MEDICAL CARE
Volume 40, Number 6, pp 477–484

Common Methodological Terms in Health Services Research and Their Symptoms

MATTHEW L. MACIEJEWSKI, PHD,*† PAULA DIEHR, PHD,†‡
MAUREEN A. SMITH, MD, MPH, PHD,§ AND PAUL HEBERT, PHD¶

OBJECTIVE. Health services researchers use methods and terminology from a variety of disciplines to understand individual and organizational behavior related to health, health care, and health insurance. Although this diversity benefits the process of research, the resulting differences in terminology can occasionally lead to confusion among even the most experienced researchers. The purpose of this paper is to clarify different methodological terms used to represent common study designs and statistical concepts in health services research.

METHOD. Faculty from the departments of Health Services, Biostatistics and Epidemiology at the University of Washington were asked to provide a list of terms commonly used in health services research that had one or more methodological terms with the same or similar meaning. Consistent with the Delphi method, the initial list was then circulated to internal and external methodologists and additional terms were added for completeness. Terms related to study designs, equation specification, and estimation problems with synonymous terminology were included.

RESULTS. A "methodological thesaurus" is developed and described in this paper. Many terms were found which described essentially identical concepts. In other cases, terms had slight but important differences in meaning across disciplines, leading to potential confusion in their use.

CONCLUSION. In light of these differences in terminology, it is not surprising that health services researchers with a particular disciplinary orientation currently find literature within their own discipline entirely readable and literature from another discipline considerably less accessible. The "methodological thesaurus" presented in this paper should be of value to health services researchers who are interested in exploring the full array of methods developed to address a particular problem. Teachers and students of health services research should also find this article of use in relating similar concepts across disciplines in quantitative health services research publications.

Key words: Research methods; terminology; statistics; study design. (Med Care 2002;40: 477–484)

There are many disciplines from the social and health sciences that have contributed to the emergence of health services research as an interdisciplinary field of study. Theories from economics, sociology, psychology, political science and other social sciences have improved our understanding of individual and organizational behavior related to health, health care, and health insurance. The

*From the Northwest Center for Outcomes Research in Older Adults VA Puget Sound Health Care System, Seattle, Washington.

From the †Department of Health Services, and the ‡Department of Biostatistics, University of Washington, Seattle, Washington.

§From the Department of Preventive Medicine, University of Wisconsin-Madison, Madison, Wisconsin.

¶From the Department of Health Policy, Mount Sinai School of Medicine, New York, New York.

Address correspondence and reprint requests to: Matthew L. Maciejewski, PhD, VA Puget Sound Health Care System, 1660 S. Columbian Way, Seattle, WA 98108. E-mail: mlmaciej@u.washington.edu

Received May 18, 2001; initial review June 26, 2001; accepted December 28, 2001.

application of research methods from biostatistics, epidemiology, economics, and other social and health sciences has enabled the quantitative analysis of health services research problems.

Although this diversity benefits the process of research, differences in terminology between disciplines can occasionally lead to confusion among even the most experienced researchers. For example, the important statistical concepts in a study of the effect of a drug intervention on mortality may be labeled differently by sociologists, economists, biostatisticians, and epidemiologists despite the fact that they are interested in the same relationship, namely the causal effect of a drug on mortality. These researchers may use identical statistical methods that incorporate both a dependent variable (eg, mortality) and independent variables (eg, the drug, age). However, an economist might refer to mortality as an endogenous variable, whereas a biostatistician might label mortality an outcome or response variable. Similarly, economists may call the drug an exogenous variable or a regressor, whereas biostatisticians and epidemiologists call the drug a treatment or exposure. The analytic method is identical, so these methods should be understandable to individuals from any discipline grounded in statistical theory and mathematics, even if behavioral theories explaining the causal relationships may not be immediately understood across disciplines.

In light of these differences in terminology, it is not surprising that health services researchers with a particular disciplinary orientation currently find literature within their own discipline entirely readable and literature from another discipline considerably less so, leading to much less cross-disciplinary communication than is optimal. Differences in assumptions and statistical theory of specific methods, and differences in scientific philosophy, that inhibit communication across disciplines are not discussed here. The purpose of this article is to cross-reference the various statistical terms used in health services research that represent similar study designs or statistical concepts. Most of the terminology commonly used in health services research and covered in this review is derived from economics, biostatistics and epidemiology, although there are other disciplines (eg, sociology) that have terminology that may or may not vary from that presented here.

