Medical Informatics

ISSN: 0307-7640 (Print) (Online) Journal homepage: http://tandfonline.com/loi/imif18

The General Practice Morbidity Database Project Wales–a methodology for primary care data extraction

J. Evans, C. Rogers & S. Kaul

To cite this article: J. Evans, C. Rogers & S. Kaul (1997) The General Practice Morbidity Database Project Wales–a methodology for primary care data extraction, Medical Informatics, 22:2, 191-202, DOI: 10.3109/14639239709010891

To link to this article: https://doi.org/10.3109/14639239709010891

MED. INFORM. (1997), VOL. 22, NO. 2, 191–202

The General Practice Morbidity Database Project Wales—a methodology for primary care data extraction

J. EVANS†, C. ROGERS‡ and S. KAUL†

† Welsh Health Common Services Authority (W.H.C.S.A.), Crickhowell House, Pierhead Street, Capital Waterside, Cardiff CF1 5XT, UK

‡ North Wales Health Authority, UK

(*Received November 1996*)

Abstract. There is a fundamental need for accurate, timely and relevant information for health service planning; the increasing focus on a primary care led NHS has made the collection of information from general practitioners a priority. A collaborative project between general practitioners and public health physicians in Wales has developed methodologies for the extraction and analysis of routinely collected data from general practices across Wales. Four commonly used computer systems have been investigated. This paper outlines the methodologies used and describes the problems encountered and their solutions.

Keywords: General practice morbidity database systems; Data extraction methodologies; Analysis preparation.

1. Objectives

The objectives were to determine whether it is feasible to capture morbidity and other data from commonly used general practice computer systems and to investigate the preparation of such data for analysis.

2. Introduction

Several articles have described the use of data obtained from hospital medical records, but few have addressed the use of primary care records [1–4]. In the mid and late 1980s many software developers created systems and databases that general practitioners (GPs) in the United Kingdom could use in the management of their practices. The advantages of the new technology have been well documented [5–8].

By 1995, 88% of the GP practices in Wales were using systems that could capture basic demographic information, consultation data and prescribing records [9]. Although the systems were designed and implemented for administrative rather than analytical purposes, several investigators have attempted to use them to assess health care needs, for epidemiological research and for health care planning [10, 11]; however the retrieval and manipulation of such data is not easy. The systems themselves, although able to produce simple, administrative reports, do not incorporate statistical or analytical programs. Extracting the data from these different systems and placing it in a statistical software package for analysis has also proved difficult [12]. When this project began in 1993, there were no published methods for the systematic collection of data from several different GP computer systems.

2. Methods

2.1. *Data extraction*

Four out of the 35 different GP database systems used by general practice in Wales in 1993 were selected for extraction; these four systems, AMSyS, AAHMEDITEL, VAMP and EMIS, represented 70% of the computerized systems used by practices throughout Wales [9]. Selection criteria for participation in the project included a fully computerized age sex register, collection of diagnostic information for more than 12 months using a standard, recognized computerized clinical classification system (Read codes or Oxmis codes) and the recording of a reason for each consultation. The range of data recorded by practices is summarized in table 1.

Initial data retrieval methods and ideas were explored. Different approaches were considered to map available practice data with the variables required for the project. These included selecting data fields by using the report writing facilities incorporated within the four selected database systems to produce the desired output, to either screen or data file in ASCII or any usable format. This approach proved difficult due to the complexity of the systems, each with their own database structure for storing and accessing data. The operating environments used by the systems were also different and made creating a universal program to extract data impossible; VAMP runs in the BOS operating system, AMSyS in DOS and AAHMEDITEL and EMIS databases use XENIX.

Following discussion and collaboration, VAMP and AMSyS provided the project with customized software for extraction. The Morbidity Information Query and Export Syntax (MIQUEST) was used to extract data from AAHMEDITEL and EMIS systems. MIQUEST was developed by the Department of Health and Northern Regional Funded Project to collect data from general medical practice computer systems in a common computer-readable format and Wales became one of the regional pilot sites. Figure 1 summarizes the requirements and considerations for using these three methods.

