

2 Methods

This report details findings from a secondary analysis of data collected in the BEACH (Bettering the Evaluation and Care of Health) program between April 1998 and March 2004. During this period 6,019 GPs recorded data about 601,900 encounters.

The methods applied in the BEACH program have been described in detail elsewhere.¹⁵⁻¹⁷ In summary each year a random sample of approximately 1,000 recognised GPs each records details about 100 doctor-patient encounters of all types. The information is recorded on structured encounter forms (on paper). The GPs are recruited across the year in a rolling sample, approximately 20 GPs participate each week, 50 weeks a year.

2.1 Sampling frame

The source population includes all GPs who claimed a minimum of 375 general practice A1 Medicare items in the most recently available 3-month Health Insurance Commission (HIC) data period. This equates with 1,500 Medicare claims a year and ensures inclusion of the majority of part-time GPs while excluding those who are not in private practice but claim for a few consultations a year. The General Practice Branch of the Australian Government Department of Health and Ageing draws a sample on a regular basis.

2.2 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 3 to 4 weeks ahead. A research pack is sent to each participant about 10 days before the 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.

Participating GPs earn Clinical Audit points towards their quality assurance (QA) requirements. As part of this QA process, each GP receives an analysis of his or her results compared with those of nine other de-identified GPs 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 consume alcohol at hazardous levels.

2.3 Data elements

BEACH includes three interrelated data collections: encounter data, GP characteristics, and patient health status. An example of the forms used to collect the encounter data and the data on patient health status is included as Appendix 1. The GP characteristics questionnaire is included as Appendix 2.

Encounter data include date of consultation, type of consultation (direct, indirect), Medicare/Department of Veterans' Affairs item number (where applicable) and specified other payment source (tick boxes).

Information about **the patient** includes date of birth, sex and postcode of residence. Tick boxes are provided for Commonwealth concession card holder, Repatriation health card holder, non-English-speaking background (NESB), an Aboriginal person (self-identification) and Torres Strait Islander (self-identification). Space is provided for up to three patient reasons for encounter (RFEs).

The **content of the encounter** is described in terms of the problems managed and the management techniques applied to each of these problems. Data elements include up to four diagnoses/problems. Tick boxes are provided to denote the status of each problem as new to the patient (if applicable).

Management data for each problem include medications prescribed, over-the-counter medications advised and other medications supplied by the GP. Details for each **medication** comprise brand name, form (where required), strength, regimen, status (if new medication for this problem for this patient) and number of repeats. **Other management** techniques recorded for each problem, include counselling, procedures, new referrals, pathology and imaging ordered.

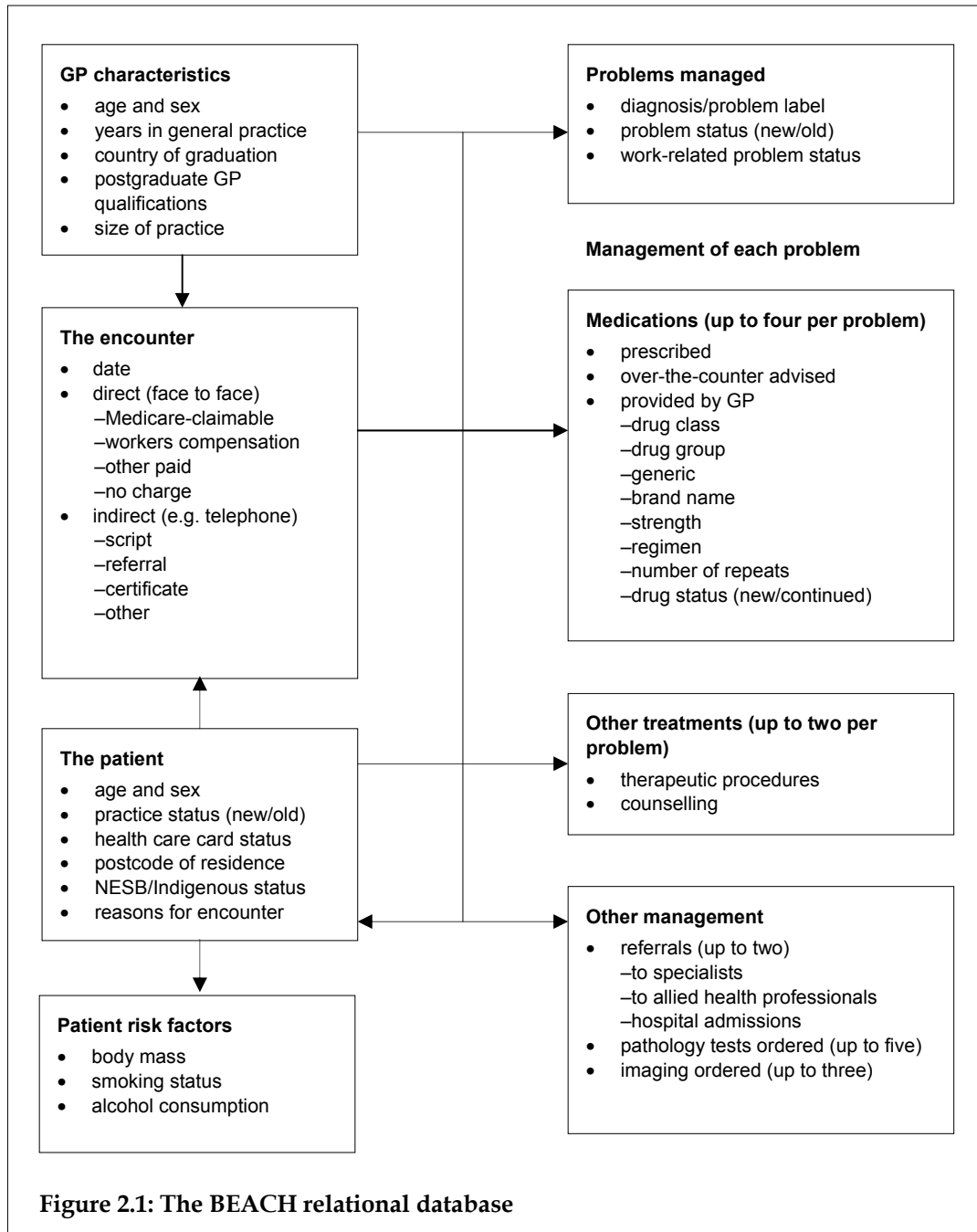
GP characteristics include age and sex, years in general practice, number of GP sessions worked per week, number of GPs working in the practice (to generate a measure of practice size), postcode of major practice address, country of graduation, vocational general practice training, Fellowship of the Royal Australian College of General Practitioners (RACGP) status, after-hours care arrangements, use of computers in the practice, whether the practice is accredited and whether it is a teaching practice, work undertaken by the GP in other clinical settings, hours worked in direct patient care and hours on call per week.

