# REPORT OF THE UNIVERSITY OF ADELAIDE EDQUM DATA CONSULTANCY

## Report on EDQUM Data Sources

**May 2003**

Andre Duszynski
**Department of General Practice**
**The University of Adelaide**

# CONTENTS

# INTRODUCTION

This report aims to document all available Australian health data sources which may be used within the context of the Enhanced Divisional Quality Use of Medicines (EDQUM) project, with the understanding that issues of accessibility of data is in response to queries made by participating Divisions of General Practice of firstly what data sources are available and secondly how such health data sources could be used. It additionally forms part of the requirements of the University of Adelaide Data Consultancy to address available datasets, specifically:

- Documenting the range of data captured,
- Corresponding constraints and usefulness of each source, and,
- The potential of improving the quality of the data source if applicable.

Documenting the range of available data sources is the main aim of this document, whereby it will allow further debate on the caveats and implications on using, and how to use each identified data source. Feedback from data source stakeholders was obtained and will continue to inform data within the context of EDQUM per se.

The initial format and direction of this document, has been kindly provided through perceived sources of data table as determined and developed by Karin Orlemann (Canning Division of General Practice); inclusion of this table occurs within Appendix A – Table of Data Sources for EDQUM.

# BACKGROUND TO DATA SOURCES

### HIC Datasets

In Australia, routinely gathered data associated with general practice activity is gathered by the Health Insurance Commission (HIC). It includes prescribing data derived from the Pharmaceutical Benefits and Repatriation Pharmaceutical Benefits Scheme (PBS/RPBS), as well as service utilisation data comprising the Medicare (MBS) claims system and Department of Veterans' Affairs (DVA) claims systems. In addition there is information associated with the Department of Veterans' Affairs (DVA) that administers a range of care services provided to eligible members of the veteran community.

### Pharmacy Data

Two sources of data exist derived from the dispensing activities of pharmacists; that data collected at a local (i.e. community pharmacy level) and that aggregated to estimate and inform volumes of medications within non-subsidised categories (i.e. Pharmacy Guild Survey).

The Pharmacy Guild Survey began in 1989. Each month, dispensing information is collected from a random sample of approximately 250 to 300 pharmacies belonging to the Pharmacy Guild of Australia. The pharmacies are stratified into three equal ranges by prescription volume. The survey is administered by Chemdata; a major pharmacy computer software supplier. Each month, participating pharmacies send their dispensing records to Chemdata in Canberra where data is summarised by drug code and category. Although the Pharmacy Guild Survey collects information on all drugs dispensed, the Department of Health and Ageing only uses data to calculate estimates of prescription volumes for drugs in the non-subsidised categories i.e. private prescriptions and prescriptions priced under the general patient co-payment (prescription count data aggregated at national level). No patient identifying or socio-demographic data are available, nor is data available on the cost of prescriptions or the medical condition for which a drug is prescribed.

## General Practice Desktop Data

The implementation of general practice clinical software in Australia has been a recent phenomenon with incentives provided by the Federal Government to increase the uptake of such software.

Current sources of clinical software are typically derived from Australian software houses and in some respects New Zealand publishers, reflecting similar needs in both countries. Such packages and the scope of data collected (and the ease of collection) vary considerably between software packages and by the activities of the end-user entering data. Such entry by GPs of clinical patient data ultimately relies on how the software is used; in that the level of use of clinical software is ultimately a function of GP training, motivation and need to collect such data.

Several clinical software packages enjoy a majority share in the Australian market. As such, the most utilised clinical practice software is Medical Director 2 with "excess of 90% of computerised Australian GPs using its software" (Health Communication Network (HCN) media release 19/7/2001).

Data however can be derived from clinical encounters (or software utilised as a prescribing tool), which would enable inferences to be made regarding the management of patients at either the patient level or services provided by the GP to their individual populations of patients. The use of aggregated data to this point to inform population management is a recent phenomenon, understandably undertaken by those with an interest and time to “mine” such a rich source of data.

## Bettering the Evaluation and Care of Health (BEACH) Data

The *Bettering the Evaluation and Care of Health* (BEACH) survey is the continuous national survey of general practice activity within Australia, with the most recent report of 2000-01 relating to paper-based surveys conducted in April 2000 to March 2001. It includes nationally aggregated and presented trend analyses against previous BEACH surveys, inclusive of information collected on:

- General practitioners and their practices

- Participating patient details
- The general practitioner / patient encounter
- Problems managed
- Management techniques,
- Patient risk factors.

**The Drug Utilisation Sub-Committee (DUSC) Prescribing Data**

The Drug Utilisation Sub-Committee (DUSC) of the Pharmaceutical Benefits Advisory Committee also collates prescribing data, which is presented in an annual publication known as the Australian Statistics on Medicines (ASM) - production of the 1999 and 2000 Australian Statistics on Medicine has been delayed, no mention of later data collections and their releases is avaliable. The collation of this information allows the DUSC to undertake "pharmacosurveillance and the targeting and evaluation of quality use of medicines initiatives"[1]. The data reported in this publication relates to estimates of community utilisation (non public hospital) of prescription medicines derived from two data sources[1]; the HIC collection of PBS/RPBS data and sampling community pharmacies.

**National Prescribing Service (NPS) Pharmaceutical Utilisation Data.**

*1. Internal sources*

a) NPS self-audits and case studies:

The NPS as an organisation gains a degree of 'prescribing intelligence' through the participation of Australian GPs and pharmacists in NPS self-audits and case studies. The information gathered from responses to these audits and case studies is aggregated to the national level and then reported back to GPs and pharmacists. This allows individual GPs and pharmacists to compare their responses with that of their peers as part of a behavioural Intervention.

The 'intelligence' gathered in this way has been utilised by the NPS to inform the development of future NPS programs. For example from self-audit or case studies, actual clinical or hypothetical data on prescribing rate or selection of drugs within a class is obtained. However usefulness as a pharmaceutical utilisation data source is

limited, for example in self-audit responses there is self-report bias and selection bias.

This data is not regarded as useable in the context of EDQUM due to it's characteristics, mode of collection and additional access restricted primarily to the NPS only.

b) Focus group and survey data:

The NPS convenes focus groups and commissions surveys to establish knowledge and behaviour relating to medication use by health care professionals and consumers. These allow evaluation of NPS communication activities which impact not only on that of the provider of medication but on consumers of medication within the Community. Surveys have also evaluated consumer knowledge, attitudes and behaviours regarding specific NPS programs, for example the appropriate use of antibiotics for Upper Respiratory tract Infections. Such data is made available in the NPS Evaluation reports.

*2. External sources*

The NPS utilises data from external sources to inform program choice, delivery and evaluation. These sources include the HIC, AIHW, DUSC and electronic medical record databases (i.e. the HCN's GPRN and the University of Adelaide MEDIC-GP datasets). This information is included here to illustrate how the NPS as an organisation utilises external data sources.

a) Prescribing data feedback:

Data from the Health Insurance Commission (HIC) pharmacy claims database (i.e. PBS dataset) is utilised by the NPS for prescribing feedback purposes. The NPS has no access to the data and simply has a contract with the HIC to transpose the data, sight unseen. The data is not intended for any regulatory purpose and NPS surveys have confirmed that doctors value this personalised, comparative information.

GPs receive their individual prescribing data together with educational messages on relevant therapeutic topics to assist them to personally review their prescribing habits in the context of their own practice characteristics. This data is provided, with the

Prescribing Practice Review on the same therapeutic topic, to all general practitioners and specialists if relevant. Such feedback has always been as one aspect of a multi-faceted approach.

The data is presented as prescribing rates (per 1000 Medicare consultations) to adjust for volume of service. It is not possible to normalise the data for patient demographics. However, a profile of the age of patients in the GPs practice and their concession card status is provided. Data is presented in comparison with doctors located in a similar geographical area using the Remote, Rural and Metropolitan classifications. The data is confidential with each GP seeing their own data. The limitations of the use of HIC data in this way are recognised and points of assistance are provided to assist the GPs to interpret the information presented.

For the purposes of EDQUM, the NPS has developed expertise in the presentation of data from the PBS dataset to GPs as an educational tool Experience and feedback gained can be utilised in the development of resources for EDQUM.

b) Other external data sources:

NPS may obtain data from other sources for specific evaluation questions, program planning and to convey an educational message i.e.:

- Data from AIHW from the BEACH program for antibiotic use by indication.
- De-identified PBS data to evaluate changes in drug utilisation, which may result from NPS program delivery.

# EDQUM DATA SOURCES

**HIC Datasets**

The HIC data collections (inclusive of PBS and MBS statistics) are regarded as the most comprehensive currently available in Australia but have limitations as a consequence of their primary role as a source of data within the health service administration context.

- The Medicare claims data for example does not include any information relating to 'non-fee-for-service' components in general practice as well as including incomplete data on services provided to the Indigenous population[2],
- Data related to pharmaceuticals derived from the PBS, is available for those medications receiving subsidy by the Commonwealth Government. The PBS purchases about 90% of all prescription medicines[3,4].
- Community subsidisation that is, reimbursement of pharmaceuticals outside of those prescribed in the hospital environment, represent 75% of all pharmaceuticals[5].

Under the PBS, patients contribute towards the price of subsidised pharmaceuticals, otherwise known as the 'copayment'[6]. General patients presently pay a maximum of $22.40 on purchases of PBS-listed medications under the co-payment arrangement, while pensioners and Commonwealth concession card holders are allowed a reduced co-payment, presently $3.60 per prescription[7]. In addition there is a safety net which takes into consideration those people with increased medication requirements. Therefore for general patients and their families who have already paid $686.40 on the PBS within a calendar year, prescriptions for the rest of the year are charged at $3.60[7]. Pensioners and concessionals who have paid $187.20 in PBS-listed pharmaceuticals, receive all subsequent prescriptions at no cost[1,7] For patients who pay the full cost of a medication, no record of the prescription is generated within the HIC data.

Many medications such as the antibiotics, diuretics, tricyclic antidepressants and non-steroidal anti-inflammatory medications are priced less than $22.40. As a consequence, only concessional patients or general patients who have reached the safety net and who have a prescription dispensed for these medications, will have a

record of the dispensing included within the Health Insurance Commission data[8]. As the co-payment for general patients increases, the dispensed price for less expensive medications falls into this range and patients are increasingly required to pay the full cost of the medication. Again, the PBS will have no record of a claim for payment for these lower cost medications[1]. It has been estimated that general prescriptions under co-payment represents almost 20% of all community prescribing[1].

While there exists the potential to link de-identified pharmaceutical data with patient characteristics (derived from Medicare data) such as age and gender, this is not performed in a routine manner nor readily available. Following the requirement to produce a Medicare card to obtain a script under the PBS, from 1 May 2002 patient characteristics such as age and sex are available for all PBS records. Investigations are underway within the HIC to explore concurrent medication usage and / or definition of indications associated with a particular medication.

**Table 1: HIC MBS and PBS Data**

| **Data Source** | **HIC MBS services** |
|---|---|
| **Information Provided** | Number of consultations and patients, by provider number |
| **Data Fields** | • Age<br>• Sex<br>• Month of service<br>• Provider number<br>• Consultation item number |
| **Use of Data** | Denominator for calculation of prescribing rates |

| **Data Source** | **HIC PBS services** |
|---|---|
| **Information Provided** | Total number of subsidised scripts prescribed (original and repeats), by general and concessional beneficiary entitlements for drugs over the co-payment |
| **Data Fields** | • Drug item code<br>• Script volume for item code by general (safety net, non-safety net) and concessional (safety net, non-safety net) status<br>• Month (or year) script processed<br>• Derived gross price and net benefit<br>• Drug strength, quantity, manufacturer, date script written, number of repeats, dispensing pharmacy code |
| **Use of Data** | Evaluate changes in prescribing, economic modelling |

## Pharmacy Data

*Pharmacy Guild Survey Data*

The inception of this mechanism that began in 1989, allows collection of monthly dispensing information from random samples of approximately 300 pharmacies who belong to the Pharmacy Guild of Australia.
Prior to presenting the results of this aggregated de-identified data which aims to be representative of the population of operational pharmacies on PBS dispensing volume (including geographical location and similar stratification), a weighting factor is calculated for each of the strata by comparing the number of pharmacies in the survey with the total number of pharmacies in Australia. Volumes of non-subsidised drug use are calculated by multiplying the survey estimate by the weighting factor, which is assumed to apply equally to both the subsidies and non-subsidised prescription volumes. Such data is also utilised to inform the publications of the Drug Utilization Sub-Committee (DUSC) within their annual Australian Statistics on Medicine publications.

*Individual Pharmacy Data*

Although not expected to be an indicative source of prescribing trends nor comparable against other pharmacy data due to statistically limited dataset sizes from such individual pharmacies (except at aggregated levels across regional areas or Divisions), pharmacy data could albeit offer a rich source of prescribing data which would cover all script presentations regardless of concessional or co-payment status.

Contact has been made with the main supplier of pharmacy software (Pharmacy Computers Australia Pty Ltd) regarding the process of approaching pharmacies for such prescribing data. Payment to participating pharmacists who contribute data is assumed (but may depend on the context of such a data request), in hand with the requirements to have good co-operation with such pharmacists (personal communication Mary Collins, Southern Tasmania DGP). Contact details for Pharmacy Computers Australia - who have developed both the “FRED Dispense” and “Winifred” software for the Pharmacy Guild of Australia - is listed below:

| | |
|---|---|
| **Pharmacy Computers Australia Pty Ltd** | **P**: +61 3 9810 9962 |
| **Contact:** Glen Gooma (CEO) | **F**: +61 3 9810 9991 |
| Guild House, 40 Burwood Rd, | **E**: glen.gooma@pharm.com.au |
| Hawthorn, VIC 3122 | **W**: www.pharm.com.au |

The below graphic represents the type of data which can be inputted by the pharmacist upon presentation of a script by a patient. Its only use is to inform Divisions who may have not approached individual pharmacies, the extent of data capture which can occur at the pharmacy level and particularly the mechanisms of capture within the electronic data paradigm.