2.2. *AAHMEDITEL/EMIS extraction using MIQUEST*

MIQUEST provided the extraction software and training necessary to obtain AAHMEDITEL data. The software package is in two parts, enquirer and respondent. The enquirer software is a Windows based application, easily installed and run. The package enables the user to create programs containing a set of data extraction queries specifying which variables require collection. Respondent software is loaded onto the XENIX system where it attaches to the AAHMEDITEL database, allowing loading and running of the created enquirer program(s). The MIQUEST programs, installation and programming documentation are available from the MIQUEST project board [13].

MIQUEST was not specifically designed to extract large amounts of data. Project tests indicated that MIQUEST was able to collect the Welsh project's specified data by the application of two created extraction programs. One program extracts patient and lifestyle information, the other diagnostic information. This is necessary as the patient/lifestyle program examines the whole database, creating one record per patient. The medical program extracts diagnostic information from a defined year and patients attending more than once during a selected period will produce more than one saved record (this data is lost when using the patient/lifestyle

Table 1. Data specifications of the GPMDP.

Practice information	FHSA identifier, practice identifier, GP identifier
Age sex register	Unique patient identifier, age, sex, registration status
Lifestyle information	Smoking, alcohol, height, weight, blood pressure, screening (cervical cytology, breast screening), immunizations, contraception, child health surveillance
Diagnostic information	Nature of visit, diagnoses, referrals, prescriptions, surgical procedures, disability

program) and patients who did not attend during a selected period would be lost to the medical program where they are picked up by the patient/lifestyle program. Running a single program extracting all the specified data into one file for a defined period results in substantial data distortion and loss.

Running the MIQUEST programs after they are installed requires no supervision and practice staff are able to access the database while the extraction process is working. The data files produced can be copied to a diskette by MIQUEST, but a practice with 12500 patients produced a patient file containing 1·47 megabytes and a medical file containing 7·18 megabytes of data. The data can be transferred to a tape streamer or portable hard disk.

The same extraction programs and procedures were used for EMIS as for AAHMEDITEL practices. On-line support was given by EMIS.

General problems encountered using both systems included the time the programs took to run. This is dependent on the number of users on the system while extraction takes place, the size of the practice population and type of computer processor available. Extraction of 1993 data on a practice base of 12500 patients, during a weekend, with a 486 processor required approximately 36 h. The systems backup facility has to be switched off to prevent a conflict between the two packages.

For AAHMEDITEL practices. MIQUEST could not extract Date Registered or Date Left fields and the security recommendations imposed by MIQUEST caused several problems. These problems were brought to MIQUEST's attention. They were flagging/programming errors and were corrected in the new MIQUEST version.

Problems occurred during EMIS extraction due to bugs and programming errors, as would be expected with any new software being placed on an established database structure. These were corrected by EMIS. Another problem was due to EMIS systems allowing practices to develop their own medical coding system alongside Read codes; data loss occurred initially as the programmes looked only for Read code classifications. Re-programming MIQUEST programs to identify EMIS practice codes as well as Read codes eventually produced an output file with no loss of information, though this required a great deal of time.

2.3. *AMSyS extraction*

AMSyS, following consultations with the project group, wrote the extraction program used with their system. The program requires no preparation before use. It is menu driven providing the user with a choice of two programs, patient or medical (for the same reasons as for MIQUEST programs). While the extraction

