

# TRANSFORMATIONS

## From Health Claims Data to Business Intelligence

June 2004

### Creating a Reliable Dataset for Bottom Line **IMPACT**

If your organization is at-risk for the cost of health care services, you know that accurate data analysis and reporting is essential when setting appropriate payment schedules. An error in calculating cost per member per month could mean losses of hundred of thousands or even millions of dollars.

Can you trust the data on which such critical reporting is based? Yes—but only when it's been very carefully prepared and edited.

Raw claims data present a multitude of opportunities for drawing mistaken conclusions about claims experience. These apparently insignificant glitches can have significant bottom-line impact. Typical examples:

- A self-insured employer had diagnosis data that contained non-standard periods within the diagnosis codes. Impact: Material understatement of the number of patients with certain chronic diseases, leading to incorrect conclusions about the viability of targeted disease management programs.
- A pharmacy benefit manager combined retiree and active employees for data analysis. When a contract ran a monthly loss of over \$1 million, it was impossible to identify the cause. Finding and correcting the data flaw let the client pinpoint the primary cause of the deficit, and adjust premiums to reflect a dramatic increase in the proportion of retirees.
- Common systematic errors in gender and date of birth lead to incorrect patient counts, resulting in over- or underestimation of per capita costs by age and gender.
- Some data are missing from the publicly available national Medicaid drug utilization dataset, and some existing values are off by orders of magnitude. As a result, analyses and cost projections based on the unedited national dataset will be wrong.

As these examples make clear, when critical decisions or settlements are riding on the results of your analyses, you must develop a health claims dataset that you can trust. DGA Partners has found these five steps to be key to success in establishing such a dataset:

1. *Have all required datasets available.* For comprehensive claims analyses, it is ideal to have all medical claims available: inpatient, outpatient, and professional services claims; all pharmaceutical claims; and both current and historical data on enrollment, membership, and eligibility.
2. *Validate data against other internal reports.* Check cost and service unit totals for your claims dataset against other internal data sources,

*(Continued on back)*

### DGA Profile:

George Chalissery,  
*Director, Business  
Intelligence Solutions*



George Chalissery puts a lot of effort into cleaning—of client data, that is.

“My job is to get our clients reliable business intelligence. Because health care transaction data is rarely of high quality, proper data preparation and editing is crucial when there's a lot riding on analysis of that data.”

In eight years with DGA, George has worked with insurers, policy groups and health care providers. He has developed payment systems, helped manage financial risk, created outcome measurement programs, and supported disease management initiatives. George has extensive experience with Medicare and Medicaid datasets, and with reimbursement systems such as DRGs, APCs and RBRVS fee schedules. He uses all of what he's learned in developing clean, integrated datasets from multiple sources.

Before bringing his talents to DGA, George worked at Lewin-VHI/Value Health Sciences, was a software engineer at IBM and was a faculty member at DEC. He has a masters degree in Computer Engineering from Case Western Reserve University. ■

such as accounting data. Significant discrepancies are a call to investigate the relevant data fields very carefully. DGA has used this kind of cross check, for example, to find major errors in the field that flags Medicaid/Medicare dual eligibility.

3. *Separate data and metrics by line of business.* For financial and clinical analyses, data should be broken out by payer class (Commercial, Medicare, Medicaid), as well as more detailed analyses within the commercial category (indemnity, HMO, PPO). A breakout by major payer class may be important for clinical analyses, as well.
4. *Enhance data with clinical classifications.* To make data more meaningful from a clinical perspective, add DRG codes. Severity scores can be assigned to each inpatient case using a commercially available patient severity system such as DxCGs. These enhancements will also permit case mix analysis and severity adjustments to utilization and cost figures.
5. *“Scrub” and prepare data.* The first task in preparing data is to run an extensive series of edit checks, which often include the following:
  - Load only relevant claims into the dataset. Exclude claims for non-members, for the wrong time period, or for non-relevant benefits (e.g. dental claims).
  - Validate all codes (e.g. ICD-9-CM diagnosis and procedure codes, National Drug Codes). New codes may need to be mapped to old ones to allow trend analysis. Since the status of codes changes constantly, be certain you are using the most recent reference database.
  - Categorize claims by major category—inpatient, outpatient, etc.—based on place of service and type of provider.
  - Screen pharmaceutical claims for inconsistent quantity coding. What one provider may count as one vial, another may count as 10 one—ml. doses. Develop “patches” to correct any problems.

The second phase in data preparation is transforming the transaction-based (adjudicated) claims data to a dataset that can be used for reporting and analysis. This requires:

- Compressing interim claims into a single claim for the entire admission or service (prevents overcounting of inpatient cases and understating of average LOS).
- Resolving claims adjustments, reversals, denials and duplicates. Typically, a re-adjudicated claim will generate three records—initial entry, entry to zero that out, and a new entry. Unless the first two are deleted, average cost will be understated.

Similar difficulties occur when the computer system sends the same claim to accounting twice. Correcting these kinds of errors can provide an order of magnitude return on investment when the data is being used to support a negotiated settlement.

- Ensuring that member months are counted based on when each member enrolled and disenrolled, not on a snapshot of how many people were enrolled at one point in time.
- Making certain that dependents are included in counting member months.

Creating a trustworthy dataset is a complex and resource-intensive process that requires guidance from industry, clinical, financial and actuarial experts. It’s a process that’s vitally important to your bottom line, and it’s worth the investment. ■

## DGA News:

*We’ve Moved:* You won’t have trouble finding us—we’re across the street from our old offices. The new suite provides quiet and functional work space and room for growth.

*Fully HIPAA compliant:* Our new offices feature server room and office security, allowing DGA to enter into business associate agreements required by HIPAA.

## About DGA Partners:

For 10 years, DGA Partners has been providing crucial business intelligence derived from health claims and related datasets. We serve risk-assuming entities, self-insured employers, insurers and payers, pharmaceutical companies and consulting firms. Our sophisticated data management and reporting services also provide the basis for our actuarial and clinical analytics services. DGA’s health care management consulting group provides strategy, business planning and related services to health care providers.

Let’s talk about how we can make your data work for your organization.

**DGAPARTNERS**  
Healthcare Strategy • Finance • Data

Two Bala Plaza, St. 301  
Bala Cynwyd, PA 19004  
610.667.8782

[www.DGAparters.com](http://www.DGAparters.com)  
[info@dgapartners.com](mailto:info@dgapartners.com)