29

Information technology in research

RITA AXFORD, GARY GRUNWALD AND ROB HYNDMAN

This chapter discusses computers and information technology (IT) in relation to the conduct of health care research. Research is a cyclic process which varies in terms of the number, type and complexity of activities depending upon the question(s) posed, methodology used and outcomes expected. In this chapter, discussion is limited to two major groups of research activities—identifying the background knowledge base for a research project and understanding the data that result from the undertaking. The many complex issues relating to IT and the use of client data for research are beyond the scope of this chapter and are only briefly mentioned.

The first section explores the use of IT in relation to how researchers find and access previous work upon which their thoughts and activities build. Some examples of electronic literature searching tools are explored.

The next section looks at collecting, organising and analysing data. These activities may be thought about in terms of how researchers change data into information. Specific types of computer-based data analysis packages are described and some questions to consider when selecting such packages are posed. The chapter concludes with some general comments about computer assisted analysis and research outcomes.

The use of IT in accessing the literature

Health care information exists in both paper-based and computer-based recorded forms. Health care information appears in formally published literature in journal articles, books, and conference papers. Other kinds of documents that help researchers access the literature include indices to and abstracts of these published documents.

Health care information also consists of important, but more difficult to access, informal literature. This includes unpublished, noncopyrighted reports, databases from other studies and surveys, and patient care protocols or facts relating to practice, education, research or management in health care.

Bibliographic and library information retrieval systems are important research tools and also exist in both paper (book) and computer form. Computer based retrieval systems may be

available on-line or stored on CD-ROM or magnetic tape. They are extremely powerful tools because of their ability to be updated rapidly and accessed widely. Computer networks allow many users to access the same tools and information from a variety of geographic locations at virtually the same time. These networks may be local in nature (local area networks or LANs) or may be extensive, even world wide and include use of the Internet (see Chapter 10).

Electronic bibliographic retrieval systems enable researchers to identify documents and the location of these publications. Some systems contain more than just the citation and allow the user to access an abstract or even the entire text of a document via their computer terminal or printer. These tools vary according to the scope and content of their databases. The most common ones contain citations to published journal articles, books and monographs. Others contain references to newspaper articles, technical reports, unpublished theses, audiovisual materials, computer software packages and other media forms.

Bibliographic databases vary in their focus. For example, citations indexed in a given database may cover a specific topic, such as gerontology, bioethics, or toxicology. The collection may be limited to a specific type of literature (eg, journal or newspaper articles only), a specific time period (publications since 1980), or certain languages (such as English only or English, German and Russian). Similarly, databases vary in terms of their organisation. One may be able to search according to an author's name, a key word or topic of interest, titles of articles or books, citations (or references to a given writer), or other such categories.

When searching a database for a topic of interest, knowledge about the specific vocabulary used in the database is essential. Databases require the user to describe their topic using only words specified in a thesaurus of terms. For example, information relating to cancer may be located only through use of the designated term "oncology" or "neoplasm" depending upon the terms in the thesaurus for the database.

Searching the database is best undertaken using a methodical search strategy. This might include the following steps: (1) decide the topic; (2) check the thesaurus for the term or heading; (3) if found, check for related topics and conduct the search; (4) if not found, cross check for related terms and depending upon success at this step, either proceed, rephrase the heading, or reconsider the topic. This search strategy can be used for both paper and computerised databases. Although paper databases can also be searched in non-methodical ways by browsing through the pages, computer based searches are powerful tools because they are methodical. Users enter commands which appropriately broaden or restrict the search to identify the most relevant literature for their purpose. For example, the words "and" and "or" are common commands for many bibliographic databases. A command which asked the database to search for "neoplasms or adolescents" would identify all publications in the database which relate to neoplasms and all which relate to adolescents. A command which stated "neoplasms and adolescents" would locate only those references which referred to both terms, ie adolescents with neoplasms. In this case, a much more specific list of resources would be generated.