**FRED Dispense (Pharmacy Computers Aust.) pharmacy dispensing software**

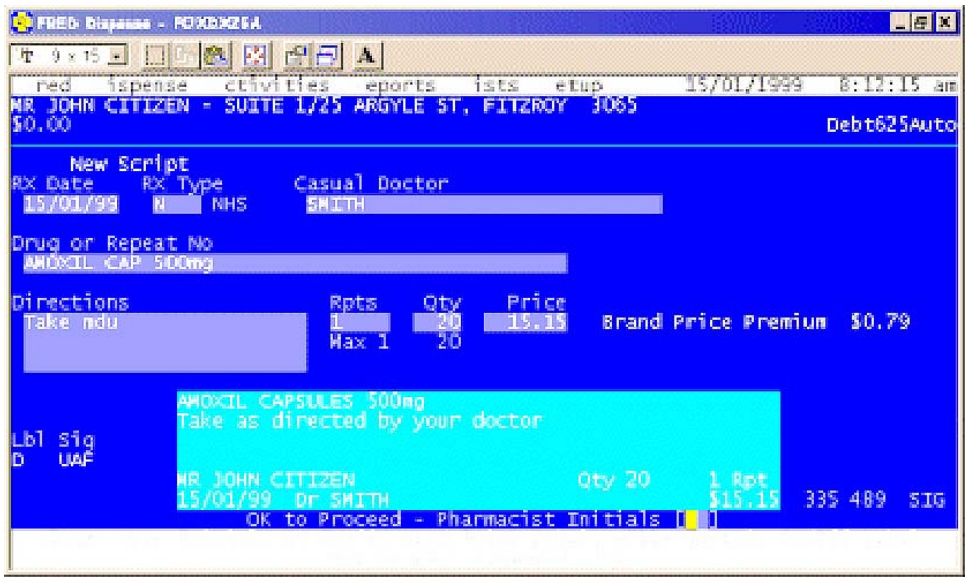



**Table 2: Pharmacy Software Data**

| | |
|---|---|
| **Data Source** | **Pharmacy software e.g. FRED Dispense, Winifred (Pharmacy Computers Australia)** |
| **Information Provided** | Pharmacy level prescribing / dispensing data (but not OTC preparations) |
| **Data Fields** | • Prescribing doctor (inc. specialists & others),<br>• Script date and type (e.g. concessional),<br>• Drug type, drug repeat, directions of use, repeats, quantity,<br>• Price (inc. brand price premium),<br>• Patient name (not provided within de-identified dataset)<br>• Pharmacy postcode (not provided within de-identified dataset)<br>• Derived defined daily dose (DDD) |

| | |
|---|---|
| **Use of Data** | Mechanism to collect prescribing data in both subsidised and non-subsidised categories, inclusive of private prescriptions and PBS prescriptions priced under the general patient co-payment and those that would appear normally within the HICs PBS/RPBS records. |

*Other sources of Pharmacy Data*

Pharmacy data can also be derived from additional sources, however their context within the EDQUM project is debated, particularly at the levels of both Division and GP usage. Such sources include:

- Sales Data obtained from sources such as that provided by commercial market research organisations that collect sales and surveys data from general practitioner prescribing. Data is retrieved initially as unit sales by form, strength and drug class and reported as DDD / 1000 population / per day. (cf. Intercontinental Medical Statistics (IMS), Melbourne)[10].

**General Practice Desktop Data**

The use of medical record databases have a number of advantages, the most apparent being that clinical information is readily available and accessible. Clinical data retrieval from computerised medical records is regarded as:

- An ideal source for conducting investigations in a more rapid, efficient and economical manner compared with de novo collection of data,
- Representing medical records information recorded in real-life practice,
- Not contrived or limited by restrictive inclusion/exclusion criteria or the requirement for specific patient populations, elements characteristic of information derived from a clinical trial environment,
- Having methodological advantages relating to clinical information being documented as modifications to patient medication regimen or clinical history occur over time. This ‘immediate’ recording of data reduces considerably, the occurrence of information bias such as recall or observer bias, which could otherwise compromise an investigation’s validity,
- As having additional methodological benefits include the ability to calculate incidence rates (where the eligible population is stable and can be verified) and the relative ease in which appropriate controls can be selected for either

case-control or cohort investigations if interventions were conducted in the context of EDQUM.

General practice databases however, also have a number of limitations including:

- One of which includes the stability of the population comprising a longitudinal data set. That is, in the circumstance of medical records, patients may leave their medical practice resulting in no further events being recorded for patients after a period of time and the information available from existing records finite and untimely,
- Relevance in investigations involving comparisons between case and control subjects or unexposed and exposed subjects and where person time from time of enrolment, needs to be similar in two groups of subjects. A more specific example may relate to determining the influence of lifetime exposure to a particular medication on an outcome of interest such as death. Failure to adequately determine exposure would bias the investigation towards a null result, underestimating the effect of the medication on the outcome,
- Medical record databases generally contain a finite longitudinal data. As a consequence, the availability of restricted information on life time exposure for example to a particular drug and simple matching on age will not ensure that person time is well controlled. This has particular relevance to databases such as the GPRN (General Practice Research Network; HCN Ltd.) which have been collecting data relating to a relatively short period of time,
- The usage of the Medical Director software by practitioners has changed over time. It has been demonstrated that doctors contributing to the GPRN have utilised the software in a more comprehensive manner over time, changing from generating solely prescription information to making use of other features of the Medical Director software such as recording diagnoses and progress information.

While prescription data included in medical record databases are generally regarded as true examples of script (and therefore medication) provision, Australian medical record databases suffer from:

- The disadvantage that there is no straightforward method of ascertaining whether a prescription has actually been dispensed or subsequently consumed,
- Furthermore pathology results and diagnostic outcomes are often poorly recorded and previous attempts at validation of these data have shown that it is necessary to go back to the original paper records to validate this information.
- A failure to document potential confounders such as nicotine or alcohol consumption or even the consumption of over-the-counter medications. This makes subsequent attempts to adjust for these confounders difficult.
- Representativeness of the individuals who comprise the database population and whether statement about the generalisability of results from an investigation can be made.

It is not the intention of this document to explore in detail, the relative epidemiological advantages and limitations of medical record databases. From an Australian perspective however, the application of electronic medical record databases to research particularly in the realm of pharmaceutical prescribing, offers distinct advantages over current Australian data sources collecting general practice information. Data sources such as those maintained by the Health Insurance Commission (Pharmaceutical Benefits Scheme and Medicare claims system) have ostensibly been developed for administrative purposes that is, forwarding payments of subsidies for services rendered or prescriptions dispensed and consequently, have limitations to the extent of information collected. Drug utilisation data collected by the Drug Utilization Sub-Committee also have shortcomings.

Medical record databases such as that associated with the General Practice Research Network which is described below, offer an opportunity to fill in detail missing from other data sources.

For example in terms of evaluating prescribing practices, prescribing data from medical records can be used to examine:

**Table 3: General Practice Desktop Software Data**

| Data Source | GP Desktop Software e.g. HCN Medical Director |
|---|---|
| **Information Provided** | • Under co-payment prescribing,<br>• Appropriateness of prescribing,<br>• Patient cohorts based on specific medication classes or medication combinations and prescribing linked to patient characteristics (such as age, gender and diagnosis),<br>• Medications can be assessed from an prescribing perspective as opposed to what has been dispensed, thereby providing information on the quality use of medicines related to appropriateness of prescribing and changes in prescribing behaviour following interventions, educational or administrative in nature. |
| **Data Fields** | Refer to Appendix C for a MD2 subset relating to data fields required for EDQUM activities, in the context of a supplied GPRN extract. |
| **Use of Data** | GP electronic record of clinical activity dependent on level of software use. |

*MEDIC-GP General Practice Research Datasets*

Since 1996, the Data Analysis Unit at the Department of General Practice, University of Adelaide has been collaborating with general practitioners who utilise computerised record systems to document all their clinical information. These general practice data sets were derived from disparate systems, and were consequently processed and unified to form a single integrated database. All subsequent validation (coding) and analyses are therefore maintained in a computerised form. The collaboration has resulted in an anonymous data collection containing almost 59,000 patients and their associated (more than a million) detailed clinical records. Records include pathology investigations, diagnostic imaging, allergies/adverse reactions, referrals as well as details relating to doctor-patient encounters e.g. Subjective, objective, assessment, plan and medications prescribed. This pharmaco-epidemiological database has been named the Medical Enquiry Drug Information Centre – General Practice, otherwise known as Medic-GP.

Software comprising Medic-GP is customised for the research answers sought and the data 'validated' (coded) appropriately. Software has been further developed to incorporate complex vocabularies and unique concepts such as episodes of care. Once all data has been compiled it is analysed and reported on by the Unit. The Data Analysis Unit has established a solid foundation of analytical methodology for reporting on information in the database. Investigative tools have included time-series analysis, 'episodes of care' and case-control studies.

The Medic-GP database is used by the Data Analysis Unit to answer complex clinical research questions, usually in the domains of clinical epidemiology, pharmaco-epidemiology and pharmaco-economics. Since the Medic-GP database offers an integrated longitudinal perspective of patient clinical data in general practice it is an efficient and ideal means of clinical investigation. Measures of effectiveness in general practice have also been piloted within the Medic-GP database.

*Health Communication Network (HCN) GP Research Network (GPRN) Dataset*

The General Practice Research Network comprises a national network of Australian general practitioners (GPs) and participating practices, who take part in quality assurance and research activities through the patient management and script-writing software mechanism known as Medical Director. This initiative has been coordinated by the Research Division of HCN who have developed Medical Director[12], with the GPRN program launched in August 2000 with a view that doctors signing up to the program would initially be involved until June 30th 2001. Data has thus been extracted from the participating doctors systems from January 1, 1999 - if data has been recorded within the GPs desktop systems from that date, with ongoing extraction appending data since the last extraction. While it was envisaged that data extraction would occur on a regular automated basis, the doctor has maintained control and is the initiator of the process; doctor consent is thus preserved with each new extraction. The installation of the GPRN data extraction tool further implements a software routine that automatically obliges doctors to record a reason for prescribing and reason for the visit for encounters where no prescription is generated.

Practices with more than one general practitioner are not required to sign up all doctors within the practice. By collaborating in the GPRN, doctors have been able to take advantage of two Royal Australian College of General Practitioners (RACGP) endorsed clinical audit activities – the completion of both audits attributing 50 Clinical Activity points. The participating doctors also receive feedback on their use of Medical Director for breadth of functionality and data recording.

It must be emphasised that the GPRN is an ongoing dynamic entity; one data extraction will never be the same as the next - not only from a longitudinal

perspective, but also because of the continuous signing up of and loss of additional doctors to the Network.

*Description of GPRN data set provided to the University of Adelaide*

The General Practice Research Network has provided data to the University of Adelaide comprising clinical information dated from the 1st January 1999, until the 24th July 2001. The GPRN has continued to expand since the receipt of this data. GPRN data were provided as a series of 27 SAS files each of which had been abstracted from database files comprising the Medical Director software program - the current version of Medical Director (v2.7.1) has more than 130 data tables. Note that the SAS files comprising the GPRN abstract represent modified versions of the original files from which they had been derived, having been standardised, cleaned and in some cases, renamed. Some files such as the predominant scripts file (SCRIPTS_ALL) included merged variables from other Medical Director database files not provided with the GPRN data abstract. Also note that requested extracts of the GPRN dataset may include subsets of the complete Medical Director database dependent on end-user requirements. A comprehensive description of the 27 modified SAS files and their associated variables are provided in Appendix C. The list of the 27 SAS files is recorded in the table below.

Note that since the Medical Director software undergoes periodic updates and modifications, there potentially will be changes to some of the GPRN database files; inclusion or deletion of variables as well as the addition of new database tables. However it is understood that the structure of the GPRN data abstract will largely remain stable since the essential clinical information required for research has been determined. However it is important that in any analysis involving GPRN data that as a first task, all database files and their corresponding fields/variables should be catalogued.

**Table 4: GPRN data tables and corresponding MD2 data tables**

| | GPRN SAS file name | MD2 table name | Additional Notes |
|---|---|---|---|
| **1** | **ACTIONS*** | **actions.dbf** | |
| **2** | ACUPUNCTURE | | |
| **3** | ACUPUNCTURE_REASON | | |
| **4** | ANTENATAL_VISIT | antenat.dbf | Stores details of Antenatal visits. |
| **5** | DIABETES | | |
| **6** | **DIAGNOSIS*** | **diagnosis.dbf** | |
| **7** | **ENCOUNTERS*** | **progress.dbf** | Progress notes |
| **8** | **GPs*** | **(from GP Questionnaire)** | |
| **9** | HEALTH_ASSESS | | |
| **10** | **HISTORY*** | **history.dbf** | Patient's Past Medical History list |
| **11** | IMMUNISATION | **imm.dbf** | Patient immunisation database |
| **12** | **MDUSAGE*** | **mdw.dbf** | |
| **13** | **MEASURES*** | **measures.dbf** | Stores all patient clinical measurements, e.g. height, Weight, BP, Pulse etc. |
| **14** | PAP_RECALL | **paprec.dbf** | Date of last Pap smear used for generating list of patients due for smears. |
| **15** | PAP_SMEARS | **pap.dbf** | Patient's pap smear database |
| **16** | **PATHATOM*** | **pathatom.dbf** | Atomised pathology values from HL-7 files. |
| **17** | **PATHOLOGY*** | **pathol.dbf** | Patient's full text pathology results |
| **18** | **PATIENTS*** | **patients.dbf** | Patient demographic information |
| **19** | **PATIENT EDUCATION*** | **ped.dbf** | List of patient education leaflets given to patients |
| **20** | PRACTICES | (from practice questionnaire) | |
| **21** | PREGNANCY | pregnanc.DBF | Details of patient's pregnancies. |
| **22** | **REACTIONS*** | **reaction.dbf** | Details of Allergies/adverse reactions |
| **23** | **RECALL*** | **recall.dbf** | List of outstanding recalls. |
| **24** | **REQUESTS*** | **requests.dbf** | Store details of all requested tests. |
| **25** | **RISK FACTORS*** | **allergy.dbf** | Contains ancillary patient information not stored in PATIENTS.DBF. It originally stored the allergy list, but this is now stored in REACTION.DBF. It now stores secondary demographic information and clinical information where there will only be one entry per patient. There should only ever be one record per patient in this file. |
| **26** | **SCRIPTS_ALL*** | **scr_file.dbf** | |
| **27** | **SCRIPTS _CURRENT*** | **rx.dbf** | patient's current medication list |

*Indicates that detailed description of this file and its accompanying variables is included in Appendix C.

**Bettering the Evaluation and Care of Health (BEACH) Data**

The BEACH survey of general practice activity within Australia that collects and reports national figures is based on methods as outlined below[13].

- A random sample of GPs who claim at least 375 general practice Medicare service items in the previous 3 months (i.e. 1500 Medicare claims per year),
- GPs are approached and are asked to complete 100 consecutive patient encounters on structured paper encounter forms,
- In the 2001-02 BEACH data year survey; 983 GPs took part providing information on 98,300 GP-patient encounters.

Results are reported in aggregate form (i.e. national level / all GPs), with assumptions that it would be possible to derive state and RRMA level outputs). Results include the following items[13]:

- General practitioner and patient characteristics,
- Patient reason(s) for encounter,
- Problem(s) managed,
- Management techniques used including prescribing, non-pharmacological management, referral & admissions and investigations,
- Health risk factors in sub-samples of patients including BMI, smoking and alcohol consumption information,
- Other sub-sampling results.

