

Common Information Management Problems which can be solved with the use of Information Technology (IT)

by Robin Beaumont

e-mail: robin@robinbt2.free-online.co.uk

Monday, 01 May 2000

Contents

1. INTRODUCTION.....	2
2. YOU WISH TO WRITE A PAPER FOR A POPULAR JOURNAL.....	2
3. YOU WISH TO SEND A REPORT BY E-MAIL.....	3
4. YOU WISH TO COLLECT AND ANALYSE SOME DATA FROM A SET OF PATIENT NOTES FOR A ONE OFF STUDY.....	3
5. YOU WISH TO COLLECT AND ANALYSE SOME DATA PROSPECTIVELY FROM SEVERAL CASUALTY DEPARTMENTS	6
6. YOU WISH TO OBTAIN AND ANALYSE SOME ELECTRONIC DATA FROM YOUR DEPARTMENT	8

1. Introduction

During your time working as a SpR and consultant you will come across situations involving some aspect of information management that could be aided by the use of a computer. Often, because of circumstances you will prefer to do things manually. However, while this may offer a temporary solution in the long term, it will be very time consuming.

For example, deciding that it would take too much time to learn to use a spreadsheet you collect all the data from a set of questionnaires on cards. While you manage to fill in the cards quite quickly, you spend an inordinate amount of time sorting them into various piles on your dining room floor working out tallies. Eventually when you see a statistician you realise that she will only accept the data in an electronic format.

This handout looks at several clinical scenarios where appropriate technology may help you become more efficient and effective.

2. You wish to write a paper for a popular journal

This will undoubtedly be done using Word for windows or the word processor with which you are most familiar. Some journals specify a particular format for example the BMJ provides detailed instructions, some of which are given below:

Our preferred format is a 3.5" floppy disk with the file written in WordPerfect 6.1 for Windows. If you can not provide your file in our preferred format we should still be able to process it provided that you supply us with sufficient information to enable us to convert it into our format.

We can read DOS, Windows, and Apple Mac disks in many word processing formats including Word, WordStar as well as plain ASCII. If you have to submit your article in any of these formats (that is, not WordPerfect for Windows 6.1) then please provide us with one copy of your file in its "native" format and one saved as a plain text file with no formatting information added (that is, as an ASCII text file).

Please use these guidelines when preparing your file:

- Put all text (including references, legends, abstracts, etc.) in one file and all tables and figures in another
- Prepare your article in the simplest possible form. We shall put in the correct font, type sizes, column measures, etc. later. DO NOT use the automatic formatting features of your software, such as hyphenation, endnotes, footnotes, headers, footers, etc.
- Layout and punctuation of references are very important and follow the Vancouver style:
1 Surname AB, Surname CD. Article title. *Journal Title* 1994;2:165-8.
2 Surname EF. Chapter title. In: Surname GH, ed. *Book title*. Town:Publisher;1999.
- For journal references do not use any spacing between date, volume, and page numbers. Journal titles should be in italic and abbreviated as in *Index Medicus* (if not in *Index Medicus* they should be written out in full).
- See the BMJ for styles of headings within text.

Label your disk with: the number of the article (allotted by BMJ) the article's title and name of first author; the computer type (IBM compatible or Apple Mac); the word processing program (and version); and the filename(s) to be found on disk.

Etc.. etc....

Notice that the above says very little about figures of any type, if you are planning on including such things as charts / graphs etc., it is essential that you contact the publishers directly.

3. You wish to send a report by e-mail

Most research teams now a days pass drafts of documents by e-mail to one another. This is achieved by sending them as 'attachments', which can be of any type. Because such documents may be large it is sensible to remember that if colleagues are using modems at home to collect the attachments, particularly Word documents containing images, charts etc, it can result in a long, expensive phone call. As a guide, you should avoid sending anything as an attachment that is bigger than about 500k (about half the capacity of a floppy disk).

If you are planning to pass documents between each other make sure that you all agree upon using a standard set of software packages **and versions right from the start**. A common problem is people having different versions of Word for Windows. Suppose you have a more recent version of Word for Windows than your colleagues. This means you can read a word document created in an earlier version but if you update it your less trendy colleagues, with older versions of the software, will be unable to read them. The solution is to use the 'save as' option to save the document as a version appropriate to your less techi colleagues.

This problem occurs with most software so beware.

