

# Pharma Marketing News™

[www.pharmamarketingnews.com](http://www.pharmamarketingnews.com)

July 2003

Vol. 2, No. 7

Published by  
VirSci Corp.

[www.virsci.com](http://www.virsci.com)

## Reprint

### Patient-Level Data: A New Level of Analysis

by Jean-Patrick Tsang, PhD and MBA (INSEAD)



Back in the early 1990's, prescribing information was gathered at the *pharmacy outlet level* and the pharmaceutical market analyst's nightmare was traveling scripts. The data were captured where the script was filled as opposed to where the prescription was written. So, if Dr. John Smith, who is based in Chicago, writes a script for a patient who fills the script in New York, sales rep Jane Doe, who calls on Dr. John Smith, gets no credit. In short, that muddies the link between promotion and results, puzzling marketers as to what really happened.

*Physician-level data* was the next level and soon established itself as the data source of choice for a whole array of analyses. Today, physician-level data is so deeply ingrained in our analytical reflexes we can hardly conceive of doing virtually any physician analysis (e.g., targeting, segmentation, promotion response, etc.) without having recourse to IMS Xponent or NDC Source Prescriber. We are hooked!

#### Patient-Level Data—the New Kid on the Block

The new kid on the block is *patient-level data*. Unlike physician-level data that describes the wholesale activity of Dr. John Smith as writing say 50 scripts of Lipitor, patient-level data zooms in on the patients of Dr. John Smith and specifies patient by patient who actually gets Lipitor.

Because of privacy provisions of HIPAA, the actual identity of the patient is not revealed, only an encrypted id and general profile

information such as age and sex. It turns out this is sufficient for most of our sales and marketing purposes, since a patient is associated with one encrypted id that is used all along.

This means while we do not know the name of the patient (actually we could not care less), we can follow the prescriptions filled by any patient over time, or "longitudinally" in patient-level data parlance. This is terrific news for the sales and marketing department since it is now endowed with additional levers to close in on Dr. John Smith's prescription habits and modify the latter's behavior for the better!

Patient-level data means even more data to analyze for an industry that is already replete with data.

Physician-level data, which is currently the mainstay of the industry, pertains to about three quarters of a million prescribers. Patients, by contrast, can easily run as high as a hundred million even under the most conservative of assumptions. In short, patient-level data means Terabytes instead of Gigabytes.

#### Patient-level data is not a passing fad

The interest in patient-level data has picked up significant steam lately. June 2003 saw the first conference entirely dedicated to anonymous patient-level data—which I had the honor to chair—with a turnout of 130 attendees, a real crowd compared with the 40-50 headcount such events normally fetch. The feedback of the attendees was so enthusiastic the producer of the event, the Center for Business Intelligence (CBI), responded by putting up an "encore" event as early as

November 2003, six months ahead of the June 2004 initial schedule for the follow-up event.

Either to gain a competitive leg up or for fear of being left behind in the dust, many companies have already started to use patient-level data for sales and marketing ends, or have initiated the shopping process.

While the marketer marvels at the findings patient-level data unlocks, others in the organization closer to the operational side may be less sanguine considering the heaps of data tapes to analyze or the gaping holes such purchases leave in the budget. It is probably this vague sense of uneasiness that emboldens detractors to portray patient-level data as inept for serious analyses in a bid to save us both effort and money.

But make no mistake. Patient-level data is not only here to stay but also to blossom. Very soon, it will pervade virtually all strains of analyses including the bread-and-butter physician segmentation, targeting, sales force sizing and structure, performance evaluation, incentive compensation, and the like. Indeed, there are already signs patient-level data is reshaping the very way we conceive of sales and marketing in the pharmaceutical industry.

### What Are the Benefits?

What's the fuss about knowing the prescription activity at the patient level?