This "methodological thesaurus" should be of value to health services researchers who are interested in exploring the full array of methods developed to address a particular problem. In particular, econometricians, biostatisticians, and epidemiologists interested in learning one another's methods will be able to grasp more quickly an analytic approach and compare approaches if terminology is clarified. Finally, teachers and students of health services research should also find this thesaurus of use in relating the same concepts across disciplines in quantitative health services research publications.

The thesaurus provides a set of terms related to study designs, elements of a question, equation specification, problems in estimation, and types of estimators that are found in many health services research publications. Each set of related terms includes an example to illustrate the relationship between a given concept and its synonymous terms. In the end, it should be clear that health services researchers from different disciplines have more in common (methodologically) than might appear to be the case initially.

Materials and Methods

All regular, research, clinical, and affiliate faculty from the departments of Health Services, Biostatistics, and Epidemiology at the University of Washington were contacted by email and asked to submit a list of terms that refer to the same statistical concept (eg, unmeasured confounding and omitted variables bias). Consistent with the Delphi method, the initial list was then circulated to internal and external methodologists, and additional terms were added for completeness. Terms related to study designs, equation specification, and estimation problems with synonymous terminology were included. Terms, such as cross-sectional studies and censoring, which are universally understood across disciplines were excluded. To ensure conceptual consistency, selected econometric texts,[1,8,9,12,13,25] epidemiologic texts[29] and biostatistical texts[3,4,6,16,18,19,24,32] were consulted.

The terminology discussed in this paper is organized into five types in Table 1. Terms used to describe study designs are presented first, followed by a section on the elements of an equation. The next three sections explain terminology related to equation specification, types of bias and types of estimators. The paper concludes with a brief discussion.

TABLE 1. Study Design and Statistical Terms in Health Services Research

Row No.	Common Term	Synonymous Terms
	STUDY DESIGNS	
1	Panel data study	Longitudinal or cohort study
2	Time series study	Longitudinal study
3	Cross-section, time series	Longitudinal study
4	Choice-based sampling	Case-Control study
	ELEMENTS OF AN EQUATION	
5	Dependent variable	Outcome, response, endogenous variable
6	Explanatory variable of interest	Dose, treatment, exposure, intervention, Exogenous variable of interest, predictor variable
7	Explanatory variable	Confounder, independent variable, regressor, exogenous variable, covariate
8	Interaction	Effect modification
9	Parameter estimate	Beta, regression coefficient, treatment effect
	REGRESSION MODELS	
10	Partitioned model	Stratified model
11	Multiple regression	Multivariate regression
12	Qualitative analysis	Categorical data analysis
13	Logit (or probit) model	Binomial logistic regression, logistic regression
14	Ordered logit regression	Ordinal logistic regression, ordinal log-linear regression
15	Multinomial logit regression	Polytomous logistic regression
16	Conditional logit regression	Conditional logistic regression, McFaddens logit
17	Survival analysis	Cox regression, hazard model, duration model failure-time analysis, event history analysis
	TYPES OF BIAS	
18	Omitted variable	Unmeasured covariate, unmeasured confounder, unobservable
19	Sample selection bias	Censoring, selection bias, incidental truncation
20	Selection bias	Unmeasured confounding, omitted variable bias, confounding by indication or contraindication
	TYPES OF ESTIMATORS	
21	Simultaneous equations	Multiple multivariate regression

Results

Study Designs

This comparison of terminology focuses on a particular statistical problem: examining the relationship between the variable X and the variable Y from a sample of individuals or organizations. Notice that we purposely have avoided using the expression "the causal effect of the variable X on the variable Y." This distinction is taken up later in the paper.

There are three general types of study designs that are commonly used in health services research studies to examine the relationship between X and Y: cross-sectional studies, observational longitudinal (ie, prospective or cohort) studies, and randomized longitudinal (ie, experimental) studies. In cross-sectional studies, all data for a group of individuals or organizations is observed at one point in time without a follow-up period or repeated measurements.

