A Model for Developing Data Mining Capacity at the Coal Face

Peer reviewed research paper

Dianne Pelletier
Senior Lecturer and Deputy Director, Centre for Health Services Management, University of Technology, Sydney.
Dianne.pelletier@uts.edu.au

Donna Diers
Adjunct Professor, Faculty of Nursing, Midwifery and Health, University of Technology, Sydney donna.diers@yale.edu

Annie W. Goodrich
Professor Emeritus, Yale School of Nursing, Lecturer, Department of Epidemiology and Public Health, Yale School of Medicine and Senior Clinical Coordinator – Finance, Yale New Haven Hospital, USA.

Abstract

A strategy in New South Wales to develop the capacity of health professionals, particularly nurses, to mine hospital administrative data to support their clinical and managerial decision making was reported in 2001. (1) Health service delivery will keep its promises of efficiency and effectiveness only to the extent that how service delivery works is appreciated and understood. That has meant pushing "using data" down to the people who really understand what the data mean. This four stage model involved developing an appreciation of the potential for data use through public presentations of completed studies; using data workshops; two tiers of postgraduate subjects on methods for and considerations in using administrative and other health data; and establishing local hospital or Area Health Service based groups of users. The model involves industry-academic partnering. Evidence that this model is working is reported here. Factors that impede or facilitate the process at various stages are discussed.

Key Words:

Data mining, information management education, health data,

As the importance of information for management, planning and evaluation in the health care system is increasingly articulated (1) the development of an information culture continues to be on the agenda of New South Wales (NSW) Health Department (3). Undoubtedly, a key factor in the increasing awareness of the data available (or about to be more accessible) has been the introduction of the Health Information Exchange (HIE). While the ease of access promised is not yet achieved equally throughout the State, large amounts of the state's health data are now housed in the HIE and many Areas are using it to support their work processes.

The HIE is an enormous data depository with data down to the particular (de-identified) patient's diagnoses and treatments and demographics. Recently, the NSW Health Department has released to the Area Health Services a Microsoft Access database on CD ROM with the latest Hospital Costing Data for the majority of hospitals in the state. Data to the hospital, DRG and product line are made available. A range of reports by hospital and DRG allows the user to easily drill down into the data. The Department seems willing to share this widely and sought consultations with groups of clinicians to identify anomalies and demonstrate its functionality. Annual releases are planned. This may signal the Department's commitment to development of a more positive information culture. As greater awareness of this facility spreads, operational managers will appreciate the potential of using these data.

The dialogue about the Electronic Health Record and the Electronic Medical Record increases as the National Health Ministers Council's

HEALTHConnect initiative was recently publicized by the Canberra office at a series of forums throughout Australia and through their internet site (http://www.health.gov.au/healthonline/connect.htm). The goal of this is a national electronic health record. In NSW, the electronic record initiative NSW EHRNET works closely with HEALTHConnect and two live trial sites bring the reality of this IT project closer to the minds of clinicians and administrators.

In NSW another priority initiative is the introduction of appropriate POCCS, Point of Care Clinical System(s). As this process evolves both the availability and accessibility of information should increase and those with an interest in data mining eagerly anticipate this increased functionality.

The Data Mining Model

We have described elsewhere the early concepts that led to seeding the nursing information management capacity. (1,4) Briefly, we sought opportunities to engage nurses in using hospital administrative data for operational and clinical analyses as a way to push data-based decision support to the arenas where nursing is actually practiced. This use of administrative data is not common either in Australia or in other developed countries.

Here, we report the fully developed model with assessment of outcomes to date.

Step 1: Inspiration. The model we have developed to encourage interest and develop the skill to use administrative data at the coal face has four more-or-less sequential components. First, we have made use of established relationships between the Centre for Health Services Management, Faculty of Nursing, Midwifery and Health at the University of Technology, Sydney to provide entry to nursing groups for (hopefully) inspirational presentations of completed studies. Prof. Diers is well known in Australia and accepted all invitations to speak to local groups. We have primarily used studies brought to Australia from the RIMS (Resource Information Management System) at Yale-New Haven Hospital. (5,6) As detailed below, now that there are completed projects produced locally, it is expected that they will become the backbone of these sorts of presentations. The forums have most often been nursing inservice sessions, "professional days", or other special sessions.

Their deliberate purpose has been to open nurses' eyes to the possibilities of using administrative data. Since 2000, there have been approximately 50 of these sessions.

Step 2: Short Courses. The second part of the model is intended to provide a short – one day – introduction into the basics of health care information management with emphasis on disease and procedure coding and ARDRGs (Australian Refined DRGs). These workshops are voluntary (a fee is charged) and they also provide hands-on experience in a computer lab in working in Microsoft Access on a small de-identified patient data set. Handouts provide background information for further reading. These kinds of sessions are sufficient to equip participants with basic understanding and skill. They are especially encouraged to seek out the data managers in their home institutions and begin to establish the collaboration that allows them to ask questions of administrative data and search for answers.

