

Harvard Medical School

Faculty researchers bring drug safety and effectiveness studies to enhanced levels with IBM® Netezza® technology

Overview

The need

The lab at Brigham and Women's Hospital was looking to find a platform for computational pharmacoepidemiologic analytics that would address rapidly emerging trends.

The solution

IBM Netezza data warehouse appliance

The benefits

By utilizing IBM Netezza technology in their pharmacoepidemiology research studies, the Harvard Medical School division will be able to:

- Increase the speed of computationally-intense analysis of claims data
 - Accelerate testing of new, more sophisticated algorithms
 - Facilitate automation of continuous drug safety and effectiveness monitoring
-

Introduction

The Harvard Medical School® Division of Pharmacoepidemiology and Pharmacoeconomics at Brigham and Women's Hospital in Boston, Massachusetts, is a globally recognized leader in drug safety and effectiveness research. Created in 1998, the division is led by Dr. Jerry Avorn, Professor of Medicine at Harvard. The division is part of Brigham and Women's Hospital's Department of Medicine, and performs advanced analytics on patient health claims data. Through computationally intensive 'big data' analytics, this research brings to light insights into how drugs compare to each other in terms of safety and effectiveness.

The charter of the Division of Pharmacoepidemiology and Pharmacoeconomics is to:

- Study the use of medications in very large populations in order to determine which drugs are most likely to effectively treat patients' medical conditions, and which are more likely to cause adverse events.
- Develop advanced, data-intensive methods in epidemiology, biostatistics and informatics to detect and evaluate important early signals of potential drug safety problems, as well as to compare the therapeutic results of alternative treatment strategies.
- Define patterns of physicians' prescription and patients' use and compliance of medications.
- Analyze the influence of health system factors (policies, coverage and reimbursement differences) on the quality of medication use.
- Design, implement, and test innovative programs to improve the appropriateness of medications prescribed by physicians and used by patients.



“We as a society cannot afford to be told, ‘I can’t run your algorithms... because our technology can’t support it.’ People need and want to know the information that our studies generate. We can’t afford not to carry out these studies. It’s too important to get these answers for drug developers, regulators, physicians, policy makers, payers and most importantly, for our patients.”

Dr. Schneeweiss
Director for Drug Evaluation and
Outcomes Research
Brigham and Women’s Hospital

The challenge

The nature of Harvard’s pharmacoepidemiology research has been highly complex and computationally intensive since its inception. Today, however, the sheer volume of information available – and the increasing detail in which the information is recorded – has dramatically increased the analytic and computational challenges.

The expansion of data – both in its depth and breadth – has also presented opportunities for more complex and robust drug safety and effectiveness studies. Dr. Avorn and his colleague, Dr. Sebastian Schneeweiss, Director for Drug Evaluation and Outcomes Research, in their 2009 editorial “Managing Drug-Risk Information-What to Do with All Those New Numbers” (*New England Journal of Medicine*, July 27, 2009), reveal some of the diverse challenges facing epidemiological research in the current era. The challenges include applying new, rigorous analytic techniques to enhance the levels of confidence in observed drug safety results, as well as challenges in organizational politics, patient privacy and administration. This diversity is precisely why the research group includes experts in fields such as specialty medicine, health policy, sociology, law and more.

With access to the markedly larger amounts of patient data, the more traditional ‘reactive’ approaches to adverse drug events can no longer suffice. Examining data after the fact to confirm a suspected drug safety problem will always be a crucial element of safety research. Powerful computation simultaneously opens up avenues for identifying issues after a drug has entered the market, but before it is in widespread use. The Harvard team envisioned the potential of proactive research techniques using very large databases. This way, they could identify “signals” of risk associated with the particular drugs themselves or with combinations of drugs.

The lab at Brigham and Women’s Hospital was looking to find a platform for computational pharmacoepidemiologic analytics that would address these trends. Specifically, Dr. Jeremy Rassen, Assistant Professor of Medicine at Harvard, identified certain needs as critical:

- Rapid execution of computationally intense analysis of structured medical claims data.
- Capabilities for parallelized in-database analytics.
- Accelerated development and testing of new, more complex algorithms.
- Facilitated automation of continuous drug safety event monitoring.
- Capabilities for complex analytic SQL queries.
- Simplified database administration.
- Ability to work with semi-structured electronic healthcare record databases.
- Security features for the continued protection of patient privacy.