Many electronic bibliographic databases are designed to be searched by the researchers themselves. Others require the skills and experience of specialist librarians. Many times it is some combination of both that results in the best access to the research literature. One can improve ones searching skills by trial and practice. Similarly, one should not be afraid to ask for help.

Health-related electronic bibliographic databases

The following electronic bibliographic retrieval systems have been selected either because of their wide use or their unique capabilities. This is neither a comprehensive list, nor a recommendation to use only these. An introduction to some common and some unique databases will hopefully spark further inquiry.

MEDLARS (Medical Literature Analysis and Retrieval System) is one of the largest databases in the world. It is a computerised system of databases and databanks offered by the National Library of Medicine (NLM) in the United States. The NLM enters into bilateral agreements with public institutions throughout the world to serve as International MEDLARS Centres. These Centres assist health professionals to access these databases. The National Library of Australia in Canberra, A.C.T., is the International MEDLARS Centre for Australia. MEDLARS is a set of over 20 databases and was begun as MEDLINE in 1966 as a means of preparing the *Index Medicus*. It contains over 7 million references to the journal literature from over 3,800 biomedical journals from 1966 to the present and is updated weekly. The system is searched using a unique vocabulary called MeSH (Medical Subject Headings). This consists of a hierarchical tree structure with broad categories and a large number of headings and subheadings (some 17,000). The database search provides a citation that lists author, title, publisher, key words and in approximately 65-70% of the cases also provides an abstract.

GRATEFUL MED is a software package that allows one to access MEDLARS directly from their personal computer any time of the day or night at very low cost. It is complemented by an electronic ordering system so that subscribers can obtain the full text of many articles. This system is called LOANSOME DOC.

CINAHL refers to the Cumulative Index to Nursing and Allied Health Literature and is another bibliographic retrieval system that provides on-line searching. This database is housed in paper form and electronically on-line, on magnetic tape, and on CD-ROM. It contains references to over 580 serial publications on nursing and allied health dating back to 1982. It also provides access to health care books, nursing dissertations, selected conference proceedings, standards of professional practice, the serial publications of the American Nurses' Association, the National League for Nursing, the U.S. Department of Health and Human Services' Division of Nursing, and nursing associations from all 50 states in the U.S.A. CINAHL is updated bimonthly. It can be accessed through the Internet and is also available commercially through two vendors, Silver Platter and Paperchase. As from June 1994 CINAHL is including the bibliographic citations for the references from articles in selected journals.

HealthROM is the Australian reference source on public and environmental health in electronic form for PC (DOS and Windows based) and Macintosh computers. It contains information on public health policy, HIV/AIDS, drug and alcohol use, food and nutrition, casemix, therapeutic drugs, family health, clinical and nursing practise, aboriginal health, health economics, and sports medicine. It is produced by the Department of Human Services and Health in collaboration with the Australian Institute of Health & Welfare, the National Health and Medical Research Council (NH&MRC), the National Library of Australia, and the Alcohol and Other Drugs Council of Australia. HealthROM contains bibliographic citations plus the full text of over 100 Australian health publications plus the following journals:

Australian Prescriber, Communicable Diseases Intelligence, Australian Casemix Bulletin, and Health Expenditure Bulletin.

ERIC stands for Educational Resources Information Center and is a paper-based and electronic bibliographic retrieval system developed by the U.S. Department of Education. It consists of two databases. One covers periodicals indexed in the *Current Index to Journals in Education* and the other, *Resources in Education*, includes other printed documents on education including complete coverage of all educational literature, including that relating to a number of health care professions. It is available through the Internet and vendors such as DIALOG and BRS.

SOCIAL SCISEARCH (Social Sciences Citation Index) is an electronic bibliographic retrieval system prepared by the Institute for Scientific Information. It consists of a database with citations to over 3,700 journals in the social, behavioural, and related sciences (which include health care disciplines). In addition to the usual author, title, and source listings, SOCIAL SCISEARCH produces a list of references cited in the indexed journals. This database is also accessed via commercial vendors such as DIALOG and BRS.