4. You wish to collect and analyse some data from a set of patient notes for a one off study

That old war horse? There are two overriding principles here:

- Parsimony / Ockham's razor - That is keep it simple.
- Familiarity - I cease to be amazed how people push particular software packages beyond what I would ever have imagined them capable of. For example someone who is only very familiar with Excel would use this software package rather than a statistical (i.e. SPSS) or database package (i.e. Access) for tasks that might be more easily accomplished in them.

Considering the above problem and using a 'systems framework', that is thinking about the various aspects of the information Inputs, Processes and Outputs we will need to decide:

Inputs - What? - Assume that it is data from a set of approximately 200 case notes,
Who? - Assume it is you, who will also interpret it.
When? - Most likely whenever you can grab the time or do you have a deadline?
How? - Do you have keyboard skills or availability of a suitable computer / laptop while your looking at the notes or off site?

Processes - Do you just want descriptive statistics, basic inferential statistics or something special (i.e. Bayesian forecasting)?

Will you be passing the data on to anyone - managers, statisticians. Is it safe to assume not?

Is it the first stage in something else? Is it safe to assume it's not?

Outputs - Is it just for you to re-organise into a paper or the Trust Management Board or an audit committee etc.

Considering the above scenario it is important to realise that there is no one correct answer. Often the overriding factor is familiarity with a particular method /software for the person carrying out the task. For example if you were a public health doctor you would most likely choose an excellent free software package called 'Epi Info' for the above problem. However, we will assume you have just been on an introductory course in Health Informatics.

You are therefore familiar with Excel, Access and SPSS. Lets consider some important points from the above:

The amount of data to enter is small (200 records). So the first question is, which software package will cope with this amount of data. The table below shows the limitations of the various common packages.

Software package	Number of records	Number of fields	Comments
Excel	16,000 maximum in one worksheet	250 (approx.) maximum in one worksheet	You can create user friendly front ends if you want to. It also contains a number of statistical functions. Good for graphs (but not for full EDA). Not very good at preventing you from entering duff and unstructured data
Access	no limit	250 (approx.) maximum in one table	Can easily link tables together. Few statistical functions. Good at preventing duff data from being entered. Can produce good reports to give back to interested parties.
SPSS	No limit	No limit	More difficult to get a copy? New version will have a 'forms' front end. Good at forcing the entry of structured data, by defining variable types and labels. [hint; choose Menu option Utilities -> value labels. then right mouse click cell.] Must use this if you want to do sensible statistical analysis, encourages a sensible approach to data analysis (i.e. EDA then full statistical output of inferential statistics).

How easy does it need to be to enter the data is another consideration. As you are the only one entering the data, you don't really need a pretty front end, such as Access forms, to make data entry user friendly.

Taking the above considerations into account, You initially think Excel would be the easiest, partly because you already have it on your lap top.

However, after entering a few records you begin to realise that you are entering too much textual information. You realise that you need to develop a paper Data dictionary/ coding frame (see notes on main web site: <http://www.robinbt2.free-online.co.uk/virtualclassroom/contents.htm>) to allow you to enter more numerical information which you will be able to analyse afterwards. You start to do this and after two weeks you come up with something, you think will allow you to collect all the data you want. Part of the data dictionary is given below.

'field'	Database data type	Unit/Coding scheme	Example	Comments
ID	Number Integer	-		My own number to cross reference with Hospno
Hospno	text (size 12)	Local hospital number		
Diagnosis1	text (size 5)	ICD-9 999 = missing		Keep book with me.
Height	Number Integer	- Unit: Centimetres 999 = missing		
Sex	Number Integer	- Coding scheme: 1=male 2=female 999=missing		
X 165 to make 270 fields in total				

Because you now have 170 fields defined for each record you decide that it would be better to enter the data into Access so that you can set it up to force you to enter particular values into each field. You create a table with 170 fields and proceed to enter the data directly into Access.

You now begin to start entering the data. You set up your desk with:

- your laptop
- the ICD-9 book
- a paper copy of your data dictionary
- several sets of the notes

You work like this for the next month and collect data from over a hundred sets of case notes by which time you rarely refer to the dog eared, commented, data dictionary. At the start you added a few extra fields to the table for textual comment but otherwise nothing else has changed when disaster hits! Your husband tells you that he needs to have the laptop back at work. At least you feel moderately happy as you have been taking regular backups so have everything on floppy disk. You remove the database from the laptop before giving it back, as your husband travels about a lot and has already had one laptop stolen.

