2. Methods

The methods adopted in the BEACH program have been described in detail elsewhere (Britt et al. 1999). This section describes the development of the BEACH process and provides a brief summary of the methods adopted.

2.1 Development of the BEACH methods

The 1990–91 AMTS provided a dataset on which to develop a sample size model for future National studies (Meza et al. 1995) and on which to test new statistical analytical techniques (Britt et al. 1996a; Sayer and Britt 1996; Sayer and Britt 1997). With the agreement of the RACGP, the AMTS process was also used as a quality assurance option (the Morbidity and Therapeutic Index) for GPs between 1991 and 1996. Over 4,000 GPs took up this option and further tested the process on a wider general practice population.

In 1996 the Western Sydney Division of General Practice provided funding for a local morbidity and treatment survey. The Division required current data concerning the health needs of the population in its region and the activities of GPs practising in Western Sydney in order to plan future projects and educational programs. This provided the opportunity to test a more detailed encounter form, inclusion of patient based questions on health risk factors, the application of extensive and detailed coding systems for diagnoses, pharmaceutical treatments and other management techniques, a comprehensive database 'front-end' and direct computer assisted secondary data entry.

As 85% of the population visit a GP in any one year and over 90% visit at least once in any two year period, there was also interest in the possible use of the GP patient population to measure aspects of population health. This led to the addition of a new section on each form concerning patient-based risk factors and health assessment. The revised program was approved by the RACGP as a quality assurance option (audit) for participants.

All of these aspects of the research method were found to be viable, with the exception of the layout of the prescription details to be recorded on the forms. GPs had been asked to record the daily dose of medication. Feedback from participants alerted the researchers to the fact that this involved the doctors making a calculation they would not normally make when writing a prescription or writing up their records. To eliminate this extra burden on busy GPs, daily dose of medication was replaced with drug strength and regimen fields (eg 500mg; 4 daily) to bring these fields into line with the usual manner in which a doctor records that information. The recording form was revised after the Western Sydney pilot study.

In 1997 the Department of Human Services, Victoria, commissioned a study of general practice activity in that State. The objective was to measure any changes in activity since 1990–91 (AMTS) and provide a new baseline for the measurement of future change. The revised recording form was used in this statewide study.

When establishing the consortium of government departments and instrumentalities and industry participants to support the BEACH program there were some additional data elements identified for which there was a current lack of national information. As a result a number of fields were added to the form for the national BEACH program. These included patient Veterans' Affairs card status, indication of problems regarded by the GP as likely to be work related and space for GP recording of specific orders for pathology tests and imaging.

2.2 BEACH methods

In summary, a random sample of approximately 1,000 recognised GPs per year each records details about 100 doctor–patient encounters of all types on structured paper encounter forms.

The source population includes all recognised GPs who have claimed a minimum of 375 general practice Medicare items (items 1–51) in the most recently available three-month Health Insurance Commission (HIC) data period. This equates with a cut-off of 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 every three months.

The randomly selected GPs are approached by letter with 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 ten days before their planned recording date. The research pack contains a covering letter, a project information sheet, a GP profile questionnaire, a pad of 105 recording forms (to allow for some error), a detailed set of instructions, a height and weight measure conversion (to metric) chart (for body mass index), a sample completed form with explanation, a pictorial 'standard drinks' chart to help patients answer questions on alcohol intake, additional instructions for completing supplementary questions on each form, a reply-paid envelope and several copies of a patient information sheet. The patient information sheet gives patients the choice to 'opt out' and not have details of their consultation included in the study by informing their GP of this decision. A telephone reminder is made to each GP participant in the first days of the agreed recording period. Non-returns are followed up by regular telephone calls.

Each participating GP earns 25 audit points towards their quality assurance (QA) requirements. As part of this QA process they receive an analysis of their own results compared with those of nine other unidentified practitioners who recorded at approximately the same time. Comparison with the national average and with targets relating to the National Health Priority Areas is 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.