We have presented seven of these one-day workshops involving around 120 participants. Participants have not been restricted to nurses and we have had a number of health information managers and others interested in using data enrol.

Step 3: Courses. The third part of the model includes two masters level academic subjects: Using Healthcare Data for Decision Making, a prerequisite for the second: Information Management Application. The first subject takes students through a workshop-like session similar but more substantial than that described above. It then expands on that knowledge base by presenting methodologies as well as discussions of healthcare data systems such as operating room information systems, emergency department systems, nursing acuity systems etc. Students are supplied a small dataset (1000 cases) and assigned work leads them to answering questions written by the faculty. Students are introduced to basic costing methods and data cleaning/auditing as well. The final assignment is oral and written presentation of their assigned projects, which they may complete individually or in groups of up to four. One additional feature of this subject is an assignment to search out an area of their institution that uses data and is relatively unfamiliar to them. They find out how data get there, and what use can be made of standard healthcare data as well as the strengths and weaknesses of the particular system. A brief written report of this exercise

leads to course revision for the next iteration. Student reports also give us a good sense of what kinds of interesting data mining opportunities might be out there.

Building on the Using Data subject is the advanced subject which includes advanced material on population definition, basic summary statistics, graphing and data presentation strategies, and writing for publication. The focus of this subject, however, is helping students craft a project they can complete in their home institutions, with a local mentor. Class sessions use group discussion to focus the projects, anticipate data acquisition problems, suggest data analysis strategies, etc. Students present their projects to an invited group of faculty and members of the various data mining groups that are the last part of the model.

Step 4: Practice-based implementation. The Centre for Health Services Management has entered into formal partnerships with four Area Health Services in NSW. One facet of this relationship is that we have worked with hospitals and Areas to create regular forums, some supported by UTS faculty in which data mining can be fostered through mutual support and interest. By 2002, formal groups with names and acronyms, regular meeting schedules and minutes are well established in Central Coast and Southeast Sydney Area Health Services. New ones are emerging at Royal North Shore, Prince of Wales Hospital and Wentworth AHS. Groups have been set up at the hospital level as well. Working with the operational managers at the coal face is of course ideal as this places the data in the hands of those who understand the nursing unit the best. Although these groups often have a representative on the area level group, the hospital groups can suffer from severe work pressures, lack of understanding and capacity (or access) to mine the data, lack of support from administration or perhaps lack of a local champion. A member of one of the Area groups conducted an Area wide exploratory study of nurse managers' experience of receiving and using casemix data. The results were not encouraging and highlighted a need for both education and greater dissemination of nursing unit (ward) level case mix data. (7)

One of the critical factors that has helped develop the capacity of the participants and spread the enthusiasm has been the annual "network feedback day". Representatives of all the groups and all other interested parties are invited for at

least one formal presentation, refreshments and lunch at minimal cost. A full day of their own presentations allows participants to showcase their projects and networking activities are always vigorous. Often there are outcomes beyond the project itself. For example, a detailed presentation of data analysis of data on pressure sores that was extracted from the Excelcare system at John Hunter Hospital was not only valuable in itself but offered another valuable spinoff. The group intends to document their findings in detail to the ACCCH, the Australian Centre for Coding and Classification in Health to lobby for development of new ICD codes that would enable more effective data on this important clinical problem.

In addition to the Network Day , network participants and the students in the Using Data subject are invited to hear the public reports of students' projects from the Applications course.

What makes it work – or not

The four aspects of the model have evolved differently. The inspirational presentations have grown fewer, understandably as the work has become visible locally. These sessions continue to catalyze interest in the process. The one day workshops are now attracting fewer numbers which may indicate that saturation of the market for that level of skill development has occurred. We are also finding that while there may be considerable interest in attending (or by administrators in sending staff) the fiscal constraints and more importantly the staffing constraints are precluding attendance for many.

Where real evidence of "cutting edge" information management development is showing is at the graduate subject level. The first subject in the series continued for the third year to attract about 20-25 students, most with an active interest in using health care data. Increasingly, the students come from institutions which are part of the networks described above. Indeed, we are finding that the Using Data subject is beginning to attract nurses who have an interest in data use to enter formal postgraduate degree perhaps for the first time. The hands-on experience of using the UTS hospital data base to answer meaningful questions works well. Students have often used web resources to provide the context for their projects and to flesh out what is admittedly a very small database for concluding anything. Some students have become very creative both in searching out

tangents to their chosen topic and in developing clever presentation strategies.

