Issues in Pooling Administrative Data for Economic Evaluation

Richard T. Meenan, PhD, MPH; Michael J. Goodman, PhD; Paul A. Fishman, PhD; Mark C. Hornbrook, PhD; Maureen C. O'Keeffe-Rosetti, MS; and Donald J. Bachman, MS

Managed care, in particular the health maintenance organization (HMO), now dominates US healthcare delivery, and economic evaluation is receiving increasing attention as a management tool that can be tailored to its perceived business needs. This encourages use of HMO administrative data as an efficient source of resource utilization and cost measures. Use of administrative data coincides with growing research interest in multisite analyses that increase external validity. The best alternative to a nationally representative data set is to pool administrative data from multiple sites within one database. However, pooling administrative data is problematic because HMO data sources reflect differences in systems of care, costing, and coding. This paper describes issues inherent in the pooling of HMO administrative cost data for use in multisite economic evaluations. We describe the attributes of administrative data that are relevant to costing and discuss their implications for multisite economic evaluations. We then briefly describe our experience with pooling multisite cost data, discuss lessons learned, and offer suggestions for researchers working with such data, followed by concluding comments. Multisite administrative data provide unique opportunities to conduct population-based clinical and economic research.

(Am J Manag Care 2002;8:45-53)

Economic evaluations are steadily gaining acceptance by policy makers responsible for healthcare resource allocation. They inform issues ranging from program implementation and formulary composition to treatment guidelines. Emphasizing population-based medicine and evidence-based practice, managed care, in particular the health maintenance organization (HMO), now dominates US healthcare delivery, and economic evaluations are increasingly tailored to its perceived needs, eg, pharmacoeconomic studies targeted toward formulary committees. This has encouraged use of HMO administrative data as an efficient source of resource utilization and cost measures. Although not designed for research, these data are population based and increasingly accessible through informatics.[1,2]

Many economic evaluations arise in connection with a randomized clinical trial at a single site or targeted at a specific subpopulation. Randomized clinical trials are internally valid and are still considered the "gold standard" by some government agencies, including the US Food and Drug Administration. However, their utilization and cost implications are relevant primarily to groups (including delivery systems) that resemble the study population. Use of

From the Center for Health Research, Northwest and Hawaii, Kaiser Permanente Northwest, Portland, OR (RTM, MCH, MCO-R, DJB); HealthPartners Research Foundation, Minneapolis, MN (MJG); and the Center for Health Studies, Group Health Cooperative of Puget Sound, Seattle, WA (PAF).

This study was funded in part by Cooperative Agreements 17-C-90433/9-01, "Development of Global Risk Assessment Models (GRAM) to Support Risk-Adjusted Community Rating," and 17-C-90864/9-01, "Reforming Medicare Risk Payment Through Competitive Market Forces," from the Centers for Medicare and Medicaid Services, Baltimore, MD; grant 33039 from the Robert Wood Johnson Foundation, Princeton, NJ; Kaiser Permanente, Oakland, CA; and Post-Doctoral Fellowships T32 HS00069 (RJM) and F32 HS00072 (MJG) from the Agency for Healthcare Policy and Research (now the Agency for Healthcare Research and Quality), Rockville, MD.

The statements contained in this report are solely those of the authors and do not necessarily reflect the views or policies of the Health Care Financing Administration, the Agency for Healthcare Research and Quality (formerly the Agency for Healthcare Policy and Research), the Robert Wood Johnson Foundation, Princeton, NJ; Kaiser Permanente, Oakland, CA; HealthPartners Research Foundation, or Group Health Cooperative of Puget Sound. The authors assume full responsibility for the accuracy and completeness of the information contained in this report.

Address correspondence to: Richard T. Meenan, PhD, MPH, Investigator, Center for Health Research, Northwest and Hawaii, Kaiser Permanente Northwest, 3800 N Interstate Avenue, Portland, OR 97227. E-mail: richard.meenan@kpchr.org.

administrative data coincides with increasing interest in multisite studies that can enhance external validity. This interest is reflected in the creation of the HMO Research Network and its extension, the National Cancer Institute–supported Cancer Research Network. These consortia are intended to facilitate clinical and economic healthcare studies among millions of patients in 10 to 12 predominantly nonprofit HMOs that collectively reflect national diversity of age, sex, income, education, culture, and location. Through their larger populations, multisite studies are more likely to detect rare but important events and address ethical concerns about the inclusion of appropriate subpopulations. Ideally, a nationally representative data set would exist from which all studies would extract relevant standardized elements.[3] The next best alternative is to pool existing administrative data in one multi-HMO database, which requires common data elements with standard variable definitions.

