Data Provider Relationships: Pros, Cons, and Considerations

Andrew Kress

Surveillance Data Inc.

Surveillance Data Inc. (SDI), Plymouth Meeting, PA, has been providing disease surveillance data to pharmaceutical manufacturers, consumer products companies and retailers for over thirteen years.

SDI data is generated from a variety of sources and includes direct data collection from laboratories and physicians by SDI as well as the re-licensing and aggregating of data from existing sources who generate useful surveillance data as a byproduct of their primary businesses.

Objective

What are the key issues relating to relationships with potential data providers?

What are the challenges in working with similar data?

SDI

SURVEILLANCE DATA INC.

Who do we consider "data providers"?

Examples of SDI data generated as either standard transactional data or program-specific data:

Data Type	Coverage	Source
Outpatient Physician Office Visits	>400 million visits/year	Insurance claims
Prescriptions	65% of retail prescriptions (>1 billion Rxs per year)	Insurance claims
Clinical laboratory results	130+ hospital labs, national commercial labs	Lab resulting systems
Sentinel physician offices	500+ sites	Direct data collection
Hospital discharge data	900+ sites	Insurance claims

SDI has also worked with OTC sales data, online consumer-panel data, electronic medical-record data, ER discharge system data.

SDI

What's available?

The level of detail available depends upon the source, but most SDI data is available at localized geospacial level (five digit patient, physician, or pharmacy zip code) and both retrospectively and prospectively. Within the transaction-based data streams discussed, there are transaction standards and code sets that are used, however in other data streams, the variables may be significantly less standardized.

SDI

Evaluating Potential Data Providers

Differentiate between clinical and transactional data

How complete is the data?

What variables are provided (clinical vs. transactional)?

How much geographic information is available?

How quickly is it available (system vs. behavioral latency)?

Data cleanliness?

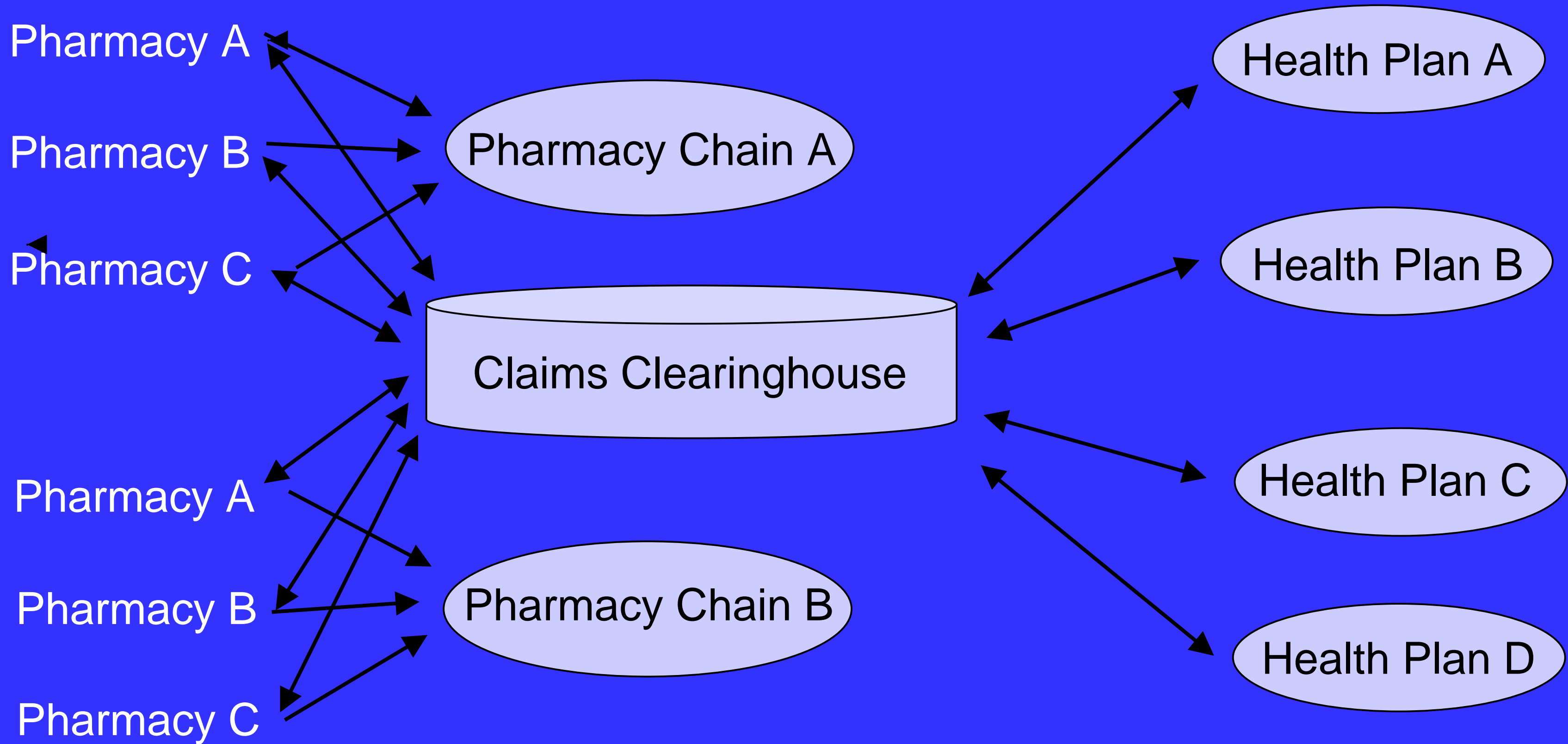
Is it re-identifiable if necessary?