Encounter data is weighted for age (stratified by sex) and GP activity level (i.e. indicator of number of patient encounters per time period). Patient age and sex distribution comparisons against Medicare demonstrate that encounters are representative and are precise[13].

*Results for 2001-02 BEACH survey*

96,973 encounters were included for analysis in the final 2001-2002 dataset (collected between April 2001 and March 2002), which comprise[13]:

- 144,654 reasons for an encounter i.e. 149 reasons per 100 encounters,
- 139,092 problems managed i.e. 143 problems managed per 100 encounters,
- 101,350 medications i.e. 105 medications prescribed per 100 encounters, or a rate of 73 per 100 problems,

- 50,308 non-pharmacological treatments (clinical and procedural).

Each of the main components of the reports' results are divided and tabulated below.

**Table 5: BEACH Encounter Data**

| **Data Source** | **BEACH Encounter Data** |
|---|---|
| **Information Provided** | • Information regarding the general practitioner and patient encounter or consult |
| **Data Fields** | **Encounter Data**<br>• Date of consult,<br>• Type of consult (direct or indirect)<br>• Medicare / Veterans Affair item number<br>• Specified other payment<br>**Patient Data**<br>• Date of birth and sex,<br>• Postcode of residence,<br>• Health care card status and type,<br>• Whether from non-English speaking background<br>• Whether patient self-identifies as Aboriginal or Torres Strait Islander<br>• Patient reasons for encounter (up to three)<br>**Content of Encounter**<br>• Problems managed,<br>• Management techniques,<br>• Diagnosis,<br>• New or existing problem,<br>• Whether problem is work-related |
| **Use of Data** | Provide information regarding the particular cycle of care for the encounter inclusive of patient presentation, diagnosis of problem and method of patient management. |

**Table 6: BEACH Management Data**

| **Data Source** | **BEACH Management Data** |
|---|---|
| **Information Provided** | • Information regarding the general practitioner management options provided to the patient |
| **Data Fields** | **Management Data**<br>• Medication prescribed for each problem (if any), including over the counter medication advised or supplied by the GP<br>• Medication details:<br>  ○ brand name, form, strength, regime, new medication to patient status, number of repeats,<br>• Non-pharmacological management:<br>  ○ Counselling and procedures (including attempts to effect a change in a patients lifestyle)<br>  ○ New referrals,<br>  ○ Pathology and imaging orders. |
| **Use of Data** | Provide information regarding the particular cycle of care for the encounter inclusive of patient presentation, diagnosis of problem and method of patient management. |

**Table 7: BEACH GP Characteristics Data**

| Data Source | BEACH GP Characteristics Data |
|---|---|
| **Information Provided** | • Information regarding the general practitioner |
| **Data Fields** | **GP Characteristics Data**<br>• Age and sex of GP<br>• Years in general practice<br>• GP sessions per week<br>• GP employment profile<br>• Consultancies other than within English<br>• Main practice postcode<br>• Training details<br>• After hours care arrangements<br>• Computer usage |
| **Use of Data** | Establish a profile of participating GPs, utilised as a comparator against other data sources. |

*BEACH data limitations*

Several items have been identified in the use of BEACH data and the below caveats apply to interpretation and use of BEACH in comparative analyses, including[13]:

- Participating BEACH GPs may not represent the Australian GP profile, inclusive of activity levels; post-stratification weighting is performed however on both of these factors for more precise national estimates,
- Some cluster sampling assumptions are violated,
- Prescription rates per encounter does not take into account the number of repeats as provided as part of the prescription,
- There is a maximum of four medications which can be recorded for each of the four problems presented by a patient i.e. a maximum of sixteen medications (although only 2.6% of encounters had four or more medications prescribed),
- The number of repeats ordered for a prescription was skewed towards the maximum allowable through PBS provisions which allow this to occur i.e. one months supply and five repeats. Therefore 26.9% of prescriptions had five repeats ordered. Additionally 33% of ordered prescriptions did not have the number of repeats defined, making extrapolation of intended prescription difficult.
- It is recognised that some prescriptions are not presented for dispensing, non-redemption rates for prescriptions vary between 5.2% in the UK[14] and 13% in the NZ experience[15],

- As the PBS does not record medications which fall below the subsidy threshold and private prescriptions, suggestions are made that PBS data cannot be used to monitor significant areas of GP medication management nor be used as a comparator to BEACH data,
- Shifts of some forms of management (i.e. education, patient reviews, etc.) to other health providers (e.g. Division clinics), may not be reflected in collections of data which relate to these management types. For example data on patient education could be perceived to have fallen as management of patients shifts from the direct domain of general practice and general practitioners.

*Using BEACH as a comparative prescribing dataset to PBS data*

The following items have been documented which highlights the problems faced in making assumptions of the BEACH dataset against the HICs prescribing data. They include[13]:

- Medications within BEACH include prescribed, those supplied by a GP (including samples, medications unavailable at short notice or are restricted), and those medications which are over the counter advised. These items would be included within the total count of medications recorded as prescribed including repeats.
- Each script within BEACH is counted once, irrespective of repeats ordered, whether or not the medication is covered by the PBS or patient Health Care Card and co-payment status.
- BEACH datasets do not indicate whether a prescription has been dispensed to the patient.

*The production of BEACH data specific to the EDQUM drug target groups*

Recently the GPSCU had been approached to provide baseline information on the specific areas that the current EDQUM pilot is delving into; namely data regarding the usage of statins, PPIs and antibiotics. The production of these reports has been sponsored by the Department of Health and Ageing and as a result summary reports of BEACH encounter data relating to these subjects for the 2001-02 BEACH collection year have been made available. Such summary data in included within the following section “Data sources in the context of informing EDQUM”.

**The Drug Utilisation Sub-Committee (DUSC) Prescribing Data**

The Department of Health and Ageing has derived a combined database on the prescription estimates for the non-subsidised sector (which includes general co-payment and private prescriptions), and counts of those prescriptions categories that are submitted to the HIC for payment of a subsidy. The original datasets that are combined include data from respectively, surveys conducted in a representative sample of community pharmacies which is used to ascertain non-subsidised use of prescription medicines in the Australian community (i.e. Pharmacy Guild Survey as discussed above) and secondly the HIC's collection of prescriptions subsidised by the PBS/RPBS[1]. Data is presented at an aggregated national level only

Drug utilisation is measured both in absolute terms (script volumes) and with the defined daily dose per 1000 population per day (DDD/1000/day) calculated as:

$$\frac{N \times M \times Q \times 1000}{DDD \times P \times D}$$

Where N is the number of prescriptions dispensed in the year, M is the mass of each dose (eg, in milligrams or grams and needs to be expressed in the same unit as DDD), Q is the average dispensed quantity per prescription, P is the mid-year Australian population and D is the number of days in a year. The DDD/1000/day can therefore be calculated over other time periods such as month or quarter[1]. For PBS items the mass amount (M) is the amount of active contained in an individual dose unit eg, tablet, capsule, suppository etc, whereas, because prescription estimates from the Pharmacy Guild/Chemdata survey (see below) do not include quantity information, the mass amount for non-subsidised items is the total amount of active contained in the pack. For prescriptions forwarded for subsidy, the average quantity dispensed Q, is available from the HIC data. For prescriptions that are priced under the general co-payment, this quantity is assumed to be the average quantity of the subsidised prescriptions for that drug (i.e., as concessional, safety net and Veterans Affairs prescriptions). For private prescriptions the quantity dispensed is assumed to be the retail pack size.

The defined daily dose was developed by the World Health Organisation's Collaborating Centre for Drug Statistics Methodology and uses the assumed average

dose per day of the medication for the main indication in adults[1]. Again, information on dispensed medications not appearing within the PBS are not obtainable. Limitations associated with the ATC/DDD methodology have been detailed and include:

- Dispensed medications not necessarily representing consumption of medications,
- Defined daily dose determined for the population as a whole and may differ for particular age categories or gender,
- Difficulty in assigning a DDD (and main indication) for multiple constituent products,
- Medications may have a number of indications and therefore it may be difficult to establish the principal indication of usage.
- The DDD is derived from international knowledge and may not represent the prescribed Australian adult dose.

Additional limitations regarding what is not captured within the combined dataset apply to:

- Over the counter medications (except for S3 Recordable medications),
- Information on prescribing within public hospitals,
- Highly specialised medications for out-patients dispensed through public hospital pharmacies and aboriginal medical services,
- Effects seen by stockpiling of medications due to reaching the safety net level, although this has been rectified with a limit to re-dispensing chronically used drugs to not less than 20 days.

**Table 8: DUSC Combined Database Data**

| **Data Source** | **DUSC Combined Database (PBS & Pharmacy Guild Survey)[14]** |
|---|---|
| **Information Provided** | Complete coverage of the community use of prescription drugs at the national level |
| **Data Fields** | **Combined Database (PBS & Pharmacy Guild Survey)**<br>• Anatomical Therapeutic Chemical (ATC) code (anatomical main group, therapeutic main group, therapeutic subgroup, chemical / therapeutic subgroup, generic drug name),<br>• Quantity dispensed,<br>• Number of prescriptions; and<br>• Defined daily dose.<br>**The DUSC database also includes the following data items for PBS/RPBS subsidised prescriptions only:**<br>• Pay category of recipient (i.e. general, concessional, repatriation, safety net, doctor's bag, etc.),<br>• Benefits paid (i.e. cost to Government), |

| | |
|---|---|
| | • Patients' contributions,<br>• Total cost (i.e. cost to Government plus patients' contributions),<br>• Postcode of dispensing pharmacy,<br>• State and Territory of dispensing pharmacy,<br>• Statistical Local Area (SLA) of dispensing pharmacy. |
| **Use of Data** | Inform the usage of all community prescribed medication as a pre-defined measure i.e. defined daily dose, comparative against past surveys allows trend analysis on usage and comparisons against international utilisation data. |

**National Prescribing Service (NPS) Pharmaceutical Utilisation Data**

The two sources that the NPS has used for the collection of health information has been previously characterised as that derived from patients' health records – directly or from external sources - and secondly data derived from NPS conducted and/or endorsed educational activities.

Review of the data sources of the NPS and personal communication has determined that specific data collection activities are not undertaken to inform EDQUM, besides those as reported within evaluation reports of the NPS. Such reports include those educational activities that have targeted consumers within areas of antibiotic usage for example. Further information regarding this area will be provided upon feedback from the NPS on sources of data, which may impact EDQUM activities. For example, does consumer education allow prescribing patterns to be influenced at the GP level, and methods that allow such interventions to be evaluated?

## Data Sources in the context of informing EDQUM

### HIC datasets in the context of EDQUM

The use of PBS data within the context of informing GPs regarding the volumes of drugs prescribed above the co-payment (i.e. $22.40), is regarded as a complete source of data. Further linkage of this complete source with MBS data, would allow prescribing (ultimately dispensing) rates to be calculated by drug by patient age / sex, albeit constraints exist - as mandated by the Privacy Commissioner - ultimately preventing linkage of MBS and PBS records at the unit level.

As the record of a GPs' services recorded within the Medicare services files (particularly in the context of PBS records) is reliant on whether the relevant drug exist above the patient contributed co-payment or concessional status of the drug; the ability to derive meaningful data relies on:

- Assumptions about under co-payment activity based on Australian Statistics of Medicine (ASM) data to derive all prescribed records,
- The selection of an appropriate and comparative drug dosage / strength standard i.e. defined daily dose (DDD) or other,
- That temporal trends are associated with a fixed co-payment status i.e. the drug must continually remain above or below the co-payment to exist as an electronic record,
- Records of the PBS are a volume construct i.e. volume of prescribing and not an indicator of prescribing quality,
- Accepting that the drug is utilised for its prescribed indicator,
- The roles of the consumer in regards to:
    - Actual script dispensing i.e. patient presenting to dispenser,
    - The number of repeats and the requirements for obtaining the maximum number (influenced by GP educational activities),
    - Consumer education relating to usage of medications, which may influence their decision to fulfil dispensed scripts.

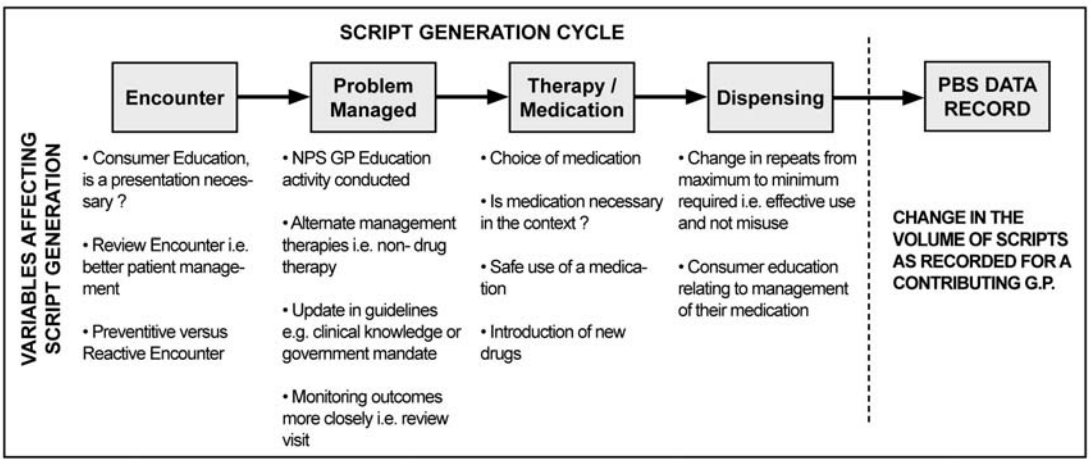
In relaying information back to individual GPs about their particular prescribing activities derived from the HIC datasets, additional caveats are also applied in respect to:

- Those GPs whose prescribing profiles are most likely to be unreliable (in a comparative sense), because of the small sample sizes involved i.e. limited number of scripts generated by the general practitioner,
- That any analysis that involves under co-payment drugs, must involve a stable subset of the concessional population i.e. require patients who have a continuous concessional status throughout a given year (personal communication, Gordon Calcino).
- The patient demographic profile which is serviced by the GP which would influence the ratio of concessional / to non-concessional scripts generated, leading ultimately to whether there is more complete capture of prescribing activity through the PBS mechanisms for those GPs who practice within low SES communities against others, albeit noting that concessional scripts form only a small percentage of total scripts dispensed i.e. $1.9 \times 10^6$ concessional scripts versus $154 \times 10^6$ non-concessional patient scripts.
- The extent of linking MBS data (i.e. Medicare number, age and gender) with the PBS records (concession and co-payment status), and whether such patient linkages are in an aggregate format or by individual practice detail. This further leads to:
    - More complexity in interpretation of patient profiles against benchmarks at peer, state or national levels,
    - Whether there is an increase in the amount of generated reports which may not be informative at a single-patient level, bearing in mind that patient-to-patient variability exists in those, for example, with co-morbidities and specific management requirements. This additional problem of patient-to-patient variability will pose problems in the analysis of derived GP desktop software records – albeit it is recognised that with bounds on the prescribing guidelines for a drug, particular groups of patients could be selected who may not conform to the prescribing indicators (i.e. algorithms / decision support); such patients should be reviewed on a case-by-case basis.
- Whether resources exist to allow linking and reporting of both the MBS and PBS data by the HIC, back to those interested parties.