2.2.1 Data elements

The BEACH recording forms build on those used in the earlier work of the Family Medicine Research Centre but with considerably more detail about each encounter. In particular BEACH includes details of all types of encounters whether paid by Medicare, by another source, or unpaid. Indirect consultations for which GPs are not remunerated are also included and there are more details about the characteristics of patients attending general

practice. For the first time information was collected about pharmacological management including that for drugs advised for over the counter purchase (OTCs) and drugs supplied by the GP. More specific details of dosage and regimen have also been added to the data collection process. Unlike the AMTS, orders for pathology tests, imaging referrals to specialists, allied health professionals and emergency departments, and hospital admissions are related to the specific morbidity under management. The specific types of pathology tests and imaging ordered at the encounter are also included, whereas in earlier studies these orders were only broadly grouped (i.e. blood, urine, culture tissue).

BEACH includes three inter-related data collections: encounter data, GP characteristics, and patient health status. An example of the form used to collect the encounter data and the data on patient health status is included as Appendix I. The GP characteristics questionnaire is included as Appendix II.

Encounter data includes information about the consultation itself: date of consultation; type of consultation (direct, indirect); Medicare item number (where applicable); specified other forms of payment; clinical services provide at indirect encounters.

Information about **the patient** includes: date of birth; gender; status to the practice (new/seen before); postcode of residence; health care card status (yes/no); Veterans' Affairs status (Gold/White); non-English speaking background (yes/no); Aboriginal (yes/no) (self-identification); Torres Strait Islander (yes/no) (self-identification); patient reasons for encounter (up to three).

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; the status of each problem (new to patient/managed before) and whether it was thought to be work related.

Management data for each problem include: medications prescribed, over the counter drugs advised and other drugs supplied by the GP. Details for each **medication** comprise: brand name; form (where required); strength; regimen; status (new drug this problem this patient/continuation) and number of repeats. **Non-pharmacological management** of each problem includes counselling and therapeutic procedures, new referrals and pathology and imaging ordered.

GP characteristics include: age and gender; years in general practice; number of GP sessions worked per week; number of full-time and part-time GPs working in the practice (to generate practice size); consultations in languages other than English; postcode of major practice address; country of graduation; postgraduate general practice training and FRACGP status; membership of professional organisations; brand substitution behaviour (Appendix II).

Supplementary analysis of nominated data (SAND): A section on the bottom of each recording form investigates aspects of patient health or health care delivery in general practice not covered by the consultation-based information (see Appendix I). The year-long data-collection period is divided into 10 blocks, each of five weeks and designed to include data from 100 GPs. Each GP's recording pack of 100 forms is made up of: 40 forms which contain questions about patient well-being, height, weight and alcohol intake; 40 which have a single question about the patient's smoking status together with questions on other subjects nominated for that block; and 20 forms with other nominated questions.

The results of the SAND sub-studies will be reported in a separate publication.

2.2.2 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.