	AAHMEDITEL/EMIS	AMSyS	VAMP
Preparation Requirements	Create extraction programs within MIQUEST enquirer package. Copy onto 3.5" disk with appropriate security code for practice(s).	None Required - programs already on 3.5" disk.	Format 30 + spare 5.25" disks into BOS environment within GSM (Global System Management) package.
Average Time Taken	30 minutes	None	60 - 90 minutes
Extraction Requirements	Follow on-screen MIQUEST respondent package set procedures to import programs onto the database. Select which programs to run and execute. Running programs require no supervision. Requires user from practice to initially log-on to dB	Follow AMSyS on-screen menu to obtain program from disk. Select which program to run, begin and end dates, the directory extracted file to be written to and execute. Running programs require no supervision. Requires user from practice to initially log-on to dB	Follow VAMP on-screen menu to execute program from disk. Running programs require constant supervision. *Disks require to be replaced once full of data* every 30-45 minutes (can use up to 30 disks). Requires constant supervision to replace disks Requires user from practice to initially log-on to DB
Main Considerations	More than one program may be stored and run one after each other. *Allows use of the database while extracting* though slows down the system depending on the number of used terminals.	Allows use of the database while extracting though slows down the system depending *on the number of used terminals.*	No other users are allowed on the database while the extraction process is running. *The program extracts data from the whole database* (could cover many years)
Average Time Taken (for a (practice with 12,000 patients)	Approximately 36 hours	Approximately 1 hour 20 minutes (Depending on processor speed)	Approximately 22 hours (extracting over 6 years of data)
Media Used to Convey Data from the practices	Copy onto disks or tape streamer in UNIX format. (unless through MIQUEST which converts to DOS format onto disk).	Copy onto disks or tape streamer in DOS format.	Automatically copied onto disks by the extraction process in BOS format.
Ease of Overall Extraction	Moderate to Complex - Requires proficiency in MIQUEST programming and MIQUEST techniques. - Takes rather a long time to extract the data though not under supervision. - Best done over weekends.	Straightforward - Can be done anytime.	Straightforward to Moderate - Requires formatting a large number of disks for each practice. - Takes up a lot of time as the extraction process requires constant supervision while running. - Best done during evenings or weekends.

Figure 1. Summary of requirements and considerations for using three methods to extract data from the four database systems selected.

program is running, other users are able to access the database. Extraction time depends on size of the practice population, the number of users on-line and on the type of computer processor available. A practice with a 386 processor took 4 h to extract 1993 data for around 4500 patients. The same practice and population for 1994 data took 15 min following an upgrade to a Pentium processor.

Initial problems were quickly sorted out by AMSyS, on-line support was also available for the project.

2.4. *VAMP extraction*

VAMP provided the extraction program, a copy of the British Operating System (BOS) within a Global System Manager (GSM) and training in the preparation of disks and in data conversion from BOS to DOS. The GSM is required as it allows BOS to run on a DOS machine.

Approximately 30 5·25 inch disks have to be prepared using BOS formatting, including spares in case of disk errors. The data collecting program is copied onto the first disk. At the practice site, running the extraction programm is straightforward following the on-screen menu. The program copies the extracted data onto the 5·25 inch disks, requiring an operator to be present to change the disks once full. Other users cannot access the system while the extraction program is running.

VAMP extraction cannot be limited to annual or any other date/time periods; it extracts the entire database. Fortunately the program places a flag on the database when it has run allowing subsequent extraction to begin at the flag. This reduces duplication of data retrieval and time spent on extraction. A practice covering seven years with an active population of 6000 took 14 h to extract.

2.5. *Confidentiality of data*

The extracted data is in an anonymized form. The practice databases create an unique identifier which does not contain information that can be linked to individuals; this identifier is used to trace and structure information concerning individuals through the database. Variables containing general practitioner and practice identity information are also maintained in coded form and are only available to the practices themselves.

The only address variable picked up by the data extraction is the postcode. MIQUEST internal security measures allow the retrieval of only the first four digits of a postcode for external requesters. AMSyS extraction permits obtaining the whole 7 digit postcode. No postcode variable is extracted from VAMP practices. Analyses concerning postcodes are implemented for individual practices only if requested by them. For general use, the postcode variable is eliminated from the database.

The extracted data is processed and stored in a secure environment according to the normal high levels of security enforced by the Healthcare Management Information Services, implemented by a series of ISO 9000 procedures [14].