“IBM Netezza will accelerate our ability to devise, test and publish new computationally intensive algorithms applied to large longitudinal healthcare databases that we hope become the ‘gold standard’ for researchers globally.”

Dr. Schneeweiss
Director for Drug Evaluation and
Outcomes Research
Brigham and Women’s Hospital

The solution

IBM Netezza partnered with the lab and installed an IBM Netezza data warehouse appliance. IBM Netezza data warehouse appliances are purpose-built to make advanced analytics on data simpler, faster and more accessible.

The IBM Netezza data warehouse appliance is:

- Designed specifically for running complex analytics on very large data volumes, at orders of magnitude faster than competing solutions.
- Simple to maintain – The appliance architecturally integrates database, server and storage into a single, easy to manage system that requires little on-going maintenance.
- Simple to deploy – As a purpose-built product, the IBM Netezza data warehouse appliance comes pre-tuned with the optimal architecture for fast, advanced analytics, making it ready to go right out of the box.
- Simple to try – The IBM Netezza TestDrive makes it easy to experience the high-performance data warehouse appliances on-site, with real data.
- Unmatched in value – The IBM Netezza appliance offers faster time-to-value and lower total cost of ownership than competing products in the industry.

IBM Netezza data warehouse appliances deliver the proven performance, value and simplicity organizations need to dive deep into their growing data. Customers are able to easily and cost effectively scale their business intelligence and analytical infrastructure, to leverage deeper insights from growing data volumes, throughout their organization.

The results

According to Dr. Rassen: “The IBM Netezza data warehouse appliance put us ahead of the curve. With almost no system administration or index-building, we quickly saw orders of magnitude performance improvement. It also allowed us to conduct basic analytic processing at two to three times previous speeds, with no change in code. Furthermore, it enabled one of our novel algorithms, the high-dimensional propensity scoring, to run 20-30 times faster than in our previous relational database environment. The appliance also gave us the ability to explore previously inconceivable new research avenues, with the availability of highly parallel data access and analysis.”

As a result, Dr. Rassen's team has been able to:

- Load data into the database more quickly.
- Run all previously developed SQL and SAS code with minimal changes.
- Know that their efforts to migrate certain SAS analytics and processes to the open-source R language would be easy.
- Dramatically increase query response.

The research team noted that one of the very significant benefits of the technology was its simplicity of use. Dr. Rassen remarked that the system was situated in the data center, installed, and up and running in less than 48 hours. In addition, the team required no outside administration time, and system maintenance was handled by the researchers and analysts themselves. Other systems typically require weeks of set-up and tuning, not to mention frequent or constant administration.

The impact of the analytics

The stakes are high. In the U.S. alone, half of the population is taking one or more prescription drugs. Nearly four billion prescriptions are being filled annually. U.S. prescription drug sales are well over \$300 billion per year, and according to IMS Health, global sales are expected to reach \$1.1 trillion by 2014. Additionally, the Kaiser Foundation is reporting that prescription drugs account for well over 10 percent of the total U.S. health expenditure (over \$2.5 trillion). These figures highlight the positive impact that Harvard's advanced research can have on more rapid identification of unsafe and less effective drugs. The absence of more rapid risk awareness can be devastating to patients and continuously crippling to healthcare costs. Without steadily advancing complex analytics and tremendous computing power, the exceptional potential of Harvard's vision in pharmacoepidemiology research could face delays that global healthcare cannot afford.

The Division of Pharmacodpidemiology and Pharmacoconomics at Brigham and Women's Hospital has developed research relationships with several *Fortune 500* companies to perform intervention studies on their medication benefit programs offered to employees. The lab is also a critical drug safety and effectiveness research training destination for many doctoral and post-doctoral students from all over the world. Division faculty have produced well over 300 notable papers on drug safety, drug costs and health policy. The quality and amount of accessible patient data for the group's research is critical.

The high-dimensional data analysis approach for healthcare databases is based on the assumption that there is more information in such databases than is generally utilized. Such information may not immediately look relevant but our algorithms manage to identify previously-undetected risk factors in those data and improve adjustment to make treatment groups more comparable.

High Dimensional Pharmacoepidemiology
www.hdpharmacoepi.org

The lab calls a major portion of what they do “High Dimensional” Pharmacoepidemiology. It seeks to take advantage of all the high-dimensional data space available in longitudinal insurance claims data, registries and electronic healthcare record databases. This will improve the validity of the research in the frequent cases where running a randomized trial is not possible or feasible.

