

2 Methods

This study is a secondary analysis of data from the *BEACH* (Bettering the Evaluation and Care of Health) program, a continuous study of general practice activity. The data period investigated is from April 1999 to March 2000 inclusive. The *BEACH* methods are summarised below. A more detailed description of the methods pertaining specifically to the analyses of orders for imaging is provided in Section 2.2.

2.1 The *BEACH* program

The methods adopted in the *BEACH* program have been described in detail elsewhere (Britt et al. 1999b; Britt et al. 1999c; Britt et al. 2000). In summary, each of the recognised GPs in a random sample of approximately 1,000 per year records details about 100 doctor-patient encounters of all types. The information is recorded on structured encounter forms (on paper). It is a rolling sample, being recruited approximately three weeks ahead. Approximately 20 GPs participate each week, 50 weeks a year.

Sampling methods

The source population includes all GPs who claimed a minimum of 375 general practice A1 Medicare items (items 1-51, 601, 602) in the most recently available three-month HIC data period. This equates with 1,500 Medicare claims a year and ensures inclusion of the majority of part-time GPs whilst excluding those who are not in private practice but claim for a few consultations a year. The General Practice Branch of the Commonwealth Department of Health and Aged Care (DHAC) draws a sample on a regular basis.

The sampling methods adopted by the General Practice Branch of the DHAC aim to provide a series of researchers with a random unbiased selection of GPs while minimising overlap with past samples. The method is a modification of Classic Synchronised Sampling and has been described in detail elsewhere (Britt et al. 2000; Calcino 1993).

Recruitment methods

The randomly selected GPs are approached initially by letter, then by telephone follow-up. GPs who agree to participate are set an agreed recording date approximately three to four weeks ahead. A research pack is sent to each participant about 10 days before their planned recording date. A telephone reminder is made to each participating GP in the first days of the agreed recording period. Non-returns are followed up by regular telephone calls.

Each participating GP earns 25 Clinical Audit points towards their quality assurance (QA) requirements. As part of this QA process, each receives an analysis of his or her results compared with those of nine other unidentified practitioners who recorded at approximately the same time. Comparisons with the national average and with targets relating to the National Health Priority Areas are also made. In addition, GPs receive some educational material related to the identification and management of patients who smoke or who consume alcohol at hazardous levels.

Statistical methods

The analysis of the *BEACH* database is conducted with SAS version 6.12 (SAS Institute Inc. 1996) and the encounter is the primary unit of analysis. Proportions (%) are only used when describing the distribution of an event that can arise only once at a consultation (e.g. age, sex or item numbers) or to describe the distribution of events within a class of events (e.g. problem *A* as a per cent of total problems).

Rates per 100 encounters are used when an event can occur more than once at the consultation (e.g. RFEs, problems managed or medications). Rates per 100 problems are also sometimes used when a management event can occur more than once per problem managed. In general, the following results present the number of observations (*n*), rate per 100 encounters and the 95% confidence intervals.