The Virginia Henderson International Nursing Library is a recently developed "electronic library" accessible through the Internet. This unique computerised collection of databases and knowledge resources incorporates the Directory of Nurse Researchers and includes biographical data about registered nurse researchers, projects funded by Sigma Theta Tau International, and the *Online Journal of Knowledge Synthesis for Nursing*.

Research data

Thorough access to the relevant literature is fundamental to identifying good research questions which build upon previous knowledge. With the widespread introduction of electronic databases into libraries, researchers now have worldwide access to the literature.

Collecting and interpreting data for new research undertakings is the second major area of the complex process of research that this chapter explores. It should be noted that many data collection tools and instruments can be located through electronic databases. Research data may be collected expressly for a given study, or they may be extracted from data housed as client data or as care protocols within currently existing clinical information systems. There are many important issues relating to IT and the conduct of research that are beyond the scope of this chapter. These include issues about organisation of data, data standards, data sets, language and taxonomies, integrity and quality of data, confidentiality and the ethical use of health data for research. (see Chapters 8 and 25).

The purpose of this chapter is more technical. This discussion will therefore focus on how research data are handled in order for them to assist researchers in answering their questions. This is not to detract from the importance of these issues in health informatics.

Software for handling research data

There is a range of computer software packages designed to help researchers store, manage, analyse and interpret data. For quantitative studies, a statistics package can play an important part in determining the value and effectiveness of the study. While a more recent

development, computer programs for handling qualitative data likewise are proving to be valuable research tools.

Data entry

Bioinstrumentation may allow for quantitative data to be collected and entered directly onto a computer, such as with cardiac monitors and blood chemistry analysis. Quantitative data are also often collected in the form of questionnaires, surveys or in experimental log books. In these cases, getting the data from paper into a computer is a necessary step before much use of the data can be made. A popular approach is to use a spreadsheet package (see chapter ??) for data entry. However, data may then have to be transferred to a statistics package for analysis. If this approach is followed, it is important that the statistics package be able to input data stored in the format of the spreadsheet. It is very frustrating to have data stored in one format but not be able to access it with other programs.

Alternatively, some statistics packages have their own spreadsheet-like interface which can be used for data entry. While this saves the problem of transferring data between programs, it has its own problems. The spreadsheet interfaces offered by many statistics packages have limited facilities compared to full spreadsheet packages. For example, one sometimes needs to record the same value (e.g. dose of drug) many times, but some packages require the value to be entered repeatedly rather than allow automatic repetition.

The best data entry facilities allow error checking as data are entered. For example, entering an age of four years for a married man should trigger a warning message. Most packages do not have these useful checks, however.

Qualitative research consists of data gathering strategies that generate narrative as opposed to numerical data. Qualitative data can take the form of verbatim transcripts of interviews, field notes, and reflective journal entries. Most of these activities can be greatly facilitated by data entry using electronic word-processing capabilities.

Data management

Qualitative research methodologies often require concurrent data collection, entry, coding and analysis. Coding of qualitative data consists of examining the text line by line, and in some instances word by word, to abstract meanings or categories that describe situations, processes, or concepts. As more data are collected, or current data are re-examined, new categories may be identified and prior categories reconceptualised, eliminated, or subsumed within new categories. Analysis for theory development involves an additional data management activity, examining the relationships between categories. Management techniques for qualitative data are complex as they must permit categories to be examined in reference to the original text, by prior coding schemes, and by comparing and contrasting selected data clusters.

When managing quantitative data it is equally important to be able to access the data from a variety of vantage points. Having entered the data, it is usually necessary to modify it in some way before producing graphs, tables and statistical analyses. For example, it may be useful to look at subsets of the data—only results for females or only results after a certain time. A statistics package should make such data management tasks easy.

Another common requirement is to allow more informative labelling than the stored data contains. For example, instead of treatment levels 1 and 2, one may wish to recode 1 as "Experimental drug" and 2 as "Placebo".