In the most common type of observational longitudinal study conducted in health services research, data on the exposure to an intervention (and possibly the outcome) for a group of individuals or organizations is collected at the beginning

of a follow-up period. Additional measurements of outcome (and possibly exposure) are then collected at several points in time, typically for more than 1 year. Since this type of study tracks the same group of individuals or organizations over time, an observational longitudinal study is also referred to as a cohort, panel, follow-up, incidence, or prospective study.[27] An observational longitudinal study may be completely prospective, retrospective, or both.

Economists also analyze data from two other variations of observational longitudinal studies: time-series studies and cross-sectional time-series studies. In time-series studies, data are collected at several points in time on a single observational unit (eg, US economy, one hospital). In a cross-sectional time-series study, data are collected on two or more cross-sectional samples of observations over time. Unlike a panel study, each cross-sectional sample in this study may include completely different individuals or organizations than those examined in earlier and later cross-sections. The National Medical Care Expenditures Survey (NMCES) and its follow-up, the Medical Expenditures Panel Survey (MEPS) are well-known health services surveys that employ a cross-sectional, time-series study design.

If patients are purposely included in an observational study by the researcher on the basis of their outcome status, this sampling approach results in a case-control design. In this design, the presence or absence of a disease (Y) is measured first to identify cases (individuals with the disease) and controls (individuals without the disease). Data on the exposure of interest (X) are collected retrospectively. Economists refer to a case-control study design as choice-based sampling, because they are often interested in an outcome variable that represents an individual's choice among numerous options. Case-control or choice-based sampling studies typically yield more biased estimates of a treatment effect than cross-sectional and longitudinal observational studies, which typically yield more biased estimates than randomized studies.[29]

Randomization is used widely in many epidemiologic and biostatistical studies to determine causal treatment effects because it is an effective strategy to minimize bias that can arise if unmeasured confounders (also referred to as omitted variables) are systematically different between treatment and control groups at baseline in the absence of randomization.[28] Randomized biostatistical and epidemiologic studies are most often clinical trials, which examine the effectiveness of treatments for patients who already have a disease.[28] In these studies, patients are randomly assigned to a treatment or comparison group (X) and then subsequent values of Y are observed. In many clinical trials, repeated observations of X and Y are taken over time on the patients randomized into the trial to track changes in outcomes.

A special type of randomized clinical trial—a field trial or community intervention trial—is designed to evaluate interventions intended to prevent disease onset in the first place. Few economic studies randomize participants, with the notable exceptions of the Negative Income Tax Experiment[5] and the RAND Health Insurance Experiment,[21] because it is infeasible to randomize people to educational tracks, jobs, or incomes. The types of bias that may arise in observational (nonrandomized) studies are considered in the next section.

Elements of an Equation

Suppose we would like to estimate the causal effect of a new heart valve on mortality in a nonrandomized sample of 5000 heart surgery patients treated by 50 physicians from 20 large academic medical centers who were followed for 1 year. A typical multiple regression equation would have the following form:

$$Y_i = \beta_o + X_i\beta_1 + Z_i\beta_2 + X_i{*}Z_i\beta_3 + \varepsilon_i \qquad (1)$$

Mortality is indicated by the term, Y, and is referred to as the outcome or response or dependent variable (Table 1, rows 5–9). The outcome is also referred to as the endogenous variable in the social sciences (eg, economics, sociology).

Allocation of people into the treatment group that receives the new heart valve or the control group that receives the standard heart valve is indicated by the term, X, and typically is referred to as the dose, intervention, explanatory or exogenous variable of interest, or treatment. The other, nontreatment factors related to mortality (Z) are referred to as independent, explanatory, exogenous, or predictor variables, or covariates. Covariates might include demographic or socioeconomic characteristics, characteristics of the patient's provider, characteristics of the patient's hospital, and characteristics of the local market or catchment

area. A covariate would be classified as a confounder if its inclusion changes the regression coefficient for the relationship between X and Y (ie, the treatment effect).[29] In other words, all confounders are covariates but not all covariates are confounders.

The relationship between the outcome (Y) and treatment (X) is typically referred to as the treatment effect and is captured by the β_1 term. The relationship between the outcome (Y) and the other covariates (Z) is typically referred to as the beta, parameter estimate or the regression coefficient and is captured in the β_2 term.[9,13] The intercept is indicated by β_0. The error term in the equation is indicated by the term (ϵ) and *i* indexes each individual in the sample, ranging from 1 to 5000.