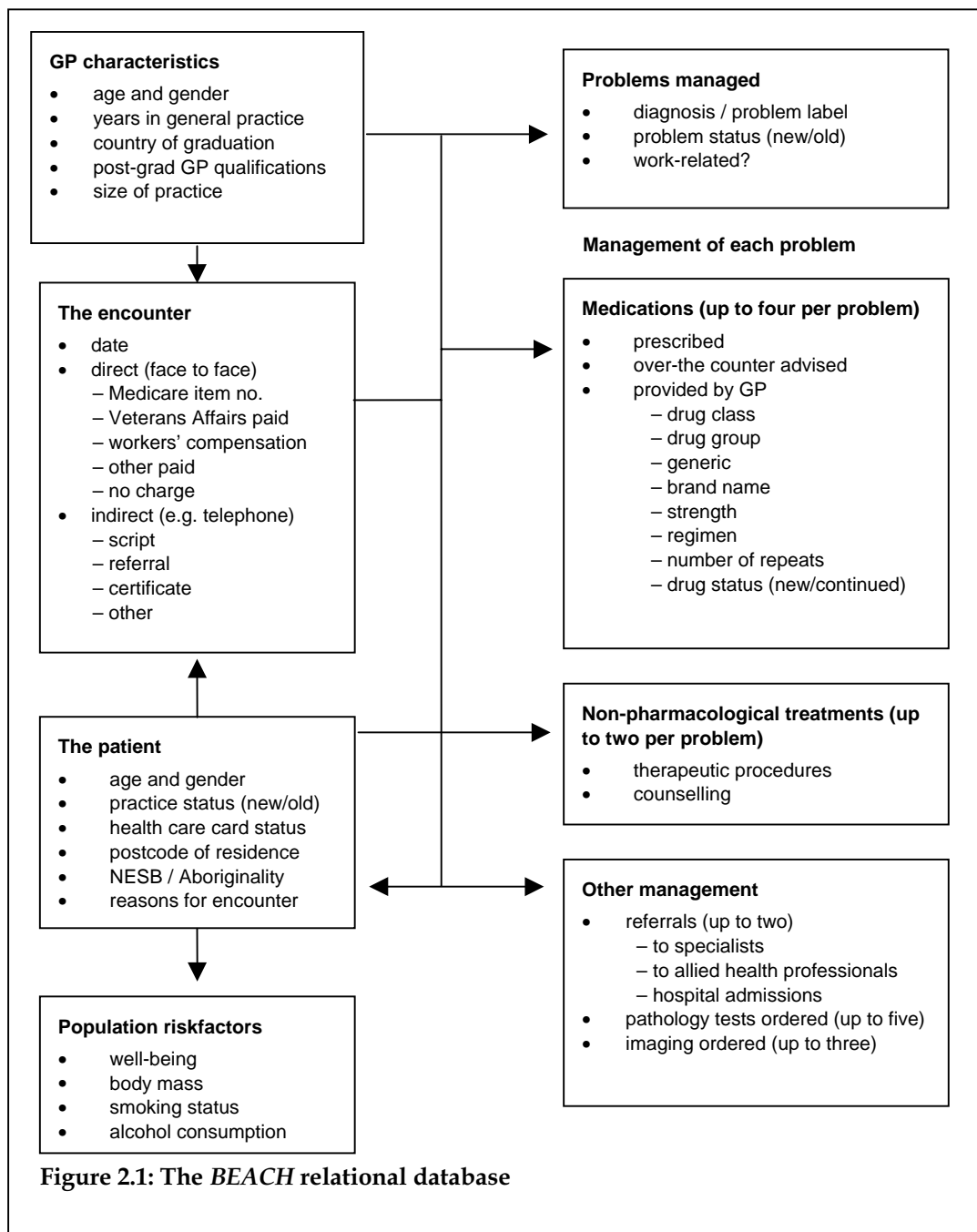


Figure 2.1: The *BEACH* relational database

The *BEACH* study is essentially a random sample of GPs, each providing data about a cluster of encounters. Cluster sampling study designs in general practice research violate the simple random sample (SRS) assumption because the probability of an encounter being included is a function of the probability of the GP being selected (Sayer 1999).

There is also a secondary probability function of particular encounters being included in the GP's cluster (associated with the characteristics of the GP or the type and place of the practice) and this increases the likelihood of sampling bias. In addition, there will be inherent relationships between encounters from the same cluster and this creates a potential statistical bias. The probability of gaining a representative sample of encounters is therefore reduced by the potential sampling and statistical bias, decreasing the accuracy of national estimates.

When a study design other than SRS is used, analytical techniques that consider the study design should be employed. In this report the standard error calculations used in the 95% confidence intervals accommodate both the single-stage clustered study design and sample weighting according to Kish's description of the formulae (Kish 1965). SAS 6.12 is limited in its capacity to calculate the standard error for the current study design, so additional programming was required to incorporate the formulas. Post-stratification weighting was also applied to the raw data before analysis (Britt et al. 2000).

The *BEACH* relational database

The *BEACH* relational database is described diagrammatically in Figure 2.1. Note that all variables can be directly related to GP and patient characteristics and to the encounter. Reasons for encounter have only an indirect relationship with problems managed. All types of management are directly related to the problem being treated.