Data summary

Quantitative data can be summarised in various ways. Numerical summaries usually consist of averages, medians, standard deviations and other statistics. Any statistical package should be able to produce a range of summary statistics easily for each variable of interest. Some packages will churn out dozens of largely meaningless statistics, which is not very helpful. The best packages will give a few key statistics by default and others if specifically requested.

Most statistics packages will compute cross-tabulations of categorical data. For example, one may have data from 500 people indicating their sex and blood-type. A cross-tabulation of the data would show how many people of each sex fell into each blood-type. An extension of this idea is to compute the average weight of people in each sex by blood-type combination. Many statistics packages will generate cross-tabulations with more than two categorical variables but these are harder to display. Not only should a good statistics package be able to compute such tables, but it should display the results in a format that is easy to read and understand.

Graphical summaries

An alternative method of summarising data is via graphs. A statistics package should be able to produce histograms of numerical variables, scatterplots showing one numerical variable plotted against another, and so on. As with numerical summaries, it is easy to generate many graphs of limited value. A good statistics package should be able to provide highly informative graphics by default.

When exploring data and trying to uncover interesting structure, it is more important to be able to produce informative graphs quickly than very high quality graphics slowly. Some packages concentrate almost entirely on producing high quality graphics and so are obviously better suited to presenting rather than investigating or exploring data.

Data analysis and modelling

Data analysis and modelling are the areas most people think of as "real statistics". Producing a statistical model for a set of data is an important part in drawing conclusions from the data. Using a statistical model, one can calculate the probability that data could have arisen given a certain hypothesis. This means more objective inferences can be drawn than are possible from graphical or numerical summaries.

There is a huge range of statistical tests and models that have been used and are available in statistics packages. All statistics packages should be able to do simple tests such as t-tests (testing whether the mean of one group of data is significantly different from the mean of a second group of data). Likewise, statistical packages should be able to fit simple models such as straight line regression models (modelling one variable as a linear function of another variable).

Most packages can also do a wide range of standard but more complicated tests and models. Some packages aim to cover almost all of the methods that are widely used in

applied statistics. One or two packages are designed to enable the user to develop and implement new methods. (See section 5).

Presentation—text, graphs, and tables

At the end of a study, it is usually necessary to present the results by giving a talk or by writing a report or research paper. Word processing is an extremely useful tool for developing text, but graphs and tables will be needed to summarize some aspects of the study more succinctly. Often these graphs are the best of those found when exploring the data in the earlier stages of the study, but tidied up by using informative headings and axis labels. Legends are often added and, if available, colour may be used.

In the past, the quality of graphics in statistics packages has varied enormously. These days, most packages make some attempt at producing graphs suitable for publication. Nevertheless, there are still packages which produce graphs which are hard to read because of clotted lettering and shading by diagonal lines.

If a statistics package does not produce sufficiently high quality graphics, it may be necessary to produce the final graphs in another specialised graphics package. In this case, it is important that the data are stored not only in a format suitable for the statistics package, but are also easily accessible by the graphics package.

If a report is to be written, it is usually desirable to be able to enter the graphs directly into the word processed document rather than sticking them in with glue later on. Therefore, it is useful if graphs can be saved in a format that the word processor can input and print.

Tables of summary statistics are often also required for presentation. Many statistics packages do not present tables in a form suitable for presentation. Often they have to be heavily edited in a word processor or simply re-typed.

Examples of statistics packages

There is a huge range of programs that can make some claim to being statistics packages. In order to illustrate the wide range of packages available, brief descriptions of several types of computer packages that do some degree of statistical analysis are presented in this section. An indication of their usefulness to professional statisticians has been made in order to benchmark the use of these tools. Examples of each type are given, but inclusion (or exclusion) is not meant to be a recommendation (or otherwise) of the package, though packages mentioned are among the standards of that type.

