

Methodological Issues in Health Human Resource Planning: Cataloguing Assumptions and Controlling for Variables in Needs-Based Modelling

Gail Tomblin Murphy

Les modèles de Planification des ressources humains en matière de santé [Health Human Resource Planning] (HHRP) reflètent les futures exigences infirmières et sont élaborés en fonction d'un éventail de facteurs liés au modèle utilisé. Il existe un urgent besoin de comprendre davantage les sources de gauchissement quant à la modélisation statistique, dans le but de nous doter de solides formules précises qui nous serviront de guides. Cet article traite de ces questions telles qu'elles s'appliquent dans le contexte de la recherche infirmière HHRP basée sur les besoins. Il présente une critique et discute de la documentation pertinente portant sur : (1) la mise à l'épreuve des formules, (2) l'élimination d'erreurs écologiques et atomistiques, (3) les liens directs ou indirects entre les besoins et les soins de santé, et (4) les solutions de rechange permettant de regrouper les analyses portant sur l'évaluation de la relation entre les besoins en matière de santé et le recours aux services infirmiers. L'article conclut que la modélisation à niveaux multiples permet d'effectuer une analyse de simulation des individus et de leurs écologies, et que la modélisation de variation pour secteur réduit s'avère prometteuse quant à l'évaluation des liens entre les besoins en matière de santé et l'utilisation des services de santé.

Health Human Resource Planning (HHRP) models approximate future nursing requirements based on a variety of factors specific to the model being employed. There is an urgent need to develop a better understanding of the sources of bias in statistical modelling in order to ensure that we are guided by accurate and robust formulae. This paper addresses these issues as they apply in the context of needs-based HHRP research for nursing by presenting a review and discussion of the relevant literature as it relates to: (1) the testing of assumptions, (2) avoiding ecological and atomistic fallacies, (3) how need is directly or indirectly related to health care, and (4) alternatives to aggregate analysis for assessing the relationship between health needs and utilization of nursing services. The paper concludes that multilevel modelling is useful for the simulation analysis of individuals and their ecologies, and that small area variation modelling holds promise for assessing the relationship between health needs and utilization of nursing services.

Gail Tomblin Murphy, PhD(c), is Associate Professor, School of Nursing, Dalhousie University, Halifax, Nova Scotia, Canada.

Introduction

Canada and other countries are experiencing a nursing shortage that is expected to become severe. Policy-makers, health-care leaders, and researchers are interested in developing a meaningful, long-term approach to health human resource planning in nursing. This will require collaborative efforts to develop effective mechanisms and policies for establishing, monitoring, and predicting nursing requirements to meet the health needs of the population (Dickson, 2001; Sechchrist, Lewis, & Rutledge, 1999). It is critical that these resource allocations be based on evidence. Health human resource planning (HHRP) models are intended to provide such evidence by approximating planning requirements on a variety of factors specific to the model being employed. Regardless of which model is used, however, sound methods of study design and data analysis are essential if the evidence is to be meaningful. To complicate matters, HHRP research is a nascent discipline for which a firm scientific basis has not yet been fully worked out and tested. Methods for predicting human resource requirements are few in number and continue to be plagued with methodological and conceptual difficulties (O'Brien-Pallas, Birch, Baumann, & Tomblin Murphy, 2000). Until very recently, researchers in HHRP have not focused on the development of a theoretical knowledge base favourable to the creation of theoretical frameworks. Furthermore, there has been little consideration of research design and empirical specification of relationships among variables. As a result, health researchers often remain blind to the hierarchical structure of the conceptual models and data sets with which they are working. This has led to the development and dominance of statistical models that fail to capture the importance of the interface between individual-level and community-level data (Birch, Stoddart, & Beland, 1998). There is an urgent need to develop a better understanding of the sources of bias in statistical modelling, in order to ensure that we are guided by accurate and robust formulae (Carr-Hill et al., 1994).

Understanding the sources of bias in HHRP research requires first that we identify relevant assumptions, and second that we develop effective strategies to control for variables associated with these assumptions, the ultimate goal being to strengthen the generalizability of findings in order to more accurately forecast the need for nurses across a range of practice settings and jurisdictions. I will address these general considerations as they apply to needs-based HHRP research for nursing. Specifically, this will require a discussion of: the improper testing of anticipatory assumptions, codependency assumptions, and

assumptions related to design and analysis; the need for the development of multilevel modelling methods for the simulation analysis of individuals and their ecologies, in order to avoid problems associated with both the ecological fallacy and the atomistic fallacy; strategies to overcome the limitations of past research caused by a failure to adequately account for how need is directly or indirectly related to health care; and the use of small area variation as a promising alternative to aggregate analysis in enhancing our understanding of the relationship between health needs and utilization of nursing services.