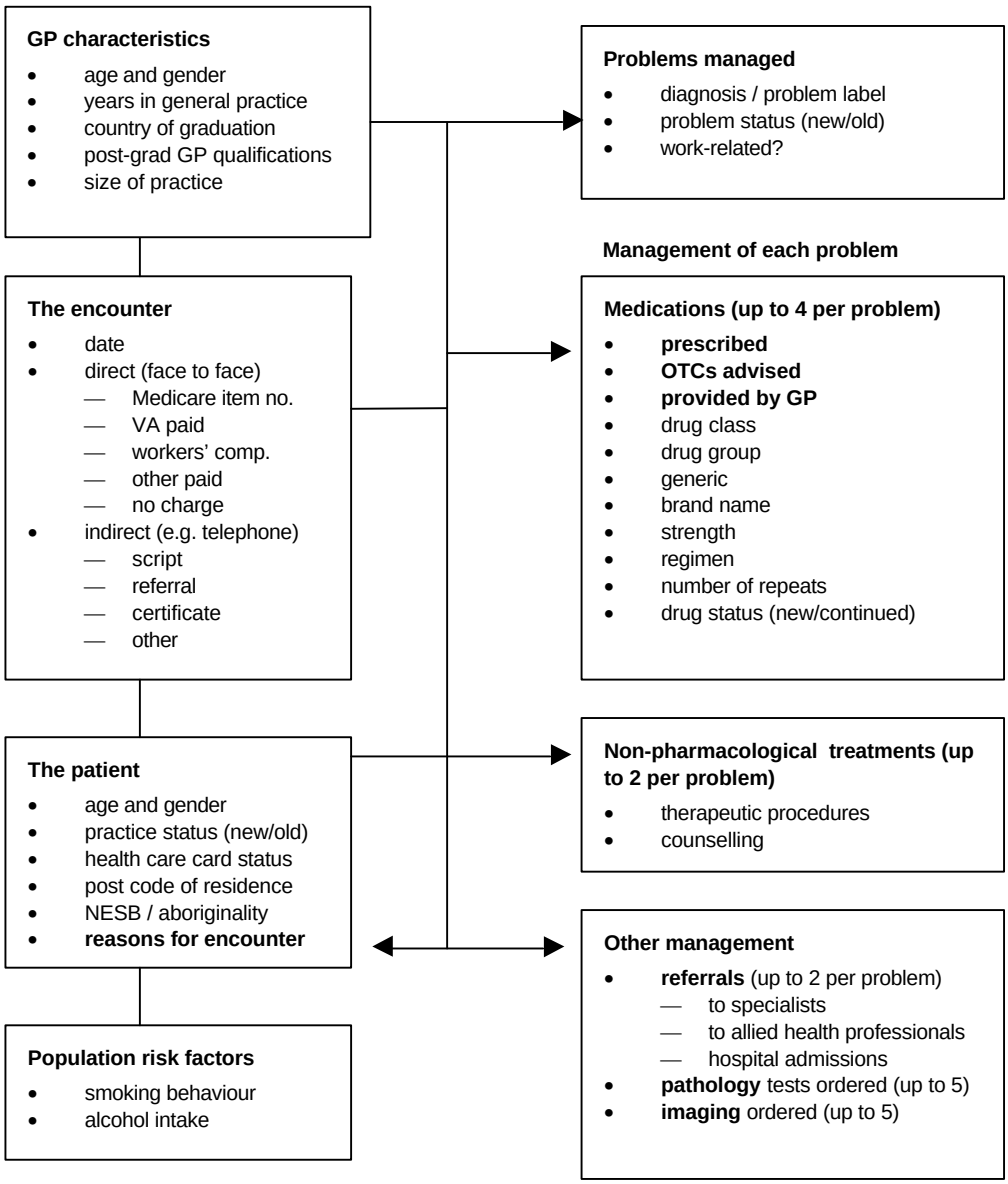

Figure 2.1: The BEACH relational database

2.2.3 Statistical methods

The analysis of the BEACH database is conducted through SAS version 6.12 (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 an consultation (e.g. age, gender or item numbers) or to describe the distribution of events within a class of events (e.g. problem *A* as a % 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 used when a management event can occur more than once per problem managed (e.g. prescribed drugs; orders for pathology). In general, the following results present the number of observations (n), rate per 100 encounters and the 95% confidence intervals.

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 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 statistical bias. For example, female GPs tend to see more female patients than their male counterparts; a group of patients of one GP may receive different treatments to those received by patients of another GP, reflecting different practice styles. 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 an investigator violates the SRS assumption, analytical techniques that consider the study design should be employed. In this report the standard error calculations used in the 95% confidence intervals incorporate both the single-stage clustered study design and sample weighting according to Kish's description of the formulae (Kish 1965). SAS is limited in its capacity to calculate the standard error for the current study design, so additional programming has been required to incorporate the formulae.