Spreadsheets

Many spreadsheet packages include some statistical and/or graphical functions. The range of functions available is usually quite limited, but it is increasing with newer versions so that the line between interactive statistics packages (see 5.2 below) and spreadsheets is becoming less defined. Often spreadsheets do not provide the tools needed for exploratory analysis or checking model assumptions, but rather function as statistical calculators. A working statistician would not get far with these packages, but someone interested in a whole range of spreadsheet operations, presentation graphs and tables, and a bit of statistical analysis, all easily interfacing with other packages, could find them quite useful. Examples: Excel, Lotus 1-2-3.

Interactive statistics packages

These packages have a simple structure, with data stored and entered in spreadsheet format. They are easy to learn and use and are good for many kinds of standard data analyses and graphs. They contain many standard statistical methods as built in functions. Professional statisticians doing general consulting could perhaps do somewhat more than half of their work with one of these packages. Some packages allow some customisation and programming, but users familiar with computer programming (see below) may find these interactive packages somewhat tedious in this regard. Because of simplicity, general applicability, and fairly low cost, these packages are very popular in a variety of settings, including teaching. Examples: JMP, Minitab, Statgraphics, Systat.

Large statistics packages

These large-scale packages were originally written for mainframe computers, though now some PCs run them. They do a very wide variety of analyses and do them fast. They can handle very large data sets (millions of cases), which would cause havoc with most other packages. They allow some level of programming and customised analysis, and have good graphics, particularly for presentation purposes. They tend to be less interactive and flexible than some of the other packages discussed here, and can be tedious for those used to working with more interactive packages or languages. Examples: BMDP, SAS, SPSS.

Statistical programming languages

These are interactive programming languages designed especially for statistical applications. They allow extreme flexibility in data handling and application of statistical methods to data. They contain many built-in standard statistical functions, including very modern methods, and also have a language that allows virtually any sort of calculation to be done. Users can write specialty programs to perform the desired analyses, using the built-in functions as building blocks. These languages are most useful to experienced statisticians who are developing new statistical methods, applying existing methods in new or non-standard ways, or customizing existing methods to new situations. All of this flexibility and power has its advantages, but these languages are not for the novice. They require good computer skills and a high level of statistical knowledge. Even the simplest analyses require a large initial investment in learning the basics of the language. Example: S-Plus.

Specialty quantitative data analysis packages

Sometimes packages are written to do a certain type of analysis, rather than a full range of statistical analyses. These packages often include analyses not available in other packages. These are not generally substitutes for statistics packages but can be indispensable to those who need the particular analyses. Some of these analyses are now available in the large statistics packages or in S-Plus. A few examples are:

EGRET (Epidemiological GRaphics, Estimation and Testing) performs analyses common in epidemiology and other medical research areas. For instance, logistic and Poisson regression and variations of them, exact contingency table analysis, and survival analysis are available.

GLIM (Generalized Linear Interactive Modelling) is an interactive package for fitting Generalized Linear Models (McCullagh and Nelder (1989). These models are very

commonly used in social sciences and medical research, and require a fairly high level of statistical knowledge.

TIME SERIES: There are various packages such as ITSM, Forecast Pro and TSP specifically designed for the analysis of time series data.

Qualitative data analysis packages

Qualitative research is not a single methodology, but includes a range of methodologies grounded in the humanities and social sciences such as sociology and anthropology. The methods include among others, ethnography, content analysis, and grounded theory construction. Qualitative data management and analysis software packages are a recent development and those available are designed to suit the specific methodology employed. The Ethnograph was developed in the late 1980s and is particularly suited to ethnographic methods as the name implies. NUDIST (Non-numerical Unstructured Data Indexing, Searching and Theorizing) is a qualitative data analysis system developed in the early 1990s. It was developed to support a range of methods for the analysis of unstructured data, with emphasis on the building and testing of grounded theory. HyperQual2 is an Apple Macintosh program which incorporates the features of a common Macintosh program called HyperCard. It is designed to store, manipulate and present data in the form of written reports. It frees the mind of the researcher as much as possible from the mechanics of qualitative data analysis so that the analyst can concentrate on the conceptual aspect of data analysis (Padilla 1993).