Before exploring the main themes of this paper, I will describe briefly three HHRP methods (utilization-based, needs-based, and demand-based) and their related assumptions, questions, and methods (Lavis & Birch, 1997; O'Brien-Pallas et al., 2000). Table 1 presents an overview of the three methods. By proceeding in this fashion I hope to make clear that the needs-based approach I am advocating has evolved from the experience gained working with utilization-based frameworks. In utilization-based models, the quantity, mix, and population distribution of current health-care resources are adopted as a baseline for estimates of future requirements. As shifts are predicted in the basic (i.e., age and sex) demographic characteristics of the population, these are compared to current baseline factors in order to predict future requirements. This approach, however, is limited by the fact that utilization rates are dramatically affected by factors other than the population characteristics typically included in utilization-based models (Lavis & Birch; O'Brien-Pallas et al.).

As the name suggests, needs-based planning models estimate future nursing requirements based on an empirical assessment of the levels of health need in the relevant population. Unlike utilization-based approaches, however, needs-based planning begins from the premise that the current distribution of nurses and nursing services is not necessarily optimal for addressing the needs of the population, and that nursing human resources must be redistributed if health needs are to be met (Evans & Mustard, 1995; Markham & Birch, 1997; Roos et al., 1995).

According to Eyles, Birch, and Newbold (1995), there are two broad approaches to needs-based planning: the absolute needs approach and the relative needs approach. The absolute needs approach attempts to quantify the health-care resources required to produce a target level of health for the general population. The relative needs approach considers the variation in health status among communities together with available resources in order to meet the needs of communities. The

Approach	Assumptions	Question	Method
Utilization-based	<p>Current level, mix, and distribution of services are appropriate.</p> <p>Appropriate level, mix, and distribution remain constant.</p> <p>Estimated future demographic profile is accurate.</p>	How many nurses are required to serve the future population in the way that the current population is being served?	<p>Uses population-based utilization rates as a baseline.</p> <p>Applies rates to demographic profile of future population.</p>
Needs-based	<p>(All) health needs can and should be met.</p> <p>Cost-effective methods can be identified and implemented.</p> <p>Unmet needs caused (only) by inadequate supply.</p> <p>Non-needs/non-cost-effective use of resources can be eliminated.</p>	How many nurses are required to meet the health needs of the population?	<p>Population-based rates of health needs replace population-based use of services.</p> <p>Identifies nurse requirements for serving population's health needs.</p> <p>Applies rates to demographic profile of future population.</p>

<p>Demand-based</p>	<p>Health needs are only one of a set of social priorities.</p> <p>Definitions of need are less than precise.</p> <p>There are clear possibilities for resource trade-offs.</p>	<p>How many nurses are required to support society's commitment to health care?</p>	<p>Estimates the size of the economy in supporting health care.</p> <p>Estimates the proportion of the economy devoted to health care.</p> <p>Estimates the proportion of health-care expenditures allocated to nursing.</p> <p>Estimates the number of nurses that could be employed using these resources.</p> <p>Provides a fiscal-resource context for needs-based (or utilization-based) methods.</p>
<p>Alternative approaches address different questions that arise from the basic assumptions of each model and that ultimately determine the method. Each approach is informed by the experience of the previous one. The assumptions associated here with each model are not intended to map precisely onto the assumptions and strategies discussed in the text.</p>			

down side of the former approach is the risk that resources will be allocated based on the demand for services (rather than actual need), thus placing too great a claim on social resources. It is primarily because of this worry that the relative needs approach is preferable. Data obtained through the relative needs approach can better be used to plan distribution of the human resources workforce in which the overall cost of increasing the supply is explicitly acknowledged. Nursing resource planning must therefore consider the relative needs of communities rather than simply the demand for services as expressed as a function of the utilization of nursing and related services. Otherwise the use of resources will in fact be based on demand rather than need, with the consequence that HHRP will continue hopelessly endeavouring to provide an ever-increasing number of workers. Both approaches, however, require assessment of the health needs of populations and variations among communities in order to match resource allocations to the needs of the population.

The type of needs-based model being advocated treats nursing requirements as independent of current service utilization while recognizing that they are interdependent with the requirements for other health human resources and various other factors. Moreover, to the degree that current needs are not all met, unmet needs will be included in the estimation process. Thus, the main virtue of this approach lies in its ability to avoid perpetuating existing inequities and inefficiencies in the deployment of nursing services (O'Brien-Pallas et al., 2000). Effective demand-based approaches to HHRP research attempt to incorporate economic considerations into the epidemiological principles of the needs-based approach. The starting point is to estimate the size of the economy from which nursing and competing services will be funded, in order to then estimate the proportion of total resources that will be allocated to health care and the share of those resources that will be devoted to nursing (Lavis & Birch, 1997). Notwithstanding the welcome inclusion of general economic considerations brought to the mix by effective demand-based approaches, it remains clear that epidemiological information on the level and distribution of needs in the population, and the role of nursing resources in meeting those needs, will continue to be at the heart of any plausible statistical model employed in HHRP. Furthermore, measuring population health needs will help identify when these needs can reasonably be met through the distribution of health-care resources and when they may best be met through various other social mechanisms. Although the health-care system may not have the resources to directly impact all of the determinants of health, it is nonetheless often left to deal with the conse-

quences of a more general social failure to address them. It is therefore imperative that the measurement and understanding of population health needs be an integral part of HHRP.

