Big data:
Raising the table stakes for master data management
Big data: The new normal

There’s so much media attention devoted to “big data” these days that it can be difficult to separate reality from hype.

At IMS, we view big data as very much a reality; it’s been part of our business for decades. However, that’s not to say that there aren’t new challenges arising everyday from novel data sources and advances in information technology.

We believe that big data—and the technology ecosystem surrounding it—is not a passing fad, but is becoming the new normal. If pharmaceutical companies aim to treat information as a corporate asset, they must have an information management strategy that can accommodate big data and the “big analytics” that make sense of it.

Here we’ll share pragmatic advice on how pharmacos can prepare their organizations to take advantage of big data in their business analytics and, as appropriate, in their business operations.

Read “Super Size Me: Big Data, Big Challenges, and Big Rewards” as it appeared in Pharmaceutical Executive magazine.
What’s the big deal?

As the marketplace evolves, pharmaceutical companies will need to avail themselves of non-traditional data sources in order to:

- **Prove the value of therapies.** This goes hand in hand with supporting pay-for-performance contracts with payers and recognizing patients as healthcare decision makers.

- **Track adoption.** With the rise of specialty distribution channels, traditional measures of sales are no longer sufficient for understanding influence on prescribing.

- **Comply with regulations.** New insights are required to support REMS programs and will be needed to comply with any track and trace mandates.

- **Reach consumers.** As patients assume greater responsibility for the cost of their healthcare and as transparency grows in the delivery and quality of care, companies have a pressing need to understand and address consumers.

- **Dissect the complexities of healthcare networks.** The marketplace is now very tangled, with a variety of stakeholders influencing and making healthcare decisions.

Companies that lag behind their peers and other stakeholders in making use of big data risk being at a disadvantage at the research bench, in the market and at the negotiation table.
“Digital exhaust” from the healthcare system

With the rapid increase in the use of Electronic Medical Records (EMRs), the creation of Accountable Care Organizations (ACOs), and the rise of health information exchanges, there’s interest in allowing information systems from across all points in the delivery of care to communicate with one another.

Transactions recorded in these information systems can be captured as “digital exhaust” and used to guide decision making. New data sources of interest include:

- **Electronic Medical Records**—for a detailed clinical picture within the site of care and an overall view of a patient’s treatment pathway.
- **Social Media**—a wellspring of rich insights into the patient or prescriber perspective.
- **Real-World Evidence**—a wealth of information on treatment practices, patient persistence and compliance, and health outcomes.
- **Personalized Medicine**—a means not only to understand what works and what doesn’t, but also to predict what will work, person by person.
- **Track-and-Trace Systems**—a record of product movement through the distribution system, with a data point being generated every time a package “hits a node” on the distribution system.
New questions for better profiling

Big data will provide answers to the questions in the three right-hand columns—questions that previously could only be answered with primary research but that paint a well-rounded picture of your audience.

<table>
<thead>
<tr>
<th>BASICS</th>
<th>NUMBERS</th>
<th>OPINIONS</th>
<th>BEHAVIOR</th>
<th>INFLUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is their state License #?</td>
<td>What do they write?</td>
<td>What is the theme of their practice?</td>
<td>Why did they become a physician?</td>
<td>What specialists do they refer to?</td>
</tr>
<tr>
<td>Do they practice in more than one state?</td>
<td>What is their TRx?</td>
<td>Do they use compliance programs?</td>
<td>How much do they enjoy it?</td>
<td>Why do they refer to that specialist?</td>
</tr>
<tr>
<td>Where do they practice?</td>
<td>What is their Rx?</td>
<td>Are they concerned about cost of therapy?</td>
<td>Where do they vacation?</td>
<td>Where are these specialists located?</td>
</tr>
<tr>
<td>Do they have more than one office?</td>
<td>What is their NRx?</td>
<td>Are they big on generics?</td>
<td>How often do they take vacations?</td>
<td>Do they work in an ER or hospital?</td>
</tr>
<tr>
<td>Where are their offices?</td>
<td>What is their competitive TRx profile?</td>
<td>Do they focus on side effect profiles?</td>
<td>What type of clothes do they wear?</td>
<td>Who are their local key opinion leaders?</td>
</tr>
<tr>
<td>Do they see reps?</td>
<td>Why do they write the way they write?</td>
<td>Do they actively check efficacy?</td>
<td>What type of watch do they wear?</td>
<td>Who are their global key opinion leaders?</td>
</tr>
<tr>
<td>How long is an average visit?</td>
<td>What disease areas do they treat?</td>
<td>Do they report adverse reactions?</td>
<td>Do they wear glasses or contacts?</td>
<td>Who do they influence?</td>
</tr>
<tr>
<td>Do they use other channels to get information?</td>
<td>What is their adoption pathway – early adopter, laggard?</td>
<td>Do they sample?</td>
<td>What type of glasses do they wear?</td>
<td>What hospitals are in their geographic area?</td>
</tr>
</tbody>
</table>
**Big data and 21st century technology**

