

POCT data mining – a practical approach

March 2005



James H. Nichols

PhD, DABCC, FACB

Associate Professor of Pathology

Tufts University School of Medicine

Director, Clinical Chemistry

Baystate Health System

759 Chestnut Street

Springfield, MA 01199

USA

Simple data-mining algorithms can be applied to point-of-care testing (POCT) data to document compliance of quality control, operator training and to identify potential preanalytical errors.

Data can be manipulated to automate manual review and other laborious processes for identifying data trends, verifying regulatory compliance, troubleshooting technical problems, and improving laboratory efficiency and patient care.

Once familiar with simple data manipulations, institutions can progress to more sophisticated management algorithms and data-mining techniques.

Laboratory results contain a wealth of information, if only this resource could be tapped for use by the clinician. Laboratory testing is estimated to influence more than 70% of all medical decisions [1].

With over 7 billion laboratory tests conducted each year in the United States, there is a growing volume of test

records that are an integral part of patient management.

Unfortunately, most test results sit in a Laboratory or Hospital Information System (LIS or HIS) and are rarely utilized beyond the next patient visit or subsequent result for a given analyte, because of the time-sensitive nature of laboratory information.

Collectively, the laboratory data sitting in these information systems is a potential warehouse of vast historical patient and population statistics, if a clinician has the key to unlock the underlying information.

The key is mathematical and lies in grouping, trending and otherwise manipulating, combining and sorting the data to reveal the underlying patterns of information.

Once revealed, this information can be used to better define a test's sensitivity and specificity, the test limitations, appropriate clinical applications, cost-effective pathways of care, and result interpretation that can benefit patient outcomes.

The technique of extracting useful information from vast amounts of data is termed “*data mining*” because of the parallels of excavating the earth to find the hidden gems and wealth under the surface.

Many believe that sophisticated software and powerful computers are required to mine data, but even the simplest combinations of data can be considered a form of data mining.

Institutions can start with simple data manipulations and progress to more sophisticated management algorithms and data-mining techniques.

Automate processes

The most fundamental data algorithms simply automate processes that are currently being conducted manually. It is impossible to identify trends amongst thousands of POCT results generated by hundreds of devices and operators manually.

Using the computer to automate manual processes can both improve the efficiency and the ability to find trends and problems. Computers also have the ability to automate communication.

Compliance issues and other problems can be sent to the medical units through e-mail messages or automated faxes/pager messages, and suggestions can even be made to assist the clinical staff in troubleshooting.

The combination of POCT data is thus the first step towards automating processes of data review and increasing staff efficiency in finding and resolving errors and trends.

Automated POCT data manipulations assist the documentation of institutional compliance and adherence of clinical practice to regulatory guidelines. POCT control and patient data can be sorted and grouped based on date, lot numbers, device serial number, testing location, operator and other parameters.

After sorting, results can be plotted, and means and standard deviations calculated. Levy-Jennings plots of quality control results are just one example of data sorting where quality control data is collected from POCT devices, sorted by date, lot number, concentration level, and plotted to assist the identification of trends and biases over time.

Automating regulatory compliance documentation

Combinations of POCT data with other data, like core-laboratory results, can provide further evidence of regulatory compliance in an automated fashion.

Laboratory accreditation by the College of American Pathologists requires verification of accuracy semi-annually by comparison of patient results between POCT and laboratory methods [2].

While this can be accomplished by analyzing a few patient specimens across the various laboratory instruments and POCT devices, the laborious process of accuracy verification can be automated by comparing results in the POCT database with results from a core laboratory on the same patient conducted at the same time.

In routine practice, clinicians sometimes want to verify the POCT results by collecting a venous specimen sent to a core laboratory.

By searching for results of patient specimens collected at nearly the same time and analyzed by the two methods in both the POCT database and laboratory information system, institutions can automatically prove method accuracy without having to expend the labor to find appropriate patient specimens, then manually calculate and review the comparisons.

The only limitation to implementing this process is the ability of an institution to combine data from POCT devices with core-laboratory data.

Those institutions that transmit POCT results to a laboratory or hospital information system will have the

results of both methods already combined in a single database, and just need to set up the calculations.

Monitoring effectiveness of patient treatment

Similar manipulations can be accomplished on individual patients to monitor trends and even the effectiveness of diabetic treatment regimens.

Comparable statistics can be performed on the patient's home-monitoring device to determine compliance with daily monitoring and treatment recommendations between clinic visits.