2.2.4 Classification of data

Patient reasons for encounter, problems managed, therapeutic procedures, other non-pharmacological treatments, referrals, and pathology and imaging ordered are coded using ICPC-2 PLUS (Britt 1997a). 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 1997). The ICPC is regarded as the international standard for data classification in primary care.

The International Classification of Primary Care

Until the mid 1970s most morbidity data collected in general practice research were classified according to the International Classification of Diseases (ICD). The many symptoms which present in general practice were difficult to code with a classification originally designed for application to mortality statistics and with a disease-based structure.

Recognising this problem, the Classification Committee of the World Organization of Family Doctors (WONCA) developed the International Classification of Health Problems in Primary Care (ICHPPC), first published in 1976 with a second edition in 1983 (WONCA 1983). Although this provided a section for the classification of some undiagnosed symptoms, it was still based on the ICD structure and the available symptom rubrics were inadequate. A new classification was needed to encompass both the patient's reasons for encounter (RFEs) and the patient's problems.

A RFE classification with a structure different from the ICD-9 framework was first tested in 1983 in an international field trial involving nine countries including Australia. After revisions and additions, the International Classification of Primary Care (ICPC) was first published in 1987. It included the majority of the diagnostic rubrics from ICHPPC and a series of process rubrics (drawn from IC-Process-PC) (1986) to describe the care provided. The second edition, ICPC-2, incorporates inclusion and exclusion criteria and was published in 1998. ICPC has been translated into more than 35 languages and is being used to classify patient reasons for encounter and/or problems managed in Norway, Denmark, Canada, the Netherlands, Belgium, France and the United States (Brage et al. 1996; Lavoie et al. 1995; Viner et al. 1994; Dupuits and Hasman 1995; Jamoulle et al. 1994; Klinkman and Green 1995; Vijlbrief et al. 1995; Zaat et al. 1995).

	Chapters																
Components	**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 (ICPC)

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 while **Component 7** covers diagnoses. These are independent in each chapter and either can be used for patient RFEs or for problems managed.

Components 2 to 6 cover the process of care and are common throughout all chapters, each rubric being equally able to be applied to any body system (Figure 2.2). 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).

ICPC-2 PLUS

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 measuring 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 is 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.

Each term has its own extended code. For example, while the ICPC code A77 is 'Other viral illness', the PLUS terms provide a list of some 33 specific viral illnesses under A77 (e.g. Ross River Fever—A77 001). This allows far greater specificity in data entry and ensures high inter-reliability between staff. It also facilitates analyses of information about more specific problems when required (Britt 1997a).

In this report some grouping of ICPC-2 rubrics has been made to overcome differences in the level of specificity recorded by GPs in describing patient RFEs or ascribing problem labels. The issue of variance in labelling is discussed below. For example, results are reported for the problem label 'rash'. Individual analysis of 'localised' and 'generalised' rash may have meant that the relative frequencies of each were insufficient to report. Another example is osteoarthritis. There are multiple rubrics into which this problem may fall depending on its body location (i.e. osteoarthritis of the knee has a different ICPC-2 code to osteoarthritis of the shoulder). Osteoarthritis of the back is only a small part of a broader rubric. In this case the grouper here reported as 'osteoarthritis' includes all the ICPC-2 PLUS terms associated with osteoarthritis rather than a number of ICPC-2 rubrics. The codes included in each grouped label are listed in Appendix III.

Classification of pharmaceuticals

Pharmaceuticals prescribed or provided and over the counter drugs 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, for example, drug class, drug group, generic composition and brand name. CAPS is mapped to the Anatomical Therapeutic Chemical classification (ATC) (WHO Collaborating Centre for Drug Statistics Methodology 1998) which is the Australian standard for classifying drugs at the generic level. Strength and regimen are independent fields which, when combined with the CAPS code, give an opportunity to derive prescribed daily dose for any drug or group of drugs.