For starters, patient-level data gives pharmaceutical marketers the ability to discern between new brand and therapy starts, to identify therapy switches and identify which drugs are favored to the detriment of drug X and which drugs are shunned in favor of drug Y, to track down changes in dosing levels, and the like. This is something that is just not possible with physician-level data.

Parenthetically, patient-level data provides an accurate portrayal of new starts by looking back over a certain period of time at the prescription-filling activity of the patient. If you believe this is something you already had with physician-level data, you've have been had like everyone else in the industry. The truth of the matter is although "NRx" has been construed to denote a new prescription, it represents nothing more than a new pad of paper, which may just as well be a refill.

Patient-level data also shines an insightful light on the practice of the physician such as the size of the practice—the number of patients treated—currently vs. at some specific point in time, demographics of the patients, profile of patients for whom the physician ventures a brand new therapy, considers changes in dosing, brand name, or mechanism of action, or sticks to status quo, etc.

### Choose the Right Data for the Job

"Give me your best wine!" commanded a guest in a very fine restaurant, to which the sommelier responded apologetically, but pointedly, "Pardon me, monsieur, you must first tell me what you are having for dinner".

In the patient-level data marketplace, there is no best data source. It all depends on the analyses that need to be run and the business questions that need to be answered.

Patient data is collected at different points of interaction with the healthcare system (e.g., pharmacy/switch, payer, PBM, employer group, provider), each spawning a different strand of data: prescription data, claims data, discharge data, stay data, and so on.

Discharge data, for instance, is more detailed and verbose than claims data. In a situation where a patient is diagnosed with three ICD-9's and just one is attended to, the claims data will report only the relevant ICD-9, the one that triggers the payment, while the discharge data will give a full account of the patient's condition as established by the institution, regardless of payment.

Employer data tends to be longitudinally more robust than payer data. This is because when an employer changes plan – and this happens more frequently than what one expects – the plan loses track of the patient but the employer does not.

Claims data capture prescriptions and medical resources utilized by the patient as long as they are paid by the plan or PBM in question. This means claims data have wide blind spots as soon as payment is made in cash, by Medicaid or by Medicare. The flipside is claims data can ascertain that when a patient drops out of a therapy this is not due to lack of coverage but for medical reasons. Indeed, the insurer has intimate knowledge of coverage status of each individual patient.

Just like the sommelier in a restaurant, a market researcher must have a very clear idea what analysis is to be undertaken prior to ordering the data.

For analyses focusing on Oxycontin, Viagra, or lifestyle drugs, the data must capture cash transactions since many people try to dodge the system by paying in cash. For persistency analyses, the analyst must be sure patients do not drop off because they do not have coverage, underscoring the importance of eligibility. Payer or PBM data are very relevant in that regard.

When patients need to be tracked over protracted lengths of time, it is advised to deploy employer group data as opposed to plan data. If sales force questions are at issue (sizing, alignment, compensation and the like), make sure the database has great coverage and is not geographically skewed. This is because when the data is chopped into 500 territories, you don't want to end up with the "hole of the Swiss cheese" for some territories!

### **The Patient-Level Data Uncertainty Principle**

The ideal patient-level database has breadth, depth, and is timely.

One of the vexing problems with patient-level data, at least today, is that you cannot have simultaneously in one database both breadth and depth, let alone timeliness. Either you have a partial view of a large population or a full view of a sliver of the population, but not both. This is a bit like the famous Heisenberg Uncertainty Principle in quantum physics where it is impossible to know both the position and momentum of a subatomic particle at the same time.

Currently, patient-level data give you a partial view of a large population, whereas patient-centric data give you a more complete view of a thinner slice of the population.

The most comprehensive of patient-level databases captures a whopping 2 billion scripts out of the 3.2 billion scripts that are written annually. Pharmacy activity is solely what this database is about.

By contrast, patient-centric data not only document the pharmacy activity but also hospitalizations, outpatient visits, lab tests, etc. However, these databases tend to include much smaller populations, in the range of

millions of people, an iota compared with the total U.S. population of 290 million people.