Yet, pooling data across HMOs is problematic: care systems have different service-specific "production functions,"[4] costing systems,[5] and coding incentives.[6] This paper highlights issues surrounding pooling of administrative data, particularly for cost estimation in the context of multi-HMO economic evaluations. We begin with an overview of study types and costing methods. We describe the attributes of HMO administrative data that are relevant to costing and discuss their implications for multisite economic evaluations. We then briefly describe our experience with pooling multisite cost data, discuss lessons learned, and offer suggestions for researchers working with such data, followed by concluding comments.

··· BACKGROUND ···

The spectrum of economic evaluations can be categorized into 2 broad types: trial-based studies and modeling studies.[4] In trial-based studies, clinical and economic data are collected concurrently. Modeling studies use data from many sources, including, eg, efficacy data, surveys, epidemiologic studies, and administrative databases. These types are not distinct—"pragmatic" effectiveness trials have been conducted,[7] and even trial-based studies often use external cost data. Data requirements for any study are similar: treatment effectiveness (or efficacy), quality adjustments, resource quantities, and unit costs.[4] We necessarily emphasize resource quantities and unit costs, with less emphasis on treatment effectiveness, which matters if clinical improvements are measured by fewer strokes, for example. We focus on estimating direct healthcare costs, defined as the dollar-valued change in healthcare resource use from an intervention, its adverse effects, and its current and future health consequences,[5] which can include ostensibly unrelated future costs.[8] Costs include laboratory tests, drugs, supplies, healthcare personnel, and facilities and are normally measurable to some degree using administrative data. Direct nonhealthcare costs (eg, patient travel), productivity losses, and quality-of-life information normally are not available. Our focus on direct costs also means that we are not concerned with correctly evaluating societal intervention costs[9] but with estimating direct intervention-related costs to the health plan or provider—health plan "liability."

··· PRODUCTION AND COST FUNCTIONS ···

The analyst's initial task is to identify the intervention's "production function," including its implementation process, and expected treatment and health status effects, given disease progression and epidemiologic findings.[5] The analyst then outlines intervention resources and applies a simple conceptual model: resource cost is the quantity of the resource consumed during a relevant period multiplied by its unit cost, which ideally is the resource's opportunity cost, its value in the best alternative use.

Organizations can differ in production functions of care. Even controlling for patient case-mix, organizations might treat the same condition differently, leading to variance in treatment cost.[4] In a trial with balanced randomization, this variance will not affect the relative risk reduction from the therapy but could affect the absolute size of the economic effect. Drummond[4] concludes that subject randomization would ideally occur across settings as well as treatments.

Administrative data systems also highlight other organizational differences. For example, even with similar production functions, overhead allocation methods can differ. Alternatively, some HMOs might capitate physician payment, whereas others might use a relative value scale-based fee schedule; this might cause similar procedures to differ in reported cost. Another source of variation is the effect of financial incentives on coding within data systems,

eg, fee-for-service (FFS) reimbursement vs capitation. Coding system variation implies that within and across HMOs, representations of their production and cost functions in the data differ from their "true" functions—the plan actually does one thing (at one intensity) but reported data suggest that it does something else (perhaps at a different intensity).

Hence, administrative data indicate organizational differences in utilization and cost in 3 ways: real variance in the approach to care ("care system" variation),[4] costing method variation, and coding system variation, the latter 2 of which can either mask or amplify the first. Differences in systems of care, costing, and coding manifest themselves in various ways given the type of study, eg, single site or multisite.

··· COST ANALYSES ···

Cost Estimation

Two cost estimation methods are commonly used with administrative data. In "microcosting," patient-level data are collected on the exact number and type of resources consumed. For example, a hospital stay can be separated into its various components (eg, physician time, nursing time, tests, supplies, and room), which are then assigned unit costs and the total summed. In fact, microcosting directly applies the conceptual model described earlier.

