intervals from the oft-computed, and generally much tighter, "confidence intervals" about the expected value (mean) of the dependent variable. In evaluating the precision of a given staffing estimate at a particular VAMC, it is the (larger) prediction interval that is more relevant. See Moses (1986) for the relevant formulas.

of output. But in this VA context, what is the criterion for distinguishing among feasible input combinations?

One approach is to use mathematical programming to select the "optimal" combination of providers for the patient care area of interest, subject to meeting specified constraints. For example, one can determine the cost-minimizing combination of physician and nonphysician providers for handling the projected patient workload for a PCA, subject to meeting certain quality-of-care requirements–for example, that the ratio of attending physicians to residents lies within certain bounds. This appealing, though computationally intensive approach is illustrated in detail in Institute of Medicine (1991, 299–320). (For a related application comparing VA medical centers based on the cost efficiency of input use, see Sexton, Leiken, Nolan, et al. 1989).

As a practical alternative, we adopted the following simplifying strategy: Given projected workload W^* for a PCA, determine what equal-proportionate change in all provider variables (from their current levels) would yield new values just sufficient (in concert) to produce W^*. This "proportionate adjustment" method makes the strong, though not unreasonable assumption that input *ratios* will not vary significantly with output (although input *levels* will vary, of course). The results of applying this method to determine physician requirements in internal medicine at our selected VAMC is summarized in Table 2. Parallel analyses have been performed for five other specialty groupings[12] (surgery, psychiatry, neurology, rehabilitation medicine, and spinal cord injury) and for two clinical program areas (ambulatory care and long-term care); see Institute of Medicine (1991, 114–19).

As with the IPF application, divergences between actual and projected staffing do not *ipso facto* mean inappropriate staffing–the data do not "speak for themselves." Rather, they must be interpreted in light of (1) additional information that the VAMC or others may wish to bring to bear and (2) the relative weight accorded an empirically based approach in the overall physician staffing strategy.

The PFs and IPFs can be deployed similarly to derive physician requirements for future fiscal years, once the applicable patient care workloads have been projected. For illustrative calculations pertaining to fiscal years 2000 and 2005, see Institute of Medicine (1991, 100–103, 124–135).

Table 2: For Internal Medicine, CDR-Based Actual Physician FTEE and PF-Derived Projected FTEE for Direct Patient Care at a Selected VAMC, FY 1989

	Physician FTEE	
Patient Care Areas	*Actual*	*Projected*
Inpatient		
Medicine	10.95	12.13
Surgery	3.52	3.94
Psychiatry	0.57	0.00
Neurology	0.28	0.40
Rehabilitation medicine	0.06	0.06
Spinal cord injury	0.51	0.00
Ambulatory		
Medicine	8.05	12.28
Surgery	0.28	0.00
Psychiatry	0.23	0.00
Neurology	0.00	0.00
Rehabilitation medicine	0.06	0.00
Other physician services	12.47	9.59
Long-Term Care		
Nursing home	0.68	0.60
Intermediate care	0.11	0.18
Total	37.77	39.18

COMMENTARY

These analyses demonstrate that statistically strong, clinically plausible models for calculating physician staffing for patient care and resident education can be developed with currently available VA data. As noted, *total* physician requirements (by specialty) for a facility is the sum of these model-derived FTEE and (separately assessed) FTEE for research, continuing education, and other activities (Institute of Medicine 1991, 375–84). While making these empirically based models a central component of its proposed staffing strategy, the IOM committee also concluded that such models could be strengthened substantially by (1) improving the accuracy of the staffing data in the VA cost distribution report, (2) developing new variables (especially for measuring workload and for capturing certain physician FTEE currently omitted, such as for consulting-and-attending physicians), and (3) continuing to refine the specification of the PFs and IPFs and the statistical methods used to estimate them (Institute of Medicine 1991, 394–96). The committee recommended that VA should test, evaluate, and revise (as needed) the empirically based models on an ongoing basis. With the VA system in dynamic transition, this important component of the overall physician staffing strategy should not be regarded as a static construct. Rather, what is required is an *evolutionary* process, with the flexibility to adapt to changing times and to learn from its own discoveries.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the special contributions of W. Daniel Culver, VA program analyst; Gabriel O. Manasse, VA project officer; the VA Liaison Committee, chaired by Elizabeth L. Rogers; and Karl Yordy, director of the IOM Division of Health Care Services during the period of this project.