Unfortunately, there are no suitable computers at work so you arrange with him to borrow the laptop at weekends. You realise that you need to develop a paper form to fill in to abstract the relevant information from the notes. Such a form will allow you to enter the data quickly into the computer at the weekends. Once again because you're the only one filling in the form you decide that just a series of columns on a sheet of paper is sufficient. You discover a quick way of doing this using the Report wizard in Access of which details are provided below.

Tip: How to create a paper based data entry form for an Access Table

This consists of two stages, firstly creating a new table based upon the structure of the one with that contains data and secondly creating a report for you to be able to use for collecting the data:

1 Creating the new table

Copy the table by selecting the table, (you only have one in the database) using the menu options edit -> copy then edit -> paste Choosing paste options structure only. And give the new empty table a different name.

2 Creating the paper data collection form

Now to create the report your going to use as the paper data entry form:

Choose the report tab and click the 'new' button. Select the new empty table to base the report on.

Choose the report wizard option.

Choose the single column report type. Click OK.

Select all the fields to go into the report. Click next.

Ignore the sort option. Click next.

Choose report style as ledger. Click next.

Accept the title or change it. Click finish.

The report will now appear

When you try to close the report a message appears 'Save changes to report1' Click OK else you will loose all your work. You will then be presented with a 'save as' dialogue box in which you should type an appropriate name.

5. You wish to collect and analyse some data prospectively from several casualty departments

You are part of a particularly keen A & E regional set-up and have been asked to take on the co-ordination of a multi-centre prospective study concerned with head injuries.

The study is ultimately concerned with the development and assessment of a tool for the management and prediction of acute head injuries. Such a device will support doctors decision making regarding rationale management of these patients including, investigations, admission or discharge along with prognostic details.

The first phase of the study, which is mainly your responsibility, is to devise appropriate data collection techniques, which will be used in the next phase by the statisticians / computer developers to develop the management / predictive device.

The project has many political agendas including:

- Several sites, two of the five, have only agreed to be included if data is regularly fed back to them in appropriate format to allow them to use it as part of an internal audit mechanism.
- Your boss, who is more interested in the next phase of the trial, is concerned that the sites are bothered as little as possible during the 'necessary' data collection phase. He don't want to find any of them dropping out.
- The money has already been awarded to the project team that consists largely of pure scientists rather than clinicians. Their contracts run out in two years.

Considering the above problem and using a 'systems framework':

Inputs - What? - Data from possibly 50 patients at each site (considering a period of 6 months data collection).

Training will be given to; 1) casualty officers in the correct use of the forms. 2) casualty administrative staff concerning obtaining / processing completed forms.

Who? - Casualty officers then you.

When? - Whenever a head injury turns up and the appropriate casualty office is on duty.

How? - Casualty officers on specially prepared paper forms. Returned to you to interpret and enter into a central database. Database query results are sent to the statisticians.

Processes - Do you just want descriptive statistics, basic inferential statistics or something special (i.e. Bayesian forecasting)?

Will you be passing the data on to statisticians yes

Paper (and electronic?) reports will be developed to provide appropriate reports back to the casualty departments.

First stage -> must keep people on board.

Outputs - Statisticians, Casualty departments.

This is a complex project with many inter-linked activities. Clearly, the reports that the various casualty departments have requested could become a project in itself. Yet, although this is an insignificant component to the academic success of the project it is very important for keeping the departments on board. Furthermore, those clinicians involved in the project may feel that the ultimate aim, the development of a system which will relieve them of some tasks, is inappropriate.

The software you decide to use is Access. The reasons for this are:

- Your familiarity with the software package
- You don't know the number of records you might end up with at the end of the data collection period,
- Your not sure if in future the data collection process may change significantly.
- Your boss is very interested in collecting the data via PCs on site which would require the development of a user friendly front end by creating forms etc.

You decide to ignore initially the last reason above. This is because of time constraints and, at least initially, you will be the only one entering data. Therefore entering data directly in 'datasheet' view of the table(s) will do for now.

You begin developing a set of paper data collection forms that you pilot in each of the casualty departments for a month.

Over the same time period you refine a paper data dictionary which you use as the basis for the Access database.