It is at least in part for this reason that I have chosen to use needs-based modelling as the backdrop to the present discussion of assumptions and variables.

Cataloguing Assumptions

Anticipatory Assumptions

The success of forecasting nursing requirements varies with the ability to accurately predict the future (O'Brien-Pallas et al., 2000). The accuracy of projections will in turn depend on the completeness, accuracy, comprehensiveness, and availability of information (O'Brien-Pallas, 1998; Pong, 1997). Unfortunately, however, many jurisdictions in Canada and elsewhere lack the management infrastructure and data on which to base projections. Projections, therefore, often are made on the basis of a set of anticipated conditions. These "anticipatory assumptions" concern both conditions that are likely to remain static and conditions that are likely to undergo change. Typically, however, they concern only those relevant background conditions that change relatively slowly, if at all. Because of this resistance to change, anticipatory assumptions are unlikely to be susceptible to small, rapid, or short-duration changes in other conditions — for example, a change in the demand for nurses. Only a relatively large-scale and lasting change in other conditions is likely to produce a noticeable alteration in these anticipatory assumptions. Thus, anticipatory assumptions provide a suitably stable assumptive basis upon which to conduct needs-based HHRP. This is not to say, however, that unanticipated shocks to the system that are both rapid and severe in magnitude might not jeopardize the stability of the planning platform. Beyond this, projections will invariably be influenced by the values and political commitments that underlie the policy framework of the health-care system (Eyles, Birch, & Newbold, 1993).

As an example of what I call anticipatory assumptions, policies and strategies for nursing resource planning are often guided by the following assumptions: that the current distribution of nurses is not optimal for population health needs; that the distribution of nurses will reflect human and fiscal realities; that fiscal resources are allocated based on their availability relative to competing demands and the vagaries of political will; that practice conditions will remain

unchanged for nurses in the short term; that the health-care system will be dramatically affected by demographic changes in society; and that the health-care system will continue to be based on a “medical model” and remain substantially under the control of physicians (Hughes, Tomblin Murphy, & Pennock, 1998).

Other anticipatory assumptions relate to areas such as health services, nursing utilization and trends, and health needs. Current approaches to projecting nursing resources, for instance, are often based on a recognition that the population is aging. According to O’Brien-Pallas et al. (2000), it is reasonable to assume that average per capita age distribution will continue to move towards an older median age. At the same time, it must be assumed that the proportion of other age groups will decrease. The net effect of this shifting age distribution must be considered in the analysis of data used to predict nurse resource requirements if the predictions are to match future need.

Codependency Assumptions

Not all relevant assumptions are related to anticipated conditions. Another set of assumptions recognizes a priori that the supply and distribution of other health services and personnel are also likely to influence the demand for nurses. These “codependency assumptions” are less fixed than the anticipatory assumptions described above and are therefore subject to relatively rapid fluctuations. They are also influenced by many of the same factors that affect the demand for nurses. For example, while nursing demand is unlikely to be independent from the supply and distribution (both geographical and in terms of specialty) of beds, services, physicians and other providers, it is also the case that as the factors that create the demand for nurses change, so too will the supply and distribution of these resources. It is also likely to be the case that there will be simultaneity between the demand for nurses and these other resources — that is, the demand for nurses itself (not just the factors that create the demand) may have a measurable impact on the demand for and availability of other resources. Such changes may in some instances be quite precipitous when changes in nursing demand occur.

The distinction between anticipatory and codependency assumptions is an important one. The former provide a suitably stable assumptive platform from which to study nursing demand, thus adding strength of generalizability to the statistical models that incorporate them. Making explicit these anticipatory assumptions may also assist researchers in assigning bias. Codependency assumptions serve an

entirely different theoretical and pragmatic function: they force researchers to consider more closely the full range of codependent variables and how best to control for them.