How committed is the data provider?

SDI

What do we consider a data provider?

Data can be accessed at several points within the network.



Benefits of working with outside data providers

Size, connectivity, time-to-deployment:

Existing programs and data streams

Large sample sizes

Broad geographic coverage

Passive data collection

Data timeliness

Data standards

What might you come up against?

Concerns that the entities providing the source data may have include:

- Complexity of data extraction from online transaction processing systems for analytics use
- Concerns about patient privacy and confidentiality
- Increased resource requirements (both personnel and IT) to generate and support ongoing daily data
- Limited financial upside
- Sensitivity towards public perceptions relating to release by provider of identifiable information
- Release of competitively sensitive information
- Ongoing duplicative data requests from various external organizations

SDI

"The goal is worthwhile, and we'd love to help but…."

Legal/Regulatory

- "Will we get sued for this?"
- "How do we know for sure that you'll protect our data from our competitors?"

Corporate/Senior Management

- "Will we make money at this? Are we diluting the value of our data?"
- "Will this make us look bad to our customers, patients, or shareholders?"
- "Are we allowed to do this based on existing contracts?"
- "What else are you or could you do with this besides what we're agreeing to?"

IT

- "How much time and system resourcing is this going to take me to set up and run?"
- "Are you going to call me every time something happens?"

All three stakeholders have differing concerns, risk levels, decision-making (or blocking) authority, and approval processes.

The easy answer is always – NO!

SDI

What do they really think about all this…

"I see only downside to (us) from cooperating. While your stated goal is a good one, I'm concerned that the only possible publicity we would get from this is negative. Consumers are already concerned that (we) know too much about them and what they do and when they do it. Also I do not readily see how our privacy policy would be read to allow us to provide this data to you. In order to do so, I will need to see if we can narrow the scope of the data we provide to fall within the language of our privacy policy before we could turn anything over to you.

If we were to provide you with data, would you be willing to:

indemnify us for our damages, including consequential damages from adverse publicity, as well as attorney fees to defend any action brought against us by consumers; and

agree to not name us in any way, identify the data you receive from us as coming from (us) and use code names for all locations and take other measures so that the data would not be traced back to us?