Classification of data

The imaging tests ordered, patient reasons for encounter, problems managed, therapeutic procedures, other non-pharmacological treatments, referrals, pathology and imaging are coded using ICPC-2 PLUS (Britt 1997). This is an extended vocabulary of terms classified according to the International Classification of Primary Care (Version 2) (ICPC-2), a product of the World Organization of Family Doctors (WONCA) (Classification Committee of the World Organization of Family Doctors (WICC) 1997). The ICPC is regarded as the international standard for data classification in primary care.

The ICPC has a bi-axial structure, with 17 chapters on one axis (each with an alphabetic code) and seven components on the other (numeric codes). Chapters are based on body systems, with additional chapters for psychological and social problems.

- Component 1 includes symptoms and complaints.
- Component 7 covers diagnoses.

These are independent in each chapter and both can be used for patient reasons for encounter or for problems managed.

- Components 2 to 6 cover the process of care and are common throughout all chapters. The processes of care, including referrals, non-pharmacological treatments and orders for pathology and imaging, are classified in these process components of ICPC-2. Component 2 (diagnostic screening and prevention) is also often applied in describing the problem managed (e.g. check-up, immunisation).

The ICPC-2 is an excellent epidemiological tool. The diagnostic and symptomatic rubrics have been selected for inclusion on the basis of their relative frequency in primary care settings or because of their relative importance in describing the health of the community. It has only about 1,370 rubrics and these are sufficient for meaningful analyses. However, reliability of data entry, using ICPC-2 alone, would require a thorough knowledge of the classification if correct classification of a concept were to be ensured. In 1995, recognising a need for a coding and classification system for general practice electronic health records, the Family Medicine Research Centre (then Unit) developed an extended vocabulary of terms classified according to the ICPC. These terms were derived from those recorded in more than half a million encounter forms by GPs participating in the quality assurance option mentioned earlier. The terms have developed further over the past six years in response to the use of terminology by GPs participating in the *BEACH* program and in response to requests from GPs using ICPC-2 PLUS in their electronic clinical systems. This allows far greater specificity in data entry and ensures high inter-coder reliability between staff. It also facilitates analyses of information about more specific problems when required (Britt 1997).

2.2 The study of imaging orders

This study is based on the imaging orders recorded on the encounter forms completed in the second year of the *BEACH* program. In that year a random sample of 1,047 GPs participated, providing details of 104,700 GP-patient encounters. After post-stratification weighting for GP age, sex and activity level, there were 104,856 encounters. An overview of these encounters is provided in Section 3.

An initial analysis of the relative rates of imaging orders for each GP was undertaken in order to identify low, medium and high ordering groups of GPs. The characteristics of the GPs in each group were compared and the results are reported in Chapter 4 (Section 4.4). The factors that affect the imaging order rates of GPs were investigated using analysis of variance and linear regression at both the univariate and multivariate levels (Section 4.5).

All encounter records that included a record of at least one order for an imaging test were identified, and a sub-file was created. An overview of the characteristics of the encounters at which imaging was ordered, and the patients for whom they were ordered is provided in Chapter 5. The remainder of the report investigating the types of imaging tests ordered by GPs and the problems for which they were ordered includes more detailed analyses of selected specific test types and selected problems.

Statistical reporting

This study uses the annual weighted *BEACH* data from the 1999-00 collection period. The weighting process (described earlier) leads to raw figures which are not round numbers. SAS generates the raw results to two decimal places. The raw figures in this report have been rounded for simplicity. This means that individual rounded frequencies do not always summate exactly to the reported total. Further, where results are reported in terms of rates, these have been calculated using the more specific two-decimal place raw data rather than the reported rounded frequencies. This may result in slight inconsistencies between results from different analyses.