Differential response within the treatment group by an important characteristic may also be important to evaluate. For example, it might be useful to determine if the same relationship between treatment and outcome holds for white and non-white patients. This differential response is called effect modification or an interaction.[29] In Equation 1 above, it is represented by the X*Z term. Alternatively, the subgroup of whites and the subgroup of non-whites could be analyzed separately. Such an analysis is called a partitioned model or a stratified model.

Regression Models

One issue on which health services researchers of all stripes tend to agree is the value of multiple or multivariate regression, which controls for the effect of observed confounders on the outcome.[13,25,29] There are several decisions about equation specification that any health services researcher must make, including how to characterize the outcome, whether to consider the treatment as possibly biased by systematic unobserved differences between the treatment and control groups, and how to address bias if it is present (Table 1, rows 10–17).

Mortality can be characterized or specified as a discrete (noncontinuous) outcome, such as a dichotomous outcome of dead/alive, or as a continuous outcome like number of days survived. The modeling of outcomes as two or more discrete groups is called categorical data analysis[19] or qualitative data analysis.[20] A dichotomous outcome can be estimated using generalized linear models (GLM), which is a flexible model with many possible functional forms that is commonly used to model mortality when only patient characteristics are used. GLM can also be used to estimate models based upon Gaussian, Poisson, and Gamma distributions.[24] Common statistical models of dichotomous outcomes include logistic, logit or probit models. Probit and logistic models have different distributions, particularly in the left-hand and right-hand tails, and are not strictly interchangeable. Models that characterize an outcome as three or more discrete groups with a natural ordering (eg, health status as excellent, good, fair, poor) and covariates of patient characteristics are referred to as an ordinal logistic[29] or ordered logit[20] regressions.

Models of three or more unordered outcome groups (eg, physical therapist, nurse, chiropractor) and covariates are characteristics of the patient are referred to as polytomous logistic[29] or multinomial logit regression.[20] Finally, models in which the outcome is characterized by two or more discrete choices with covariates that are characteristics of the choice or both the choice and the individual (eg, chooser) are referred to as conditional logit or conditional logistic or McFaddens logit regression.

If mortality is modeled as a continuous process, survival analysis is appropriate because of censoring in the date of death for people who were still alive at the end of the study period. Survival analysis is synonymous with failure-time analysis, Cox regression, hazard model, duration model, or event history analysis, which are parametric or semi-parametric methods. Survival analysis also includes nonparametric methods.

Unmeasured Confounding or Omitted Variables Bias

The next decision in an analysis is whether to be concerned that the treatment effect may be biased by unobserved variables that are correlated with both the treatment (X) and outcome (Y) in some way (Table 1, rows 18–20). Bias from differences in observed characteristics can be statistically controlled, but systematic differences in unobserved differences cannot. This general type of bias is called omitted variables bias, selection bias or unmeasured confounding. In observational studies which do not randomize patients, systematic differences between treatment and control groups cannot be ruled out and may arise because:

(1) The variable is known by patient but unmeasured by researcher because of an oversight in data collection or missing data (eg, severity of illness).
(2) The variable is known by patient but known and unmeasurable by researcher (eg, tastes and preferences for care).
(3) The variable is unknown by both patient and researcher.

Selection bias can also occur in randomized studies if randomization was not effective or crossover between treatment and control groups occurs longitudinally. Equation 1 could be modified to reflect omitted variables or unmeasured confounders (U):

$$Y_i = \beta_o + X_i\beta_1 + Z_i\beta_2 + X_i{*}Z_i\beta_3 + U_i\beta_4 + \varepsilon_{it} \qquad (2)$$

In the example of our observational study of heart surgery, selection bias may arise if there are systematic differences between the patients that received the new valve (treatment group) and the standard valve (control group). It is likely that physicians decide whether to use the new heart valve. It is possible that patients with the most severe heart disease seek treatment from the highest quality physicians, these patients receive the new valve, and (independent of valve received) these patients have higher mortality than less severely ill patients. This type of selection bias is also referred to as confounding by indication. The finding that black patients are less likely to receive the new valve and more likely to die whether they get the valve or not is an example of confounding by contraindication, which is another form of selection bias.[29]

Selection bias in typical health services research problems is a more general case of the original sample selection bias or incidental truncation or censoring model first developed by Heckman.[11] Economists use this term to refer to a specific situation in which treatment is observed on both treatment and control groups, but outcome is only observed on the treatment group. Incomplete observation of outcome in both groups may be caused by the sampling design or incomplete follow-up of one group. This phenomenon is also called selection bias in some literature.