I need to also know … why you believe (we) should provide its valuable data without compensation when we charge others for subsets of the same data.

Once we resolve these issues, the next step would be to have your technical people speak to ours about what exactly you need and how much work it would take to provide it."

THE WALL STREET JOURNAL ONLINE

E-COMMERCE/MEDIA

Responding to Privacy Concerns, JetBlue E-Mails an Explanation

By SUSAN CAREY and STEPHEN POWER

WIRED NEWS

Senators Want JetBlue Probe

02:00 AM Oct. 16, 2003 PT

Three powerful U.S. senators are pressuring the Pentagon to reveal more about its role in the JetBlue privacy scandal, pointedly asking whether a key privacy law was violated when a defense contractor acquired 4.9 million passenger itineraries from the upstart airline for a security study without giving notice to the passengers.

Governmental Affairs Committee chairwoman Susan Collins (R-Maine), ranking member Joe Lieberman (D-Connecticut) and Armed Services Committee ranking member Carl Levin (D-Michigan) wrote Defense Secretary Donald Rumsfeld, asking whether the Pentagon complied with the Privacy Act, which requires that government agencies and their contractors notify the public when a system of records is created.

The Pentagon's contractor, Torch Concepts, acquired more than 1 million passenger records from JetBlue and then augmented the data with

The New York Times — Business

DealBook

Senators Press Inquiry on Privacy Issue

Are you paying too much

Archives — The New York Times

COMPANY NEWS; JETBLUE SUED OVER RELEASE OF PASSENGER DATA

What data has business sensitivity?

Personally-identifiable information

Total volumes of patients, tests, procedures or sales

Store or site locations

Any financial or cost data

Overall business size measurements

Patient demographic mix

Number of members

Provider names and addresses

SDI

Why would a data provider participate?

Reasons for participation in a surveillance system by a data provider generally relate to:

1. Commercial benefit
2. Low and defined expectation of resource requirements
3. Corporate goodwill
4. Clearly defined and mitigated financial or P.R. risk

SDI

Assuming a data provider says yes….

Who is responsible for technical issues? For example, de-identifying data in a HIPAA-compliant fashion, extracting relevant records, obtaining and keeping master files of NDC or UPC product codes up-to-date.

Other issues to consider in working with these types of sources include:

- non-trivial data processing, quality control, and storage requirements when working with non-aggregated records
- necessity of maintaining provider relationships
- methodological and process complexity
- lack of easily available proven analytic techniques for use in data interpretation.

SDI

Assuming a data provider says yes….

Who writes or modifies, installs, and maintains de-ID and data transfer/load processes?

Who obtains and keeps master files of NDC or UPC product codes up-to-date?

How do I keep track of stores opening and closing, physicians moving, or products being added or deleted by a specific retailer?

Who handles data processing?

SDI

Assuming a data provider says yes….

What kind of quality control is required?

How much storage is needed?

Who at the provider approves changes or fixes problems?