*Can HIC data be utilised to determine savings?*

It has been perceived that by utilising HIC PBS data relating to the total volume of scripts dispensed by an individual GP, that activities of EDQUM which generate a decrease in the number of scripts could be attributed to that GP – generating a revenue based on such differentials of volume. It must however be noted that "measuring" such changes to provide an indication of the success of EDQUM activities is in itself fraught with challenges. It is thus suggested that the use of HIC PBS data to demonstrate perceived savings cannot be made with any accuracy with many a variable influencing what constitutes a change in volume; be it changes in patient profiles, education activities at both the consumer and GP spectra's, the addition of new drugs to the PBS, the shift of drugs from over to under the co-payment, the time-period utilised to demonstrate change, consumer behaviours etc. The below flowchart of an anticipated stream relating to patient presentations and eventual creation of an electronic PBS record, highlights some of the variables which could be encountered, which would indeed be reflected in a change in the PBS volumes recorded against a particular GP. The number of variables discounts attempts to attribute changes in the volume of scripts against future EDQUM activities.

**Table 9: Variables affecting script volumes recorded within the PBS dataset**




*Mechanisms of analysis*

Personal communication between Gordon Calcino of the HIC, has highlighted several mechanisms which exist outside of the standard prescribing (dispensing) ratios that could be conducted on current HIC GP subsets of data i.e. number of

patients who have an encounter resulting in a script which exists above the co-payment (denominator), by the number of scripts written by the GP (numerator) to derive a prescribing / dispensing ratio:

$$\text{Dispensing rate} = \frac{\text{Scripts written by GP}}{\text{Patients receiving script}}$$

Methods include integration with MBS data (as discussed above), to allow dispensing rates to be derived by drug by patient age / sex. Further detail could be integrated within reports through linking with individual patient IDs (captured within scripts), to allow patient level comparisons i.e. allows a “truer” indication of number of scripts per patient inclusive of age, sex and even socio-economic status. However it is unsure whether such data is released to parties outside of the HIC and whether such analysis is currently being conducted internally by the HIC i.e. a matter of resources and priorities.

The use of observed versus expected ratios of drug utilisation is a further method, which moves beyond simple descriptive statistics of drug usage. Whereby the expected usage of a drug could be derived from the clinical guidelines associated with a particular drug (plus requisite date which may include population susceptibilities for particular disease states derived from external sources e.g. BEACH data for example) and reported against a standard i.e. the defined daily dose. The observed utilisation of the drug would be that, that is derived from current observations recorded within the HIC datasets (requiring both the MBS and PBS datasets to allow patient sex / age standardisation). Ratio’s derived from comparisons of observed versus expected drug utilisation, could inform whether changes in utilisation is occurring over time; factoring in external variables which undoubtedly influence prescribing decisions.

With regards to PBS data which informs usage of under the co-payment drugs, the use of PBS concessional patient data could be utilised to inform such usage. Bearing in mind that there is a need to look only at those patients who have a continuous concessional status throughout the year – linkages with the script ID must be exploited to determine which concessional patients exist within this domain. Again application of the observed versus expected ratio of drug utilisation, would enlighten

under the co-payment activity; the comparative method being that which utilises ASM statistics to derive pharmacy dispensing activities of under the co-payment drugs.

**Pharmacy datasets (inclusive of DUSC data) in the context of EDQUM**

Data derived from pharmacies exists at two levels i.e. national aggregated sales versus dispensing data, and that data derived from community pharmacies; each level of data could inform standard utilisation rates for particular drugs. It is perceived that any further detail on utilisation at the community level - i.e. singular patient utilisation - could only occur if barriers to identifying singular patients could be removed (or ascribing identifiers that allow analysis). Statements however should be made that if access to de-identified dispensing data were occurring at the Division level, descriptive utilisation rates for particular drugs value would be obtained.

*Nationally aggregated pharmacy data*

It has been perceived that the role of nationally derived pharmacy data will be of limited value in informing utilisation within EDQUM activities. Although we broach the issue of the relative importance of such volume data, such statements can be easily be applied to that derived from the other source of drug volume data – the HIC. It is there important to state that such aggregated national datasets may provide useful in determining utilisation ratios for the subset of EDQUM target drugs.

It has been recorded that one source includes sales data (Intercontinental Medical Statistics); recognition that such a source is firstly inaccessible to Divisions, may not be particularly informative and is an aggregate of nationally derived statistics which cannot be unpacked to inform local activities. But additionally may inform of the ratio of under the co-payment to over the co-payment drug utilisation, this being the value of such a source of data – access to such a source is however the issue. It is thus anticipated that Divisions of General Practice are unable to use such a source within the context of EDQUM.

*The Drug Utilisation Sub-Committee (DUSC) Prescribing Data*

Other activities that aggregate sources of pharmacy data, principally the activities of DUSC, are informative with respect to standard utilisation of drugs. DUSC data through its use of Pharmacy Guild Survey data, informs the use of non-subsidised use of prescription medicines (i.e. general under co-payment statistics) against HIC collections of prescription data within the domain of subsidised and over the co-payment medications. Derivation of the ratios of non-subsidised to subsidised and additionally under co-payment activity ratios (through linking of PBS concessional records), provides comparative usage indicators for data sources that may capture subsets of volume prescribing i.e. HIC over the co-payment records.

The additional quality aspect of this source of data is through the use of a "standard" measure for drug utilisation: the defined daily dose (DDD). Therefore derivation of drug utilisation from other sources of data, could feasibly utilise the same methodological construct as the DDD; thus providing a comparative meaningful figure.

*Community pharmacy data*

Although sources of data exist which inform the dispensing of medication derived from general practitioner script generation, considerable limits exists which impede the flow of such data back to the GP or organisations which service GPs. Limits that do exist and which would inform EDQUM to the greatest extent, are those that hinder information exchange from the dispenser (the pharmacist) back to the GP or organisation – particularly patient identifying information which allows more detailed reviews of patient consumption of medication post-management of their illness. For example pharmacist derived data could inform at the GP-level, what is being dispensed through the GP provider number which is generated from a dispensing electronic record; providing information about all prescribed drugs regardless of co-payment or concessional status. Additional to this, mention has been made of the capture of over-the-counter (OTC) medication capture by pharmacists; however the limited involvement of such drugs in the subset of chosen EDQUM foci negates the need to capture these additional "transactions". Albeit mention should be made in regards to ulcer treatment conducted by GPs versus dyspepsia treatment by pharmacists - two disparate sources of data that would inform prescribing and dispensing activities for the treatment of "peptic ulcers".

*Interpretation of Community Pharmacy Derived Data*

Possible problems faced in interpretation of pharmacy data include:

- Marked variation in capturing over the counter medication, limited data links to enable this to occur,
- Problems of access to some pharmacy computer systems and pharmacy data in general, cost involved and co-operation with participating pharmacists,
- Local GPs may not have written the majority of medication scripts dispensed by the pharmacy; unless data was linked back to prescriber name i.e. GP ID,
- Maximum quantity of repeats often prescribed by GP, with no mechanism to determine the extent of patient consumption; require link between patient name at pharmacy level data and comparative prescribing data at the general practice level to determine "actual" consumption rates[11],
- Drug strength and dose would have to be converted into a standard format i.e. defined daily dose (DDD).

**General practice desktop datasets in the context of EDQUM**

Three sources of general practice data have been identified which would inform EDQUM activities of drug utilisation and which would act as a comparative source against locally derived and aggregated Divisional data. These sources include:

- Locally derived collections of individual GP desktop data, that which occurs concurrently through typical day-to-day interactions by GP within their practices, through clinical software usage with analysis at the GP-level,
- Locally derived collections of individual GP desktop data, aggregated by third-parties (inclusive of HCN's GPRN collection activities) and reported at the national level,
- Data derived from practices that provide "quality" data for research purposes e.g. the University of Adelaide's MEDIC-GP dataset.

Each source of aggregated data can be used as a comparative source (to some extent), to that derived at the local level. Emphasising that the smaller the size a dataset is, the more limited it is, in comparing activities contained within it against the aggregated datasets. The variability of small desktop derived datasets is an important factor in efforts aimed at comparing like-sized datasets i.e. practice against practice or Division aggregations against other Divisions aggregations. Therefore the

statement that has to be made, is that aggregated datasets (i.e. GPRN and MEDIC-GP) are useful in informing in a “general” sense drug utilisation against condition and is ultimately dependent on multiple factors which influence information recorded within local collections of data:

- Whether data is firstly recorded by the GP, the increased use of a software product over time and its inherent increasing “quality”,
- Stability of the practice population if data is used for longitudinal analyses,
- Representativeness of the individuals who comprise the database population e.g. patient age / sex mix, co-morbidities, socio-economic status, etc..

Further discussion of the relative advantages and disadvantages of using desktop data - recognised as *the* most important source in informing activities of EDQUM - is included within the previous section that describes EDQUM Data Sources.

*Utilising GPRN in the context of EDQUM*

A suggestion has been made which may allow aggregation of EDQUM practice data into a singular dataset, through a similar mechanism of the GPRN i.e. the signing up of EDQUM practices who utilise the Medical Director clinical software.

This would occur through the GPRN tool which extracts data for participating EDQUM Divisions i.e. the data is extracted to HCN (or Divisions) with subsets of data created based on a Division code or postcode for example (analysis or reports would be completed by Division stratified by GP/practice/etc.) - as the main interest is in pharmaceutical data, it would fit within with current GPRN activities requiring no real modification to the process besides signing up to the GPRN.

Alternately there is the option of approaching HCN with regards to a customised GPRN-like tool which extracts firstly a subset of data which informs EDQUM (i.e. diagnosis, history, progress notes plus script files) and secondly extracts it to individual Division safe house - noting that this option has been done in the past for other parties utilising a customised GPRN extraction tool; all that changes is the destination of the extracted file.

Each of these options albeit, rely on GPs utilising the requisite software of HCN; other Divisions would have to proceed on their own or utilise other options. One other

option includes that which packages the derived clinical software data within a data standard “wrapper” such as HL7 to be emailed in an easily accessible format cf. ARGUS and alternate messaging products (e.g. intrahealth profile www.intrahealth.co.nz/profile.html).

**BEACH datasets in the context of EDQUM**

The survey results of general practice activity within the context of the BEACH data collection, informs aggregated national figures of encounters, management and GP characteristics. In informing the activities of EDQUM, limits are placed on utilising this source of data in that no direct comparisons could be made against individual Divisional activity, albeit possible comparisons could be elucidated against an aggregated set of EDQUM Division data; this currently is not the case, nor would be an aim of EDQUM and again would be of limited value.

Thus data items, which would be of some value for EDQUM, are those that provide detail of patient management within the GP-patient encounter i.e. BEACH Management Data (see Table 6). Importantly this could include information on those medications that are not captured within the PBS dataset:

- Under co-payment drug use; captured within BEACH as medications prescribed for each problem (if any), including over-the-counter advised,
- Ratio of subsidised to non-subsidised encounters and hence medication ratios within these two groups, allowing possible extrapolation of PBS data; captured by BEACH through the Patient Data item of health care card status and type,
- “Standard” utilisation rates; medication details recorded for each medication prescribed i.e. brand name, form, strength, regime, new medication to patient status, number of repeats. Any comparisons against other data would have to utilise and be standardised against defined daily dose (DDD) for example.

With the caveats of comparing BEACH data against HIC prescribing data, inclusive of:

- BEACH data includes prescribed, GP-supplied (including samples, medications unavailable at short notice or are restricted), and over the counter advised medications. Items would be included within the total count of medications recorded as prescribed including repeats.

- Each script within BEACH is counted once, irrespective of repeats ordered, whether or not the medication is covered by the PBS or patient Health Care Card and co-payment status,
- The BEACH datasets does not indicate whether a prescription has been dispensed to the patient.

An additional important data item collected is that which includes non-pharmacological management, particularly attempts to influence patient behaviour. Thus GP-mediated patient education could inform for example, how the NPS or other providers of educational activities are ultimately reaching there targets i.e. both the GP and the consumer. It is also important to note, that the BEACH activities may indirectly capture the shift of patient education from the domain of general practice to other providers of such education or management. These providers include Division clinics in the form of asthma & diabetes clinics and the role of pharmacists. It is uncertain at this stage, the degree of analysing the trends involved within non-pharmacological management, how it is influenced (by whom) and whether it could be directly used to measure activities that enhance the quality use of medicines i.e. through education.

**PPIs in general practice**[16]

**An overview of peptic ulcers and proton pump inhibitors**

In this section two separate reports of GP activity are provided from the BEACH data 2001-02. The first is a summary of selected data pertaining to the relative rate of management and the GP management of peptic ulcers. The other presents selected data about Proton pump inhibitors (PPIs).

**Peptic ulcers:**

- Were managed at a rate of 0.28 per 100 encounters, or 2.8 per 1,000 encounters,
- Extrapolated to the approximately 103 million GP-patient consultations across the country we estimate there were about 300.000 consultations that involved the management of peptic ulcer,
- Medications were prescribed/advised/supplied at a rate of 82.6 per 100 contacts with peptic ulcer,
- Exactly 50% of these medications were PPIs,
- Extrapolated national estimate is therefore 123,900 occasions on which a GP prescribes PPIs for peptic ulcer,
- Remember when looking at the rate of prescription of triple therapy for peptic ulcer, that this rate is influenced by the one-off nature of triple therapy, compared with the continuing nature of PPI treatment.

**Proton Pump Inhibitors:**

It is worth noting from the PPI report that:

- The National estimate of the annual number of PPI scripts from GP = 1.73 million,
- Peptic ulcers accounted for only 7.2% of the problems for which PPIs were prescribed,
- Oesophageal disease accounted for 65% of the problems managed with PPIs.

**Further information**

More detailed information is available from the following PDF documents:

- http://www.fmrc.org.au/divisions/Peptic_ulcer_summary.pdf and,
- http://www.fmrc.org.au/divisions/PPI_summary.pdf

**Statins in general practice**[17]

**Medication summary:**

- Statins were prescribed at a rate of 2.46 medications per 100 encounters and at a rate of 1.72 per 100 problems managed.
- Account for 2.36% of all medications prescribed for all problems.
- National estimates suggest a total of 2.59 million scripts/supplied medications in 2001-02 for statins.