3. Preparation before analysis

The software package selected to prepare and analyse the data is SPSS (Statistical Package for the Social Sciences). The SPSSx (the SPSS version used by the project) package is situated on a Data General UNIX machine, allowing multiple program to run at once with a database memory and storage capacity large enough to contain

EMIS

```
$1,"UUFFTTYUUT","1971","18/06/1986","99/99/9999","F","CF01","344","R","120","0","19930209","1371","19930209","","","166","19930209","76.5","19930209"
$1,"UUFFTTWUTT","1987","30/04/1987","99/99/9999","F","CF01","344","R","","","","","","","","","","","","","","","",""
$1,"UUIFXTYUVT","1968","30/09/1987","99/99/9999","F","CF01","383","R","120","0","19910516","1371","19910516","","","","","","",""
```

AAHMEDITEL

```
$1,"WW344","1953","F","","G3231568","3","130.00","80.00","19940503","137A","19940503","1362","19940503","1.67","19940503","70.00","19940503"
$1,"D3SW4","1947","F","","G3231568","3","130.00","80.00","19931103","","","","","","","70.00","19950103"
$1,"396S4","1975","F","","G3268700","5","","","","","","","","","","","","","","","",""
```

AMSyS

```
RJ2407520,24/07/52,F,CF01 5XT,15/10/92,,R,124,78,03/02/94,MODERATE 10-20,11/10/90,LESS THAN 15 UNITS,11/10/90,1.65,11/10/90,68.93,11/10/90
EE2501390,25/01/39,F,CF01 5XT,01/10/89,,R,,,,,,,,,,,,
MM2809550,28/09/55,M,CF01 5XT,20/01/78,,R,,,,HEAVY >20 CIGS/DAY,28/03/94,15-20UNITS,30/09/93,,,,,
```

VAMP

```
016300633319130101  31101 288010413617937   0950306TO    0DSPOSPOS0  P40
016300834119480101  45350 691072603901966   092090714    0 SPOS      0  P23
016300620719860101  28671 188020203617937   091042322Y617937AEPOS   0  P11
```

EMIS, AAHMEDITEL and AMSyS separate the extracted data fields by comma delimiters. The variables are in the same order for each case, but not necessary in the same locations. More than one case can be entered within the same record. Values are separated by quotation marks and/or commas.

VAMP arrange the data fields retrieved in a fixed format. Each variable is recorded in the same location on the same record for each case in the data.

Figure 2. Examples of the format of the extracted raw data retrieved from the four database systems.

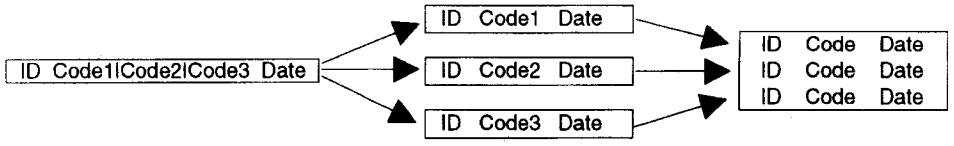

Figure 3. Converting raw AMSyS practice data to prevent diagnostic data loss.

and analyse all the files. UNIX is also a very secure and password sensitive system for security reasons with frequent back-up facility guarding against data loss.

3.1. *Preparation of AAHMEDITEL/EMIS data*

AAHMEDITEL or EMIS data, if saved on tape or disk from XENIX, can be copied straight onto the project's Data General UNIX system, as both systems are similar. Saving onto disk through MIQUEST requires copying the data onto the computer's DOS hard drive, then onto UNIX using FTP (File Transfer Protocol), part of TCPIP (Transmission Control Protocol Internet Protocol) UNIX standard for DOS/UNIX conversions. The resulting records are in a comma delimited format as shown in figure 2. SPSSx programming converts the delimited fields from raw data flat files into usable SPSSx system files. SPSSx allows naming variables and defining fields as numeric or alphanumeric. The quotation marks included posed no problems as they are identified with the commas as delimiters.

AAHMEDITEL and EMIS consultations are recorded in Read 4. To simplify analysis, Read 4 codes are converted too Read 5 using a conversion file obtained from the NHS Centre for Coding and Classification/Computer Aided Medical Systems Limited (CAMS).