Assumptions Related to Design and Analysis

There are other kinds of assumptions that are more directly related to research design and data analysis. Assumptions underlying regression analysis, for example, can cause biases and distortions in data analysis. These assumptions relate to: model specification, measurement, fixed and random variables, and residuals (Pedhauzur & Schmelkin, 1991). Researchers are challenged to find the appropriate statistical models to link with theoretical frameworks and research designs. Assumptions around model specification and measurement are in most urgent need of attention, since they have not been well attended to in health services and nursing research until quite recently. According to experts in measurement and design, researchers need to be reminded that the statistical models used should reflect the theoretical models (Pedhauzur & Schmelkin). However, because the theoretical underpinnings of HHRP remain relatively inchoate, the link between statistical methods and theoretical models is often not apparent. Unless researchers pay attention to these assumptions, their findings will be either meaningless or misleading. For example, the regression equation must reflect the formulation regarding the effects of the independent variable(s) on the dependent variable(s). Failure to include relevant variables in the research design can also lead to biased parameter estimation. Conversely, inclusion of irrelevant variables can affect the significance testing of coefficients for the relevant variable. Specification errors can also result when the manner in which the variables in the model affect the dependent variable is incorrectly specified.

Investigators may also assume that the measurement of independent variables is immune to error. However, biased estimation of regression coefficients can result from random errors in the independent variable (Pedhauzur & Schmelkin, 1991). Random measurement errors in multiple regressions may lead to either overestimation or underestimation of the regression coefficient. The biasing effects can also affect the variables that correlate with the target variable in question.

A common approach to controlling the magnitude of errors and their impact on the estimation of model parameters is Structural Equation Modelling (SEM). This approach takes into account measurement errors when estimating model parameters (Pedhauzur &

Shmelkin, 1991). The literature suggests that SEM holds promise for health services and nursing research. At present, however, researchers are struggling to find the appropriate statistical models to link with theoretical frameworks and research designs. The paucity of such models highlights the importance of developing analytical strategies that make assumptions explicit and control variables within a coherent theoretical framework.

Controlling for Variables

Recall that the goal of devising meticulous and effective strategies to control for variables is to strengthen research findings and enhance their generalizability. In considering the appropriate theoretical framework, research design, and statistical methods to control for variables, it will be helpful to consider some important issues in needs-based HHRP. How, for instance, can simultaneity (i.e., mutual dependence) between health status and nurse resource allocation be addressed? Will generic measures of health status be more closely related to nurse resource utilization or to need? How will the approach to controlling variables lead to appropriate projections of need?

Researchers need to consider the following issues associated with HHRP assumptions: (1) Researchers typically introduce community data into statistical models in the same way they enter individual data, causing confounding effects. (2) Aggregate analysis may be an inappropriate methodology for studying place differences (regional variation). (3) Researchers do not consistently use specifications that allow the model to reflect the underlying conceptual framework. (4) The Ordinary Least Squares (OLS) regression techniques often used by researchers in health services research do not consider selectivity bias in utilization studies. Moreover, if supply affects needs, it is inappropriate to use conventional OLS methods in determining utilization of nursing and other health services.

Some researchers consider these issues in their research designs. In the examples that follow, the researchers stress the importance of using statistical models that reflect the nature of conceptual models and data sets.

In their work in estimating relationships using models that reflect the hierarchical nature of both conceptual models and data sets, Birch et al. (1998) have recently developed creative statistical methods for addressing the complexity of the individual-community interface. Based on methods used in educational research (Goldstein, 1994), they

applied methods to demonstrate “how separate statistical models for variation in health between communities and between individuals can be combined to provide a multilevel model for the determinants of the health of populations” (p. 402). A community effect — “a variation in an individual level dependent variable embodied within communities” (p. 402) — is a critical consideration when using a statistical model. Specifically, this means that although it is defined at the community level, the community effect is concerned with influences on or through individuals according to the community in which they live.

Multilevel Modelling

Studies using multiple units of analysis in which comparisons among or between communities are desired require statistical methods that can address the different levels of analysis and at the same time deal with the complexities of the community-individual interface. Researchers typically introduce community data into statistical models in the exact same way that they enter individual data, thus causing a confounding of the effects. In order to properly estimate the relationship between a particular community-level variable and a particular individual-level variable, one should allow for the possibility that the community-level variable does not fully explain between variation in individual-level relationships. To adequately explain between variation, slopes and intercept terms for each community would need to be included (Birch et al., 1998). Birch et al., using an alternative approach based on the work of Goldstein (1994), suggest a multilevel model that reflects the hierarchical structure of the individual community data to explore the variation at the community level and covariation between community-level and individual-level variables.