**Encounter summary:**

- Ninety four percent of encounters at which a statin was prescribed/provided were conducted in the surgery,
- The patient receiving statin was more likely to be male (50.8%, 95% CI 48.3-53.3) when compared with the usual gender distribution in encounter data in Australia (42.6, 95% CI 41.43.3),
- 95% of the patients at encounters generating a Statin script were over 45 years, 42% being aged between 45 and 64 years, though this age group only accounts for 26.3% of total GP encounters (i.e. they are over-represented in the statin group). All other older age groups are also over represented but not to the same degree,
- Health care cardholders were also over-represented in this Statin sample, accounting for 62.3% of these encounters (compared with 41.9% of total encounters),
- Patients receiving a statin at the encounter described more reasons for encounter than average. These two results would be related to their age distribution,
- Very few of these patients were new patients to the practice.
- Only 7.5% of the problems managed with a statin were new problems to the patients,
- Ninety five percent of the statins were prescribed and 5% were supplied by the GP direct to the patient.

**Problems managed with encounter summary:**

- "Lipid disorder" was the most common problem for which statins were prescribed/supplied (75.2%) and the remaining 25% of morbidity managed with statins included a very broad range of problems with low relative frequency rates of specific conditions.
- IHD accounted for 5.7% of the problems for which statins were prescribed/supplied, hypertension 4.2% and diabetes 1.3%.
- 63% of GP contacts with the problem 'lipid disorder' generated a script/supply of a statin whereas only 11.18% of contacts with ischaemic heart disease and 1.14% of contacts with hypertension resulted in a statin script/supply.
- It is unsure whether contacts with the problem "Blood test metabolic/endocrine" also included the cohort of diabetes patients (14.2% of such contacts resulted in a script/supply for statins). Diabetes as a separate contact only resulted in 1.04% of such contact receiving a script or supply for statins.

Further more detailed information including the average Prescribed Daily Dose (PDD) and the number of repeats for statins, are included within the following link:

- http://www.fmrc.org.au/divisions/Statins_summary.pdf

**Antibiotics in general practice**[18]

**Medication summary:**

- Antibiotics (excluding anti-fungals) were prescribed at a rate of 14.1 per 100 encounters and at a rate of 9.8 per 100 problems managed,
- They account for 13.5% of all medications prescribed/supplied for all problems,
- National estimates suggest a total of 14.8 million occasions on which GPs prescribed/supplied an antibiotic during 2001-02,
- Amoxycillin accounted for 21.7% of these medications (alone accounting got almost 3% of all medications for all problems). This was followed by Cephalexin (15% of all antibiotics), roxithromycin (10.4%), Cefaclor monohydrate (8.0%), doxycycline (5.8%) and Erythromycin (4.4%). Together these made up approximately two thirds of all antibiotics recorded.

**Encounter summary:**

- Ninety four percent of encounters at which an antibiotic was prescribed/provided were conducted in the surgery,
- The sex distribution of patients at these antibiotic encounters was normal (approximately 40% being male). Children were over-represented at these encounters. And about 11% were aged less than 5 years (compared with about 7% in the total data set, and 12% were aged between 5 and 14 years (compared with 6.4% in the total data set. Young adults (15-24 were also somewhat over represented. The older age groups were under-

represented in this antibiotic sub-set. This suggests that children and young adults are more likely to receive an antibiotic when they visit the GP than are adults (perhaps reflecting more acute presentations and less chronic conditions in children),
- The characteristics of the patients receiving an antibiotic reflected those of the total sample in terms of their health care card status, their non-English speaking background status and their status to the practice. They reported an average number of reasons for encounter,
- 96.5% of the antibiotics were said to have been prescribed and the balance were supplied directly to the patient by the GP.

**Reason for encounter summary:**

- The most common reasons for encounter (RFE) recorded at those encounters at which an antibiotic was prescribed were:
    - Cough (23.4 per 100 antibiotic encounters),
    - Throat symptoms/complaints (15.4 per 100 antibiotic encounters),
    - Fever (7.0 per 100 antibiotic encounters) and,
    - URTIs (including 'cold') (6.8 per 100 antibiotic encounters).
- The top 20 RFEs accounted for 67% of all RFEs recorded at these encounters.

**Problems managed with encounter summary:**

- The most common problem for which antibiotics were prescribed or supplied include:
    - Acute bronchitis/bronchiolitis (accounting for 15.7% of all problems managed with an antibiotic),
    - URTI (14.7% of all problems managed with an antibiotic),
    - Urinary tract infections (8.4% of all problems managed with an antibiotic),
    - Sinusitis (8.1% of all problems managed with an antibiotic),
    - Acute otitis media (7.8% of all problems managed with an antibiotic) and,
    - Tonsillitis (7.4% of all problems managed with an antibiotic).
- URTI was far less likely to generate a script for antibiotics (33% of all problems managed with an antibiotic) when compared with:
    - Acute bronchitis (80%),
    - UTI (73%),
    - Sinusitis (82%),
    - Otitis media (83%) and,
    - Tonsillitis (93%).
- Three quarters of the participating GPs prescribed an antibiotics for acute bronchitis on at least one occasions and 59.3% prescribed one for URTI at least once during their 100 recorded encounters.

**Further information**

Further more detailed information including the average Prescribed Daily Dose (PDD) and the number of repeats for antibiotics, are included within the following link:

- http://www.fmrc.org.au/divisions/Antibiotics_summary.pdf

## References


1. Commonwealth Department of Health and Aged Care. Australian Statistics on Medicines 1998. Canberra: AusInfo; 1999. pp2-17

2. MacIsaac, P. Calcino, G. and Lea, T. Chapter 13- General practice data issues In: General practice in Australia: 2000. Commonwealth Department of Health and Aged Care. 2000 Pirie Printers: Fyshwick pp463-66.

3. Public Health Association. Medicare fact sheet 10: The Pharmaceutical Benefits Scheme [online] 2001. (Available from URL: http://www.drs.org.au/new_doctor/75/fact_sheet_10.htm)

4. Harvey, K. The Australian Pharmaceutical Benefits Scheme: Background Paper. (Available from URL: http://www.pha.org.au/friends_of_medicare/PBAC%20Background.html)

5. Productivity Commission. International Pharmaceutical Price Differences, Research Report. Canberra: AusInfo; 2001.

6. Public Health Association. The Australian Pharmaceutical Benefits Scheme and the Pharmaceutical Benefits Advisory Committee controversy. (Available from URL: http://www.pha.org.au/friends_of_medicare/PBS%20PRIMER%20&%20QA.html)

7. Pharmaceutical Benefits Scheme. Schedule of Pharmaceutical Benefits, May 2002, Patient Charges. May 2002. (Available from URL: http://www1.health.gov.au/pbs/contents/explain4.htm)

8. Waters A-M, Armstrong T, Senes-Ferrari S 1998. Medical care of cardiovascular disease in Australia. AIHW cat. no. CVD 4. Canberra: Australian Institute of Health and Welfare (Cardiovascular Disease Series no. 7). pg 2 to 17.

9. QUM Roadmap Department of Health and Ageing. Health Insurance Feedback (ID: 766). Quality Use of Medicines Database. (Available from URL: http://qummap.health.gov.au/plist.asp?project=766).

10. McManus et al 1997. Antibiotic use in the Australian community, 1990-1995. MJA 167; 124-127.

11. Gill G, Loone K 2002. What database should GPs use for asthma medication usage audit? 2002 General Practice and Primary Health Care Research Conference. Melbourne, Australia; Primary Health Care Research and Information Service September: Adelaide, SA, 2002.

12. Health Communication Network. General Practice Research Network. (Available from URL: http://www.australiandoctor.com.au/gprn.asp Cited November 20, 2001).

13. Britt H, Miller GC, Knox S, Charles J, Valenti L, Henderson J, et al. 2002. General practice activity in Australia 2001–02. AIHW Cat. No. GEP 10. Canberra: Australian Institute of Health and Welfare (General Practice Series No. 10).

14. Beardon PH, McGilchrist MM, McKendrick AD, McDevitt DG & MacDonald TM 1993. Primary non-compliance with prescribed medication in primary care. BMJ 307(6908):846–8.

15. Gardner TL, Dovey SM, Tilyard MW & Gurr E 1996. Differences between prescribed and dispensed medications. N Z Med J 109(1017):69–72.

16. General Practice Statistics and Classification Unit, University of Sydney. Summary problem report for GP divisions: Peptic ulcers in general practice 2001-2002. (Available from URL: http://www.fmrc.org.au/divisions/Peptic_ulcer_summary.pdf. Cited May 1, 2003).

17. General Practice Statistics and Classification Unit, University of Sydney. Summary problem report for GP divisions: Statins in general practice 2001-2002. (Available from URL: http://www.fmrc.org.au/divisions/Statins_summary.pdf. Cited May 1, 2003).

18. General Practice Statistics and Classification Unit, University of Sydney. Summary problem report for GP divisions: Antibiotics in general practice 2001-2002. (Available from URL: http://www.fmrc.org.au/divisions/Antibiotics_summary.pdf. Cited May 1, 2003).

.

# Appendix A

## Table Of Data Sources For EDQUM

| Data Source | Features | Limitations | Potential | Recommendations for use for EDQUM | Contact for further info |
|---|---|---|---|---|---|
| **HIC** | • Detailed PBS data available quarterly<br>• Divisional data soon to be available on HIC website:<br>http://www.hic.gov.au/statistics/imd/forms/gpStatistics.shtml | The following data not captured:<br>• Below co-payment<br>• OTC (over the counter) drugs<br>• Private prescriptions<br>Data currently not available trough website ➔ special requests via HIC or DoHA ➔ long wait<br>For more detailed limitations, see:<br>http://www.mja.com.au/public/issues/176_09_060502/rob10519_fm.html | Potential for:<br>• Monthly data<br>• Linkage with age and sex data | • this is the data that the government is interested in – any savings will be calculated from this data | Gordon Calcino<br><br>gordon.calcino@hic.gov.au |
| **Pharmacy Data** | Includes:<br>• Data below co-payment<br>• OTC data<br>• Private Scripts | • Data sits with individual pharmacies<br>• Data no easily accessible – depends on cooperation of local pharmacies ➔ cost?, privacy issues, not practical in most instances<br>• Cannot be easily linked with geographic Divisional data | | This may be feasible for some Divisions (eg. Tasmania), but not others. Talk to Mary Collins about how they got hold of the data. | Mary Collins initially, then local pharmacies |
| **GP desktop data** | Prescribing data<br>Links patient data with scripts ➔ only place where quality indicators can be implemented | • Prescribing data gives no indication of whether script was filled and medication taken.<br>• Requires GPs to input all patient data electronically<br>• Coding issues ➔ lack of consistency across software and GPs | Potential for:<br>• Better data query capabilities<br>• better data extraction | For use within practices and with individual GPs | Divisional IM/IT Officers |
| **BEACH data** | Random sample of 999 GPs with 99,900 patient encounters examining:<br>• GP & patient characteristics<br>• Reasons for encounter<br>• Problems managed<br>• Management techniques used<br>Produced annually<br>Available at: http://www.aihw.gov.au/publications/gep/gpaa00-01/index.html | • National, not local data set. | Local data could be made available if a Division is big enough (or two or more Divisions combine) ➔ at a cost | Useful to put pharmaceuticals / prescriptions into perspective with the consultation. | Helena Britt<br><br>helenab@fmrc.med.usyd.edu.au |
| **DUSC data** | • Data is presented in defined daily dose per 1000 population per day (DDD/1000/day).<br>• Data used from PBS and survey of community pharmacies ➔ includes below co-payment data<br>• facilitates international comparisons of drug utilisation profiles<br>1998 Australian Statistics on Medicine available at: http://www.health.gov.au/pbs/pubs/asm.htm | • Currently available data (1998 data) is very old. On 25th October 2002 the 1999 and 2000 Statistics were still at printers – it should be available in about 6-8 weeks<br>• Data relatively difficult to interpret and use | | May be useful to estimate the proportion of scripts below co-payment | DUSC Secretariat: (02) 6289 7293 |
| **NPS data** | | • | | | |

# Appendix B

## Subset Of DUSC Data Sources Table

- The following matrix documents a subset of data sources that are utilised to inform DUSC activities i.e. data sources which are accessible by the participating Division's of General Practice.

- Sources of data are presented which provide information on what is collected and how they may be utilised within the context of EDQUM activities, noting that caveats will apply to what each of the selected data sources may offer.

- A perceived checklist of "Items useful in context of EDQUM" has also been included as a separate column.

| | | Bettering the Evaluation and Care of Health (BEACH) | General Practice Research Network (GPRN). Health Communication Network (HCN) | MEDIC-GP | HIC claims processing PBS/RPBS subsidised prescriptions | Pharmacy Data | Item Useful in EDQUM Context |
|---|---|---|---|---|---|---|---|
| **Based on** | Drugs prescribed | x | x | x | | | x |
| | Drugs dispensed | | | | x | x | x |
| | Drugs provided/sold through the supply chain | | | | | | |
| | Population self-reports/patient surveys | | | | | | |
| **Focus of drug usage information collected** | General Practice | x | x | x | | | x |
| | Community pharmacy | | | | x | x | x |
| | Hospitals | | | | some public hosipitals dispense PBS scripts | | |
| | Community and hospital | | | | | | |
| | Nursing home / hostel | | | | | | GP serviced |
| | Population survey | | | | | | |
| | Other (specified) or further details on the above | | | | Some private hospitals supplied using the PBS | | |
| **Coverage of the database (of focus identified in response to above)** | Full capture | | | | x | | x |
| | Full capture of restricted sub-set (specified) | | | | | | Sub-set specified |
| | Based on sample (reference for sampling methodology or N/A if not available) | Britt et al, BEACH. A study of general practice activity | Methodology is available | Based on a set of GPs who utilise computerised record systems | | | |
| **Outline of regional analysis that is possible (or N/A if not possible)** | | Analysis by Divisional Areas of General Practice is possible on one years data fro large Divisions e.g. West Sydney | State by rural / remote vs. urban | N/A | State file; postcode of pharmacy; postcode of doctors major practice, postcode of patient | Postcode of pharmacy, prescribing doctor, postcode of prescribing doctor | x |
| **Coverage of OTC medicines** | | Collection of all OTCs advised for purchase by the GP and on ay drugs rived to the patient directly by the GP. GPs can also recommend "alternative medicines" | Not explicitly | If recorded by the GP | No | Yes, if prescribed by the GP | Limited requirements |
| **Unit of analysis** | Person (allowing medication profiles on individual patients to be assembled over time) | | able to follow patient within practice | x | | | x |
| | Person (but not allowing medication profiles on individual patients to be assembled over time) | | | | | x | |
| | Encounter (between doctor and patient) | x | x | x | | | x |
| | Drug | x | x | x | x | x | x |
| | Other (specified) | | | | can track individuals | | |