Grouping home-monitoring data with laboratory results, like HgbA1c, on groups of clinic patients can estimate the effectiveness of the overall clinic management in meeting practice recommendations, like the frequency of HgbA1c testing, office visits and other patient outcomes.

Greatly simplified operator competency checking

Beyond simple data sorting and combination is the creation of prediction algorithms. Here, retrospective data is sorted and analyzed to identify trends that can then be utilized to predict future problems.

One example of a prediction algorithm is the use of quality control data to verify operator competency and to automatically update periodic operator certification without requiring the operators to go through retraining or competency checks [3, 4].

In this example, operators perform two levels of quality control each quarter (**Table I**). The mean of each concentration level for all quality control tests conducted each quarter is calculated.

This is then compared to the mean of all quality control tests performed on the same lot for the entire institution. A z-score or standard deviation index is calculated as the difference of the operator's mean from the institution's mean as the target goal.

If the operator performed within two standard deviations of the institution's mean, the operator is deemed to be competent and their training updated. If the operator's performance is outside two standard deviations, their name is flagged for retraining or manual recertification.

Example of a data-mining algorithm to reduce the labor involved in manually certifying operator competency. Prior to implementation of this algorithm, the institution would have a trainer watch every operator perform testing.

With over 2,000 operators, this required over 16 weeks of labor (assuming 10 minutes of trainer time plus 10 minutes for each operator). With this algorithm, operators are automatically updated based on their quality control performance in routine practice.

Only those operators with different performance (more than two standard deviations from the group mean) are retrained, significantly reducing the required labor and utilizing data from routine practice.

Statistics also allow detection of operators with greater imprecision that could be a source of error (adapted from references 5 and 6).

This data algorithm has several advantages. Operator competencies are verified more efficiently through actual performance statistics rather than having the added labor of requiring someone to watch every operator perform a test periodically.

This labor could be significant when considering the hundreds of operators that perform POCT in the average hospital. The data algorithm is also biased towards operators that do more testing, since their control results contribute more to the institutional group mean.

Operators who do infrequent testing run a higher risk of generating an outlier, so retraining is targeted towards operators performing fewer tests. This is exactly the group of operators that require more practice at performing the test and need closer supervision.

Step 1 – Calculate group statistics

Calculate mean, standard deviation (SD) and coefficient of variation (CV) of quality control results sorted by concentration and lot number.

High control N=1016 Mean=283.4 SD=15.1 (lot# 5A367)
(CV=15.1/283.4=5.33%)

Step 2 – Calculate individual statistics

Calculate mean and standard deviation index (SDI) of quality control results for each operator sorted by concentration and lot number.

(SDI=operator mean–group mean/group SD)

High control (lot# 5A367)

Nurse S. Sullivan N=4 Mean=291.6 SD=17.5 CV=6.0%
(=291.6–283.4/15.1=0.54)

Nurse J. Smith N=3 Mean=241.2 SD=25.6 CV=10.6%
(SDI=241.2–283.4/15.1= -2.79)

Nurse J. Miller N=24 Mean=273.6 SD=12.8 CV=4.7%
(SDI=273.6–283.4/15.1= -0.65)

Step 3 – Identify operator outliers

Sort and flag operators with SDI greater than 2.0

Certified nurses

Nurses requiring follow-up

Nurse S. Sullivan: SDI=0.54 **Nurse J. Smith:** SDI= -2.79

Nurse J. Miller: SDI= -0.65

Step 4 – Send results to medical unit managers

E-mail results to unit managers for follow-up and retraining of those nurses with outliers.

TABLE I: Use of quality control data to document operator competency

POCT quality control data can thus be manipulated or data-mined to identify performance trends and automatically target operators for retraining, increasing efficiency.

Identifying preanalytical error

Another prediction algorithm is the “*delta check*” that can be utilized as a means of identifying potential preanalytical errors. A delta check is a comparison of two laboratory results by calculating the difference between two test results for the same analyte and estimating if the difference is clinically significant.

In a core laboratory, delta checks are utilized to identify sampling errors, like clots or bubbles in a specimen aliquot that could generate incorrect results.

For POCT, a current result can be compared to a previous POCT result to determine if there might be a sample-collection problem. Common errors with POCT include inadequately mixed specimens or delays in testing that could result in specimen clotting before analysis.

Identification errors are also common when the operator manually types an incorrect digit in the patient identification field.

