Ask an Economist

How do I address outliers in cost data?

Most researchers will eventually encounter “outlier” cost observations. At HERC, we have found hospital stays costing millions of dollars. In most analytical samples, that observation would be viewed as expensive, and many would classify it as an outlier in the distribution of costs.

However, when many people say outliers, they not only use the term to refer to high costs, but also costs that are likely to be erroneous. The question of data accuracy is more difficult because it is a question of truth when the truth isn’t known. Triangulation, i.e. using two or more sources of data to check for accuracy, is one method for finding the truth about the data.

One source of triangulation is comparing MCA cost data to HERC cost data. Researchers can also compare the MCA cost to the median MCA cost for that product or service, making sure to identify homogenous services where the median is meaningful.

Anesthesia and pharmaceuticals are two domains that may contain inaccurate data. In both of these domains, total costs are computed by multiplying units by unit costs. If either of these is incorrect, it will affect the total cost. An example of this type of error occur is when the unit cost is based on one measure (a liter of IV fluids) and a particular patient’s utilization is measured in a different unit (mL). For pharmaceuticals, researchers can easily compute the median MCA cost for the VA drug class and use that as the triangulation reference. The supply cost of drugs can also be found on the Federal Supply Schedule.

(continued on page 2)
Ask an Economist

How do I address outliers in cost data? (continued from page 1)

For inpatient medical/surgical care, researchers can use the length of stay and the Diagnosis Related Group (DRG) weight to estimate if the cost is reasonable. Each hospital stay is assigned a 3 digit code for DRG, and each DRG has a weight that is the basis of Medicare reimbursement. More complex DRGs are assigned a higher weight, resulting in higher reimbursement. DRG weights are updated each year and the value placed on the web site of the Centers of Medicare and Medicaid Services. If the average cost per day of inpatient stay per unit of DRG weight is very high, greater than $15,000-$18,000 per day per DRG weight, the researcher should look more closely at the observation. Remember: it is important to examine low cost outliers in addition to high cost outliers. Both ends of the distribution can include incorrect data and create problems for the statistical analysis.

How do I identify provider location in Fee data?

It is difficult to determine the geographic location where purchased care was provided for some claims. The vendor location may not be the provider location because the fee basis vendor is not necessarily the provider. The SAS fee vendor file has the zip code of the place where the remittance was sent, which is not necessarily the place where care was delivered. Payment may be sent to headquarters of a multi-hospital chain, to the vendor of a national contract, or to a billing service.

For hospitals, researchers can use the 6-digit provider identifier and link it to the street address in the Hospital Compare dataset. This identifier is called MDCAREID in the fee inpatient claims data, PROVIDID in the Hospital Compare dataset, and is commonly known as the OSCAR identifier in Medicare documentation. In the fee files, not all inpatient records have a value for MDCAREID. It is only entered if payment is based on the standard Medicare reimbursement method (called the pricer). This means that MDCAREID is blank in fee records if the hospital was paid on a contract. Researchers can fill these blanks by analyzing non-blank entries. Look for claims from that same vendor processed by the same VA on the same year that did include a hospital identifier. This populates most entries, but an assumption must be made when there is more than one MDCAREID per VEN13N per VA per year. One possible assumption is to use the MDCAREID that is used most frequently by that site for that vendor ID.

For non-hospital providers, there is a national provider identifier (NPI) in the fee data. Most fee claims lack this identifier. There is a huge national data set of NPI and provider data that could potentially be linked, but this is a very difficult analysis to undertake, as there are so many different providers. It is often difficult to distinguish hospitals from the physician groups that practice at that hospital.

It is important to know that there are multiple records for each vendor identifier. The SAS vendor files have multiple records with the same vendor ID (VEN13N), and no key field to link vendor records to utilization. In most cases, researchers can merge by year, station, and VEN13N. However, a record must have both a claim and a vendor for this method to be effective. At HERC, when we had a claim without a vendor, we looked in the prior year(s) to identify the vendor. The Fee Payment Processing System (FPPS) may have information on the actual provider, but because there is no way to assign “read only” access, only purchased care staff can access to those data. Each claim must be looked up one at a time, so these data are probably not very useful for most health services studies.

Because the SQL files lack the full set of diagnosis and procedure codes and may be insufficient for certain projects, we discuss the SAS files. For more information on Fee Basis data in both SAS and SQL, please visit the Fee Basis Guidebook on the HERC website.
Cost Effectiveness Analysis Course Happening Now

Course Overview
Interested researchers can still register for any or all of the remaining cost-effectiveness analysis course sessions. This introductory course, hosted by HERC, covers cost-effectiveness and budget impact analyses and VA economic data sources. Cost-effectiveness analysis is a method to determine if interventions provide sufficient health benefit to justify their cost. Budget impact analyses complement cost-effectiveness analyses, providing decision makers with information on the effect of adopting cost-effective interventions. This HERC course will provide an overview these and other types of decision analysis and how to conduct these analyses with VA data.