Econometric and biostatistical approaches to this problem attempt to estimate selection bias in simultaneous equations, in which the outcome is modeled in an equation and the treatment choice is modeled in a second equation, which is also referred to as a multiple multivariate regression. Standard econometric methods include Heckman two-step estimation, two-stage least squares and instrumental variables.[2,7,10,11,14,22,23,33] Standard biostatistical methods include marginal structural models and inverse-probability-of-treatment weighting.[17,26,30] Propensity scores are another biostatistical method for estimating causal effects in observational studies.[28]

These methods are not discussed in detail here, because they are different methods with different terminology, not the same methods with different terminology. Each method requires different assumptions about the nature of the data and correlations between individuals and time. All these methods have been subject to criticism and are bound by important limitations. Unmeasured confounding remains an issue in health services research because no discipline has solved the problem and all biostatistical and econometric methods require some variant of the assumption that statistical corrections are not affected by unmeasured confounding. This clarification of terms should assist the interested reader in identifying methods across disciplines that tackle the same problems, but defers discussion of specific methods to the published literature.

Discussion

Statistical methods from economics, biostatistics, and epidemiology have contributed to the advancement of quantitative analysis of health services research problems. There is clearly significant variation in the terminology that these disciplines use for many of the same statistical concepts and methods. Some of these differences reflect different terms used to describe essentially identical concepts (eg, omitted variable and unmeasured confounder). In other cases, terms have slight but important differences in meaning across disciplines, leading to potential confusion in their use (eg, confounder and exogenous variable).

This paper will assist health services researchers and students new to the field in approaching the diverse methods literature. In addition, this "methods thesaurus" can assist in translation of methods by researchers already familiar with one discipline's literature but unfamiliar with methods literature from other disciplines. We limited our

study to situations where the underlying statistical method was the same but the terminology used to describe aspects of this method differed. Because alternative methodological approaches exist among disciplines, a direct comparison of these methods would also enhance interdisciplinary understanding. First, the assumptions behind the advanced methods from econometrics and biostatistics and the generalizability of results from these methods should be presented in a unified framework. For example, distributional and other assumptions, requirements for identification, implications for associational/causal interpretation, prediction, and model fit and diagnostics should be explained. In addition, settings in which each method is appropriate could be clarified.

Second, empirical work comparing the treatment effects that are estimated when applying different econometric and biostatistical methods to the same data should be conducted. Simulation work would enable health services researchers to understand differences in treatment effects that are caused by distributional assumptions, model specification, treatment effect estimation, and optimization algorithms. Further methods development might result from the limitations realized in these comparisons. This higher-level integration of methods would support the efforts of this paper to clarify some of the methodological isolation and confusion caused by semantic differences for similar statistical concepts in economics, biostatistics, and epidemiology.

With a common language for understanding methods that may be more similar than the terminology demonstrates, the field may become more interdisciplinary instead of isolated into disciplinary silos. With a consistent methodological terminology, or at least clear understanding of the varying nomenclature, researchers can broaden the number of methods that might be brought to bear on a particular empirical problem. Economists familiar with biostatistical terminology can explore the tradeoffs in estimating marginal structural models and other methods, instead of just the fixed effects, selection correction, and instrumental variables models typically employed. With terminological barriers removed, biostatisticians can more easily assess a variety of econometric methods developed specifically to address selection bias. Advances in methods for estimating causal effects can be pursued by economists and biostatisticians in concert.

Integration of methods and terminology will also contribute to efforts by health services researchers to develop unified or interdisciplinary theories to explain the behavior of individuals and organizations.[32] Interdisciplinary theories that incorporate constructs from various social sciences into models of causal relationships may also simplify the translation of theory into tractable methods. This paper represents a first step to align the efforts of the various disciplines engaged in health services research. Evolution of theory and methods in health services research will be hastened by a common understanding of terminology.

Acknowledgments

The authors thank the faculty in the University of Washington departments of Health Services, Biostatistics, and Epidemiology for input. Comments from Fred Zimmerman, Noel Weiss, Ed Perrin, Bryan Dowd, and Jeffrey Harman and four reviewers are also acknowledged.