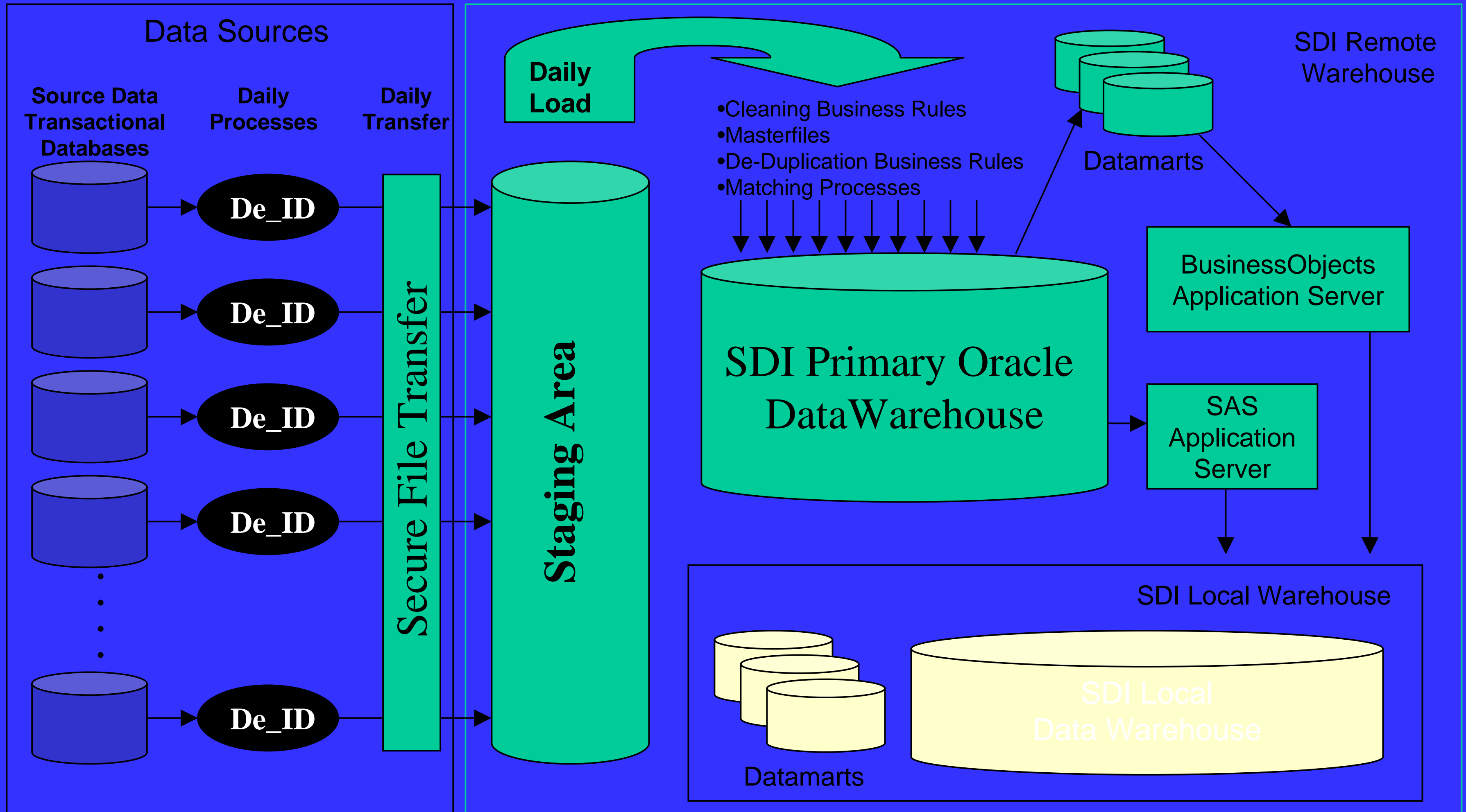
What do I do with this volume of data?

Will I be able to find the patient/consumer if necessary?

SDI

Example of Enterprise Architecture

SDI handles 2 billion transactions per year (visits & prescriptions).
Overall data warehouse currently around 12 TB.



Assuming a data provider says yes….

Formal versus informal relationships to acquire data.

•Performance-based contracts
•Data use agreements
•Non-disclosure agreements
•Service level agreements

Developmental versus deployable solutions.

Development= Informal OK, proof-of-concept
Deployment= Formal structure and stability required

SDI

What happens to the data feed if….

- A technical event occurs
- Change of data provider business strategy
- Loss of altruistic motives
- Poor technical implementation and support
- Breach of confidentiality
- Bad P.R.

