Clinical Intelligence – How clinical knowledge may be derived from medical data

Dr. Suman Bhusan Bhattacharyya

MBBS, ADHA, MBA

(2005)

Keywords

Clinical intelligence, business intelligence, data, knowledge, intelligence, sampling

Introduction

Business intelligence as applicable to clinical medicine is clinical intelligence. Business intelligence can be described as the process of enhancing data into information and then into knowledge. Data are numbers, words, images, etc., accepted as they stand. Information is simply any message the sender chooses to create. Knowledge is an appreciation of the possession of interconnected details (like those represented by information and data) which, in isolation, are of lesser value[I].

Data leads to information or knowledge that in turn lead to intelligence. Clinical intelligence is what clinicians employ to treat their patients. This is based on knowledge gathered from a variety of sources like medical education, journals, books, patients, peers, etc. They all supply both information as well as data. During patient workups loads of data are collected that are analyzed and matched up with the knowledge that the clinician's possess. Intelligence allows for accurate prediction of the possible consequences of the information that the data analyses provide.

For example, an observation leads to generation of data like a person has a blood pressure of 120/80 mmHg. It is well known that normal persons normally have a blood pressure of 120/80 mmHg – this is information. Therefore it can be concluded that the person is normotensive – this is intelligence. Each part of this example is important, without which proper decisions cannot be made.

It is plain that the intelligence and therefore the knowledge are vital for the proper discharge of a clinician's job. Unfortunately, it is these vital areas that are most ill served. The paper-based methods still in vogue in clinical medicine are loaded with fallacies that are too error-prone. Till an improved method could be located it was the best. However, with the advent of information technology, excellent tools are available to the clinicians to do a better job.

This does not mean that there are no errors at all, for improper collection of data, incorrect analysis leading to faulty information and wrong interpretations still can fail even the best of systems. With proper utilization, errors can be made to fall to a minimum and largely acceptable levels.

Methodology

The main steps are:

1. Collecting data
2. Discerning patters and meaning in the data (information)
3. Knowing what to do with it

For any knowledge to be derived, most of them have been found either due to chance findings by keen observers or due to exhaustive research. The former is largely *ad hoc*, the latter a formal process. For the formal process, the steps are:

1. Formulation of the research question at a high level, usually by a committee
2. Allocation of resources, mostly funds
3. Deciding on the question and the null hypothesis – the assumption the is to be proven or not through the research
4. Ways and means – the formulae to be used
5. Identification of target population
6. Collection of data
7. Collation of data, usually by tabulation, and after de-identification
8. Analysis of data
    a. Randomization[ii]
    b. Harvesting
    c. Calculation using the formulae
9. Reporting of data – formal publication, usually after peer review, in journals, books, etc.

Each step is time-consuming, expensive, and wherever manual, error-prone.

Let us examine the detailed steps involved.

1. Formulation of question – takes a long time due to formal processes
    a. Existing literature search – requires good library and librarian support
    b. Searching for information from peers – both local and foreign
    c. Sanctioning a budget
    d. Planning for data collection
    e. Planning, formulating and formalizing the questionnaire
2. Collection of data
    a. Logistics
    b. Appointing data collectors
    c. Actual data collection
    d. Archiving of data sheets
3. Collation of data
    a. Transforming raw data into useable format
    b. Transferring formatted data in to electronic form or paper-based form that allows easy calculations of totals, etc.
4. Analyzing the data
    a. Randomization where a statistically significant population size is considered. However, if the population-size under study is small, then all the data collected is used for study. So, calculation of a sufficient sample size with decent error margins of no more than ± 2.5%
    b. Generating a set of non-repeating random numbers that total the sample size calculated above
    c. Gather pertinent answers to the questions formulated

d. Calculate the indices as per the accepted formulae
e. Conclude on the status of the null hypothesis
f. Report on findings

5. Publication
   a. Finding the "right" vehicle for information dissemination
   b. Disseminating information at the "right" time
6. Revalidation of the findings, essential in clinical protocols and evidence based medicine

With these many steps involved, the opportunities for committing errors are substantial. With bringing in efficiency in treatment being the goal, an error-process process is the last thing that is acceptable and should be the first thing to correct as it is the step where IT can help immensely by bringing in efficiencies in the system and eliminating errors.

IT enablement

With tools provided by information technology, many of the inefficiencies that almost invariably creep in or are inherence within the process are effectively eliminated at best or very much reduced at the worst. An additional advantage is that ad hoc research becomes extremely simple. A clinician, or a group thereof, can easily form a variety of null hypotheses and draw conclusions from the electronic data bank – something that is almost next to impossible with a paper-based system.

Let us now examine the process, as it would be after IT-enablement.

1. Electronic medical records with clinical data warehousing[iii] support ready for immediate data capture
   a. Data in electronic form
   b. Data already cleaned and prepared for immediate harvesting on demand
2. Formulate questions on-the-fly, supports formal processes too by allowing the testing of the null hypothesis on demand
3. Select pre-set indices or design your own
4. Run query
   a. System will automatically perform such functions as
      i. Randomization
         1. Sufficient sample size with error margins chosen by the user
         2. Generating non-repeating random numbers
      ii. Analysis
         1. Retrieving data
         2. Performing calculations
         3. Reporting calculations as per format requested
5. Reporting – either in pre-set or customized formats
6. Quick search for online and data archive sources, even references. Requires that the application is suitably enabled to fulfill such functions
7. Revalidation can be done easily with minimum of effort vis-à-vis paper-based systems

8. Minor changes to the process or query is easily implementable, something that is difficult to accomplish in the paper-based system

Comparative Analysis

Item	Paper-based	IT-enabled
Project formulation	Formal, time-consuming, *ad hoc* analysis quite difficult	Ad hoc as easy as formal process
Observation	Manual	Manual
Data collection	Manual, entry into paper records difficult	Manual, entry into electronic records easier
Data tabulation	Manual, data retrieval challenging if entries are illegible	Automated, data retrieval easy and true
Data analysis	Manual, requires variegated analytical tools, results difficult to redo due to time taken and cumbersome process	Manual intervention, usually one application for all analytical tools and automated, results easy to redo
Results publication	Manual, difficult to reformat and redo	Automated if pre-set, interventional otherwise, easy to reformat and redo
Results review	Manual, difficult to discard and re-run	Automated and easy to discard and re-run the entire process
Process validation	Manual, time-consuming, cumbersome	Automated, easy, possible immediately

Conclusion

While it is not claimed that IT-enabled is the panacea of all that dogs the current clinical management process, its application would definitely improve the process by bring in efficiencies into the system. Important strides have already been made in almost every other industry with business intelligence being recognized the most important tool for any business aiming to thrive, if not succeed. For clinical practice to remain impervious and unconcerned regarding the existence of such a tool and continue to disregard its importance would not be in the best interests of the service neither in the short nor in the long term.

---

[i] Online Wikipedia http://en.wikipedia.org

[ii] Randomization is the process of making something random.

[iii] A data warehouse is, primarily, a record of an enterprise's past transactional and operational information, stored in a database designed to favour efficient data analysis and reporting (especially OLAP). Data warehousing is not meant for current, "live" data.