Duncan, Jones, and Moon (1993) argue along similar lines, stating that analyses of variations in health require multilevel methods that acknowledge both the ecological fallacy (i.e., transferring results from aggregates to individuals) and the atomistic fallacy (i.e., researching exclusively at the individual level, thus failing to account for the context in which the individual action occurs). In needs-based HHRP, for example, the ecological fallacy may occur if a community’s health profile is not representative of the health status of the individuals actually using health services in that community (Jones & Duncan, 1995). This concern is compounded when measures of health are derived exclusively from pre-existing databases (e.g., National Population Health Survey, Canada Census data) rather than derived also from individuals presenting for care. The ecological fallacy is problematic in this

context because it will result in the creation of an incomplete analytical model of the hierarchical relationships between utilization and individual characteristics as well as between utilization and community characteristics.

Some authors (Birch et al., 1998; Duncan et al., 1993) describe multilevel modelling as a two-step process: specifying models at different levels in a hierarchy, then combining them into a single overall model. This process consists of an individual or micro-model, in which within-place equations are represented, and an ecological or macro-model, in which the parameters of the within-place model are set based on responses from between-place models.

Aggregate analysis is an inappropriate methodology for studying place differences. When researchers are forced to work at a single aggregate level, they often resort to grouping together qualitatively heterogeneous individuals in order to ensure statistically reliable rates based on a large number of respondents (Duncan et al., 1993). This may be conceptually damaging, since theoretically important relationships between people and places are collapsed for technical reasons alone. On the other hand, multilevel modelling does allow for simulation analysis of individuals and their ecologies.

It is important that the specifications used enhance the capacity of the model to reflect the underlying conceptual framework (Birch et al., 1998). For instance, the statistical approach of Eyles et al. (1995) in their analysis of the relationship between need for care and utilization of nursing services in Canada is particularly interesting in the context of needs-based HHRP. They argue that there is a potential bias in extant utilization studies that do not recognize users of care as a self-selected group within the population sample of users and non-users. This selectivity bias violates the normality condition of the OLS regression techniques and precludes the use of Poisson regression, binomial model, and OLS. These concerns are reiterated by Newbold, Eyles, Birch, and Spencer (1998), who also point out that these methods have in fact been commonly used by researchers who fail to recognize the bias. Furthermore, Eyles et al. (1995) stress that it is critical researchers progress beyond simple use-need ratios and address the interaction of need with other explanatory variables.

As stated earlier, indirect measures of health should be used cautiously because they are in fact only markers for risks to health. Direct measures of health status (e.g., SF-36, SMR, and self-reported health status) should be used in examining the relationships between health needs and utilization. Researchers should consider the recent work of

Sutton, Carr-Hill, Gravelle, and Rice (1999), who have shown that there may be significant bias in respondents' self-reporting of health status. They have developed a model of the relationship between morbidity and health service utilization that allows for reporting errors and simultaneity. Eyles et al. (1995) have made a valuable contribution to the knowledge of statistical models by considering some of the issues associated with the assumptions in their work. Based on Andersen's (1995) theoretical framework, they studied the following variables: use (number of nurse contacts); need (general health); predisposing categories (age, sex, marital status, region of residence, employment status, activity level, degree of community contact, smoking, alcohol use, and lifestyle); and enabling categories (income, education, and household tenure). To ensure that the estimated variance of variables is meaningful, they weighted the sample data (using weighting developed by Statistics Canada). In addition, they developed the following four proxy variables for need, all of which are independent of need: individual assessment of general health, satisfaction with personal health, presence or absence of longstanding limiting illness, and number of days within a 2-week period that individuals were limited by health. They found a significant level of intercorrelation (correlation coefficients ranging from 0.24 to 0.56) among the four variables. Although not explored by Eyles et al. (1993), the number of comorbid states could also serve as a proxy variable for need.

Data analysis was carried out in two stages. During stage one probit regression was used to estimate the incidence of nurse utilization. Dummy variables were created for general health (need proxy) and entered in blocks into an equation based on need and predisposing or enabling categorization (Andersen & Aday, 1978). The variables were entered additively and forced into the equation based on prior knowledge of the relationship between use and the explanatory variable. Using Confirmatory Factor Analysis (CFA), goodness-of-fit of the model was determined through Likelihood Ratio (LR) testing, which measures overall significance of the model, and the Rho-squared test, which indicates the explanatory power of the model.

To test the relationship between nursing services utilization and need, Eyles et al. (1995) estimated utilization using a correction factor (λ). The λ variable was estimated on the full sample using probit analysis before being entered into the least squares regression of number of nurse contacts on need and predisposing and enabling variables. The statistical significance of the λ variable was an indicator of whether the correction factor was significant.

During the second stage a multivariate model was employed to test the relationship between nursing utilization and selected socio-economic indicators. The first step was to divide the sample by level of need and analyze the data to determine horizontal inequity. This partitioning was carried out in order to explore the extent to which the relationships observed at the aggregate level can be generalized to the full sample, and then to identify relationships between utilization and other variables. However, this approach resulted in a sample that was too small to run the analysis. A second approach required the model to be re-estimated for the full range of users with interaction terms between need and specific explanatory variables. Although interaction terms may not lead to a large increase in the overall explanatory power of the model, they may be of help in explaining the observed relationships.