The course is designed for researchers who would like an introduction to methods of cost-effectiveness analysis and budget impact analysis as applied to health services and medicine.

The course takes place through conference calls with web-based slide presentations. Please visit the HERC website to register.

Course Schedule

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>02/24</td>
<td>Modeling in Medical Decision Analysis</td>
</tr>
<tr>
<td></td>
<td>Jeremy Goldhaber-Fiebert, PhD</td>
</tr>
<tr>
<td>03/02</td>
<td>Estimating Transition Probabilities for a Model</td>
</tr>
<tr>
<td></td>
<td>Risha Gidwani, DrPH</td>
</tr>
<tr>
<td>03/09</td>
<td>Evidence Synthesis to Derive Model Transition Probabilities (Part I – Systematic Literature Review)</td>
</tr>
<tr>
<td></td>
<td>Risha Gidwani, DrPH</td>
</tr>
<tr>
<td>03/23</td>
<td>Evidence Synthesis to Derive Model Transition Probabilities (Part II – Quantitative Pooling)</td>
</tr>
<tr>
<td></td>
<td>Risha Gidwani, DrPH</td>
</tr>
<tr>
<td>03/30</td>
<td>Sensitivity Analyses</td>
</tr>
<tr>
<td></td>
<td>Risha Gidwani, DrPH</td>
</tr>
<tr>
<td>04/06</td>
<td>Budget Impact Analysis</td>
</tr>
<tr>
<td></td>
<td>Todd Wagner, PhD and Wei Yu, PhD</td>
</tr>
<tr>
<td>04/13</td>
<td>How can Cost-Effectiveness Analysis be Made More Relevant to US Healthcare?</td>
</tr>
<tr>
<td></td>
<td>Paul Barnett, PhD</td>
</tr>
</tbody>
</table>

Understanding Inpatient Costs

Although hospital discharge data is frequently used in health services research, these datasets often exclude the professional fee, which can account for 18-26% of inpatient costs. A new paper in the journal Medical Care by Peterson et al. presents a method for adjusting for professional fees when conducting a cost analysis using hospital discharge data. In their paper “Professional Fee Ratios for US Hospital Discharge Data,” Peterson and team calculated the professional fee costs for each DRG. They used the Market Scan data, which reports paid insurance claims and encounters for public and private payers. They report the ratio of total costs to DRG reimbursement separately for private payers and Medicaid.

“Professional Fee Ratios for US Hospital Discharge Data,” by Peterson C, Likang X, Florence C, Grosse S, and Annest J, can be found in the October 2015 issue of Medical Care (volume 53, issue 10, pages 840-849).

Error in FY 2015 Treating Specialty NDE

The Managerial Cost Accounting Office (MCAO) recently became aware of a department mapping problem in the Inpatient National Data Extract (NDE) impacting estimates of surgery cost. This affects the treating specialty and discharge files for fiscal year 2015. Total cost data were not affected.

The problem has now been fixed, and the corrected NDEs are available at the Corporate Data Warehouse. For questions or more information, visit the MCA website and log a WonderDesk ticket at MCA Help Desk.
HERC Cyber Seminars

Each hourly session begins at 11:00am Pacific (2:00pm Eastern), unless otherwise noted.

Register:  http://www.hsr.d.research.va.gov/Cyberseminars
Schedule & archives:  http://www.herc.research.va.gov/include/page.asp?id=courses-seminars

Cyber Seminars
The Health Economics Cyber Seminars feature presentations on a variety of health economics and health services topics.

February 17, 2016  Re-examining Facility-Level Effects on Diabetes Care Quality for Veterans
Adam Wilk, PhD,  
Assistant Professor of Health Policy and Management  
Emory University’s Rollins School of Public Health

Little is known about how organizational and structural factors mediate patterns of chronic disease care quality. Identifying specific facility-level determinants of diabetes care quality has proven difficult, in part because of cross-sectional study designs susceptible to bias due to reverse causality. We re-analyze the effects of facility's average diabetes quality care measure performance and available resource on individual physicians' diabetes care quality measure performance in the Department of Veterans Affairs.

Target audience: Researchers interested in primary care or diabetes care quality, how unit and facility-level inputs can improve quality, and methods for identifying facility-level effects.

Interested in presenting in the HERC Health Economics Cyber Seminar Series? Contact HERC Economist Jean Yoon (Jean.Yoon@va.gov) for more information.