3.2. *Preparation of AMSyS data*

AMSyS data requires conversion from DOS to UNIX using FTP. As with AAHMEDITEL and EMIS data, the saved format is comma delimited (figure 2). Unfortunately AMSyS allows written free-text variables. This causes numerous preparation and analytical problems since SPSS reads commas and spaces as delimiters. A raw data record containing Registration, Alcohol and Date Recorded data fields for AMSyS looks like this:

.....R, LESS THAN 15 UNITS, 11/10/90,.....

Instead of producing the three variables required by the commas, SPSS will read the spaces in the free text producing 6 variables as follows:

Var1	Var2	Var3	Var4	Var5	Var6
R	LESS	THAN	15	UNITS	11/10/90

This data structure produces numerous unspecified, non-uniform variables, causing immense data distortion.

Currently the problem is being solved by opening the AMSyS raw data file within a spreadsheet as a comma delimited file. The spreadsheet package reads and displays the data in the structure required, holding the free text as a single variable in one field. Saving the data as a dDBASE IV (or similar) file will allow SPSS for windows to import and read the file, outputting an SPSS portable file which SPSSx understands and reads with all the variables housed in the correct fields. This long and time consuming process is the only current solution to this problem.

	AAHMEDITEL/EMIS	AMSyS	VAMP
Transfer to UNIX	Move data straight from tape streamer to UNIX or Copy data from disk to DOS before moving to UNIX using FTP (File Transfer Protocol) program.	Copy from disk/tape to DOS. In DOS open files in a spreadsheet. Check that the alphanumeric variables are correctly placed in their data fields. Save spreadsheet file as DB4 and input into SPSS PC. Output after checking as an SPSS portable file. Copy onto UNIX using FTP	Convert data from BOS disks onto DOS using GSM (Global Systems Management) package. Copy the created files onto UNIX using FTP.
Average Time Taken (for a practice with 12,000 patients)	20 minutes	60 minutes	45 minutes
Preparation Requirements	Within SPSSx convert the 2 raw data files into SPSSx system files. Identify each field using the comma delimiters, name variables, classify as numeric or alphanumeric. Convert READ 4 classification to READ 5 by matching the READ 5 converter to the Medical file. Match Patient and Medical files together by the Patient Unique Identifier producing the desired practice file for analysis.	Using SPSSx programming, import the SPSS portable files into an SPSSx format classifying and naming variables as required. Define the lengths within the Consultation variable to access each consultation code separately. Add the saved outputs together creating a Medical file with all the consultation codes accessible. Match together the Patient and Medical files using the Patient Unique Identifier produces the desired practice file for analysis.	Convert the 4 retrieved data files into SPSSx files identifying field lengths for each desired variable and their numeric/alphanumeric classification. Match OXMIS to READ 5 conversion file to Medical and Prevention files. Match the Drug Classification file to the Therapy file. Join all four files together using the Patient Unique Identifier in each case to produce the single practice file for analysis.
Average Time Taken (for a practice with 12,000 patients)	45 minutes	110 minutes	90 minutes
Ease of Overall Data Preparation	Straightforward to Moderate - Most complex technique being matching the files together and converting Read 4 to 5.	Moderate to Complex - Requires extensive data manipulation constructing files and variables to prevent data loss. - Requires a large amount of time to implement changes.	Moderate to Complex - Matching different files requires careful planning and checking to prevent data loss in the actual analysis procedures. - Matching the files also takes up time to run.

Figure 4. Summary of the requirements and considerations necessary in preparing raw data retrieved from the four database systems before analysis.

AMSyS extraction permits more than one case to be entered within the same record. If a patient's consultation produced more than one code, the codes are written one after each other in the data record as shown:

Var1	Var2	Var3
Patient ID	Code1/Code2/Code3	Date Consulted

Any analytical requests requiring Var2 or Var3 codes will result in significant data distortion and loss due to the three codes being analysed as one. Splitting the fields using SPSSx programming counters the problem. Defining the length of each code then saving the output of each one before adding together the outputs saved produces the desired error free file for analysis (see figure 3).

3.3. *Preparation of VAMP data*

VAMP data requires conversion from BOS, using the GSM package which creates four DOS files containing the practice data. Altogether the four files contain 47 variable fields of which only 24 are required (figure 2). A description list of the variables, field lengths and values were given by VAMP. Converting the four fixed-format data files into SPSSx system files by naming the variables is straightforward as long as the field length and alphanumeric/numeric variables are defined.