References

1. **Baltagi BH.** Econometric Analysis of Panel Data. New York, NY: John Wiley and Sons; 1995.

2. **Basmann RL.** A generalized classical model of linear estimation of coefficients in a structural equation. Econometrica 1957;25:77–83.

3. **Brown H, Prescott R.** Applied Mixed Models in Medicine. New York, NY: John Wiley and Sons; 1999.

4. **Bryk AS, Raudenbush SW.** Hierarchical Linear Models. Newbury Park, CA: Sage Publications; 1992.

5. **Burtless G Hausman JA.** The effect of taxation on labor supply: evaluating the Gary Negative Income Tax Experiment. Journal of Political Economy 1978;86:1103–1130.

6. **Diggle PJ, Liang KY, Zeger S.** Analysis of Longitudinal Data. New York, NY: Oxford University Press; 2000.

7. **Dowd B, Feldman R, Cassou S, et al.** Health plan choice and the utilization of health care services. Rev Economics Statistics 1991;73:85–93.

8. **Greene W.** Econometric Theory, Second Edition. Englewood Cliffs, NJ: Prentice Hall; 1993.

9. **Gujarati DN.** Basic Econometrics, Second Edition. New York, NY: McGraw-Hill Publishing Company; 1988.

10. **Hausman JA, Taylor WE.** Panel data and unobservable individual effects. Econometrica 1981;49: 1377–1398.

11. **Heckman J.** Sample Selection Bias as a Specification Error. Econometrica 1979;47:153–161.

12. **Kennedy P.** A Guide to Econometrics, Fourth Edition. Cambridge, MA: MIT Press; 1998.

13. **Kmenta J.** Elements of Econometrics. New York, NY: Macmillan Publishing Company; 1986.

14. **Kyriazidou E.** Estimation of a panel data sample selection model. Econometrica 1997;65:1335–1364.

16. **Leyland AH, Goldstein H, eds.** Multilevel Modelling of Health Statistics. New York, NY: Wiley and Sons; 2001.

17. **Liang KY, Zeger SL.** Longitudinal data analysis using generalized linear models. Biometrika 1986;73:13–22.

18. **Lindsey JK. Models for Repeated Measurements.** New York, NY: Oxford University Press; 1999.

19. **Lloyd CJ.** Statistical Analysis of Categorical Data. New York, NY: Wiley; 1999.

20. **Maddala GS.** Limited-Dependent and Qualitative Variables in Econometrics. Cambridge: Cambridge University Press; 1983.

21. **Manning W, Newhouse JP, Duan N, et al.** Health insurance and the demand for medical care: evidence from a randomized experiment. American Economic Review 1987;77:251–277.

22. **McClellan M, Newhouse JP.** The marginal cost-effectiveness of medical technology: a panel instrumental-variables approach. Journal of Econometrics 1997;77:39–64.

23. **McClellan M, Newhouse JP, McNeil B.** Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality? JAMA 1997;272: 859–866.

24. **McCulloch CE, Searle SR.** Generalized, Linear, and Mixed Models. New York, NY: Wiley and Sons; 2001.

25. **Pindyck RS, Rubinfeld DL.** Econometric Models and Economic Forecasts, Third Edition. New York, NY: McGraw-Hill; 1991.

26. **Robins JM.** Causal inference from complex longitudinal data, in latent variable modeling and applications to causality. In: Berkane M, ed. Lecture Notes in Statistics (120). New York, NY: Springer Verlag; 1997:69–117.

27. **Robins JM, Greenland S.** The probability of causation under a Stochastic model for individual risk. Biometrics 1989;45:1125–1138.

28. **Rosenbaum PR, Rubin DB.** The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70:41–55.

29. **Rothman, Greenland.** Modern Epidemiology, Second Edition. New York, NY: Lippincott-Raven; 1998.

30. **Rubin D.** Estimating causal effects of treatments in randomized and nonrandomized studies. J Educational Psychology 1974;66:688–701.

31. **Searle SR, Casella G, McCulloch CE.** Variance Components. New York, NY: John Wiley and Sons; 1992.

32. **Shortell SM.** The need for interdisciplinary-based theory. Health Serv Res 1999;34(Part I):1–5.

33. **Vella F, Verbeek M.** Two-step estimation of panel data models with censored endogenous variables and selection bias. J Econometrics 1999;90:239–263.