Significant intervention-related events can also be "gross costed." For example, acute care hospitalizations can be assigned a standard diagnosis-related group (DRG) or per diem cost. Ambulatory care visits can be assigned a fee schedule value based on the resource-based relative value scale (RBRVS). Gross costing assumes that an event uses the same basic input mix regardless of location or provider specialty, jointly measuring resources and costs. A cost-to-charge ratio applied to billing or department-level expenditures is another common gross costing technique. The usefulness of such a ratio depends in part on the quality of the source data (eg, Medicare Cost Report or internal cost accounting) as well as its purpose (eg, estimating average vs marginal procedure cost).[5] Shwartz et al[10] found that cost-to-charge ratios were better at estimating average DRG cost than patient-level cost.

The choice of microcosting or gross costing depends partly on the evaluation context. The estimated total cost of a hospitalization resulting from (or avoided by) an intervention will probably be assigned a different value than the sum of its inputs, some of which might be integral to the intervention itself.[11] Hence, economic evaluations often use microcosting to value intervention resources and gross costing to value adverse effects and future consequences.

Trial-Based vs Observational Studies

Trial-based analyses are most likely to use microcosting because resource quantities can be tracked prospectively with the clinical data.[12] However, because resource data are rarely collected prospectively, utilization is often reviewed retrospectively through medical chart reviews or administrative data.[13] Unit costs can be determined externally or by applying an internal average cost. Gross costing is then applied to downstream effects. Administrative data are potentially useful here, especially if the system provides standardized costs such as RBRVS-based fees or DRGs representing the cost of future hospitalizations, for example.

Microcosting also applies to observational studies such as cost-effectiveness or "pragmatic" trials[4,7] and cost-of-illness and cost consequence studies, although gross costing is more common. Such analyses track temporal population effects to provide "real-world" information about a condition or intervention. They are useful when event rates are low, effectiveness is important, or a long time horizon is needed.[14] One type of observational study, a retrospective cohort analysis, compares utilization or costs of patients who did and did not receive a given therapy. This design requires large comprehensive data sets to facilitate statistical control of possible confounders such as age, sex, and comorbidities. Selection bias can remain problematic if not addressed during study design because those receiving the therapy might differ from those who do not in ways that cannot be controlled for statistically.[15,16] The Cancer Research Network HMOs are currently collaborating in a retrospective cohort, case-control analysis of temporal changes in utilization and costs among continuing, former, and never smokers.

Trial-based and observational studies can be performed at one site or multiple sites. A single-site study eliminates potential differences in care systems, costing methods, and coding systems, although a single-site observational study might still exhibit internal bias because potential sources are not directly controlled. Multisite studies introduce all forms of variance. Note, however, that within a multisite trial-based study, the trial protocol constrains care system variation, at least in terms of implementation, because resource use can be specified and standard protocols established. Tracking of downstream events is more likely to differ because the pattern of care

reflected in the data is not limited to the patient's biological response but is also affected by system characteristics manifested through the available data.

··· ADMINISTRATIVE DATA IN HMOs ···

Administrative data come in 2 basic types: claims and encounter. HMOs use 3 types of claims systems: inpatient (UB-92), professional and noninpatient (HCFA-1500), and outpatient prescriptions.

Claims Systems

Both FFS and capitated HMOs use claims data to track utilization. Claims systems generally use nationally standardized record layouts and coding systems for key variables: *International Classification of Diseases, Ninth Revision, Clinical Modification,* diagnostic coding; *Current Procedural Terminology 4th Edition* procedure coding; and National Drug Code for prescriptions. Both claims and encounter systems often include internal "homegrown" codes for plan-specific pharmaceutical compounds, repackaging, or similar circumstances with no obvious National Drug Code number. It can be difficult to distinguish between submitted, allowed, denied, paid, and duplicate claims. This is especially true in capitated HMOs, in which encounters do not create physical funds transfer. Claims-based HMOs often use the actual paid amount as a cost proxy; however, this usually excludes overhead and might not approximate true resource opportunity cost if structural factors influence payment (eg, market power, withholds, manufacturer rebates, and ex-post case-mix adjustments). Fixed-fee capitation contracts can complicate patient-level costing because expended funds cannot be assigned to individuals. Furthermore, linking specific services across providers to underlying events, such as stays, visits, and episodes, remains problematic.[17]