SDI
SURVEILLANCE DATA INC

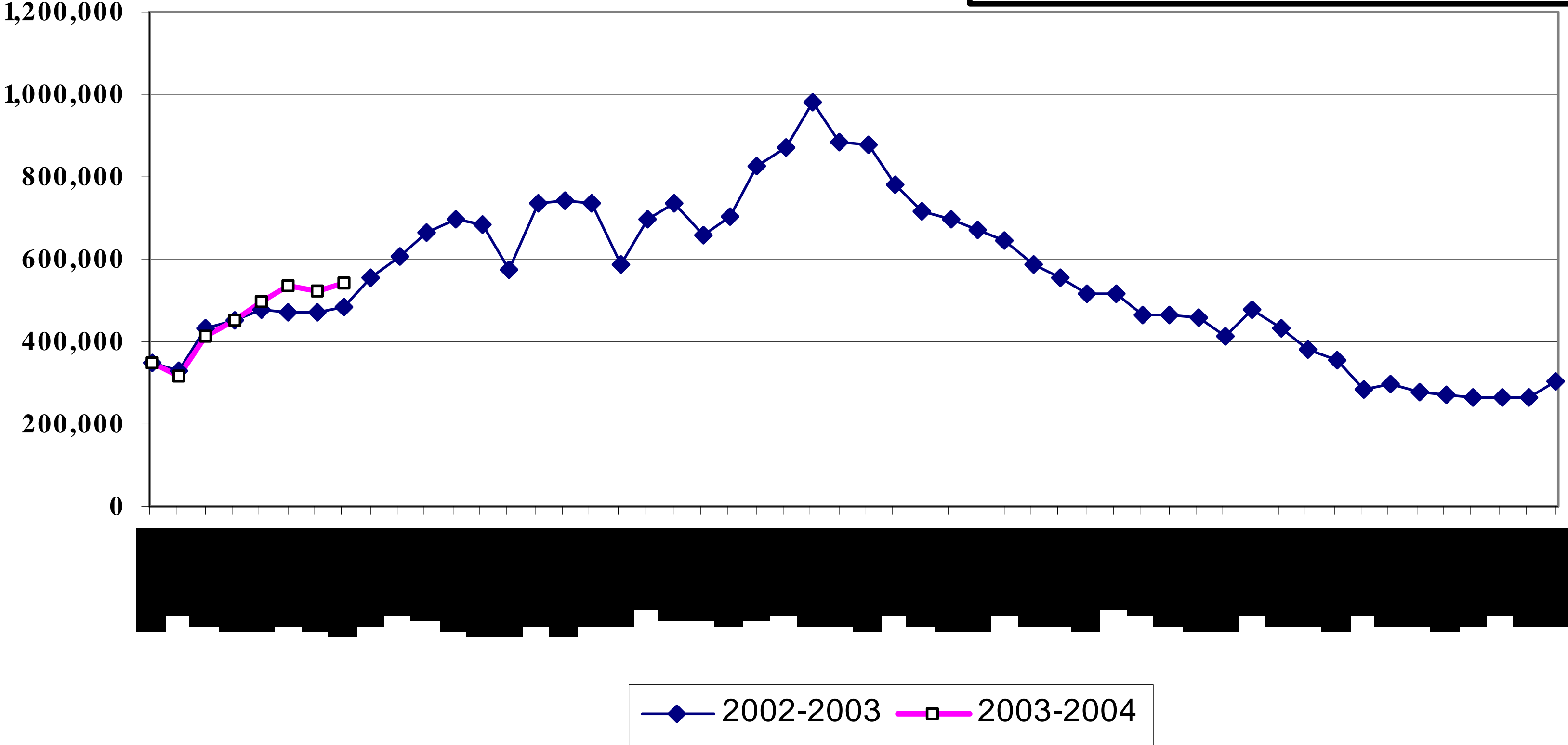
Conclusions

There are several advantages such as large sample size, timeliness, and passive collection that may be gained by using some of the types of data mentioned.

In order to do this however, end users of the data must be prepared to address the concerns of the data providers and to deal with the methodological and technical complexities involved in such an undertaking.