NOTES

1. The detailed facility-level models developed here differ significantly from the more aggregate models used typically to derive national estimates of physician requirements (Feil, Welch, and Fisher 1993; Reinhardt 1991).
2. VA supported IOM to create a "mathematical/statistical methodology, incorporating both empirically derived and expert judgment based values in the methodology's algorithms, which translate quantitative measures of. . . mission-related workload demands . . . into numerical estimates of physician staffing

requirements" (Institute of Medicine 1991, 14). A 19-member IOM study committee, chaired by David R. Challoner of the University of Florida, included experts in several physician specialties, nursing, allied health, statistics, economics, operations research, and health services research. Lipscomb was staff director of the IOM study, Kilpatrick was a member of the study committee and chaired its data and methodology panel, and Lee and Pieper were the principal statistical consultants.

3. For a detailed description of the modified Delphi processes through which these expert judgment estimates of staffing are derived, see Institute of Medicine (1991, 151–220). For a demonstration of how empirically based and expert judgment approaches can be used singly or in concert (through a "reconciliation strategy") to calculate physician staffing for a given VA facility, see Institute of Medicine (1991, 221–40).
4. Gaynor and Pauly (1990) specified a "behavioral production function" model, in which the medical group partnership selects the utility-maximizing level of work effort and profit.
5. In concept, an IPF can be estimated for each specialty–PCA combination (e.g., for neurology requirements on the inpatient medicine PCA), or for each specialty on a facility-total basis by aggregating across PCAs (e.g., for neurology requirements for all 14 PCAs combined). However, efforts to estimate PCA-specific IPFs for each specialty produced equations frequently exhibiting poor goodness of fit and coefficient estimates with algebraic signs that were counterintuitive. Hence, we focus exclusively on facility–level IPFs. Originally, the IPF also included variables for nurse and support staff. But after a number of statistical analyses, it became clear that the physician-substitutive role these providers are hypothesized to play in the production of workload could not be modeled adequately at the facility level.

   Rather, the PCA is the more appropriate level of aggregation for studying these relationships.
6. In these instances, the regression coefficient estimate for Nurse_{ij} was invariably positive and strongly significant (with t-statistics often larger than 10), while estimates for the physician variables, although generally positive, were frequently not significant. One possible explanation is that VA's extensive use over the years of nurse staffing algorithms pegged, in part, to the level and severity of patient workload serves here to create a significant causal link running from W_{ij} and Nurse_{ij}. If so, using Nurse_{ij} in the PFs would lead to biased coefficient estimates throughout the model, since this regressor would be significantly correlated with the error term. We assume here, on the other hand, that the variable $\text{Nurse}_{ij}/\text{StaffPhys}_{ij}$ plays quite a different role conceptually in the PF (see text) and, hence, that causality runs (if at all) from this "staffing intensity" variable to workload.
7. Another motivation is that affiliated and unaffiliated VAMCs face differing incentives in the allocation of physician FTEE to patient care in the VA cost distribution report. Physician research has been treated budgetarily as a "pass-through," and affiliated VAMCs have a much higher percentage of research

FTEE than nonaffiliated. If patient care FTEE are undercounted at affiliated VAMC, this would lend an upward bias to the physician productivity estimates inferred from any PF or IPF–and a resulting downward bias on the estimated physician FTEE required to handle any workload target. The hospital group variables offer a simple, albeit crude, means to correct for this possible bias.

8. As the VA health system moves toward a managed care design, it would be desirable to use workload measures in the PFs and IPFs that are defined on a capitation basis. Also, such measures would be more intuitive than the WWU, CAPWWU, and RUGWWU variables defined here. However, in FY 1989, these were the workload variables principally used by VA for budgeting and management purposes.
9. For purely computational reasons, each workload variable in Equation (4) is divided by 10,000; this affects the absolute value of each coefficient estimate, but not its sign or statistical significance.
10. Since the dependent variable in each IPF is in natural log form and is assumed to be normally distributed, the underlying physician FTEE variable is assumed to be lognormally distributed. It can be shown that the predicted value of physician FTEE is then properly interpreted as the ***median*** (not the mean) physician FTEE expected at such a facility (see Kmenta 1986). We regard this as a reasonable normative base against which to compare actual physician staffing at the facility. A similar interpretation applies to predicted workload values derived from the PFs.
11. In the least-squares regression model used to estimate these PFs, physician FTEE is assumed to be nonprobabilistic; hence, it is not possible to determine the "statistical precision" of the physician FTEE calculated to be consistent with the production of a given patient workload.
12. A PF could not be estimated for specialties (anesthesiology, laboratory medicine, nuclear medicine, diagnostic radiology, and radiation oncology) that, because of the scope and dispersion of their patient care contributions, could not be linked to a single inpatient, ambulatory, or long-term care PCA.