VAMP practices need additional data files to be added to the practice files to retrieve data. These files, provided by VAMP, convert Oxmis codes to Read 5 and converts VAMP prescription codes into usable drug text description variables. Matching the Oxmis-to-Read conversion file to practice files retrieves the consultations and lifestyle data in Read 5; equating the Drug file provides the drug description, form, strength and pack size information for the prescriptions extracted.

3.4. *Other preparation considerations*

Matching of the prepared practice data files, Patient and Medical, for AMSyS, EMIS and AAHMEDITEL files is required to obtain analysis in a standard, clear and useful format. Matching proved relatively straightforward as both files contained patient unique identifiers—the variable used to link both files together to create one file for each practice (figure 4). VAMP practices required more time and programming as four files need to be joined, again using the patient's Unique Identifier variable. Careful data handling, programming, examination and monitoring indicated that no data loss occurred during the matching process. Constant comparisons and data checks validated the produced files with the initial extracted data proving that no data loss or distortion took place during the process.

4. Results

The General Practice Morbidity Database Project in 1993 extracted patient and medical data from 13 practices throughout the eight counties of Wales, covering cities and towns in industrial, rural, inland and coastal areas. The practices yielded a permanent registered population of 111 629 patients, 3·7 % of the Welsh population and data for 818 027 encounters for the year 1993 [15]. (note that this figure does not contain the temporary registered patients also collected and stored but not analysed.)

Data was extracted from six VAMP, four AMSyS, two AAHMEDITEL and one EMIS based practice. VAMP practices, due to extraction across many years, produced reliable data covering 3 to 4 years. The raw data once collected, was configured and prepared ready for analysis within a statistical package. The results

indicate that the project could extract, and prepare for analysis, data from systems used by 70 % of the computerized practices in Wales [9]. The information extracted contains the data specified by the project—patient, lifestyle, diagnosis, referral, prescription and service information.

4.1. *Future data collection*

The advantages and disadvantages of different data collection methods are compared in figure 5. AMSyS has the quickest and easiest extraction method. Unfortunately it requires extensive data manipulation and preparation before accurate analysis is possible. AAHMEDITEL/EMIS practices require more preparation and extraction time though they have the shortest preparation time of the systems.

The flexibility of MIQUEST, and to some extent VAMP, permits obtaining and preparing other data variables in the future. AMSyS extraction has no flexibility as the program is set to the project's present requirements. Any new data variables to be collected would require AMSyS to re-write the extraction program. MIQUEST and AMSyS collection can cover many different time periods, from days up to years. VAMP extracts the whole database or from the latest flag pointer to the day of extraction.

Collecting data from prior years depends on how long the practices have been using their systems and the quality of the data stored. With VAMP, if implementing the flagging procedure is not possible in retrieving data after the initial extraction, the time taken and number of disks used to extract data will increase yearly. This will become a major problem in the future with large practices containing 10 to 15 years worth of data.

The most important problem is the time needed to prepare, extract and manipulate data. The average extraction and data retrieval time for a single practice is approximately 8 h, or one hour per 1000 patients. This is mostly dependent on the size of the practice population, the type of processor used and the database system extracted from. Currently it takes a mean time of 46 h, or 7 h 40 min per 1000 patients, to prepare the files, extract the data and configure it.

Ways of reducing this time will benefit the project, especially when adding more practices. The MIQUEST enquirer package permits sending programs electronically via modem to the practices. Setting the programs to run, through on-line capabilities, or by the practices themselves with some initial training, is possible. Transferring the collected data electronically would save time, though careful attention to security and data encryption would require close scrutiny. Copying the data from the practice computers to portable hard drives or tape streamers would reduce extraction times, especially for VAMP practices.

Information about methods and techniques produced by other projects could provide better ways of data collection for one or more of the GP systems. Co-operation and comparison between the developers of different techniques will assist all concerned in the research and development of such programs. Involving other GP systems in the project would enable expansion into areas where there are no VAMP, EMIS, AAHMEDITEL or AMSyS practices, e.g. GP PLUS on the Macintosh.