HMO claims systems may cover all delivered services, as most independent practice associations or networks do, or only outside claims and referrals, as most staff- or group-model HMO systems do. Claims usually include all business lines, including point-of-service and self-funded plans. Some HMOs have in-house processors, and others use indemnity insurers as processors (eg, pharmaceutical benefits managers). In-house processors typically have less rigorous review algorithms than sophisticated indemnity systems. Members can withhold submissions until they meet their deductible or can seek non-HMO care if they are not covered or have high out-of-pocket charges.[2]

Encounter Systems

Encounter systems track delivered services rather than billings or payments. Conceptually, data should reflect all delivered services regardless of payer; hence, the data structure should reflect actual care more closely; however, tracking might be limited to visits rather than procedures. (Point-of-service and HMO network use might be tracked separately, which can induce double counting.) These systems are used by staff- or group-model HMOs for owned facilities and by FFS providers.

Encounter systems contain many data sets, even within one HMO: inpatient, appointments, ambulatory, drugs, imaging, surgery, laboratory tests, home health and hospice, skilled nursing, durable medical equipment, and outside claims. Typically, each set has separate coding conventions, which can be internally inconsistent except for the linking variable, eg, health record number. Some HMOs have cost management information systems (CMIS) that track individual utilization and costs by linking events to average unit costs for each departmental output; however, the temporal (eg, monthly) cost measure can differ across systems.

Administrative data can be augmented by clinical information systems,[18] which are relatively new at most HMOs. These systems include computerized patient records, disease registries, and outcome tracking or case management systems.[1] They are usually free text or structured with restricted values. Updates from on-line systems to end-user files are commonly delayed, which complicates their use in costing.[17] Many HMO medical groups currently resist clinical information systems because of high financial and time costs. For an independent practice association, a clinical information system can take on "public good" characteristics that lower private investment incentives because its benefits are not completely internalized by the investing HMO but are shared by other plans with which it contracts.

··· DATA POOLING FOR MULTI-HMO ECONOMIC EVALUATIONS ···

Multi-HMO studies require pooled data in a usable research format; however, administrative data are not generated for research purposes but are organized to fulfill other functions, such as finance or to meet regulatory requirements. This complicates pooling because there are no HMO data standards. Homegrown codes are common, and data from network and internal providers are usually

inconsistent. Operating systems might be incompatible, and data might not be in end-user format.[1] These characteristics might then contradict the central assumption of pooled cost data—controlling for case-mix, no systematic relationship of plan to production cost exists.

In multi-HMO studies, intervention components must be determined, as well as their level of abstraction. One could track utilization elements and apply site-specific costs[19]; "natural" resource units would be tracked first, with detailed results provided for all settings. This assumes a trial format in which all activity is accurately tracked. Alternatively, standardized unit costs could be assigned to site-specific utilization. Another common suggestion is to use DRGs and RBRVS-based fees to standardize utilization and costs for an event such as a stroke-related hospitalization. Standardization assumes that a given event stimulates similar responses across systems. As practice guidelines and "best practices" become increasingly commonplace, this assumption might become more valid, reducing care system variation. (Note, however, that plans and medical groups in Portland, OR, have had problems with instituting even internal care management programs.[20])

A potential limitation of DRGs, especially for HMOs, is their basis in FFS experience. If relatively resource-conserving HMO policies are reflected within an average condition-related hospital stay, then FFS-based DRGs might overstate the HMO hospitalization cost (assuming equal length of stay). This can have implications for absolute cost differences, with consequences for resource allocation. A similar problem arises in costing professional services and ambulatory care visits using the FFS-based RBRVS. In fact, the relation of the RBRVS system to efficiency even within FFS has been questioned.[21] Note that relative hospital costs differ from relative total episode costs, which can still seem lower in HMOs because of fewer hospitalizations (or coding differences that influence how an episode is viewed).

This problem arises in part from well-understood differences between FFS providers and HMOs. FFS providers are paid for procedures performed, which creates strong incentives to code them. In fact, the Medicare Prospective Payment System provides an extra reimbursement incentive to code diagnoses and procedures. Financial incentives might lead to "code creep," reminiscent of DRG creep.[6] Conversely, capitated HMOs have limited (financial) incentives to code accurately because they receive the capitated amount regardless of the number of services provided. In general, organizations with owned facilities or that contract with a single group have relatively less incentive to code activity.