REFERENCES

Efron, B., and G. Gong. 1983. "A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation." *The American Statistician* 37 (1): 36–48.

Feil, E. C., H. G. Welch, and E. S. Fisher. 1993. "Why Estimates of Physician Supply and Requirements Disagree." *Journal of the American Medical Association* 269 (20): 2659–63.

Feldstein, M. S. 1967. *Economic Analysis for Health Service Efficiency*. Chicago: North-Holland, Markham Publishing Company.

Gaynor, M., and M. V. Pauly. 1990. "Compensation and Productive Efficiency in Partnerships: Evidence from Medical Group Practice." *Journal of Political Economy* 98 (3): 544–73.

Institute of Medicine. 1991. *Physician Staffing for the VA, Volume 1,* edited by J. Lipscomb. Washington, DC: National Academy Press.

Institute of Medicine. 1992. ***Physician Staffing for the VA, Volume 11,*** edited by J. Lipscomb and B. J. Alexander. Washington, DC: National Academy Press.

Jensen, G. A., and M. A. Morrisey. 1986a. "Medical Staff Specialty Mix and Hospital Production." ***Journal of Health Economics*** 5 (3): 253–76.

Jensen, G. A., and M. A. Morrisey. 1986b. "The Role of Physicians in Hospital Production." ***Review of Economics and Statistics*** 68 (August): 432–42.

Kmenta, J. 1986. ***Elements of Econometrics.*** 2nd ed. New York: Macmillan Publishing Company.

Kushman, J., R. Scheffler, L. Miners, et al. 1979. "Non-Solo Dental Practice: Incentives and Returns to Size." ***Journal of Economics and Business*** 31 (1): 29–39.

Maurizi, A. R. 1969. ***Economic Essays on the Dental Profession.*** Iowa City: University of Iowa, College of Business Administration.

Moses, L. E. 1986. ***Think and Explain with Statistics.*** Reading, MA: Addison-Wesley Publishing Company.

Pauly, M. V. 1980. ***Doctors and Their Workshops.*** Chicago: University of Chicago Press.

Reinhardt, U. E. 1991. "Health Manpower Forecasting: The Case of Physician Supply." In ***Health Services Research: Key to Health Policy,*** edited by E. Ginzberg. 234–83. Cambridge, MA: Harvard University Press.

Reinhardt, U. E. 1975. ***Physician Productivity and the Demand for Health Manpower.*** Cambridge, MA: Ballinger Publishing Company.

Scheffler, R. M. 1981. "Productivity and Economies of Scale in Dentistry: Some Recent Empirical Evidence." In ***Proceedings of a Conference on Modeling Techniques and Applications in Dentistry,*** edited by L. J. Brown and J. E. Winslow. U.S. Department of Health and Human Services Publication No. (HRA) 81-1. Washington, DC: DHHS.

Scheffler, R. M., and J. E. Kushman. 1977. "A Production Function for Dental Services: Estimation and Economic Implications." ***Southern Economic Journal*** 44 (1): 25–35.

Sexton, T. R., A. M. Leiken, A. H. Nolan, S. Liss, A. Hogan, and R. Silkman. 1989. "Evaluating Managerial Efficiency of Veterans Administration Medical Centers Using Data Envelopment Analysis." ***Medical Care*** 27 (12): 1175–88.

Stefos, T., N. LaVallee, and F. Holden. 1992. "Fairness in Prospective Payment: A Clustering Approach." ***Health Services Research*** 27 (2): 239–61.

U.S. Department of Veterans Affairs. 1990. ***Annual Report of the Secretary of Veterans Affairs, FY 1989.*** Washington, DC: DVA.