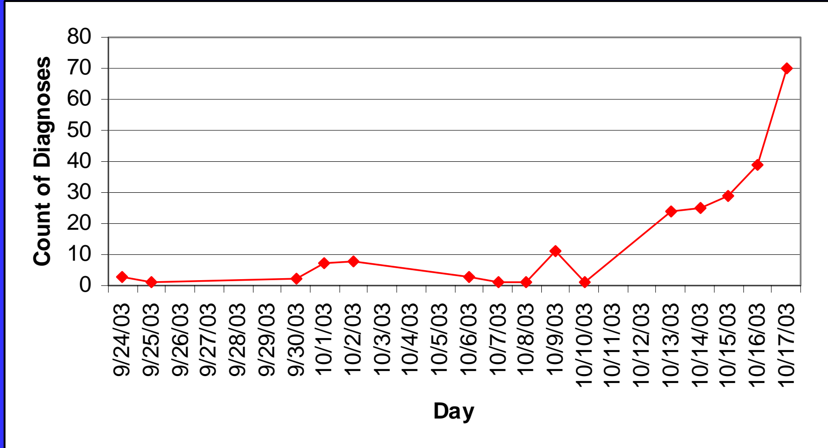
SDI

Example- National

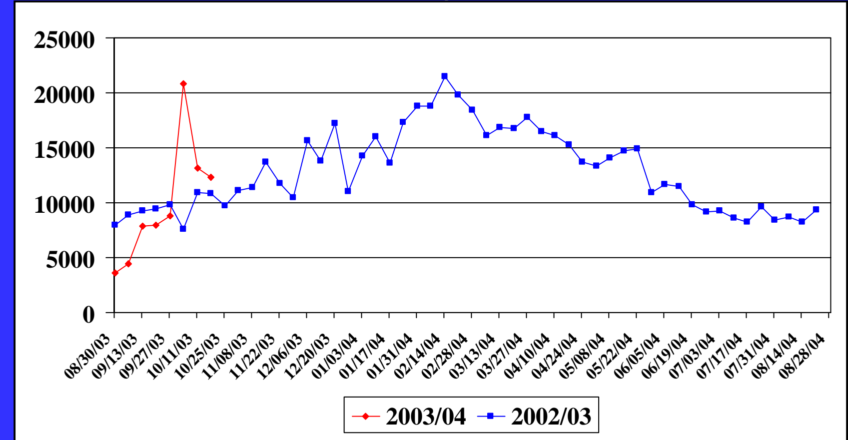


Example- Houston

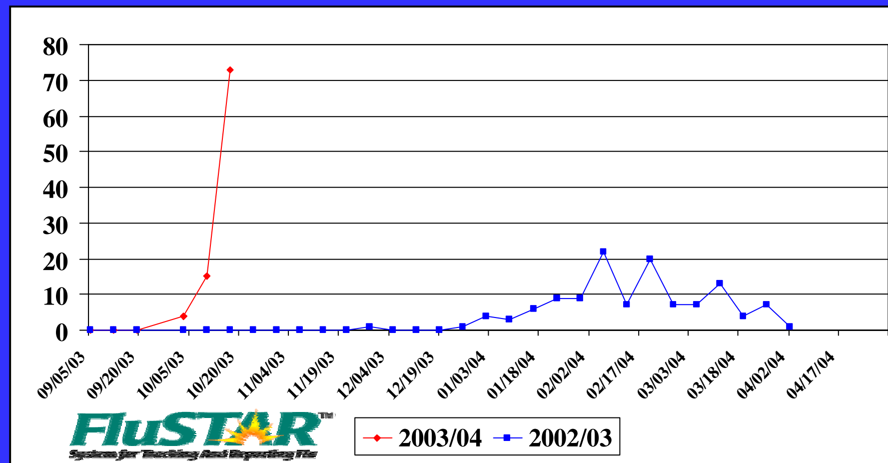
Outpatient Office Daily Visits- Influenza



Outpatient Office Weekly Visits- Total Respiratory Illness



Lab Results- Influenza Test Positives



6 Different Data Provider Relationships Required