| | | Bettering the Evaluation and Care of Health (BEACH) | General Practice Research Network (GPRN). Health Communication Network (HCN) | MEDIC-GP | HIC claims processing PBS/RPBS subsidised prescriptions | Pharmacy Data | Item Useful in EDQUM Context |
|---|---|---|---|---|---|---|---|
| **Drug details collected** | Generic name | x | x | x | x | x | x |
| | Brand name | x | x | x | x | x | x |
| | Form and strength | x | x | x | x | x | x |
| | Dosage | x | x | x | quantity, original and repeats | x | x |
| | Indication/problem managed | x | x | x | | | x |
| **Nature of collection** | On-going (start date of collection specified) | Apr-98 | Jan-99 | Jul-94 | 1987 | Dependent of pharmacy | Start of EDQUM |
| | Periodic (frequency of collection specified) | | | | | | |
| | For ad hoc projects only (details of most recent project specified) | | | | | | |
| | Other (specified) | | | | | | |
| **Patient details** | Patient details collected | Yes | Yes | Yes | Yes (May 2002) | Yes, but information not available | Required |
| | Age | x | x | x | x | | x |
| | Gender | x | x | x | x | | x |
| | Diagnosis | x | x | x | | | x |
| | Reason fro drug therapy | x | x | x | | | x |
| | Co-morbidity | x | x | x | | | x |
| | Pathology / radiology ordered | x | x | x | | | |
| | Other (specified) | HCC status, NESB, postcode, ATSI, new/old patient to practice | Sub-set of practices with pathology results and concurrent medications | Validation of clinical notes contained within the MEDIC-GP database would be required for further analysis | | Medicare Number | |
| **Prescriber details** | Prescriber details collected | Yes | Yes | No | Yes | Yes | Required |
| | Age | x | x | | x | | |
| | Gender | x | x | | x | | |
| | Specialty | All are GPs | All are GPs | | x | | |
| | Other (specified) | Country of graduation, practice size, number of sessions, years in practice, postcode, qualifications, computer use | length of computerisation, place of graduation, workload | Prescriber information is not routinely collected | Place of graduation, year of graduation | Provider Number | Provider Number |
| **Details of how ethics issues regarding privacy and confidentiality are handled** | | Identified information on prescribers is collected but kept on separate database from all encounters | Signed informed consent for doctors to participate, patient de-identified at the practice prior to transfer to HCN. Patient opt-out facility for use of de-identified information | No identified information on patients, prescribers or pharmacies is collected. | Identified pharmacies and doctors are collected by the HIC | No identified information on patients collected, pharmacy postcode is not generally disseminated. | Address current RACGP privacy mandates |
| **Coding systems used** | Drugs | ATC and CAPS | AZ-DEX mapped to ATC | Inhouse based on MIMS | PBS/RPBS, Amfac Chemdata for non-PBS, ATC, ICPC2 Plus | Proprietary systems | Comparable |
| | Morbidity / indication | ICPC2 Plus | DOCLE partial map to ICD | SOAP arrangement (Subjective, Objective, Assessment, Problem type format) | | | Comparable |

| | | Bettering the Evaluation and Care of Health (BEACH) | General Practice Research Network (GPRN). Health Communication Network (HCN) | MEDIC-GP | HIC claims processing PBS/RPBS subsidised prescriptions | Pharmacy Data | Item Useful in EDQUM Context |
|---|---|---|---|---|---|---|---|
| Number of doctors, practices, hospitals, etc. covered in the most recent 12 months or recent collection | Doctors | 1,047 | 175 | 153 | | | |
| | Practices / pharmacies | | 78 | 6 | 4,925 pharmacies | | |
| | Hospitals | | | | | | |
| | Patients | 95,000 | 300,000 | 55,187 | | | |
| | Encounters | 104,700 | 1.6 million | 613, 351 (of total of 915,773 notes) | | | |
| | Prescriptions | 98,372 (plus 9,842 OTCs and 7,218 items provided directly by GPs) | 1.5 million | Not known | 146 million PBS + 12.9 million RPBS | | |
| Strengths of database | | Validated, published methodology with continuous improvement. Continuous, large, representative of general practice and sensitive to changes in practice; Timely analysis and dissemination of results. Uses standard classifications which facilitate meaningful analysis. | Longitudinal - large scale numbers. Patient base and granularity, timeliness and prescription details. Treatment changes - initiation, continuation or switch. | Comprehensive and longitudinal clinical information. Information includes details of consultations, pathology, diagnostic imaging, referrals and prescriptions. Representativeness of patients. | Information related to drug and cost. Can track individual patients with Medicare number via reocrd of number on script. | Capture of all prescribed scripts regardless of co-payment or subsidy status; geographical capture; record of dispensing activities of scripts; | N / A |
| Weaknesses of database | | Cross-sectional, cluster sampling, does not include some very part-time GPs, does not cover non-recognised primary care providers, national sampling does not provide sufficient data from remote regions and some specific population groups such as ATSI | Need to improve the classification of terming - building blocks are there and locking doctors into reason for prescription will improve the database. Referrals and the nature of the collections doesn't capture between practice visits by patiens. | Number of patients in the database. Dataset comprises data from June 1994 to December 2000. Selected not random sample of GPs who are fully computerised and may not be representative of Australian GPs. Unable to determine whether drugs prescribed are dispensed. Not all medications are subsidised. | No systematic patient ID; subsidised (claims) data only. | Unable to determine actual dispensing against patient name - data confidentiality; variation in capturing OTC medications; access to data; linkage required back to prescriber name to determine whether locally generated scripts. | N / A |
| Ownership of database | | University of Sydney and the Australian Institute of Health and Welfare (AIHW) | Health Communication Network | Data is property of contributing GPs, who decide who has access. GP records access is not permitted beyond the MEDIC-GP project group | Commonwealth of Australia | Individual pharmacies, no real aggregation | N / A |
| Access of information and charges | | Non-commercial external organisations: cost depends on request but usual range is $13,20 to $1,550. Free analyses are conducted when it is for the public good. | Standard report on pharmaceutical $3,500. Access to academic groups and NGOs has been free | Interested parties approach the Department of General Practice, University pf Adelaide. Approval of projects by Ethics Committee. MEDIC-GP is not public domain. Usage of database determined by resources employed I.e. cost-recovery. | Database is in the public domain through ASM; request handled by Secretariat; more involved ones to committee. | Dependent on individual pharmacy and whether access is granted | N / A |

# Appendix C

## HCN GPRN DATASETS AS DERIVED FROM MEDICAL DIRECTOR TABLES

1. In the following notes, highlighted variables indicate that there is an underlying format, which has not been included as yet.

2. Date of birth has been removed and age group has been created as per five years intervals (where there is a date included in the dataset eg "date of script" then age is as of this date. Otherwise age has been merged via the patient file and is the age at the date of the file creation).

3. Note that "gp_id" is not a unique identifier.

4. The unique patient identifier is "patient_combined".

5. All datasets are from Medical Director except the datasets named "Practices" and "GPs" these are from the practice and GP questionnaire completed on enrolment to the GPRN.

## GPRN -Cleaning and standardisation of variables

*PRACTICES (from practice questionnaire)*

| Description | Data source | Variable | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | Practice questionnaire | PRACTICE | C | 6 | | practices | practice_id | C | 6 | | |
| Practice state | Practice questionnaire | | C | 3 | | | practice_state | C | 3 | | |
| Rural, remote classification | New Variable | | N | 3 | 0 = Unknown<br>1 = Capital<br>2 = Oth Met<br>3 = Lge Rur<br>4 = Sm Rur<br>5 = Oth Rur<br>6 = Rem Cen<br>7 = Oth Rem, Offshore | | rrma | N | 3 | | |
| Medical Director version | Practice questionnaire | MD_VERS | | | | | medvers | C | 10 | Must be valid MD version | |
| Other health profs in practice | Practice questionnaire | OTHER_HEALTH_PROFESSIONAL | C | 3 | Yes=1<br>No=2 | | Other_health | C | 3 | | |
| List of other health profs | Practice questionnaire | LIST_OTHER_HEALTH_PROFESSIONAL | C | 100 | Text | | list_other_health_professional | C | 100 | | |
| How long has practice been using computers for clinical functions (number) | Practice questionnaire | HOW_CLIN_FUN | N | 3 | | | how_clin_fun | N | 3 | >0 | |
| How long has practice been using computers for clinical functions(month/year) | Practice questionnaire | HOW_CLIN_FUN_DMY | C | 4 | Months=1<br>Years=2 | | how_clin_fun_dmy | C | 4 | | |

| Description | Data source | Variable | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| How many doctors using computers for clinical functions | Practice questionnaire | N_DOCTORS_USING_COMPUTERS | N | 3 | | | num_docs | N | 3 | >0 | |
| How long has practice been using Medical director for clinical functions (month/year) | Practice questionnaire | HOW_CLIN_FUN__MD_DMY | C | 4 | Months=1<br>Years=2 | | how_clin_fun_md_dmy | C | 4 | | |
| Are computers networked | Practice questionnaire | COMPUTERS_NETWORKED | C | 1 | Yes=1<br>No=0 | | computers_networked | C | 1 | | |
| Is there a file server | Practice questionnaire | DEDICATED_FILE_SERVER | C | | Yes=1<br>No=0 | | ded_file | C | 1 | | |
| Internet access | Practice questionnaire | INTERNET_ACCESS | C | 3 | Yes=1<br>No=0 | | net_acc | C | 3 | | |
| Type of Internet access | Practice questionnaire | TYPE_OF_INTERNET_ACCESS | C | 1 | Dial up=1<br>ISDN= 2<br>Cable= 3<br>Don't know = 4<br>Missing=9 | | type_of_internet_access | C | 1 | Missing=9 | |
| E-mail access | Practice questionnaire | EMAIL_ACCESS | C | 1 | Yes=1<br>No=0 | | email_acc | C | 1 | | |
| E-mail software | Practice questionnaire | EMAIL_SOFTWARE | C | 50 | Text | | email_software | C | 50 | | |

*GPs (from GP Questionnaire)*

| Description | Table | Variables | Type | Length | Data entry | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | GP questionnaire | PRACTICE_ID | C | 6 | | **gps** | practice_id | C | 6 | | |
| GP ID(note this is not a unique identifier) | GP questionnaire | GP_ID | C | 3 | | | gp_id | C | 3 | | |
| GP sex | GP questionnaire | SEX_GP | C | 1 | Male=1<br>Female=2 | | sex_gp | C | 1 | Missing=9 | |
| GP age group | GP questionnaire | DOCTOR_AGE_GROUP | N | 3 | <35 = 1<br>35-44=2<br>45-54=3<br>55+ = 4<br>Missing=9 | | doctor_age_group | N | 3 | valid date and >01/01/1900 and before today's date | |
| Year of graduation | GP questionnaire | YEAR_GRADUATE | C | 4 | | | year_graduate | C | 4 | valid date and >01/01/1900 and before today's date | |
| Vocationally registered | GP questionnaire | VOCATIONALLY_ REGISTERED | N | 1 | Yes=1<br>No=0 | | vocationally_ registered | N | 1 | | |
| ?More than 50% of consults other than English | GP questionnaire | NESB_ CONSULT | N | 1 | Yes=1<br>No=0 | | nesb_consult | N | 1 | | |
| Number of sessions this practice | GP questionnaire | N_SESSIONS_THIS_ PRACTICE | N | 8 | | | n_sessions_this_ practice | N | 8 | >0 or missing | |
| Number of sessions other practice | GP questionnaire | N_SESSIONS_OTHER_ PRACTICE | N | 8 | | | n_sessions_other_ practice | N | 8 | >0 or missing | |
| Clinical functions | GP questionnaire | CLINICAL_FUNCTIONS | N | 8 | | | clinical_functions | N | 8 | >0 or missing | |
| Clinical functions/day/month /year | GP questionnaire | CLINICAL_FUNCTIONS_DMY | C | 4 | Months=1<br>Years=2 | | clinical_functions_dmy | C | 4 | | |
| Consultation notes | GP questionnaire | CONSULTATION_NOTES | N | 1 | Yes=1<br>No=0 | | consultation_notes | N | 1 | | |
| Decision support | GP questionnaire | DECISION_SUPPORT | N | 1 | Yes=1<br>No=0 | | decision_support | N | 1 | | |
| Educational material | GP questionnaire | EDUCATIONAL_MATERIAL | N | 1 | Yes=1<br>No=0 | | educational_material | N | 1 | | |
| E-mail | GP questionnaire | EMAIL | N | 1 | Yes=1<br>No=0 | | email | N | 1 | | |
| How long using Medical director | GP questionnaire | HOW LONG USING MEDICAL DIRECTOR | N | 8 | | | how_long_using_medical_director | N | 8 | Number entered | |
| How long using Medical director days | GP questionnaire | HOW LONG USING MEDICAL DIRECTOR_D | C | 4 | Months=1<br>Years=2 | | how_long_using_medical_director_d | C | 4 | | |
| Internet access | GP questionnaire | INTERNET_ACCESS | N | 1 | Yes=1<br>No=0 | | internet_access | N | 1 | | |
| Internet access | GP questionnaire | INTERNET ACCES | N | 1 | Yes=1 | | internet access fro | N | 1 | | |

| Description | Table | Variables | Type | Length | Data entry | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| from consult PC | | S_FROM_CONSULT_PC | | | No=0 | | m_consult_pc | | | | |
| Medicare claims | GP questionnaire | MEDICRE_CLAIMS | N | 1 | Yes=1<br>No=0 | | medicare_claims | N | 1 | | |
| NESB language | GP questionnaire | NESB_LANGUAGE | C | 50 | | | nesb_language | C | 50 | | |
| Patient accounts | GP questionnaire | PATIENT_ACCOUNTS | N | 1 | Yes=1<br>No=0 | | patient_accounts | N | 1 | | |
| Patient summaries | GP questionnaire | PATIENT_SUMMARIES | N | 1 | Yes=1<br>No=0 | | patient_summaries | N | 1 | | |
| Payroll | GP questionnire | PAYROLL | N | 1 | Yes=1<br>No=0 | | payroll | N | 1 | | |
| Practice finances | GP questionnaire | PRACTICE_FINANCES | N | 8 | Yes=1<br>No=0 | | practice_finances | N | 8 | | |
| Prescriptions | GP questionnaire | PRESCRIPTIONS | N | 1 | Yes=1<br>No=0 | | prescriptions | N | 1 | | |
| Recall | GP questionnaire | RECALL | N | 1 | Yes=1<br>No=0 | | recall | N | 1 | | |
| Receiving/storing info | GP questionnaire | RECEIVING_STORING_INFO | N | 1 | Yes=1<br>No=0 | | receiving_storing_info | N | 1 | | |
| Referral letters | GP questionnaire | REFERRAL_LETTERS | N | 1 | Yes=1<br>No=0 | | referral_letters | N | 1 | | |
| Registration | GP questionnaire | REGISTRATION | N | 1 | Yes=1<br>No=0 | | registration | N | 1 | | |
| Scheduling appointments | GP questionnaire | SCHEDULING_APPOINTMENTS | N | 1 | Yes=1<br>No=0 | | scheduling_appointments | N | 1 | | |
| Stocks/ stores | GP questionnaire | STOCKS_STORES | N | 1 | Yes=1<br>No=0 | | stocks_stores | N | 1 | | |
| Word processing | GP questionnaire | WORDPROCESSING | N | 1 | Yes=1<br>No=0 | | wordprocessing | N | 1 | | |