In the extreme, the experience of an FFS group and an HMO using the same care protocol for a condition (and assuming case-mix adjustment) might still appear differently in the data because of dissimilar coding. Limiting data sources to HMOs might not eliminate coding differences because many newer HMOs, such as independent practice associations, are former FFS indemnity insurers that have largely retained previous coding conventions or data capture procedures. Hence, apparent treatment differences may simply represent coding incentives. These differences are exacerbated by organizational mergers that increase the heterogeneity of management structures and data systems. Carved-out services such as mental health often lead to missing data. Also, standardization of merged data systems might be toward the "least common denominator"—the simplest, cheapest, and, by implication, least informative system (Michael J. Goodman, PhD, unpublished observation, 1996).

Some HMOs, especially those with encounter systems, have inadequate information to implement DRGs appropriately. Even if they use DRGs, some HMOs use a proprietary "all-payer" version rather than the standard Medicare version. Data quality is especially important because HMOs often collect diagnoses from patient encounters with staff or capitated providers, which might be less complete or detailed than the diagnoses from an FFS-type provider. Some multi-HMO studies even create a dependent variable (expenditure) by applying a single relative value scale to site-specific activity differences, which assumes that relative service costs are the same across sites. Differences in patient-level cost estimation, such as health plan liability, annual per-capita expenses, and annual per-capita paid claims, have implications for economic evaluation, especially regarding perspective and cost measurement. For example, health plan *liability* is relevant to an HMO-perspective evaluation, for which administrative data are most useful. Conversely, annual per capita *expense* might include external patient expense, requiring survey or other data.

··· EXPERIENCE WITH POOLING COST DATA ···

We recently constructed a large multi-HMO administrative data set as part of the Global Risk

Adjustment: Payment Estimation and Simulation project. This data set contains complete eligibility and diagnosis profiles and cost and utilization estimates for roughly 2 million individuals for 1995-1996; this represents the total enrollee population of 5 HMOs located across the United States and part of a sixth. Although the specific research purpose of this effort is risk-adjustment modeling, our experience in pooling cost data is relevant to multi-HMO economic evaluations in general.

Three HMOs were claims based and 3 combined claims and utilization systems. All inpatient data for 5 HMOs were claims based, and the outpatient data combined claims and utilization systems. Three utilization systems were linked to a CMIS, but 3 required construction of complex ambulatory costing algorithms. (We also developed an inpatient costing algorithm for one HMO.) We validated the data by first separating them into categories of inpatient, outpatient, long-term care, long-term visit, pharmacy, radiology, and durable medical equipment; this required standard category definitions, eg, inpatient utilization required an overnight stay. We then created 5 cost "buckets"—inpatient, outpatient, pharmacy, other, and total expense—into which all costs were assigned. These buckets were examined for site-specific differences across age and sex groups, which were explored in depth with appropriate site data experts.

The cost buckets were invaluable in identifying data "anomalies." Simply creating a total expense variable would have made it problematic to uncover overhead allocation differences, for example. The cost buckets also facilitated communication with the sites by targeting our information requests, reducing the perception of being on a "fishing expedition." Furthermore, they allowed us to create categories that could be connected to each site's general ledger. This is an important form of validation because the general ledger can help researchers uncover areas ("dead-ended costs") where utilization data are missing, such as for capitated contracts or point-of-service transactions. At one site we discovered through the general ledger that we were missing utilization and cost data that were actually stored apart from other site data.

··· LESSONS ···

Integrated cost and utilization data offer significant benefits. Including multiple HMOs increases the cultural and economic diversity of study populations and medical groups, which is especially important in regions with few minorities, such as the Pacific Northwest. Larger populations facilitate study of less common medical phenomena and of complex models of utilization, cost, and other outcomes. However, standardizing multi-HMO cost data is complex, time consuming, iterative, and expensive. Claims and utilization systems typically use inconsistent definitions of events, and because such systems were not designed for research, data quality controls are often poor. Data documentation is frequently limited to an experienced analyst's memory, and even when documentation exists, it is usually not automated. In addition, documentation updates for new or reinterpreted site-specific codes are sporadic, which complicates the development of standard codes because proper linkages across sites depend on clearly understanding how codes are used and interpreted.