Identification errors can cause a result to be held because the identification number does not match an active patient record number in the laboratory or hospital information system, or in the worst case scenario, the identification error could cause a result to be reported to another patient's medical record with the possibility of adverse medical treatment based on another patient's result.

The Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) has focused on patient safety this year. Reduction of errors in patient identification is one of the JCAHO patient-safety goals being targeted [5].

Use of barcoded patient bands can help reduce identification errors, but are not fool-proof as there is still the possibility that patients can be banded with another patient's number, or for the band to contain incorrect or outdated identification [6].

By delta-checking POCT results as they are downloaded from devices against a patient's previous result, gross errors in specimen collection, possible patient misidentification or other analytical interference with the patient's specimen can be recognized.

Delta checks thus offer a means of identifying potential errors, but are only useful for analytes that do not change significantly over time, like electrolytes or creatinine.

Summary

Data mining is a technique for combining, sorting and manipulating data to extract useful information (**Table II**).

Summary of various current and future algorithms that can be utilized to manage POCT results to extract useful information in an automated fashion.

The list is not intended to be comprehensive but to give an overview of the various possibilities that exist for data mining and customization of viewing data.

Data can be manipulated to automate manual review and other laborious processes for identifying data trends, verifying regulatory compliance, troubleshooting

technical problems, and improving laboratory efficiency and patient care.

Simple data-mining algorithms are currently being utilized to document compliance of quality control, operator training and to identify potential preanalytical errors.

In the future, wireless communication with POCT devices will allow real-time detection of errors, clinical warnings and more sophisticated data management.

The application of data-mining techniques and the potential for the information that could be obtained from POCT are limited only by an organization's imagination and the ability to connect the appropriate databases.

Automated communication

- E-mail unit managers with updates and issues requiring attention
- Follow-up with reminders if response not received by deadlines

Data inquiry

- Sorting and grouping control or patient results by date, device serial number, patient, operator, testing location or other parameter
- Plotting sorted data (i.e. Levy-Jennings control charts or individual patient results over time)

Combining data

- Grouping multiple POCT or laboratory results (i.e. trending individual glucose results with HgbA1c results over time)
- Pairing individual patient POCT results with laboratory results on samples conducted at the same time (i.e. verifying device accuracy by grouping POCT and laboratory results conducted at the same time and calculating a difference)

Prediction algorithms

- Sorting quality control results from each operator or device and comparing to group statistics (i.e. predicting operator or device trends and verifying operator competency or device accuracy by comparing quality control statistics from each operator/device with group statistics)
- Delta checks (i.e. predicting preanalytical errors or analytical interference by comparing current POCT result with previous POCT or laboratory results)

Future mining routines

- Verification of patient identification in real-time by comparing identification input into POCT device with list of active patients in laboratory or hospital systems
- Detection of potential interferences by combining POCT requests against laboratory, pharmacy or medical record (i.e. checking patient's last hematocrit value before conducting glucose testing, checking medications for potential test interferences, or verifying physician orders in medical record before performing test)
- Predicting future management based on past response (i.e. recommending insulin or heparin dosage based on historical dose/POCT result response for the individual patient)
- Calculated results (i.e. estimating glomerular filtration rates from the POCT creatinine result in conjunction with the patient's age, sex and weight or race)

TABLE II: Examples of POCT data management algorithms

References

1. Silverstein MD. An approach to medical errors and patient safety in laboratory services. A white paper prepared for the Quality Institute Meeting: Making the Laboratory a Key Partner in Patient Safety. Division of Laboratory Systems, Centers for Disease Control and Prevention, April 2003.
2. Commission on Laboratory Accreditation. Laboratory Accreditation Program: Point-of-Care Testing Checklist. College of American Pathologists, Northfield, IL. 2004.
3. Dyer KL, Nichols JH, Taylor M, Miller R, Saltz J. Development of a universal connectivity and data management system for point of care testing. Crit Care Nurs Q 2001;24:25-38.
4. Nichols JH, Poe SS. Quality assurance, practical management, and outcomes of point-of-care testing: Laboratory Perspectives, Part I. Clinical Laboratory Management Review 1999;13:341-50.
5. JCAHO. National patient safety goals for 2005 and 2004. Joint Commission on Accreditation of Healthcare Organizations, Oakbrook Terrace, IL. 2005.
6. Nichols JH, Bartholomew C, Brunton M, Cintron C, Elliott S, McGirr J, Morsi D, Scott S, Seipel J, Sinha D. Reducing medical errors through barcoding at the point of care. Clinical Leadership & Management Review 2004;18:328-334.