Recognizing that the safety and effectiveness of new and existing prescription drugs plays an ever increasing role in the pursuit of improved healthcare worldwide, the research team designs, develops and executes studies to provide more reliable, actionable data that can:

- Bring greater clarity to issues of benefits versus risks.
- Evaluate the impact of prescription drug expenditures.
- Understand how medications are prescribed by doctors and used by their patients.
- Create action plans to help patients improve adherence to medications.
- Develop methods that optimize the use of prescribed drugs.
- Identify drug safety issues for specific subsets of the population.
- Assist public sector governing bodies in their decision making processes.
- Provide physicians with new insights when prescribing medications.
- Create a system of active drug safety surveillance, allowing proactive risk intervention.
- Help pharmaceutical companies bring safe, effective new drugs to market more smoothly.

The Harvard Medical School division has already achieved major accomplishments, one of which is developing new methodologies for computer-intensive analysis of very large datasets to identify drug risks. Another is identifying important “confounders” that can distort the results of non-randomized studies – these confounders can skew results and make incorrect findings appear robust. A third methodology seeks to assess the cost effectiveness of particular medications and drug-use strategies. As meaningful as these developments are, Dr. Schneeweiss envisions continuous advances in the analysis of ever-growing patient databases with even richer data. These advances will help automate and steadily improve society’s ability to monitor drug safety and effectiveness.

“Things that we could not have done before...things we would not have considered, are now easily within our reach.”

Dr. Schneeweiss
Director for Drug Evaluation and
Outcomes Research
Brigham and Women’s Hospital

What is a patient cohort?

In epidemiology, a cohort is a group of like individuals who are followed over time and analyzed for the incidence of diseases, adverse events or other health outcomes of interest.

“As global healthcare evolves toward an ongoing learning healthcare system with a need for continuous patient-centered comparative effectiveness and safety research integrated in routine care, it is imperative that research methods evolve in parallel...IBM Netezza will accelerate our ability to devise, test and publish new computationally intensive algorithms applied to large longitudinal healthcare databases that we hope become the ‘gold standard’ for researchers globally...Accurate effectiveness information generated with little delay will help healthcare decision makers to improve the health of our patients.” – Dr. Schneeweiss

The dramatic expansion of available patient data, and the ever-increased complexity of questions that need to be asked in more enriched studies, mean traditional methods have simply become too slow. With the research group’s new algorithms and goals of drug safety automation, the computational demands on processing technology have increased exponentially but are now being met with IBM Netezza’s data warehouse appliance.

Dr. Schneeweiss is fervent about the prospects of enhanced automation and more complex, large database studies that will be pursued with the IBM Netezza platform. His vision of a system that is continuously “learning” by incorporating data from ongoing drug safety and effectiveness research results is a “game changer” for the future of how prescription drugs are managed and developed.

Dr. Schneeweiss fully expects that the ongoing system will be able to generate reliable, early warning signals of adverse drug effects that have previously taken months to years to discover.

The benefits

By utilizing IBM Netezza technology in their pharmacoepidemiology research studies, the Harvard Medical School division will be able to:

- Increase the speed of computationally-intense analysis of claims data.
- Accelerate testing of new, more sophisticated algorithms.
- Facilitate automation of continuous drug safety and effectiveness monitoring.
- Allow research studies on larger databases.
- Operate with little to no outside database administration.
- Make it possible to work on multiple databases at the same time.
- Add the ability to store and analyze electronic health record databases.
- Achieve higher levels of research result-confidence.
- Continue the Division’s protection of patient privacy.
- Facilitate the development of an automated ‘learning’ healthcare system.
- Identify drug risk ‘signals’ far more quickly than previous systems could.

“Our lab is focused on the best ways to automate the continuous evaluation of the safety of medications and we require fast systems to assemble and analyze data.”

Dr. Rassen
Assistant Professor of Medicine
Harvard Medical School

- Provide drug developers and regulators with superior large-scale, post-market data.
- Facilitate in-depth, active surveillance for post-market drug safety.
- Provide extensive data to assist with public sector health decisions.
- Offer physicians more detailed data for their prescription decision process.
- Create tools to improve adherence to beneficial medications.
- Present patients with more reliable information about their prescribed drugs.