Development of standardized cost files is an interactive process, requiring regular communication with site data experts. Although institutional constraints sometimes necessitate relying on a single site liaison, this can be dangerous; support of that site might evaporate if the person is reassigned or leaves the organization. The risk is compounded, of course, if the liaison is inexperienced and faces his or her own learning curve about the data. Economic research projects should have a senior management champion at each site to support access to technical staff and one or more technical champions, preferably in both information technology and data analysis. Information technology programmers access computing and data storage resources required to download large data files, and data analysts understand the general structure and common uses of the systems. However, researchers cannot simply rely on the knowledge and judgment of the site data experts, accepting the cost data as given. Information technology staff often do not understand the institutional context of healthcare and therefore might not easily identify implausible data patterns. Analytic staff might not understand the technical requirements of downloading and transmitting large data files. Furthermore, neither group may be familiar with the needs of research. This implies that the research team should develop an explicit and comprehensive site data request to minimize potential misunderstandings over what data are needed and in what format. Whenever possible, a project should also develop and update contingency plans for inevitable personnel changes.

An important part of cultivating champions or creating "buy-in" is establishing expectations of the site early in the project (eg, will site programmers build and transmit the data files to the research site, or will research programmers retrieve raw site data and build the files themselves?). These expectations should be reflected in an appropriate site budget and timeline that accounts for the site's internal constraints. Another contributor to buy-in is the fact that cost data integration efforts have positive feedback effects on the participating HMOs. Increased scrutiny of their data by experienced outside analysts can identify problems that casual users might not discover; what the research team learns about the site data should be shared with the source site. As an example, Kaiser Permanente Northwest, Portland, is currently implementing a Decision Support System to support cost management. The system is intended primarily for administrative use, but over time it will become the primary cost source for economic studies at the Center for Health Research. To that end, Center for Health Research economists and analysts will be working with system developers during the next 2 to 3 years to test and evaluate the system. This process should lead not only to a better understanding of the system's advantages and limitations for research but to general system refinements that benefit administrative users as well by identifying gaps in data capture or information flows. Data documentation also tends to improve over time as the data are more intensively analyzed.

An interim step toward standardized data collection is the development of data warehouses that contain common data formats and conventions. When HMOs create individual warehouses, data quality is increased because analysts, managers, and clinicians all use the data. Within-site differences and duplicate records across systems can be addressed, although sometimes at the cost of omitted variables, which can be helpful by simplifying the end-user file or problematic if an important variable is ignored. Outside data requests can identify gaps in HMO data warehousing efforts and can also provide external technical design assistance. Note, however, that data contained in warehouses are usually considered proprietary, and research access may be granted only on a project-by-project basis.

Our experience in using multi-HMO data to create a pseudo-"warehouse" has been that the separate and duplicate data systems resulting from mixed models and recent mergers greatly complicate data collection. Barriers to data standardization are lowered within plans with unified information systems that use consistent data elements across product lines and models. In particular, data availability and quality are greatly enhanced by computerized patient record systems that create systematic access to automated, on-line, provider-generated clinical information. These systems link clinical information across settings and help elucidate disease etiology as well as the most effective and efficient care processes. They can facilitate standardized data collection and coding practices by involving physicians directly in coding. In general, HMO investment in informatics will depend on organizational structure, degree of integration, clinical information system improvement, and financial performance.[1] HMOs that delay installation of comprehensive patient-level cost management systems will continue to present barriers to cost standardization.

External forces also drive HMOs toward greater standardization. The Balanced Budget Act of 1997 mandates that all HMOs receiving Medicare payments submit inpatient claims data on their Medicare Risk and Cost enrollees in a standard UB-92 format and that ambulatory claims data be submitted in a standard HCFA-1500 format. Hence, Medicare HMOs must combine data across various delivery systems into a single data stream to the Centers for Medicare and Medicaid Services (formerly the Health Care Financing Administration), requiring significant system upgrades for many HMOs that are not accustomed to billing systems for internal services. Although not affecting all HMOs, the Act is considered a precursor to future regulations.