2.4 Applying RRMA and ASGC Remoteness Structure to the BEACH sample

RRMA and ASGC Remoteness categories were allocated based on the postcode of the GP's practice using concordance maps between postcode and RRMA or ASGC Remoteness categories. Where a postcode fell into more than one category it was allocated to the category that accounted for the largest proportion of the postcode. All encounters were then classified to the GP's RRMA or ASGC Remoteness category.

2.5 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.



2.6 Statistical methods

The analysis of the BEACH database was conducted with SAS version 8.2.¹⁸ Proportions (%) are used only when describing the distribution of an event that can arise only once at an encounter (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 percentage of total problems). Rates per 100 encounters are used when an event can occur more than once at the encounter (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.

The BEACH study is a random sample of GPs, each providing data about a cluster of encounters. When the encounter is the unit of inference, the cluster sampling study design violates the simple random sample (SRS) assumption of equal probability of selection of observations, because the probability of an encounter being included is a function of the probability of the GP being selected.¹⁹ Cluster samples also violate the assumption of independence of observations as there is an inherent relationship or correlation between encounters sampled in the same cluster. Therefore the certainty that the sample estimates reflect the true underlying population values is reduced by cluster sampling, thus decreasing the precision 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 data were analysed using SAS 8.2 procedures that adjust the standard error for the intra-cluster correlation of the cluster sample. In this way the 95% confidence intervals were adjusted to reflect the reduced precision of the estimates.

In this report we compared the results for each RRMA or ASGC category with the national average. Statistical significance of differences is identified by non-overlapping 95% confidence intervals. Marginal differences (where the confidence intervals meet but do not overlap) are generally not noted but can be identified in the tables. The width of confidence intervals (and hence the bounds for identifying differences between a given area and the national average) tend to differ systematically with the size of the sample from that category. Less populated areas are represented by a smaller sample of GPs and encounters than more populated areas. Confidence intervals for small samples are therefore wider, giving less precision for the estimates. This means that there will be a decreased chance of identifying real differences between the less populated remote categories and the national average.

However, since the ASGC Remoteness Structure is essentially an ordinal scale, linear trends with increasing remoteness across ASGC categories were also examined to allow more power to detect real differences related to remoteness.

2.7 Classification of data

The patient reasons for encounter, problems managed, procedures, other treatments, referrals, pathology and imaging orders are coded using ICPC-2 PLUS.²⁰ 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).²¹ The ICPC is used in more than 45 countries as the 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) (Figure 2.2). 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, other 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).

Components	Chapters																
	A	B	D	F	H	K	L	N	P	R	S	T	U	W	X	Y	Z
1. Symptoms, complaints																	
2. Diagnostic, screening, prevention																	
3. Treatment, procedures, medication																	
4. Test results																	
5. Administrative																	
6. Other																	
7. Diagnoses, disease																	

A	General	L	Musculoskeletal	U	Urinary
B	Blood, blood-forming	N	Neurological	W	Pregnancy, family planning
D	Digestive	P	Psychological	X	Female genital
F	Eye	R	Respiratory	Y	Male genital
H	Ear	S	Skin	Z	Social
K	Circulatory	T	Metabolic, endocrine, nutritional		

Figure 2.2: The structure of the International Classification of Primary Care – Version 2 (ICPC-2)

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 by GPs on more than half a million encounter forms. The terms have developed further over the past 8 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 secondary coding staff. It also facilitates analyses of information about more specific problems when required.²⁰

Classification of pharmaceuticals

Pharmaceuticals prescribed or provided and over-the-counter medications advised by the GP are coded and classified according to an in-house classification, the Coding Atlas for Pharmaceutical Substances (CAPS). This is a hierarchical structure that facilitates analysis of data at a variety of levels, such as medication class, medication group, generic composition and brand name.

2.8 Quality assurance

All morbidity and therapeutic data elements are automatically coded and classified by computer as secondary coding staff enter key words or word fragments and select the required term or label from a pick list. A quality assurance program to ensure reliability of data entry includes ongoing development of computer-aided error checks ('locks') at the data entry stage and a physical check of samples of data entered versus those on the original recording form. Further automatic logical data checks are conducted on a regular basis.

2.9 Validity and reliability

In the development of a database such as BEACH, data gathering moves through specific stages: GP sample selection, cluster sampling around each GP, GP data recording, and secondary coding and data entry. At each stage, the data can be invalidated by the application of inappropriate methods. The methods of coding, data entry and statistical analysis described above ensure maximum reliability and validity in the recording and reporting of the data provided by the GPs in the sample.