At the same time that big data are becoming available and are of interest to commercial organizations, the technological landscape is changing in a way that puts big data within our grasp. Today we have the benefit of:

- Cloud-based computing, making solutions scalable and affordable through a utility-type model
- Social media applications for business
- Enterprise-wide solutions that do away with disparate islands of data
- Connectivity that permits real-time data collection and use
- Analytical engines that make information actionable

All of these are enabling technologies are making it possible to deal with big data across its four dimensions: volume, velocity, variety and value.

View video on “Mastering the Big Data Challenge”
A changing information framework

For years, the type of information needed for commercial success, and the standards around how it was processed, remained quite static.

Now, however, we must take a fresh approach to the information lifecycle because:

- Users need data to be available on demand
- Different users require different information in their own, unique formats (there is no one-size-fits-all approach)
- New sources of information are becoming available all the time
Managing data as a corporate asset

Before the promise of big data can be realized, companies must have the technology platform and processes in place to accept the data and make use of it. This requires having a Master Data Management (MDM) Strategy that is part of a broader Information Management Strategy.

Master data are those information assets that are common to multiple functions in the enterprise. For the commercial organization, this could include data on: providers, provider affiliations, plans, products, and employees. Master Data must reflect “one version of the truth,” serving as the authoritative source, eliminating discrepancies and conflicts.

Read “Preparing for “Big Data:” Laying the Foundation with Master Data Management“ as it appeared in PharmaVoice magazine.
Step one: Sourcing challenges

Because there are no syndicated sources for many of the various types of new transactional data, companies must contend with the fact that most of it is:

- Unstructured
- Not integrated
- Unprojected
- Not necessarily complete
- Subject to special handling (to comply with patient privacy regulations, for example)

These means that not all data should be used to operate the business; only data that are complete and accurate should be applied to managing campaigns, incentive compensation, targeting and speaker programs, for example. Big data that are less than perfect should be reserved for analytical purposes such as spotting trends, identifying behaviors and categorizing segments.

Also, given the fact that data volumes are growing exponentially, a company's infrastructure must be scalable, and new source data must be subjected to common processes and control mechanisms.
Step two: Loading and cleansing data

The process of converting source files to a standardized format and of scrubbing them of defects, has, until now, been rather straightforward. However, it will become more complex with the introduction of unstructured data of unknown quality. Existing procedures will likely need to be modified in order to:

- Parse out operation-worthy data from data that can be used only for analysis
- Create operational and analytical views—even for some customer and product data
- Identify quality concerns and define the process for making corrections
- Compensate for the additional demand for quality assurance and the increased workload for data stewards
Step three: Matching, merging and enriching

At this stage, cleaned records are matched and duplicates eliminated to create a “golden record,” or the purest, most accurate record possible.

Due to the fact that some big data are unstructured, of suspect quality and incomplete, new solutions must be able to:

- Add some data to the master file and only cross reference certain other data to it
- De-identify patient data
- Determine the value of striving for the golden record
- Designate big data for analytical purposes and only permit their use for operational purposes when certain standards are met

The best procedures will allow for flexibility, as we anticipate that big data sources being in a state of flux for the first few years.
Step four: Stewardship

At this stage, all gray areas related to matching are reconciled, and quality is monitored and measured against enterprise standards and definitions.

While traditionally, stewardship and data governance has resided with IT, the task is better managed by those who use the data and who understand the industry and the healthcare environment.

With the influx of big data, those within the business who are charged with data stewardship will need to:

- Balance the resources committed to ensuring the completeness of big data with its directional use (at some point, it must be deemed “good enough”)
- Align the goals of stewardship with the business strategy
- Ensure that there are separate processes for data that are directional vs. those that need to be exact
- Create the proper links between analytical and operational data
Step five: Reporting and auditing

At this point in the MDM framework, operational and quality reports are produced, cross-references are maintained and a history is kept of all changes.