Choosing a statistics package

Since there is such a bewildering number of different packages to choose from for statistics and data analysis, many researchers have difficulty in selecting an appropriate package for their needs. The questions below may be useful to keep in mind when selecting a statistics package.

A good resource for finding out about various packages is the “Statistical Software Guide”, which is planned to be an annual feature in the journal Computational Statistics and Data Analysis. In the 1992-93 guide (Koch and Haag, 1993), 80 packages were summarised with brief descriptions obtained from vendors. Note, however, that several major packages were not included in this list.

Does the package do what you want?

Many people purchase software *before* carefully evaluating exactly what their needs are. It is important to consider the data analysis facilities one will require and then purchase a package which provides at least those facilities. If the research requires the fit of a sophisticated logistic regression model, it is unlikely that a spreadsheet will suit this requirement. On the other hand, if one is using a statistics package only for producing graphs, it is unnecessary to spend large sums of money and time learning a high-powered statistics package such as SAS or S-Plus. Likewise for qualitative research, if counting the frequencies of certain words and word groupings is what is required, word-processing may fulfill this function. If theorising is planned, more sophisticated qualitative research software should be sought.

What platforms is the package available on?

If one is a dedicated Macintosh user, it is unlikely a PC-based package will suit. Even if one has access to a variety of computing environments, it may be necessary to transfer graphics or data from the statistics package to a word processor or some other package. It is easier and more efficient if all commonly used software tools are available on the same computer platform.

Do you require customised statistical analyses?

Given there are a number of packages which do what is needed on the computer available, the choice between packages often comes down to ones own requirements about user-friendliness and flexibility. Some packages are very easy to use but are relatively inflexible—they provide standard statistical analyses quickly but do not allow the user to modify the methods used or implement new methods easily. Other packages are extremely powerful and flexible but are more time consuming to learn and to use. Ease-of-use and flexibility are not necessarily incompatible characteristics, but to some extent, greater flexibility in the possible analyses available tends to make a package more difficult to use.

Do you require interactive or repetitive data analysis?

Some packages are best suited to interactive, exploratory data analysis, while others are designed for more repetitive tasks. If the analysis involves the same operation applied to many different data sets, then a menu-driven package designed for interactive use may be tedious. A more appropriate package would enable a short program to be written to carry out these analyses. Some packages attempt to allow for both styles of data analysis by providing a choice between a menu-driven interface and a command-driven interface with some simple programming functions.

Do you have very large data sets?

Many statistical packages will be unable to cope adequately with very large data sets. With more than 100,000 observations, most small or medium size statistical packages will struggle. Be wary of advertised claims that a package can cope with very large data sets. These claims are sometimes based on the assumption that they will be run on an extremely powerful computer. This is often not the case. To analyse large data sets adequately it is usually necessary to use a large-scale statistical package such as SAS, SPSS or BMDP on a high-powered computer.

Is there any local support?

Having someone nearby to help is invaluable. Such a person may be a fellow user in the same organisation, a member of a local users group, or a contributer to an Internet news group. Similarly, good documentation is essential. Most programs are now accompanied by understandable user manuals and classes are run for many of the quantitative and qualitative packages available on the market.

General comments on using computer data analysis software packages

The wide availability of inexpensive, easy to use packages makes sophisticated analyses available to nearly everyone. The following comments relate to how best to use packages as well as some common sense suggestions that reflect many years experience consulting on research projects and analysing a variety of kinds of data.

The research methodology one uses will go a long way in identifying groups of data analysis packages best suited to ones needs. Within these groups, often, familiarity is what makes for the "best" software package. While choice is more restricted among packages which support specific qualitative research methodologies, most major statistics packages do most standard things. Rather than constantly pursuing the "best" packages and switching whenever a new one appears to offer some new advantages, a better strategy is to select a good package (or possibly a combination of two complimentary packages) and become familiar with it. One will then be able to use the package more efficiently and will feel more confident with the analyses.