The approach adopted by Eyles et al. (1993) is laudable for its consistency with the theoretical framework and its contribution to our store of knowledge in statistical methods. Most notable is the discovery that “analysis of utilization at the aggregate level of population can conceal important and policy relevant relationships while revealing others that are essentially artifacts of inappropriate aggregation” (p. 43). The upshot is that populations of users of nursing services represent neither a homogeneous group nor a random selection of the total population.

Carr-Hill et al. (1994) argue that health needs may directly influence utilization of health care and that health-care utilization is partly mediated by supply. Furthermore, some of the variation in supply is determined by the distribution of health needs, socio-economic conditions, and past utilization, primarily because funding has been based on these and other variables that have historically determined supply. Simply stated, past utilization influences supply and supply influences current utilization. According to Smith, Sheldon, Carr-Hill, and Martin (1994), no study using methods based solely on utilization can adequately capture variations in health needs that are not reflected — at least partly — in utilization. Regardless of what statistical procedures are used and how well they are adjusted for supply considerations, all utilization-based models remain vulnerable to the possibility that health needs will not be captured in the use of hospitals by patients.

Carr-Hill et al. (1994) stress the importance of the relationship between needs and utilization. They point out that needs can influence use directly or through the effects of supply. Their research shows that consulting patterns in general medical practice are influenced by the effects of age, social class, unemployment, housing status, marital status, and ethnicity. However, a similar relationship between these

factors and utilization of nursing services has not been supported and can, therefore, only be hypothesized. Moreover, in the case of physician consulting patterns, individual characteristics are much more powerful predictors of utilization than is geography. These lessons support the need to guard against generalizing findings from aggregate data to individuals (i.e., ecological fallacy). The authors propose further that the effects of individual socio-economic factors vary according to geographic area. They argue, therefore, that resource allocation methods based on area of residence will always be inferior to an approach that takes into account the characteristics of individual patients.

Controlling for Small Area Variations

In response to these concerns, Carr-Hill et al. (1994) have developed a more sensitive, empirically based model of demand for hospital inpatient facilities in small areas — that is, a method for determining health needs for small geographical areas. The aim of the study was to separate out the effect of supply on use. The authors argue that if supply affects current use and need, then it is inappropriate to use conventional OLS methods to model the determination of use of health services. They suggest that a method such as two-stage least squares is needed to deal with the simultaneous determination of utilization and supply.

To capture the effect of varying availability of health services to small populations, Carr-Hill et al. (1994) created and measured four supply variables: accessibility of public inpatient facilities, general practitioner services, provision of residential and nursing homes, and accessibility of private inpatient facilities. In addition to these supply variables, their acute-care model includes 25 need variables. The model was progressively restricted by omitting variables that were not found to be significant or that had small standardized effects on predictions (i.e., beta coefficients). This process was continued until deletion of another variable would have altered the model significantly. Checks were made to ensure that the statistical tests were well specified. Multilevel modelling techniques were used to account for the hierarchical clustering of errors and to try to distinguish between interdistrict and intradistrict variations.

This model of small area variation shows promise for studies examining the relationship between needs and nursing service utilization. Given that observed levels of utilization and the relationship between levels of utilization and population health needs may be affected by variations in supply (e.g., hospital beds and nurses) between commu-

nities, the relationship between needs and service utilization must be estimated after allowing for small area variation in supply. However, this approach has its limitations, as articulated in studies using small area variation. Fisher, Wennberg, Stukel, and Skinner (2000), for example, examined the relationship between area bed supply and hospital discharge rates. Using Poisson regression (McCullagh & Nelder, 1983), they weighted by the number of subjects in each age-sex-race and zip code stratum, and controlled for all demographic, socio-economic, and health status variables. To measure the relationship between hospital resource measures and mortality, they used logistic regression controlling for the same demographic, socio-economic, and health status variables. Similarly, Coyte and Young (1999) assessed regional variations in the rates of home-care use following inpatient care and same-day surgery in the province of Ontario. Using four measures of regional variation (range in home-care rates, extreme coefficient, weighted coefficient of variation, and systematic component or variation), they found that home-care rates vary considerably by region.

Using the Andersen and Aday (1978) theoretical framework, Hofer, Wolfe, Tedeschi, McMahon, and Griffith (1998) explored the validity of using community-level estimates of socio-economic characteristics as a proxy for individual-level characteristics in adjusting hospitalization rates. They used logistic regression for the analyses of the first two dependent variables and Poisson regression for the count of number of admissions in 1 year. Their work further supports the argument that the small area variation in hospitalization rates depends significantly on socio-economic status (SES) effects, and that community-level measures of variables appear to be reasonable proxies for individual measures. They also reinforce the need to use small area variation (adjustments) in cross-sectional comparisons of small area rates.