*PATIENTS (from patients.dbf)*

| Description | Source | Variables | Type | Length | Data entry | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | PATIENTS.DBF | | C | 6 | | patients | practice_id | C | 6 | | |
| Patient id | | RE-CODED VARIABLE (NOT IN DBF file) | C | 16 | | | patient_combinid | C | 16 | | |
| Active/inactive patient | | CURRENT | C | 1 | Y=Yes<br>N=No | | current | C | 1 | 1=Yes<br>2=No | outside range = 9 |
| Patient sex | | SEX | C | 1 | M=Male<br>F=Female | | patient_sex | C | 1 | 1=Male<br>2=Female | outside range = 9 |
| Patient age | | NEW VARI-<br>AB LE (NOT IN DBF file) | N | 8 | | | patient_age | N | 8 | valid date and >01/01/1880 and before today's date | outside range set to missing.Note this is age as at date of file creation |
| Patient age group by five year intervals | | NEW VARI-<br>AB LE (NOT IN DBF file) | N | 8 | **1** = '<1'<br>**2** = '1-4'<br>**3** = '5-9'<br>**4** ='10-14'<br>**5** = '15-19'<br>**6** = '20-24'<br>**7** = '25-29'<br>**8** = '30-34'<br>**9** = '35-39'<br>**10** = '40-44'<br>**11** = '45-49'<br>**12** = '50-54'<br>**13** = '55-59'<br>**14** = '60-64'<br>**15** = '65-69'<br>**16** = '70-74'<br>**17** = '75-79'<br>**18** = '80+'<br>**19** ='Missing' | | patient_age_group | N | 8 | | Note this age is as of date of file creation |
| State | | Derived from POSTCODE | C | 3 | | | state | C | 3 | Format: "state" | outside range =XXX |
| Rural,remote classification | New Variable | | N | 3 | **0** = Unknown<br>**1** = Capital<br>**2** = Oth Met<br>**3** = Lge Rur<br>**4** = Sm Rur<br>**5** = Oth Rur<br>**6** = Rem Cen<br>**7** = Oth Rem, Offshore | | rrma | N | 3 | Format :"pc7rara" | |

| Description | Source | Variables | Type | Length | Data entry | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Pension code | | PENSCODE | C | 1 | blank = none<br>P=Pension/HCC<br>R=Full DVA<br>L=Limited DVA | | patient_pension_code | C | 1 | 1=none<br>2=Pension/HCC<br>3=Full DVA<br>4=Limited DVA | outside range = 9 |
| Aboriginal/Torres Strait Islander | | ATSI | C | 1 | | | patient_atsi | C | 1 | 1=neither<br>2=Aboriginal<br>3=TSI<br>4=ATSI | outside range = 9 |
| Deceased flag | | DECEASED | C | 1 | | | deceased_flag | C | 1 | 1=Deceased<br>2=Not deceased | outside range = 9 |
| Date of death | | DECDATE | D | DATE9 | > dob, <=current date | | deceased_date | D | DATE9 | valid date,>01/01/1900,>dob,<current date | outside range set to missing |
| Date record updated | | UPDATE | C | 14 | not set | | date_record_updated | C | 14 | <=current date | Format is year,month, date, time |

*RISK FACTORS (from allergy.dbf)*

| Description | Source | Variables | Type | Length | Data entry | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | ALLERGY.DBF | | C | 6 | | risk_factors | practice_id | C | 6 | | Note this file has been re-named as "risk_factors" as this more accurately reflects the data it contains |
| Patient id | | RE-CODED VARI-AB LE (NOT IN DBF file) | C | 16 | | | patient_combinid | C | 16 | | |
| Patient sex | | SEX | C | 1 | M=Male F=Female | | patient_sex | C | 1 | 1=Male 2=Female 9=Missing | |
| Patient age group by five year interval | | NEW VARI-AB LE (NOT IN DBF file) | | | As above | | patient_age_group | N | 8 | valid date and >01/01/1880 and before today's date | |
| Blood group | | BLGROUP | C | 3 | | | blood_group | C | 3 | | |
| Occupation | | OCCUPATION | C | 30 | | | occupation | C | 30 | | |
| Smoking flag | | SMOKING | C | 1 | | | smoking_status | C | 1 | | |
| No. of cigarettes | | SMOKES | C | 2 | | | number_of_cigarettes | C | 2 | | |
| Alcohol intake | | ALCOHOL | C | 2 | | | standard_drinks_day | C | 2 | | |
| Marital status | | MARRIED | C | 1 | | | marital_status | C | 1 | | |
| Breast feeding flag | | BREAST FEED | C | 1 | | | breast_feeding | C | 1 | | |
| Elite sports flag | | SPORTSMAN | C | 1 | | | elite_sportsperson | C | 1 | | |
| Date of LMP if pregnant | | PREG_DATE | D | DATE9 | | | date_of_lmp | D | DATE9 | | |
| Date record updated | | UPDATE | C | 14 | | | update_flag_risk | C | 14 | | |

*REACTIONS (from reaction.dbf)*

| Variable | Description | Table | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | REACTION.DBF | | | | | reactions | practice_id | C | 6 | | |
| Patient id | | RE-CODED VARIAB LE (NOT IN DBF file) | C | 16 | | | patient_combinid | C | 16 | | |
| Doctor id | | DR_NO | C | 3 | | | gp_id | C | 3 | | |
| Deletion flag | | DELCODE | C | 1 | | | deletion_flag | C | 1 | 1=Yes<br>2=No | Missing = 9 |
| Name of substance | | ITEM | C | 50 | | | substance | C | 50 | | |
| Type of reaction | | TYPE | C | 50 | | | reaction_type | C | 50 | | |
| Severity code | | SEVERITY | C | 1 | | | reaction_severity | C | 1 | | |
| Date record updated | | UPDATE | C | 14 | | | update_flag_reactions | C | 14 | | Format is year,month, date, time |
| Patient sex | | SEX | C | 1 | M=Male<br>F=Female | | patient_sex | C | 1 | 1=Male<br>2=Female | outside range = 9 |
| Patient age group by five year interval | | NEW VARI-AB LE (NOT IN DBF file) | | | As above | | patient_age_group | N | 8 | valid date and >01/01/1880 and before today's date | Note this is age as at date of file creation |

*SCRIPTS _CURRENT (from rx.dbf)*

| Variable | Description | Table | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | RX.DBF | | C | 6 | | SCRIPTS_ CURRENT | practice_id | C | 6 | | |
| Patient id | | RE-CODED VARIAB LE (NOT IN DBF file) | C | 16 | | | patient_combinid | C | 16 | | |
| Doctor id | | DR_NO | C | 3 | | | gp_id | C | 1 | | |
| Patient sex | | SEX | C | 1 | M=Male F=Female | | patient_sex | C | 1 | 1=Male 2=Female | outside range = 9 |
| Patient age group by five year interval | | NEW VARI- AB LE (NOT IN DBF file) | | | As above | | patient_age_group | N | 8 | Valid date and >01/01/188 0 and before today's date | Note this is age as at date of file creation |
| Drug dose | | DOSE | C | 10 | | | drug_dose | C | 10 | | |
| Drug frequency | | FREQUENCY | C | 10 | | | drug_frequency | C | 10 | | |
| Quantity | | MAXQ | C | 11 | | | drug_quantity | C | 11 | | |
| Number of repeats | | REPEATS | N | 2 | | | drug_repeats | N | 8 | | |
| Reason for prescribing | | INDICATION | C | 60 | | | reason_prescription | C | 60 | | |
| Date added to current RX | | FIRSTDATE | D | 8 | | | date_first_prescribed | D | 8 | MMDDYY8 | |
| Date of last prescription printed | | LASTDATE | D | 8 | | | date_last_prescribed | D | 8 | MMDDYY8 | |
| Date removed from current RX | | DELDATE | D | 8 | | | date_prescription_deleted | D | 8 | MMDDYY8 | |
| Reason for deletion | | DELREASON | M | 10 | | | deletion_reason | C | 17 | | |
| Previous authority for use of this drug | | PREV | C | 1 | | | previous_authority | C | 1 | | |
| Authority indication | | REASON | M | 10 | | | authority_indication | C | 10 | | |
| Complex dosage instructions | | SPECIAL | M | 10 | | | complex_dosage | C | 10 | | |
| Date record updated | | UPDATE | M | 14 | | | update_flag_scripts_curren t | C | 14 | | |

*RECALL (from recall .dbf)*

| Variable | Description | Table | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | RECALL.DBF | | | | | recall | practice_id | C | 6 | | |
| Patient id | | RE-CODED VARIAB LE (NOT IN DBF file) | C | 16 | | | patient_combinid | C | 16 | | |
| Name of protocol | | PROTOCOL | C | 25 | | | protocol | C | 25 | | |
| Once only flag | | REASON | C | 25 | N=Recall once only blank= otherwise | | once_only_recall | C | 25 | | |
| Date next recall is due | | NEXT_DUE | D | DATE9 | | | date_next_recall | D | DATE9 | | |
| Interval for future recalls | | INTERVAL | N | 4 | | | recall_interval | N | 4 | | |
| Actual recall date | | REC_DATE | D | DATE9 | | | date_actual_recall | D | DATE9 | | |
| Date record updated | | UPDATE | C | 14 | | | update_flag_recall | C | 14 | | Format is year,month, date, time |
| Patient sex | | SEX | C | 1 | M=Male F=Female | | patient_sex | C | 1 | 1=Male 2=Female | outside range = 9 |
| Patient age group by fiveyear interval | | NEW VARI-AB LE (NOT IN DBF file) | | | As above | | patient_age_group | N | 8 | valid date and >01/01/1880 and before today's date | Note this is age as at date of file creation |

*ACTIONS (from actions .dbf)*

| Variable | Description | Table | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | ACTIONS.DBF | | | | | actions | practice_id | C | 6 | | |
| Doctor id | | DR_NO | C | 3 | | | gp_id | C | 1 | | |
| Patient id | | RE-CODED VARIAB LE (NOT IN DBF file) | C | 16 | | | patient_combinid | C | 16 | | |
| Patient sex | | SEX | C | 1 | M=Male F=Female | | patient_sex | C | 1 | 1=Male 2=Female | outside range = 9 |
| Patient age group by five year interval | | NEW VARI-AB LE (NOT IN DBF file) | | | As above | | patient_age_group | N | 8 | valid date and >01/01/1880 and before today's date | Note this is age as at date of file creation |
| Update | | UPDATE | C | 14 | | | update | C | 14 | | |
| Deletion flag | | DELCODE | C | 1 | | | delcode | C | 1 | | |
| Action | | ACTION | C | 50 | | | type_of_action | C | 50 | | |
| Date action due | | DUEDATE | D | DATE9 | | | date_action_due | DATE9 | | | |
| Date recall due | | RECDATE | D | DATE9 | | | date_recall_due | DATE9 | | | |
| Date reminder | | REMDATE | D | DATE9 | | | date_reminder | DATE9 | | | |
| Remind by | | REM_BY | C | 3 | | | rem_by | C | 3 | | |
| Type | | TYPE | C | 1 | | | type | C | 1 | | |

*ENCOUNTERS (from progress.dbf)*

| Description | Table | Variable | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | PROGRESS.DBF | | C | 6 | | encounters | practice_id | C | 6 | | |
| Doctor id | | DR_NO | C | 3 | | | gp_id | C | 3 | | |
| Patient id | | RE-CODED VARIAB LE (NOT IN DBF file) | C | 16 | | | patient_combinid | C | 16 | | |
| Medicare item | | MEDICARE | C | 30 | valid medicare numbers separated by commas | | medicare_item | C | 30 | valid medicare number | check each number, if invalid set to missing |
| First medicare item | | NEW VARIABLE (created from medicare item) | C | 6 | | | medicare_item_1 | C | 6 | | |
| Second medicare item | | NEW VARIABLE (created from medicare item) | C | 6 | | | medicare_item_2 | C | 6 | | |
| Third medicare item | | NEW VARIABLE (created from medicare item) | C | 6 | | | medicare_item_3 | C | 6 | | |
| Fourth medicare item | | NEW VARIABLE (created from medicare item) | C | 6 | | | medicare_item_4 | C | 6 | | |
| Fifth medicare item | | NEW VARIABLE (created from medicare item) | C | 6 | | | medicare_item_5 | C | 6 | | |
| Number of medicare items | | NEW VARIABLE (count of above 5) | N | 2 | | | number_of_medicare_items | N | 2 | | |
| Reason for visit | | REASON | C | 60 | | | reason_visit | C | 60 | | |
| Date of visit | | VISITDATE | D | 8 | <=current date | | date_of_visit | D | 8 MMDDYY | >01/01/1900, <= sysdate | Check in QA2.1 |
| Date recorded | | RECDATE | D | 8 | current date taken from system | | date_of_recording | D | 8 MMDDYY | | |
| Review date | | RVDATE | D | 8 | >current date | | date_of_review | D | 8 MMDDYY | >= date of current visit | |
| Start time of visit | | STARTTIME | C | 10 | from system | | start_time | Time | hh:mm:ss | | |
| End time of visit | | ENDTIME | C | 10 | from system | | end_time | Time | hh:mm:ss | | |
| Duration of visit | | DURATION | C | 10 | from system | | duration | Time | hh:mm:ss | >0 seconds | else set to missing |
| History structure | | HISTORY | C | 60 | blank=not completed | | history | C | 60 | | Note this field requires a coding frame (not included) |

| Description | Table | Variable | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Examination structure | | EXAM | C | 60 | blank=not completed | | exam | C | 60 | | Note this field requires a coding frame (not included |
| Deletion flag | | DELCODE | C | 1 | | | deletion_flag | C | 1 | | |

*DIAGNOSIS (from diagnosis.dbf)*

| Description | Table | Variable | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | | | C | 6 | | diagnosis | practice_id | C | 6 | | |
| Patient id | | RE-CODED VARI-AB LE (NOT IN DBF file) | C | 16 | | | patient_combinid | C | 16 | | |
| Patient age group by five year intervals | | NEW VARI-AB LE (NOT IN DBF file) | | | As above | | patient_age_group | N | 8 | valid date and >01/01/1880 and before today's date | |
| Patient sex | | NEW VARI-AB LE (NOT IN DBF file) | C | 1 | M=Male F=Female | | patient_sex | C | 1 | 1=Male 2=Female | outside range = 9 |
| Update flag | | UPDATE | C | 14 | | | update_flag_diagnosis | C | 14 | | |
| Reason for visit | | REASON | C | 60 | | | reason_visit | C | 60 | | |
| DOCLE code of reason for visit | | REAS_CODE | C | 60 | valid DOCLE code | | reason_visit_docle | | | | valid DOCLE code else most general DOCLE code |
| Diagnosis date | | DIAGDATE | D | DATE9 | | | date_of_visit | D | DATE9 | | |
| Deletion flag | | DELCODE | C | 1 | | | deletion_flag_diagnosis | C | 1 | | |