A picture is worth a thousand words. In quantitative data analysis, good graphs are essential in finding interesting patterns, checking assumptions, and understanding how analyses relate to the variables and raw data. For presentation in papers or talks, a few good graphs can summarise the results of a study clearly and concisely. As discussed above, most packages have features for making these graphs. Expect to spend time at this. It is often challenging to think of a good way to present results in graphical form, and many attempts will end up in the rubbish. Once a good graph is decided on it is often just as challenging to get the package to make the graph exactly as desired and to place it automatically in a document.

The advantages of having access to good, easy to use packages are clear; one can now perform a great many statistical analyses, and do them very quickly. With this power comes potential disadvantages; one can now make a great many mistakes, and make them very quickly. It is wise to learn the packages one uses and the analyses one chooses well enough to be able to cross check results of more difficult or unfamiliar analyses with simple data summaries or graphs. It's best to make sure all results make sense when considered from various angles.

A common misconception is that all statistical packages give the same answers. Unfortunately, mistakes have been made, even in some very well-known and widely-used packages. It is a good idea to check output from the package against published results or output from an equivalent analysis in a competing package.

Some errors result from numerical inaccuracy, particularly when very large or very small numbers are involved (see Sawitzki, 1994). This is a particular problem with most spreadsheets. Other errors result from statistical misunderstandings (programmers are usually not statisticians) or from programming bugs.

Ask for help from an appropriate consultant. Computer packages can answer many questions but are no substitute for expertise and experience. Just as most statisticians are not expert health care practitioners, neither are most health care researchers experts in statistics. Qualitative data analysis is likewise vulnerable to inappropriate analyses. Successful researchers plan and budget for the cost of a good statistical or qualitative research consultant and value this commodity.

Think. No computer package can substitute for careful thinking about the limitations of various research designs, the meaning of variables, possible problems with the data, expected or hoped for patterns, and other such issues. The best data analysis comes not from keystrokes and print outs, but from spending time thinking.

References

Baker, C.A. (1988). Computer applications in qualitative research. *Computers in Nursing*. Vol 6. No. 5. p. 211-214.

Bryant, J. & Ryder, R. (1993). Grateful Med (Version 6.0) Training Workbook. Outreach Kentucky Project, Medical Centre Library, University of Kentucky, Lexington, Kentucky, U.S.A.

CINAHL Information Systems (1993). Glendale, California U.S.A.

HealthROM information brochure, Department of Human Services and Health.

Hudgings, C. (1992). The Virginia Henderson International Nursing Library: Improving Access to Nursing Research Databases. (In *Computer Applications in Nursing Education and Practice*, Arnold J.M. and Pearson, G.A. eds) New York: National League for Nursing Publication No. 14-2406.

Koch, A. and Haag, U. (1993). Statistical Software Guide '92/93. *Computational Statistics and Data Analysis*. Vol. 5. p. 241-266.

McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models*, second edition, Chapman and Hall: London.

National Institutes of Health, National Library of Medicine (September, 1993). Factsheet. Bethesda, Maryland, U.S.A.

Norman, E.M. (1989). Applications: How to use word processing software to conduct content analysis. *Computers in Nursing*. Vol 7. No. 3. p. 127-128.

Padilla R.V. 1993 Qualitative analysis with HyperQual2. Chandler AZ, USA

Richards, T. & Richards, L. (1991). The NUDIST Qualitative data analysis system. *Qualitative Sociology*. Vol. 14, No. 4, p. 307-324.

Saba, V.K., Oatway, D.M., Rieder, K.A. (1989). How to use nursing information sources. *Nursing Outlook.* July/August Vol. 37, No. 4, pp. 189-195.

Sawitzki, G. (1994). Report on the numerical reliability of data analysis systems. Computational Statistics and Data Analysis. Vol 5. p. 289-301.

Tesch, R. (1990). Qualitative Research: Analysis Types and Software Tools. New York: The Falmer Press.