After the first month of pilot data collection you come across the following problems:

1. Far fewer are being returned than the casualty officers claim to be filling in.
2. The forms are often only partially completed.
3. One casualty department stores them for a week or so to photocopy before sending on.
4. The process of entering completed forms into the database seems to take much longer than you expected.
5. One casualty department is expecting a report on data supplied within the next two months.

You visit each of the sites and realise that they had received too little training. You also decide to visit the sites in the evening as well. One site had run out of forms and didn't realise that they should have asked for more. You also speak to the casualty sisters to try and encourage better return rates. You speak to the audit department at the hospital who is expecting the report and end up sending them a copy of the data collection form asking them to specify, by way of a drawing ('template') what type of report they would like.

To speed up data entry you decide to revise the paper data entry forms with coding boxes down the margins. You then use the paper data dictionary to decide upon what to place in each box. You code the forms during your spare moments at work. You make use of a spare laptop to enter the data using a colleague / friend to read out the values in each box, and find you can enter data in a fraction of the time.

After approximately three months your department decides to buy a scanner with the appropriate software for creating questionnaires that can be scanned. You take a look at the set of manuals that came with it and decide that it may be a possibility in the future. At the same time the audit department obtains a student on a three month placement whose project is to develop a patient satisfaction questionnaire for the outpatients department. By the end of three months the student is using the scanner successfully developing a version of your questionnaire for you as well. Unfortunately the after a few tries you realise how complex the scanning process it and decide to abandon it as the student has left and it appears that the course is unable to supply any further students.

6. You wish to obtain and analyse some electronic data from your department

You work at a progressive hospital which encourages doctors to make use of the data collected electronically from past patients. You wish to carry out an audit of patients under the age of 70 who have been admitted with some type of cardiovascular condition over the past ten years. You are interested to find out what, if any, surgical procedures they underwent, who they were admitted under and who they ended up under, how long they were in hospital and their outcome. You also intend to look at a set of biochemistry results from a subset of the patients. You also wish in the long term to develop some type of predictive model of future bed need in the hospital.

Inputs - What? - Electronic data from a unknown number of patients.
Paper based data abstracted from old case records for biochemistry results.

Who? - You.

When? - Spare time

How? - You specify electronic data required from IM&T department. Liaise with medical records for old case notes

Processes - Do you just want descriptive statistics, basic inferential statistics or something special - Your planning on doing various regression analyses, recommended by the university statistics department. You also plan to use SPSS to carry out the analysis.

Outputs - Audit department, Clinical directorate, Presentation to management board.

Your first problem is to specify your needs to the IM&T (Information Management & Technology) department. Luckily, they have developed a data request form, along with an explanatory booklet of which you make full use (an example is given on the following page).

You are lucky that the department has a computer with SPSS on it and also there is a colleague in the hospital who is a local SPSS guru.

You make use of the hospital number and name fields supplied to you to request the relevant hospital notes from the medical records department to obtain the additional biochemistry data. This data you enter directly into SPSS into another data file as the guru informs you that it is easy to join datasets on a key (i.e. hospital number).

After six months you have completed most of the data analysis and now wish to present your findings to the regular medical audit meetings. You decide to use Excel to prepare the graphs as they appear more sexy than those produced by SPSS and import them into a PowerPoint presentation. Everything goes fine. You are then asked to prepare a report for the management board which you do so using Word for Windows once again importing the Excel graphs.

Extract from a data request form:

Name and position of person requesting data:

Reason for request:

Media on which data is required:

3.5" disk / Paper / Zip disk / Tape (specify type)

If electronic please choose format:

Excel (version 4)

Dbase (version 4)

Access (version 2)

Fixed width (not recommended)

Comma delimited " = qualifier

Comma delimited, no qualifier

Tab delimited " = qualifier

Tab delimited (good for importing
into SPSS) no qualifier

SPSS (*.sav format) - not always available.

Comments:

Data details:

Time period for which data is required:

Please list up to five fields you would like the data filtered (e.g. a particular consultant ID, diagnosis etc.)

(a full list of fields can be found in the accompanying booklet)

Please list the fields you require details of (a full list of fields can be found in the accompanying booklet)

Additional details: etc. etc.

Confidentiality agreement. Signature.

Document info:

Robin Beaumont Tel:0191 2731150 e-mail: robin@robinbt2.free-online.co.uk

Source: Laptop; C:\Hlcourseweb new\chap5\s6\whichapp.doc

Date: 01/05/00