The relationships built with coders and casemix staff during student projects are critical and can be helpful to both sides. One participant took her new coder friend on a tour of the ICU and in doing so the coder came to understand the ICU flowchart which then was used to facilitate coding of ventilation hours.

In addition the students have learned about a whole new dimension of data use in their institution. This of course raises for them further questions of how it is currently used, by whom and possibly how it could be used. Interesting issues such as access to the data and gate keeping also emerge. Overall, the subject tends to demystify DRGs and casemix, terms students know from reading the newspapers if not from their own clinical experience. They come to understand that they can become expert in "reading" casemix data very quickly.

A particularly enriching learning strategy has been to bring into their classroom clinicians and managers who are actively using data to influence their practice or their work environment. These individuals are often members of the networks created by the larger model. They have graciously given of their time to help spread the word. Students can see live examples of colleagues just like themselves actually doing data-based studies in real life, with real success.

The second subject in the series understandably will always attract smaller numbers of students who wish to conduct an in depth data exploration project in their work environment mentored by an appropriate work associate. Each of the two years this subject has run, four to five students have participated. The first year's experience highlighted the importance of finding and getting the right mentor for the student. The identification of such a person lies with the student; the academic staff help with explanations and clarification. Enormous difficulties can occur for the students if access to the data is not relatively easy or facilitated by the mentor. As the HIE transition was in full swing when some of the first year students were attempting their projects, difficulties arose in some places. It was evident that the personalities and willingness to assist influenced the gate-keeping role. Some argued that students, often employees of the institutions, could not have access to de-identified

administrative data subsets without full clearance by an institutional ethics committee. Others downloaded such subsets into a data base for the students without twitching.

Examples demonstrate the diversity and the value of the projects themselves. One student explored in depth 2 years of data on the ARDRGs related to False Labour, a categorization giving some difficulty to clinicians and administrators in the obstetrical domain. Immediate local level outcomes were refinements of coding practices, greater understanding of the DRGs by the Casemix Branch (which refines DRG-related data) and a recommendation for a change of coding is being developed for submission to the National Centre for Classification in Health who are responsible for the coding classification system. They welcome input from consumers. As a result of this project this student has become involved in managing her hospital's adverse event data base. Another student had an inordinate amount of difficulty accessing data, following the usual processes. Her mentor became aware of a systematic problem in the institution with data management and, while the student's findings were compromised, the exercise revealed ways the institution could improve its data quality and management. That is a recognized spinoff of actually *using* healthcare data. The data will never improve until they are used.

The second offering of the subject, in 2002, attracted students who, with one exception, were part of the networks and were therefore quite advanced and independent. Not surprisingly, their projects were larger and deeper and more complicated, and the potential for actually influencing practice, greater. Interestingly, in this second group, the projects focused more often on systems issues – waiting time analysis, bed flow, operating theatre utilization, nursing unit casemix – than the clinical problems that had focused the students in the first subject offering. The intention is that we support these students in preparation of publications of their findings which offers a further mechanism for dissemination of the effective use of data.

The Value and Output of the Networks

There are also specific lessons to be learned from the Network groups. At the Area level the early

group RIN, Resource Information Network (South East Area Health) offers several. About two years into the work, the group was floundering as members, while finding it valuable, interesting and supportive felt that what they were doing in this context was unrecognized and just an added extra to their working day. This was resolved by re-clarifying the group's purpose and eliciting overt and strong support from senior Area nursing executive. The more focused and action- related minutes are now copied to the Area Director and she has placed this group's activities on the Area Directors of Nursing monthly agenda. Each site reports at each RIN meeting which helps to keep these activities centre-stage with management, leading to greater recognition and support from management to the members. New members or transient members join the group as projects or ideas come up in their areas. If there are useful databases available the "owner" or other key player can be invited to present its use to the group or to join the group.

The RIN group has been formally supported by a faculty member at UTS (DP) whose attention, minute-keeping, stroking and pushing have been essential to building a group whose energy now can carry them along. Participants already have full-time positions with competing demands and it has been critical to have a commitment from the University as an additional resource to this group.

It is very important that members of the casemix units are active participants in these groups as they have the best understanding of how their local institution is handling the casemix data and they can often facilitate access to it. Medical records coders can be useful members but often find it difficult to attend. Best use of them may be on an as-needed basis or through direct consultation. Certainly, it is useful for the medical records staff to aware of the group's activities and interest in data.

On occasion the RIN group itself will decide on a focus of interest and all sites, as appropriate will extract the same data for their hospitals for comparison. Procedures and protocols will be compared in line with best practice. Currently this exercise is being undertaken for acquired deep vein thrombosis, with particular interest in oncology patients. The interest lies in the consistency of "best practice" for acquired DVT management noting that this will differ for oncology patients who

may be treated for DVT outside of their specialty clinical area.