Working with a nationally representative sample, Dunlop, Coyte, and McIsaac (2000) carried out a study to explain the role of SES in the differential utilization of publicly funded primary and specialty health services. Specifically, they assessed the extent to which socio-economic barriers exist in the utilization of physician services. A two-stage least squares was incorporated to explain specialist services. In the first stage of the study they assessed access to physician services by comparing the characteristics of those individuals who saw a physician during a 1-year period with those who did not. In the second stage they assessed frequency of physician service utilization by looking at the characteristics of those individuals who saw a physician at least six times in 1 year. They then compared frequent and non-frequent users. Multivariate logistic modelling was employed for four models (at least one GP visit,

six or more GP visits, at least one specialist visit, and six or more specialist visits). Models were initially fit using stepwise regression with a cut-off for inclusion of $p < 0.05$. The Wald Statistic was then used to examine each variable and each estimated coefficient was compared with the coefficient from the univariate model containing only that variable. Coefficients were estimated using maximum likelihood estimation. Goodness-of-fit was measured by a generalized coefficient of determination (Cox & Snell, 1990). Moreover, they included only those variables demonstrating a statistically significant association with physician visits.

Conclusion

At a time when nursing shortages are critical, it is important that evidence-based HHRP become a priority of policy-makers and researchers. It is clear that these stakeholders need to consider some critical questions, including: What are the approaches to HHRP that have led to shortages and surpluses in the past? What are the assumptions associated with the different approaches to HHRP and what are their related strengths and limitations? What is the best approach to ensure that a health human resource plan is developed with meaningful numbers, types, and mix of professionals to meet the needs of the population? A population needs-based approach to HHRP would provide policy-makers with the means to develop human resource strategies for meeting the current needs of the population and responding to the changing needs of populations over time. It is imperative that needs-based planning based on solid evidence drive HHRP for nursing.

It is apparent that the theoretical basis of HHRP is underdeveloped. Moreover, the connections between statistical models and the related theoretical frameworks are often not evident. Making these connections explicit will help researchers to uncover and better understand the assumptions that both drive and constrain a particular methodology. Beyond this, it is imperative that researchers devise meticulous strategies to control for variables. In nursing research and health services research, it is essential that models be applied that reflect the hierarchical nature of both conceptual models and data sets. Furthermore, researchers should use specifications of empirical models that enhance rather than limit the capacity of the model to reflect the underlying conceptual frameworks. When working with aggregated data, researchers must be aware that the grouping together of qualitatively heterogeneous individuals to ensure statistically reliable data can be damaging to the testing of theoretical frameworks. Multilevel modelling has been

shown to be useful insofar as it allows for the simulation analysis of individuals and their ecologies. The small area variation model developed by Carr-Hill et al. (1994) holds great promise for researchers hoping to better understand the relationship between health needs and nursing service utilization.

References

- Andersen, R. (1995). Revisiting the behavioral model and access to medical care: Does it matter? *Journal of Health and Social Behavior*, 36(3), 1–10.
- Andersen, R., & Aday, A. (1978). Access to medical care in the U.S.: Realized and potential. *Medical Care*, 16(4), 533–546.
- Birch, S., Stoddart, G., & Beland, F. (1998). Modelling the community as a determinant of health. *Canadian Journal of Public Health*, 89(6), 402–405.
- Carr-Hill, R., Sheldon, T., Smith, P., Martin, S., Peacock, S., & Hardman, G. (1994). Allocating resources to health authorities: Development of method for small area analysis of use of inpatient services. *British Medical Journal*, 309, 1046–1049.
- Cox, D., & Snell, E. (1990). *The analysis of binary data*. London: Chapman & Hall.
- Coyte, P., & Young, W. (1999). Regional variations in the use of home care services in Ontario, 1993/95. *Canadian Medical Association Journal*, 161(4), 376–383.
- Dickson, G. (2001). *Colleagues in caring: Addressing nursing work force issues*. Newark, NJ: State University of Nursing.
- Duncan, C., Jones, K., & Moon, G. (1993). Do places matter? A multilevel analysis of regional variations in health-related behaviour in Britain. *Social Science and Medicine*, 37, 725–733.
- Dunlop, S., Coyte, P., & McIsaac, W. (2000). Socio-economic status and the utilization of physicians: Results from the Canadian National Population Health Survey. *Social Science and Medicine*, 51, 123–133.
- Evans, R., & Mustard, J. (1995). [Foreword.] *Medical Care*, 33(12), DS5.
- Eyles, J., Birch, S., & Newbold, K.B. (1993). Equity and health care: Analysis of the relationships between need for care and the utilization of nursing services in Canada. *Canadian Journal of Nursing Research*, 25(4), 27–46.
- Eyles, J., Birch, S., & Newbold, K.B. (1995). Delivering the goods? Access to family physician services in Canada: A comparison of 1985 and 1991. *Journal of Health and Social Behavior*, 36(4), 322–332.
- Fisher, E., Wennberg, J., Stukel, T., & Skinner, J. (2000). Associations among hospital capacity, utilization, and mortality of U.S. medicare beneficiaries: Controlling for sociodemographic factors. *Health Services Research*, 34(6), 1351–1362.
- Goldstein, H. (1994). Multi-level cross-classified models. *Sociological Methods and Research*, 22, 364–375.