Chapters 8 and 9 investigate the relationships between specific imaging test types and their related morbidity under management, or between selected morbidity and its related test order pattern. The reported rates are often based on the total number of 'problem-test combinations'. As there can be a many-to-one, one-to-many, many-to-many relationship between test orders and problems under management, each combination of a single test and a single problem is counted once. This means that the total number of problem-test combinations will only agree with the total number of tests ordered, and with the total number of tested problems, where there was always a one-to-one relationship between a single test order and a single problem under management.

In general, the results of this study are reported in rates per 100 encounters or rates per 100 problems managed. Where the relative frequency of an event is sufficient to provide statistical estimates of accuracy (such as in Section 5) the 95% confidence interval and relative standard errors are also provided. The rate is an estimate and its confidence limits suggest a 95% certainty that the true result lies between the reported upper and lower limits.

The relative standard error (RSE) is commonly used by the Australian Bureau of Statistics. It is a function of the standard error and the rate estimate and it provides a measure of reliability of the estimate. For general purposes a RSE of 0-15 indicates that the estimate is reliable, a RSE of 16-33 is slightly unreliable, one of 34-50 is extremely unreliable and a RSE of 51-100 suggests that the estimate should not be used.

In the more specific analyses of selected imaging types or selected problems, only the estimated relative frequency of the event is provided. While the relative standard errors would often suggest unreliability of the estimated ordering rates, in the absence of any Australian information about the relationship of test ordering to morbidity in general practice, the data are still useful in describing these trends. However, the size of the samples and the possible unreliability of the reported results must be kept in mind.

Coding and classifying imaging tests ordered

When the *BEACH* program began in April 1998, the coding and classification of imaging was less precise than that of morbidity managed. This is because in the first instance, ICPC-2 PLUS was developed on the basis of approximately 700,000 encounter records completed between 1990 and 1998 by GPs. However, in these encounter records, imaging orders were recorded only in terms of four tick box options: plain, contrast/special, ultrasound and other. Multiple boxes could be ticked where multiple test types were ordered. However, there was no opportunity to indicate multiple tests of a specific type (e.g. multiple plain x-rays of different sites) or to describe them in any specific way.

As a result GPs did not record the specific test label on the records. In the first year of *BEACH* a detailed terminology of test orders was therefore not available in ICPC-2 PLUS. During the first year the GPs were asked to record orders for imaging (in terms of whether they ordered a plain x-ray, US/CT/contrast, or other imaging), to nominate the body site of the test(s) and to elect the problem or problems for which the test was associated. This method provided no opportunity to define the terms usually used by GPs in describing their order for imaging for possible addition to ICPC-2 PLUS.

In response to government and the profession's interest in more detailed data on the orders placed for imaging, in the second year of the program the participants were asked to record the type of imaging test ordered (in their own words) as well as to specify the body site and the problem(s) for which each test was ordered. The free text was then classified according to the available ICPC-2 PLUS terms and codes.

This approach provided an opportunity to review the free text descriptors of the imaging orders and create more specific ICPC-2 PLUS codes to better reflect this free text and to better differentiate between different test types. In the second half of 2000 a research grant was received from the DHAC to conduct a full review of the imaging orders described in free text and further development of the ICPC-2 PLUS terms to reflect the higher specificity provided by these recorded terms.

In order to gain a timely view of the imaging ordered by GPs the newly created ICPC-2 PLUS codes for imaging orders were applied in a recoding of all imaging orders recorded in the second year of the *BEACH* program. This process was also funded by the DHAC.

Grouping codes for problems associated with imaging orders

In this report, some grouping of ICPC-2 codes and/or ICPC-2 PLUS codes has been made to overcome differences in the level of specificity recorded by GPs in ascribing problem labels. For example, we report imaging tests ordered for the problem label 'back pain'. This problem label includes all symptoms and complaints of the back (ICPC-2 codes L02) and those of the lower back (ICPC-2 code L03) together with sprains and strains of the back (which form part of ICPC-2 code L84) and selected back problems with radiating symptoms (which form part of ICPC-2 code L86).