A new hospital level network group is emerging: POWRING, Prince of Wales Resource Information Network Group. A key feature of this group which should serve it well is the careful selection of participants in terms of expertise, seniority and interest. This group has strong support from nursing management. At their recent accreditation visit, this institution was commended by the Accreditation team for having a hospital level network group.

In the Wentworth Area, both the Area Director of Nursing (DON) and the DON at Nepean Hospital so strongly support the notion of data mining that a new group WING (Wentworth Information Network Group) emerged spontaneously at the end of a one-day intensive session. As it happens, several of the members of this group were former participants either in workshops or in UTS subjects. On the other hand one local group elsewhere foundered in part because of the personal presence and commitment of the nursing executive, which made the participants hesitant to speak their minds. This group is now being resurrected under the voluntary leadership of one of the Nursing Unit Managers who is passionate about using data and who recently completed the Applications subject.

Another Area level network group whose experience is instructive is INCC, Information Network Central Coast. Again, strong support from area executive Director is invaluable. A local champion of data mining is a strong factor here and the group has a strong core of technical expertise which has served them well. A useful tip from this group is the value they have found in dumping all the data elements that they identified as commonly of interest from the HIE into their own separate Microsoft Access data base which allows them free range to mine this according to their changing interests. The database is now automatically refreshed every month.

We have always been quite clear that data mining is inherently multidisciplinary and in this instance the disciplines that are crucial to the success of efforts led by clinicians are the medical records professionals and the casemix units. It had been very touching to be present in a meeting in which the casemix unit staff turn to the nurses to interpret what "synovial fluid" means, and then to have the

nurses ask just how a particular data element reaches the end-point system. It is not unknown for hospitals to lose their casemix staff completely at times and these staff shortages affect the access and use of data particularly where the group member is relying on external expertise to assist.

Another strategy that not surprisingly is useful in moving projects forward is the allocation of resources, usually in allowing participating staff some part of their working time to be dedicated to the data mining activities. Productivity in the Network group is greater for employees whose work involves data use. None of the activities involved in this model required additional resources, apart from Prof. Diers's visits. The administrators who support the network groups are coming to realize that the half day per week a member might be allocated to data mining might pay off considerably if length of stay can be reduced or staff satisfaction increased.

Other outcomes

Three abstracts from members were accepted for the 2002 Casemix Conference. One paper has been accepted for publication (7) and two others are in preparation. One INCC member has been invited to present his cost modeling project at senior level to at least two other area health services. The Networks continue to increase in both number of groups and projects.

The industry partnerships between the University and the Areas and hospitals that are evolving here are a particular commitment of the University of Technology, Sydney. It is expected that as health services research develops, the network of relationships established here will be a strong support for more traditional academic activities. Members of the Area network groups are offered formal appointment to the Centre as Fellows and Senior Fellows, invited to professional and social occasions, and counted upon for advice and counsel. In 2003, financial support is being offered to participants who submit abstracts to conferences to showcase their results. Chris Conn, Fellow of the CHSM and RIN member is being supported to present data mining findings on monitering geriatrics at the NSW Health Informatics conference. Undoubtedly such tangible support will encourage data miners to spread their findings.

Conclusions

It has taken three-plus years for the activities described here to deliver which, given the competing demands on participants in a rapidly changing healthcare environment is not long. Participation is voluntary, sometimes barely recognized and not rewarded. The production of the group members (and of affiliated others whose stories we have not included) is extraordinary. One lesson to be taken away is that the use of casemix information at the coal face is entrancing to those who have rarely had much data to support their clinical and operational work. The excitement that follows from this work has proved to be its own reward and the healthcare delivery system is changing a little bit for the better.

References
1   Diers D & Pelletier D. Seeding information management capacity to support operational management in hospitals. Australian Health Review; 2001b; 24,2: 74-81.
2   Ayers D P L. Clinical information systems. In J. L. J. Onley (Ed.), Clinical Challenges 2000. 159-173 Allen & Unwin, Sydney.
3   Menedue J. Report of the NSW Health Council - A Better Health System for NSW, 2000. NSW Health. Available: http://www.health.nsw.gov.au/health-public- affaires/wayforward/hc_report.html
4   Diers D & Pelletier D. From I.T. to information management with casemix data. Australian Health Review; 2001a; 24.1: 62-67.
5   Diers D, Weaver D , Bozzo J, Allegretto, S & Pollack C. Building a nursing management analysis capacity in a teaching hospital. Seminars for Nurse Managers; 1998; 6,3: 108-112.
6   Diers D & Bozzo J Using administrative data for practice and management. Nursing Economic$; 1999; 17, 4: 233-236.
7   Blay N & Donoghue, J The provision and utilisation of casemix and demographic data by nursing managers in seven hospitals. Australian Health Review 2003; 26,1:209-18