When you get a partial view of a large population using patient-level data, you have access to the identity of the prescribing physician, but not when you have a more complete view of a thinner slice of the population using patient-centric data.

This either-or principle seems to surface in other instances as well. Hospital inpatient databases (provider data) provide very detailed information on the patient as long as the patient is confined to the hospital but loses track of the patient as soon as he/she is discharged. Claims data straddles both institution and retail and tracks the patient before hospitalization, after hospitalization, but gets blurry during the hospital stay. Indeed, such databases do not track drugs taken by the individual patient while the latter is in the hospital.

Since no one database quite cuts it, this means projection is extremely critical. Patient-centric data where a lot is known about a sliver of the population is a great laboratory to observe patterns and unearth relationships between prescription activity and other parameters such as prevalence and incidence rates of disease states, co-morbidity and poly-pharmacy, geography, physician profile, etc. Those relationships can then in turn be applied to project out the pharmacy-level data to the whole population.

This is why I believe data vendors such as Verispan are headed in the right direction, precisely because they strive to combine medical data with the mainstay pharmacy data. I also feel that companies like PharMetrics, which currently captures patient-centric data on over 50 million lives, will play an increasingly important role as time unfolds.

### **Is Patient-Level Data a Solution in Need of a Problem?**

Sales pitches of patient-level data vendors sound very much alike. They systematically sing ad nauseam the praises of persistency, compliance, switches, and dosing analyses, numbing even the most attentive of the listeners to the penetrating freshness of the analyses. What the industry needs at this point is a handful of vivid and powerful examples of what patient-level data can do for the industry beyond persistency and compliance.

Our research at Bayser, fueled and fortified by client projects, indicates patient-level data can be successfully deployed to identify influencers, map out spheres of influence, and build socio-grams of prescribers for virtually any therapeutic area. See the references at the end of this article for details.

Patient-level data can help us establish the true neighborhood of an institution by tracking where the patients come from for each ICD-9, procedure or diagnosis.. This is great news for analysts charged with measuring the magnitude of hospital-retail spillover since they no longer have to define the neighborhood of a hospital as arbitrary circles around that hospital.

Patient-level data allows us to classify physicians as switchers, loyalists, referring physicians, referred-to physicians, and so on. This suggests one can do better than target a high prescriber. Indeed, target instead the “system” of physicians that is made up of the high-prescriber, upstream physicians that refer patients to the high prescriber (to keep the flow of patients coming in), and downstream switcher physicians visited by patients of the high prescriber (to avert undesirable switching to competitive drugs). The situation can be likened to filling a bathtub. Opening the faucet is a good idea. Better still: make sure the tap regulating the inflow of water from the street is fully open and the drain is closed.

Patient-level data is also used to measure the impact of Direct-to-Consumer (DTC) opt-in campaigns. Indeed, it is possible to match the names of patients participating in the program against the patient-level database to ascertain if consumption of the drug is larger than that of a comparable population that otherwise was not exposed to the promotional campaign.

This process typically takes place in two steps to ensure that one never gets to see both the name and the medical records of a patient at the same time, which would be in direct violation of HIPAA. In the first step the names from the DTC campaign are matched against names and encrypted id's of patients from the patient-level database. This results in a list of encrypted id's of patients that participated in the campaign, at which point the names are stripped away. The second step runs the test/control measurement off that list to ascertain the differential impact of the DTC campaign if any.

Another area of great promise is predictive modeling for superior segmentation, targeting, messaging, and identification of opportunity. In essence, using patient-level data, it is possible to identify which physicians show the greatest promise in terms of prescription growth, speed of uptake, resistance to decline (due to competitive launches), referral generation, etc. You can even determine which physicians have a significant number of undiagnosed and/or untreated patients in their patient base!