The project is now able to expand to extract data from 35–40 practices covering a total population of around 300000 (10·0 % of the Welsh Population [15]). This will create a database with approximately two million patient encounters per year.

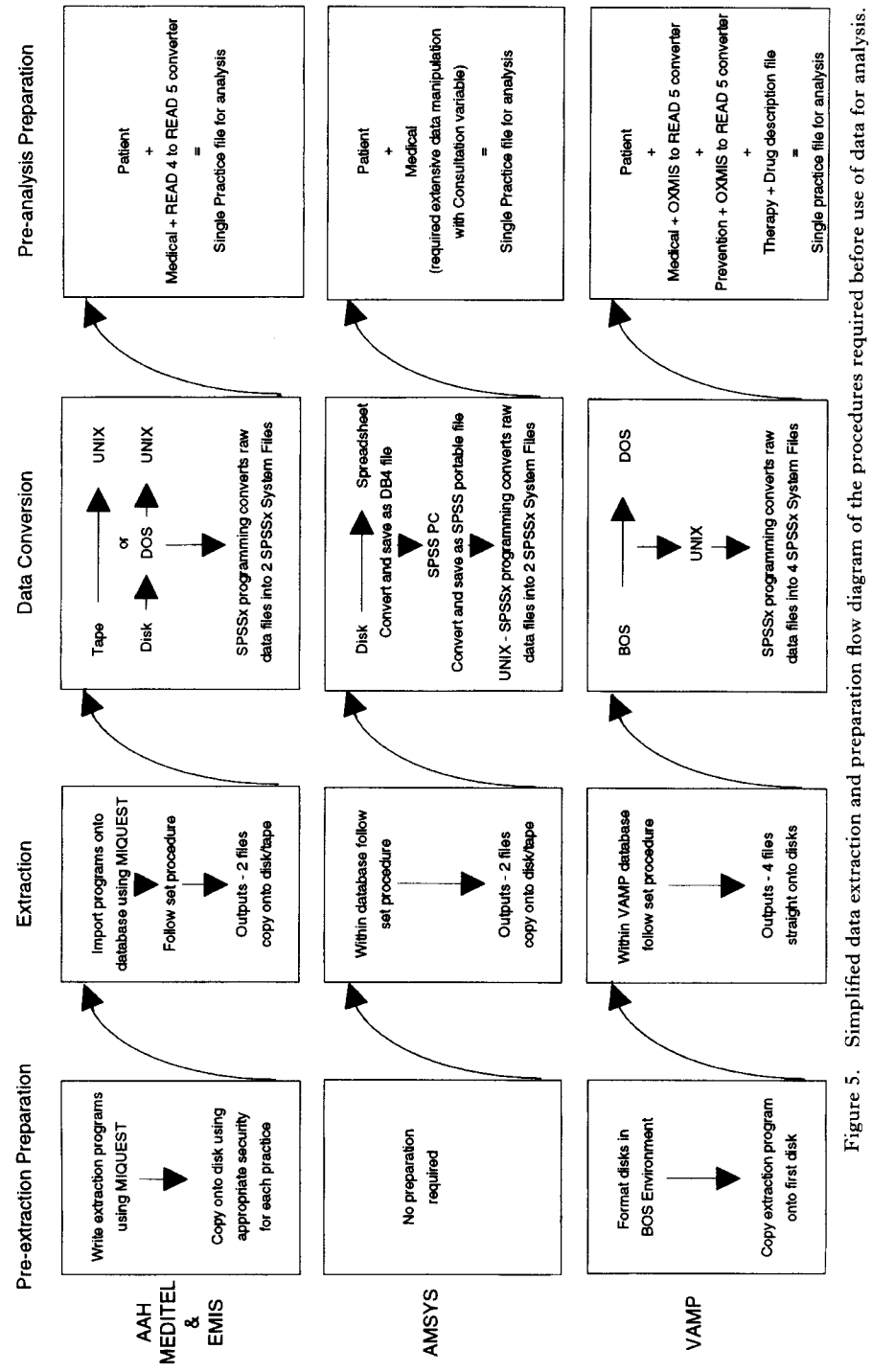

Figure 5. Simplified data extraction and preparation flow diagram of the procedures required before use of data for analysis.