The team

The Harvard team’s drug safety and effectiveness research mission takes a uniquely inclusive approach by providing a deep interdisciplinary environment of experts for its studies. The division brings together clinicians from a wide-range of backgrounds including cardiology, geriatrics, primary care, rheumatology and nephrology; with experts in the quantitative sciences of biostatistics, epidemiology and complex analytics. Professionals in other pertinent disciplines such as health policy, law, decision analysis and the social sciences are also included. In doing so, they set their research focus sharply on the ultimate care of the patient.

“Our high-dimensional pharmacoepidemiology takes an empirical, data-driven approach to solving the difficult problems of pharmacoepidemiology, drug safety and comparative effectiveness research. By making use of all the information contained in insurance claims, registries, and even electronic medical records, we seek to improve the validity of non-randomized research.” – Dr. Schneeweiss

A small sample of publications from the Harvard faculty-researchers in the Division of Pharmacoepidemiology:

Risk of death and hospital admission for major medical events after initiation of psychotropic medications in older adults admitted to nursing homes.

Huybrechts K.F., Rothman K.J., Silliman R.A., Brookhart M.A., Schneeweiss S.

CMAJ. 2011 Apr 19; 183(7): E411-9

Anticonvulsant medications and the risk of suicide, attempted suicide, or violent death.

Patorno E., Bohn R.L., Wahl P.M., Avorn J., Patrick A.R., Liu J., Schneeweiss S.

JAMA 2010; 303: 1401-9

High-dimensional propensity score adjustment in studies of treatment effects using health care claims data.

Schneeweiss S., Rassen J.R., Glynn R.J., Avorn J., Mogun H., Brookhart M.A.

Epidemiology 2009; 20: 512-22

Privacy-maintaining propensity score-based pooling of multiple databases applied to a study of biologic agents.

Rassen J.A., Solomon D.H., Curtis J., Harrington L., Schneeweiss S.

Med Care 2010; 48: S83-9

Variation in the risk of suicide attempts and completed suicides by antidepressant agent in adults: A propensity score-adjusted analysis of 9 years of data.

Schneeweiss S., Patrick A.R., Solomon D.H., Metha J., Dormuth C., Miller M., Lee J., Wang P.S.

Arch Gen Psychiatry 2010; 67: 497-506

Post marketing studies of drug safety.

Schneeweiss S., Avorn J.

BMJ. 2011 Feb 8; 342: d342. doi: 10.1136/bmj.d342

The epidemiology of prescriptions abandoned at the pharmacy.

Shrank W.H., Choudhry N.K., Fischer M.A., Avorn J., Powell M.,

Schneeweiss S., Liberman J.N., Dollear T., Brennan T.A., Brookhart M.A.

Ann Intern Med. 2010 Nov 16; 153(10): 633-40

Cardiovascular outcomes and mortality in patients using clopidogrel with proton pump inhibitors after percutaneous coronary intervention.

Rassen J.A., Choudhry N., Avorn J., Schneeweiss S.

Circulation 2009; 120: 2322-9

About IBM Netezza

IBM Netezza pioneered the data warehouse appliance space by integrating database, server and storage into a single, easy to manage appliance that requires minimal set-up and ongoing administration while producing faster and more consistent analytic performance. The IBM Netezza family of data warehouse appliances simplifies business analytics dramatically by consolidating all analytic activity in the appliance, right where the data resides, for blisteringly fast performance. Visit netezza.com to see how our family of data warehouse appliances eliminate complexity at every step and lets you drive true business value for your organization. For the latest data warehouse and advanced analytics blogs, videos and more, please visit: thinking.netezza.com.

IBM Data Warehousing and Analytics Solutions

IBM provides the broadest and most comprehensive portfolio of data warehousing, information management and business analytic software, hardware and solutions to help customers maximize the value of their information assets and discover new insights to make better and faster decisions and optimize their business outcomes.



© Copyright IBM Corporation 2011

IBM Corporation
Software Group
Route 100
Somers, NY 10589
U.S.A.

Produced in the United States of America
July 2011
All Rights Reserved

IBM, the IBM logo, and ibm.com, are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at ibm.com/legal/copytrade.shtml

Netezza is a registered trademark of Netezza Corporation, an IBM Company.

Other company, product and service names may be trademarks or service marks of others.

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates.



Please Recycle