In the sales and marketing, the might be to identify future avid prescribers of a particular drug. Companies like Dendrite are currently exploring the relevance of hidden state Markov models—a very popular predictive modeling technique along with neural nets, regression, and Bayes' rule to mention just these three—to predict which physicians are most likely to prescribe a given drug or switch from drug X to drug Y based on the sequence of scripts the physician wrote in the past 12 months.

For those who doubt that any of these techniques will ever bear fruit or simply feel they are just too far out to consider seriously, just realize that patient-level data needs to work only once to pay off.

This pearl of wisdom comes from Gerald Gallivan, a 35-year pharmaceuticals veteran, currently a Director of Pharmetrics, reflecting on the challenges he faced while managing the then newly launched Prilosec.

Sales of the drug were sluggish until it dawned on management that Prilosec was used as second line of therapy, prompting them to shift positioning of the drug to acknowledge and reinforce current practice, after which sales soared.

Of course, patient-level data did not exist at the time and this is precisely the point. Had patient-level data – which would have brought prescribing habits in sharp focus immediately – been available at the time, the situation could have been rectified much earlier, resulting in hundred of millions if not billions of dollars of additional sales.

### Obstacles to Overcome

Whether or not patient-level data is widely adopted by the industry depends on how well practitioners and vendors steer the industry in the right direction to maximize proper usage and deployment of patient-level data.

There are currently three major obstacles that need to be overcome.

First, patient-level data is foggy for many, and that translates into wholesale inertia since informed decisions cannot be made. The shot in the arm the industry needs here is quality and objective education of the ins and outs of patient-level data.

Second, the proof is in the pudding! It does not matter how intriguing or promising an emerging practice is, at the end of the day only success stories impart impetus for action! What is of relevance here is more success stories on the one hand and more detailed accounts of success stories (in terms of problem, solution, and payoff) on the other!

Third, data vendors need to be more candid and forthcoming regarding their data. It is not by hiding the data under a cloak of mystical secrecy and invoking HIPAA or other insulting pretense along the way, or by denying access to manufacturers or consultants that work on their behalf, that data vendors will gain the trust and respect upon which any healthy business relationship is predicated. A “sweep under the rug” mentality will stifle if not bury a budding industry. This is a sin we unequivocally need to eradicate now! “

**Pharma Marketing News**

---

## Subscribe to Pharma Marketing News

*Pharma Marketing News* (PMN) is the FREE monthly e-newsletter of the **Pharma Marketing Network**. Highlights are delivered to subscribers by e-mail. The full pdf version is available at [www.pharma-mkting.com](http://www.pharma-mkting.com). You also have the option to participate in 2-way, peer-to-peer e-mail discussions with your pharmaceutical marketing colleagues through the **PHARMA-MKTING list**. By subscribing to PMN, you agree to receive e-mail messages through this service as well as newsletter highlights. We do not sell or share your personal information with third parties.

RED=REQUIRED INFORMATION

FIRST NAME: \_\_\_\_\_ LAST NAME: \_\_\_\_\_

JOB TITLE: \_\_\_\_\_ COMPANY: \_\_\_\_\_

COUNTRY: \_\_\_\_\_

E-MAIL ADDRESS: \_\_\_\_\_

E-MAIL FORMAT PREFERENCE: \_\_\_ HTML \_\_\_ TEXT \_\_\_ NONE

I also wish to join the PHARMA-MKTING list to participate in 2-way peer-to-peer discussions.

SUBSCRIPTION OPTION: \_\_\_ FULL DISCUSSION \_\_\_ “LURK & LEARN” (DIGEST MODE)

BLACK=OPTIONAL INFORMATION (For our internal use only. We don't sell or rent mailing lists/labels)

ADDRESS: \_\_\_\_\_

CITY: \_\_\_\_\_ STATE/PROV: \_\_\_\_\_ POSTAL CODE: \_\_\_\_\_

PHONE: \_\_\_\_\_ FAX: \_\_\_\_\_

Mail or fax to: VirSci, PO Box 760, Newtown, PA 18940, 215-504-5739 (Fax)

---