Individual analysis of 'symptoms and complaints of the lower back' and 'hypertension with complications' and 'symptoms and complaints of the back' etc. may have meant that the relative frequencies of each were insufficient to report. Another example is osteoarthritis. There are multiple codes into which this problem may fall depending on its body location (i.e. osteoarthritis of the knee has a different ICPC-2 code from osteoarthritis of the shoulder). Osteoarthritis of the back is only a small part of a broader rubric. In this case the concept here reported as 'osteoarthritis' includes all the ICPC-2 PLUS terms associated with osteoarthritis rather than a number of ICPC-2 codes. All problem labels which include multiple codes from ICPC-2 are marked with an asterisk in the tables. Appendix 4 provides a full list of terms or rubrics that make up each problem label.

Grouping imaging order codes for analysis

Component 3 of the ICPC-2 provides for the classification of 'Diagnostic, Screening and Preventive Procedures'. This component includes a single rubric relating to imaging orders – 'diagnostic radiology/imaging' (-41). Each of these can be applied across any one of the 17 chapters in ICPC-2 (see 'Classification of data' on page 8).

Imaging orders could be naturally grouped according to the ICPC-2 chapter (body system) to which they were applied. However, imaging data are more usually reported in Australia in terms of the groupings used in the Medicare Benefits Schedule (MBS) (DHAC 1998).

The newly developed ICPC-2 PLUS imaging terms were therefore mapped to the MBS. It had been hoped that the map would be from the ICPC-2 PLUS codes to the most specific (five-digit) MBS codes or at least to the MBS sub-group level (e.g. Group I1 – Ultrasound, Sub-group 1). However, this was found to be an impossible task. The MBS classifies imaging tests for the purposes of costing and the classification bears little resemblance to clinical classifications. Some examples of the difficulties encountered in such mapping are provided below.

Matching at the MBS sub-group level

The ICPC-2 PLUS imaging order term 'x-ray; abdomen' recorded by the GP for 'abdominal pain' provides no indication of whether the cause of the pain is gynaecological, urological or gastrointestinal in nature. When the test order is received by the radiologist, s/he may use the clinical notes provided with the request to decide on the specific type of imaging to be undertaken. This allows the test to be suitably classified in the MBS. However, at the point of test order from the GP, the imaging test order of 'x-ray abdomen' could be classed in Diagnostic Radiology Sub-group 7 (Radiographic examination of urinary tract) or in Sub-group 8 (Radiographic examination of alimentary tract and biliary system).

Matching at the MBS individual (five-digit) item level

The map is even more complicated at the five-digit level. Where it is possible to allocate a specific ICPC-2 PLUS term to a single MBS sub-group, there can be multiple options for codes within that sub-group. For example, the ICPC-2 PLUS term 'breast ultrasound' could refer to a request for ultrasound of one or of both breasts. It therefore is not possible to map to the MBS at the five-digit level because there are two possible MBS items: one for a single breast ultrasound (item 55070) and another for ultrasound of both breasts (55079).

These items also provide an example of the limits placed on item selection. Many tests such as these are classified elsewhere in the MBS if they are associated with other specific test types. Further, there are many examples of a single MBS item representing combinations of other imaging tests, each of which already has a separate MBS item. If these limitations were to be considered in creating a map from ICPC-2 PLUS to the MBS, most of the PLUS terms would require multiple possible MBS maps depending on other tests ordered at the same encounter and with consideration of the problem under management.

Final map to the MBS

The ICPC-2 PLUS imaging order codes were therefore mapped to the MBS at the Group level only. These groups are:

- Group I1 – Ultrasound
- Group I2 – Computed tomography
- Group I3 – Diagnostic radiology
- Group I4 – Nuclear medicine imaging
- Group I5 – Magnetic resonance imaging.

Care process

A specific interest of this study was the extent to which imaging was ordered for investigative/diagnostic purposes versus monitoring purposes. Each test was therefore classified into one of four groups depending on the status of the problem to the patient (new/old) and the level of diagnostic specificity inherent in the label assigned by the GP to the problem associated with the test order. Four care processes were designated for imaging tests ordered: investigative imaging, management imaging, monitoring imaging, and undefined imaging. The definitions applied to each care process are described in more detail in Section 7.2.