The growing importance of confidentiality issues surrounding individual patient data and proprietary organizational data further complicates matters. Researchers who create these large databases are responsible for controlling access to and use of the data. Inappropriate disclosure of data, even inadvertently, can only damage efforts to increase legitimate research access to them. HMO data analysts—typical site experts—are used to working in environments with relatively open access to medical data and might not be sensitive to confidentiality concerns. Researchers must not simply react to governmental initiatives to protect privacy but must be proactive and creative (eg, secured Web data transfers) in establishing electronic data procedures that recognize the balance between the needs of public domain research and individual confidentiality. In basic terms, this means developing and nurturing a research culture of respect for electronic patient and organizational data.

··· CONCLUSIONS ···

Standardized measures such as DRGs and the RBRVS that capture real resource use within HMOs are feasible to produce and critical to generalizable results of economic evaluations, both within the HMO sector and beyond to FFS. Yet, our development efforts should remain sensitive to the implications of standardization. It is not enough to simply recognize that the "way of doing things" within HMOs differs from FFS. We must clearly understand the production and cost functions of our organizations so that standardized HMO-based estimates approximate the true opportunity cost of these resources as accurately as possible within measurement constraints. Of course, not all interventions will be exposed to the same degree of bias—part of the challenge is to determine where the bias matters and where it does not. This requires a clear understanding of project sites, including data system structure, and the overall organizational and care delivery structure.[17] The **Table** summarizes our major suggestions for researchers working with multisite cost data. Cancer Research Network economists are currently wrestling with these issues as they collaborate in the development of standardized cost estimates to facilitate HMO-based economic evaluations of cancer-related treatments. In general, economic evaluations based on pooled HMO data should explain data collection methods, site-specific data structure, and costing methods including devised algorithms and standardization techniques in a technical report or appendix.[22]

We end with an important distinction. It was not our objective in creating a pooled data set for use in risk-adjustment research to produce HMO-specific cost estimates or models but to limit the effect of site as much as possible. However, in economic evaluation generalizability—the applicability of the results of a given study to other populations or subpopulations—may be paramount. Multilevel statistical models (also called hierarchical linear or random coefficient models) have seldom been used in economic evaluations to date but might be especially appropriate in accounting for site-specific effects on individual patient cost variation within multi-HMO studies.[23] Such models adopt a framework of individual units, such as patients, that are nested within larger groups, such as hospitals or HMOs, and assume that overall variation is a function of both the individual and the group. Researchers have also proposed various methods of deriving country-specific cost or cost-effectiveness estimates in the context of multinational clinical trials.[24,25] The applicability of these methods to multi-HMO studies in which site-specific results are of most interest to decision makers is a logical goal for future research.

Multi-HMO administrative data offer unique opportunities to conduct population-based economic research. Their current complexity and lack of standardization will eventually be superseded by more consistent and complete organization and should not deter creative analysts.

Table. Suggestions for Researchers Using Multi-HMO Cost and Utilization Data

- Cultivate project champions at each site among senior management, information technology staff, and data analytic staff
- Develop contingency plans for inevitable personnel changes
- Manage expectations of site's role early in the project
- Provide an appropriate site budget and timeline, given the site's resource constraints
- Generate an explicit and comprehensive data request to minimize misunderstandings over what data are needed (or available)
- Recognize that data documentation will often be dated, incomplete, or nonexistent
- Cultivate buy-in by creating an information feedback loop, sharing freely with the site users what the research team learns from their use of the site data
- Create a research culture of respect for the confidentiality issues surrounding electronic patient and organizational data
- Create cost "buckets" to facilitate data validation, particularly for comparison with the general ledger, which can highlight missing utilization and cost data
- Don't underestimate the amount of time, personnel, and computing resources required to create a standardized cost and utilization file

··· REFERENCES ···

1. Hornbrook MC, Goodman MJ, Fishman PA, et al. Health-based payment and computerized patient record systems. *Eff Clin Pract* 1998;1:66-72.

2. Mullahy J, Manning W. Statistical issues in cost-effectiveness analyses. In: Sloan FA, ed. *Valuing Health Care: Costs, Benefits, and Effectiveness of Pharmaceuticals and Other Medical Technologies.* New York, NY: Cambridge University Press; 1996:149-184.