*HISTORY (from history.dbf)*

| Description | Table | Variable | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | HISTORY.DBF | | | | | history | practice_id | C | 6 | | |
| Patient id | | RE-CODED VARIAB LE (NOT IN DBF file) | C | 16 | | | patient_combinid | C | 16 | | |
| Patient age group by five year intervals | | NEW VARI-AB LE (NOT IN DBF file) | | | As above | | patient_ age_group | N | 8 | valid date and >01/01/1880 and before today's date | Note this is age at date of file creation |
| Patient sex | | NEW VARI-AB LE (NOT IN DBF file) | C | 1 | M=Male F=Female | | patient_sex | C | 1 | 1=Male 2=Female | outside range = 9 |
| Year component of date | | YEAR | C | 4 | Can be entered, or system creates it | | year | C | 4 | > date of birth | |
| Remainder of date | | MONTH | C | 10 | Can be any text entered by user, or system creates it. If system creates it, it is the full date dd/mm/yyyy that the history was recorded. | | month | C | 10 | | |
| Date condition began | | NEW VARIABLE-NOT IN DBF FILE | N | DATE8 | | | date_condition_began | D | DATE8 | new variable created from month and year | If only month and year give, then date is the first of that month. |
| Description of condition | | CONDITION | C | 60 | | | condition | C | 60 | | |
| DOCLE code | | HISTCODE | C | 60 | | | condition_docle_code | C | 60 | | Valid DOCLE code |
| Deletion flag | | DELCODE | C | 1 | | Any history items deleted, are deleted. deletion flag is not used. | deletion_flag | C | 1 | | |
| Summary flag | | SUMM | C | 1 | | | summary_flag | C | 1 | | |
| Active/inactive flag | | ACTIVE | C | 1 | | | active_flag | C | 1 | | 1 = active 2 = not active 9 = missing |
| Left or right | | SIDE | C | 1 | | | side | C | 1 | | |
| Date history updated | | UPDATE | | | | | update_flag_history | C | 14 | | Format is year,month, date, time |

*SCRIPTS_ALL (from scr_file.dbf)*

| Description | Table | Variable | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | SCR_FILE.DBF | | C | 6 | | scripts_all | practice_id | C | 6 | | |
| Doctor id | | DR_NO | C | 3 | | | gp_id | C | 3 | | |
| Patient id | | RE-CODED VARIAB LE (NOT IN DBF file) | C | 16 | | | patient_combinid | C | 16 | | |
| Script number | | SCR_NO | C | 7 | | | script_number | C | 7 | | |
| Deletion flag | | DELCODE | C | 1 | Y=deleted blank=not deleted | | deletion_flag | C | 1 | 1=deleted 2=not deleted | all others set to missing = 9 |
| Reason for deletion | | DELREASON | C | 60 | free text field | | deletion_reason | C | 60 | | depends on how SAS handles memo field after coming through SQL |
| Reason for prescribing | | PRESREASON | C | 60 | free text field | | reason_prescription | C | 60 | | ditto |
| Script date | | SCRDATE | D | 8 | generated by system | | date_script | D | 8 MMDDYY | | > current date set to missing |
| Reason for authority | | REASON | C | 60 | free text field or selected from DOCLE list | | reason_authority | C | 60 | | |
| Approval number | | APPR_NO | C | 12 | | | approval_number | | | | |
| Name of drug | | P1 | C | 34 | from AZDEX list | | drug_name | C | 34 | | |
| Strength of drug | | S1 | C | 12 | from AZDEX list | | drug_strength | C | 12 | | |
| Strength recode | | | C | 80 | | | Strength | C | 80 | | |
| Dose | | D1 | C | 10 | free text field | | drug_dose | C | 10 | | |
| Frequency | | F1 | C | 10 | from pick list | | drug_frequency | C | 10 | | |
| Frequency translated (eg qid=four times a day) | | NEW VARIABLE | C | 18 | | | drug_freq_trans | C | 18 | | |
| Dose (numeric) | | NEW VARIABLE | N | 8 | | | drug_dose_num | N | 8 | | |
| Other instructions | | SP1 | C | 60 | free text field | | drug_special_dose | C | 60 | | |
| Quantity ordered | | DR_1_Q | C | 11 | | | drug_quantity | C | 11 | | |
| Repeats ordered | | DR_1_R | C | 2 | set to 0 if non-numeric | | drug_repeats | C | 2 | | |
| Drug generic name | | NEW VARIABLE (merged from AZ-DEX file) | C | 80 | | | drug_generic_name | C | 80 | | |
| Drug class code abbreviated | | NEW VARIABLE (merged from AZ-DEX file) | C | 10 | | | drug_class_code | C | 10 | | |
| Drug class code full name | | NEW VARIABLE (merged from AZ- | C | 50 | | | drug_class_name | C | 50 | | |

| Description | Table | Variable | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | DEX file) | | | | | | | | | |
| Drug form | | NEW VARIABLE (merged from AZ-DEX file) | C | 11 | | | form | C | 11 | | |
| Drug trade name | | NEW VARIABLE (merged from AZ-DEX file) | C | 69 | | | trade name | C | 69 | | |
| Drug company name | | NEW VARIABLE (merged from AZ-DEX file) | C | 45 | | | company_name | C | 45 | | |
| Patient age group by five year interval | | NEW VARI-AB LE (NOT IN DBF file) | | | As above | | patient_age_group | N | 8 | valid date and >01/01/1880 and before today's date | Note this is age at date of script |

*MDUSAGE (from mdw.dbf)*

| Description | Table | Variable | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | MDW.DBF | PRACTICE | C | 6 | | | practice_id | C | 6 | | |
| Patient id | | RE-CODED VARIAB LE (NOT IN DBF file) | C | 16 | | | patient_combinid | C | 16 | | Will be missing in item date< 1999 |
| Patient age group by five year interval | | NEW VARI-AB LE (NOT IN DBF file) | | | As above | | patient_age_group | N | 8 | valid date and >01/01/1880 and before today's date | |
| Patient sex | | NEW VARIAB LE (NOT IN DBF file) | C | 1 | M=Male F=Female | | patient_sex | C | 1 | 1=Male 2=Female | outside range = 9 |
| Start time | | STARTTIME | C | 8 | | | start_time | C | 8 | | |
| End time | | ENDTIME | C | 8 | | | end_time | C | 8 | | |
| Screen opened by GP | | ITEM | C | 2 | See attached coding key | | item | C | 2 | | |
| Inactive field | | AD_CODE | C | 9 | | | ad_code | C | 9 | | |
| Open/Close/Cancel/Print | | BUTTON | C | 9 | | | button | C | 9 | | |
| Item date | | | N | DATE9 | | | item_date | N | DATE9 | | |

*REQUESTS (from requests.dbf)*

| Description | Table | Variable | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | REQUESTS.DBF | | | | | requests | practce_id | C | 6 | | |
| Doctor id | | DR_NO | C | 3 | | | gp_id | C | 3 | | |
| Patient id | | RE-CODED VARIAB LE (NOT IN DBF file) | C | 16 | | | patient_combinid | C | 16 | | |
| Deletion flag | | DELCODE | C | 1 | | | deletion_flag | C | 1 | | |
| Request date | | REQ_DATE | D | MMDD YY8 | | | date_request | D | MMDDYY 8 | | MMDDY Y |
| Tests ordered-note includes only first five tests (as below) | | TESTS | C | 60 | | | tests_ordered | C | 60 | | |
| First test ordered | | NEW VARIABLE | C | 20 | | | Tests_item_1 | C | 20 | | |
| Second test ordered | | NEW VARIABLE | C | 20 | | | Tests_item_2 | C | 20 | | |
| Third test ordered | | NEW VARIABLE | C | 20 | | | Tests_item_3 | C | 20 | | |
| Fourth tests ordered | | NEW VARIABLE | C | 20 | | | Tests_item_4 | C | 20 | | |
| Fifth test ordered | | NEW VARIABLE | C | 20 | | | Tests_item_5 | C | 20 | | |
| Number of tests ordered | | NEW VARIABLE | N | 2 | | | Number_of_tests _items | N | 2 | | |
| Patient sex | | NEW VARIAB LE (NOT IN DBF file) | C | 1 | M=Male F=Female | | patient_sex | C | 1 | 1=Male 2=Female | outside range = 9 |
| Patient age group by five year interval | | NEW VARI-AB LE (NOT IN DBF file) | | | As above | | patient_age_gro up | N | 8 | valid date and >01/01/188 0 and before today's date | Note this is age as at date of request |

*PATHATOM (from pathatom.dbf)*

| Description | Table | Variable | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | PATHATOM.DBF | | | | | pathatom | practice_id | C | 6 | | |
| Patient id | | RE-CODED VARIAB LE (NOT IN DBF file) | C | 16 | | | patient_combinid | C | 16 | | |
| Date of result | | RESULTDATE | D | DATE9 | | | date_result | D | DATE9 | | |
| LOINC code-coding frame used for coding pathology | | LOINC | C | 32 | | | loinc_code | C | 32 | | |
| Name of test | | TESTNAME | C | 32 | | | test_name | C | 32 | | |
| Result value | | RESULTS | C | 20 | | | results_value | C | 20 | | Missing= 99999 |
| Units used | | UNITS | C | 20 | | | results_units | C | 20 | | Missing= 99999 |
| Normal range | | RANGE | C | 20 | | | normal_range | C | 20 | | Missing= 99999 |
| Abnormal flag | | ABNFLAG | C | 3 | | | abnormal_flag | C | 3 | | |
| Path code-INACTIVE FIELD | | PATHCODE | C | 12 | | | pathology_code | C | 12 | | |
| Date record updated | | UPDATE | C | 14 | | | update_flag_pathatom | C | 14 | | |
| Patient sex | | NEW VARIAB LE (NOT IN DBF file) | C | 1 | M=Male F=Female | | patient_sex | C | 1 | 1=Male 2=Female | outside range = 9 |
| Patient age group by five year interval | | NEW VARI-AB LE (NOT IN DBF file) | | | As above | | patient_age_group | N | 8 | valid date and >01/01/1880 and before today's date | Note this is age as at date of test |

*PATHOLOGY (from pathol.dbf)*

| Description | Table | Variable | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | |
| Practice id | PATHOL.DBF | | | | | pathology | practice_id | C | 6 | | |
| Doctor id | | DR_NO | C | 3 | | | gp_id | C | 3 | | |
| Patient id | | RE-CODED VARIAB LE (NOT IN DBF file) | C | 9 | | | patient_combinid | C | 16 | | |
| Patient sex | | NEW VARI-AB LE (NOT IN DBF file) | C | 1 | M=Male F=Female | | patient_sex | C | 1 | 1=Male 2=Female | outside range = 9 |
| Patient age group by five year interval | | NEW VARI-AB LE (NOT IN DBF file) | | | As above | | patient_age_group | N | 8 | valid date and >01/01/1880 and before today's date | Note this is age at date of test |
| Deletion flag | | DELCODE | C | 1 | | | deletion_flag | C | 1 | | |
| Pathology code | | PATHCODE | C | 12 | | | pathology_code | C | 12 | | |
| Date requested | | REQDATE | D | DATE9 | | | date_request | D | DATE9 | | |
| Date collected | | COLLDATE | D | DATE9 | | | date_collection | D | DATE9 | | |
| Name of test | | TESTNAME | C | 20 | | | test_name | C | 20 | | |
| Date reported | | REPORTDATE | D | DATE9 | | | date_report | D | DATE9 | | |
| Specimen type | | SPECIMEN | C | 32 | | | specimen | C | 32 | | |
| Date checked | | CHECKDATE | D | DATE9 | | | date_checked | D | DATE9 | | |
| Notation made | | NOTATION | C | 1 | | | notation_made | C | 1 | | |
| Date patient notified | | NOTIFIED | D | DATE9 | | | date_notified | D | DATE9 | | |
| Laboratory's abnormal flag | | LABNORM | C | 1 | | | lab_abnormal_flag | C | 1 | | |
| Confidential flag | | CONFID | C | 1 | | | confidential_flag | C | 1 | | |

*PATIENT EDUCATION (from ped.dbf)*

| Description | Table | Variable | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | PED.DBF | | | | | patient_educ | practice_id | C | 6 | | |
| Doctor id | | DR_NO | C | 3 | | | gp_id | C | 3 | | |
| Patient id | | RE-CODED VARIAB LE (NOT IN DBF file) | C | 16 | | | patient_combinid | C | 16 | | |
| Date leaflet given | | PEDDATE | D | DATE9 | | | date_leaflet_given | D | DATE9 | | |
| Leaflet order from PAT_ED.DBF | | ORDER | C | 2 | | | leaflet_order | C | 2 | | |
| Leaflet category from PAT_ED..DBF | | CATEGORY | C | 20 | | | leaflet_category | C | 20 | | |
| Patient age group by five year interval | | NEW VARI-AB LE (NOT IN DBF file) | | | As above | | patient_age_group | N | 8 | valid date and >01/01/1880 and before today's date | Note this is age as at date leaflet given |
| Patient sex | | NEW VARIAB LE (NOT IN DBF file) | C | 1 | M=Male F=Female | | patient_sex | C | 1 | 1=Male 2=Female | outside range = 9 |

*MEASURES (from measures.dbf)*

| Description | Table | Variable | Type | Length | Data entry and coding rules | File | Variable | Type | Length | Coding rules | Cleaning rules |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Practice id | MEASURES.DBF | PRACTICE | C | 6 | | measures | practice_id | C | 6 | | |
| Patient id | | RE-CODED VARIAB LE (NOT IN DBF file) | C | 16 | | | patient_combinid | C | 16 | | |
| Patient age group by five year intervals | | NEW VARI-AB LE (NOT IN DBF file) | | | As above | | patient_ age_group | N | 8 | valid date and >01/01/1880 and before today's date | |
| Patient sex | | NEW VARI-AB LE (NOT IN DBF file) | C | 1 | M=Male F=Female | | patient_sex | C | 1 | 1=Male 2=Female | outside range = 9 |
| Date of measurement | | MEAS_DATE | D | DATE9 | | | measurement_da te | D | DATE9 | | |
| Measurement time | | MEAS_TIME | D | DATE9 | | | measurement_ti me | D | DATE9 | | |
| Value of measurement | | VALUE | C | 10 | | | measurement_va lue | C | 10 | | |
| Measurement description | | TYPE | C | 12 | | | measurement_ty pe | C | 12 | | |