- Hofer, T., Wolfe, R., Tedeschi, P., McMahon, L., & Griffith, J. (1998). Use of community versus individual socioeconomic data in predicting variation in hospital use. *Health Services Research, 33*(1), 243–259.
- Hughes, J., Tomblin Murphy, G., & Pennock, M. (1998). *Future demand for registered nurses and licensed practical nurses*. Halifax: Province of Nova Scotia, Department of Health.
- Jones, K., & Duncan, C. (1995). Individuals and their ecologies: Analysing the geography of chronic illness within a multilevel modelling framework. *Health and Place, 1*(1), 27–40.
- Lavis, J.N., & Birch, S. (1997). The answer is..., now what was the question? Applying alternative approaches to estimating nurse requirements. *Canadian Journal of Nursing Administration, 10*(1), 24–44.
- Markham, B., & Birch, S. (1997). Back to the future: A framework for estimating health-care human resource requirements. *Canadian Journal of Nursing Administration, 10*(1), 7–23.
- McCullagh, P., & Nelder, J. (1983). *Generalized regression models*. New York: Chapman & Hall.
- Newbold, B., Eyles, J., Birch, S., & Spencer, A. (1998). Allocating resources in health care: Alternative approaches to measuring needs in resource allocation formula in Ontario. *Health and Place, 4*(1), 79–89.
- O'Brien-Pallas, L. (1998). *Health human resources: An analysis of forecasting models*. Ottawa: Canadian Nurses Association.
- O'Brien-Pallas, L., Birch, S., Baumann, A., & Tomblin Murphy, G. (2000). *Integrating workforce planning, human resources and service planning*. Geneva: World Health Organization.
- Pedhazur, E., & Schmelkin, L. (1991). *Measurement, design, and analysis: An integrated approach*. Mahwah, NJ: Erlbaum.
- Pong, R.W. (1997). Towards developing a flexible health workforce: A conference background paper. *Canadian Journal of Medical Radiation Technology, 28*(1), 11–18.
- Roos, N., Black, C., Frohlich, N., Decoster, C., Cohen, M., Tatarzyn, D., Mustard, C., Toll, F., Carriere, K., Burchill, C., MacWilliam, L., & Bogdanovic, B. (1995). A population based health information system. *Medical Care, 33*(12), DS13–20.
- Sechchrist, K., Lewis, E., & Rutledge, D. (1999). *The California Nursing Workforce Initiative: Planning for California's work force. Phase II: Final report*. Irvine: University of California.
- Smith, P., Sheldon, T., Carr-Hill, R., & Martin, S. (1994). Allocating resources to health authorities: Results and policy implications of small area analysis of use of inpatient services. *British Medical Journal, 309*(6961), 1050–1057.
- Sutton, M., Carr-Hill, R., Gravelle, H., & Rice, N. (1999). Do measures of self-reported morbidity bias the estimation of the determinants of health care utilisation? *Social Science and Medicine, 49*(7), 867–878.

Author's Note

I acknowledge the significant contributions of my PhD committee (Stephen Birch, Irv Rootman, Dot Pringle, and Gerarda Darlington) throughout my doctoral studies. My comprehensive examinations were the basis for this manuscript. I thank my supervisor, Dr. Linda Lee O'Brien-Pallas, for her guidance and support throughout my doctoral studies. Her insight in this area of HHRP has guided my thinking. Support from the following funding agencies has been critical to my work: Canadian Health Services Research Foundation, Ontario Ministry of Health and Long-Term Care, Nova Scotia Health Research Foundation, and Canadian Institutes of Health Research. I am forever grateful to a colleague who reviewed the manuscript during its development. I also wish to thank my two anonymous reviewers for their thoughtful comments and constructive suggestions.

Correspondence may be sent to: Gail Tomblin Murphy, School of Nursing, Dalhousie University, 5869 University Avenue, Halifax NS B3H 3J5 Canada. E-mail: <Gail.Tomblin.Murphy@dal.ca>.