3. Durham ML. Partnerships for research among managed care organizations. *Health Aff* 1998;17(1):111-122.

4. Drummond MF. Experimental versus observational data in the economic evaluation of pharmaceuticals. *Med Decis Making* 1998;18:S12-S18.

5. Luce BR, Manning WG, Siegel JE, et al. Estimating costs in cost-effectiveness analysis. In: Gold MR, Siegel

JE, Russell LB, Weinstein MC, eds. *Cost-Effectiveness in Health and Medicine.* New York, NY: Oxford University Press; 1996:176-213.

6. Iezzoni LI. Assessing quality using administrative data. *Ann Intern Med* 1997;127:666-674.

7. Oster G, Borok G, Menzin J, et al. A randomized trial to assess effectiveness and cost in general practice: Rationale and design of the Cholesterol Reduction Intervention Study (CRIS). *Control Clin Trials* 1995;16:3-16.

8. Meltzer D. Accounting for future costs in medical cost-effectiveness analysis. *J Health Econ* 1997;16:33-64.

9. Russell LB, Siegel JE, Daniels N, et al. Cost-effectiveness analysis as a guide to resource allocation in health: Roles and limitations. In: Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. *Cost-Effectiveness in Health and Medicine.* New York, NY: Oxford University Press; 1996:3-24.

10. Shwartz M, Young DW, Siegrist R. The ratio of costs to charges: How good a basis for estimating costs? *Inquiry* 1995/96;32:476-481.

11. Meenan RT, Stevens VJ, Hornbrook MC, et al. Cost-effectiveness of a hospital-based smoking-cessation intervention. *Med Care* 1998;36:670-678.

12. Hall JE, LaChance PA. *An Automated Participant Tracking System Using Generalized System 1032 Data Structures: Proceedings of 10th Annual System 1032/1022 Users Conference.* Boston, MA: System 1032/1022 Users Group; 1989.

13. Schulman KA, Buxton M, Glick H, et al. Results of the economic evaluation of the FIRST study: A multinational prospective economic evaluation. *Int J Technol Assess Health Care* 1996;12:698-713.

14. Drummond MF, Davies L. Economic analysis alongside clinical trials: Revisiting the methodological issues. *Int J Technol Assess Health Care* 1991;7:561-573.

15. Friedman GD. *Primer of Epidemiology.* 4th ed. New York, NY: McGraw-Hill; 1994.

16. Torrance GW, Siegel JE, Luce BR. Framing and designing the cost-effectiveness analysis. In: Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. *Cost-Effectiveness in Health and Medicine.* New York, NY: Oxford University Press; 1996:54-81.

17. Hornbrook MC, Goodman MJ, Fishman PA, et al. Building health plan databases to risk adjust outcomes and payments. *Int J Qual Health Care* 1998;10:531-538.

18. Krall M, Mysinger T, Pearson J, et al. Integrated clinical database in a health maintenance organization. *Medinfo* 1995;8:313-316.

19. Luce BR, Elixhauser A. Estimating costs in the economic evaluation of medical technologies. *Int J Technol Assess Health Care* 1990;6:57-75.

20. Thompson R, Brown R. *Medicare Beneficiaries and HMOs: A Case Study of the Portland Market.* Prepared by Mathematica Policy Research, Inc. for the Henry J. Kaiser Family Foundation; 1998.

21. Baumgartner JR. Medicare physician-payment reform and the resource-based relative value scale: A re-creation of efficient market prices? *Am Econ Rev* 1992;82:1027-1030.

22. Siegel JE, Weinstein MC, Torrance GW. Reporting cost-effectiveness studies and results. In: Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. *Cost-Effectiveness in Health and Medicine.* New York, NY: Oxford University Press; 1996:276-303.

23. Carey K. A multilevel modelling approach to analysis of patient costs under managed care. *Health Econ* 2000;9:435-446.

24. Willke RJ, Glick HA, Polsky D, Schulman K. Estimating country-specific cost-effectiveness from multinational clinical trials. *Health Econ* 1998;7:481-493.

25. Schulman K, Burke J, Drummond M, et al. Resource costing for multinational neurologic clinical trials: Methods and results. *Health Econ* 1998;7:629-638.