To address the unique properties of big data, it is especially important that:

- Measures be put in place to proactively determine the accuracy and completeness of information, by source
- Audit trails and business rules reflect compliance standards
- Where it is possible to influence the quality of data at the source, processes and alerts should be installed to prevent/correct errors

After the quality of big data is monitored for some time and companies have a better sense of what to expect and what is possible, business rules can be refined to improve the quality and utility of the data over time.
Step six: Publishing

Once again, this step is made more complex in the era of big data because the data values and dimensions will always be changing. Companies will need to:

- Provide users with views that accommodate both analytical and operational uses
- Ensure that users understand how the data were derived and what caveats exist for its use (not all data will be “ready for prime time” and some should be provided only to certain groups for special purposes)
Step seven: Consuming

Because the trend is moving toward more customized views and reports, having a solid MDM strategy and framework in place will become critically important to ensuring that the right data are used by the right people for the right things. Increasingly, and especially with the advent of big data, companies must:

- Be able to support a diverse set of data consumers using a variety of tools
- Track how people are querying the data to stay appraised of their needs
- Be prepared to integrate appropriate aspects of analytical into operational data

All this needs to be done with a common platform that is flexible enough to accommodate ongoing growth.

View video on “Mastering Data Governance”
Planning for big data

Your company’s current level of maturity in master data management will dictate how quickly you can move to accept big data. Each company must determine its own investment priorities and development roadmap based on how it answers the key questions to the right.

It is safe to assume, though, that big data, being new and unproven, has the potential to stress your infrastructure and overwhelm your data stewards, all while providing only incomplete answers. Consequently, it is best to introduce only data that you need and according to a well-thought-out strategy.

If your current MDM environment cannot easily expand to accommodate analytical data, you can purchase MDM as a service just for “big data,” or for your existing operational processes as well. Either way, data reserved only for analytical purposes should remain separate from operational data, albeit with the ability to integrate it into your operational data when it is deemed worthy.

Key questions to ask yourself

- What is the business problem, and what question is being asked?
- What metrics are needed?
- What are the types of data that are needed?
  - Persistent vs. occasional
- Where will the data come from?
  - Internal / external
- What level of quality is needed (this will differ by application)
- How will quality be maintained?
- How will the data need to be structured?
- What system will be required to support the organization?
  - Data volume vs. variety vs. velocity
- What are the delivery mechanisms?
- Where do I build vs. partner?
Conduct a gap analysis of your MDM solution

**CURRENT STATE**

- Is it a struggle for your organization to accurately and completely identify a common view of a customer or product?

- Is there a single place for people to go for critical customer and product data? (i.e., names, addresses, ID’s, etc.)?

- Is your current MDM architected to support multiple consumers for business rules and governance?

- How would you describe the quality of the customer, product and affiliations data that you use today to drive your business (accuracy, completeness, timeliness, availability)?

- Does senior management recognize the current costs of poor information?

- Does your information have a business owner and is it being treated as a corporate asset?

- How comfortable are you that you are compliant with the Sunshine Act and regulatory requirements?

**FUTURE STATE**

- Are you currently or do you plan to focus on Account valuation and account-based selling. If so, do you have the ability to aggregate customer interactions through ownership hierarchies?

- Do you have a strategy to integrate transactional information across all touchpoints (calls, clickstream, EMR, social media, RWE, e-channels, etc)?

- Does your current technology footprint help you or hurt you?

- Have you outlined an information management strategy that will outline how new big data opportunities are introduced into the current environment?

- Are the roles and responsibilities clearly defined on how the MDM governance will extend and differential between operational and analytical MDM?
Hitting the jackpot with big data

While the “buzz” around big data is everywhere, and while big data will at some point become an essential corporate asset, you must lay the proper groundwork before you can take advantage of it.

Before rushing headlong to acquire big data, be sure that you have an Information Management strategy in place to cope with it. This begins with a best-of-breed MDM framework for operational data that includes clear roles and responsibilities around data stewardship and governance.

Then, consider how you will need to modify your MDM strategy to accommodate big data. Whether your big data are managed in-house or by a service provider, they should be quarantined from your operational data until you can be certain of their quality and completeness.

As with all new ventures, the first forays into using big data will be a learning experience. The technology ecosystem surrounding big data is still nascent, however, and we can expect every step of managing big data to become easier over time.
IMS improves the efficiency and value of clients’ commercial analytics and information processes through a broad range of services:

- Sales and marketing analytics and reporting
- KPI design and implementation
- CRM implementation
- Managed markets contract management
- Data warehousing, integration and management

Drawing from our extensive industry experience with methodologies honed over hundreds of client engagements, we recommend approaches, guide analytics, design processes and deploy systems — even manage these activities on an ongoing basis to lower operating costs, provide staffing flexibility and improve overall performance.

In turning to IMS specialists located both onsite and offshore, clients tap into our extensive global industry experience, deep domain expertise and knowledge of precisely how their commercial organizations operate. The result? Faster speed to insight and better business performance.