5. Conclusion

Extraction and preparation of data from GP computer systems is possible and practical. Such data can be used to analyse health care system utilization and support other requirements, including epidemiological research, health needs assessment, planning and monitoring health status throughout Wales.

The system described can collect data from the four main GP systems in Wales covering 70% of the computerized practices [9].

Many of the problems associated with collecting GP data have been identified and solved.

As more emphasis is being put on primary care and health needs assessment at local and national levels, the potential of a morbidity database, gathering data from a variety of GP systems, populations and geographical areas, is an important addition to the health planning process.

Acknowledgements

This project would not have been possible without the help, understanding and patience of the participating practices and our thanks are due to them.

We are grateful for the support given by MIQUEST, AMSyS, VAMP, CAMS and the NHS Centre of Coding and Classification, as well as other bodies who have been in contact and have helped the project to progress.

Thanks also to friends and colleagues for their invaluable assistance, support and encouragement, especially Dr George Pickett for his assistance with this paper. Funding was provided by the Welsh Office, and North Wales Health Authority.

References

1. McDonald, C. J., Tierney, W. M., Overhage, J. M., Matrin, D. K., and Wilson, G. A. (1992) The Regenstref medical record system: 20 years of experience in hospitals, clinics and neighborhood health centers. *M.D. Computing*, **9**, 206–217.
2. Christo, G. G., Marianus, B. V. P., Krishnan, L., Iyer, R. S., and Venkatesh, A. (1992) Computerised neonatal case records: a four year experience. *Indian Pediatrics*, **29**, 173–180.
3. Hammond, W. E. (1994) Hospital information systems: a review in perspective. *IMIA Yearbook of Medical Informatics 94. Advanced Communications in Health Care*, pp. 95–102.
4. Barrie, J. L., and Marsh, D. R. (1992) Quality of data in the Manchester orthopaedic database. *British Medical Journal*, **304**, 159–162.
5. Institute of Medicine, Dick, R. S., and Steen, E. B., eds. (1991) *The computer-Based Patent Record: An Essential Technology for Health Care* (Washington DC: National Academy Press).
6. McDonald, C. J., and Tierney, W. M. (1988) Computer-stored medical records: their future role in medical practice. *Journal of the American Medical Association*, **259**, 3433–3440.
7. Department of Health and Social Security (1995) *General Practice Computing—Evaluation of the Micros for GPs Scheme: Final Report* (UK: Her Majesty's Stationary Office).
8. Department of Health and Social Security (1996) *Micros in Practice. Report of an Appraisal of GP Microcomputer systems* (UK: Her Majesty's Stationery Office).
9. Kaul, S., Rogers, C., and Evans, J. D. (1995) *General Practice Morbidity Database Project Report* (Cardiff: Welsh Health Common Services Authority).
10. Singh, A. K., Moidu, K., Trell, E., and Wigertz, O. (1992) Impact on the management and delivery of primary health care by a computer-based information system. *Computer Methods and Programs in Biomedicine*, **37**, 55–64.
11. Isacsson, A., Koutis, A., Cedervall, M., Lindholm, L., Lionis, C., Svenninger, K., and Fioretos, M. (1992) Patient-number-based computerized medical records in Crete. A tool for planning and assessment of primary health care. *Computer Methods and Programs in Biomedicine*, **37**, 41–49.
12. Kahn, M. G. (1993) The desktop database dilema. *Academy of Medicine* **68**, 34–37.
13. MIQUEST Contact Address: Andrew Kent NHS Executive, I. M. G., Room 5E 06 A, Quarry House, Leeds LS2 7UE, UK.
14. Leake, K. (1995) HMIS Computing Security Stratergy V3 *Departments Internal Report.*
15 2960500—Wales Mid Year Population Estimate for 1993. Source: Office of Population Censuses and Surveys (OPCS).