Quality assurance

All morbidity and therapeutic data elements are automatically coded and classified by the computer as 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 one in five physical check of data entered versus that on the original recording form.

2.2.5 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; secondary coding and data entry. At each stage the data can be invalidated by the application of inappropriate methods. The methods adopted to ensure maximum reliability of coding and data entry have been described above. The statistical techniques adopted to ensure valid reporting of recorded data are described in Chapter 4.

Previous work has demonstrated the extent to which a random sample of GPs recording information about a cluster of patients represents all GPs and all patients attending general practitioners (Driver et al. 1991). Other studies have reported the degree to which GP reported patient reasons for encounter and problems managed accurately reflect those recalled by the patient (Britt et al. 1992) and the reliability of secondary coding of RFEs (Britt 1998) and problems managed (Bridges-Webb et al. 1992). The validity of ICPC as a tool with which to classify the data has also been investigated in earlier work (Britt 1997b).

However, the question of the extent to which the GP recorded data is a reliable and valid reflection of the content of the encounter must also be considered.

In many primary care consultations a clear pathophysiological diagnosis is not reached. Bentsen (1976) and Barsky (1981) suggest that a firm and clear diagnosis is not apparent in about half of general practitioners' consultations while others suggest the proportion may be even greater (Morrell et al. 1971). Further, studies of general ambulatory medical practice have shown that a large number of patients presenting to a primary care practitioner are without a serious physical disorder (Anderson 1980; Marsland et al. 1980). As a result it is often necessary for a practitioner to record a problem in terms of symptoms, signs, patient concerns, or the service which is requested, such as immunisation. For this reason this report refers to patient problems (and even 'problem' is not an ideal word) rather than diagnoses.

A number of studies have demonstrated wide variance in the way a GP perceives the patient's reasons for encounter and the manner in which s/he describes the problem under management. In a direct observational study of consultations via a one way mirror Bentsen demonstrated differences in the way practitioners labelled problems and suggested that clinical experience may be an important influence on the identification of problems within the consultation (Bentsen 1976). Two other factors that might affect GPs' descriptions of patient reasons for encounter have been identified: while individuals may select the same stimuli, some label each stimulus separately while others cluster them under one label; individuals differ in the number of stimuli they select (selective perception) (Bensing 1983).

The extent to which therapeutic decisions may influence the diagnostic label selected has also been discussed. Howie (1972) and Anderson (1980) argue that while it is assumed that the diagnostic process utilised in general practice is one of symptom → diagnosis → management, the therapeutic method may well be selected on the basis of the symptom,

and the diagnostic label chosen last. They suggest that the selection of the diagnostic label is therefore influenced by the management decision already made.

Anderson has also pointed out that the therapeutic decision may be influenced by fashion and in turn this affects the selection of the problem label. He gives the example of a rise in the occurrence of neurotic depression in parallel with a decrease in the use of menopause as a diagnosis in the United Kingdom, and suggests this may be the result of a change in the preferred treatment from oestrogen therapy to anti-depressants(Anderson 1980). This should be remembered when considering the results of Chapter 13 of this report which describes some changes in general practice over the past eight years.

Alderson contends that to many practitioners 'diagnostic accuracy is only important to the extent that it will assist them in helping the patient'. He further suggests that if major symptoms are readily treatable some practitioners may feel no need to define the problem in diagnostic terms (Alderson 1988). Crombie stated that in the second and third national morbidity surveys in the United Kingdom there was 'enormous variability in the rates at which doctors perceive and record illnesses'. He concluded that the probable cause arose from the different ways in which GPs gave priority in their perceptions and recording of certain morbidities while discounting or ignoring others. He was unable to account statistically for this variation by the effect of geographic, age, sex, or class differences in the practice populations (Crombie 1990). Differences in the way male and female GPs label problems have also been shown to be independent of such influences (Britt et al. 1996b).

These problems are inherent in the nature of general practice. Knottnerus argues that the GP is confronted with a fundamentally different pattern of problems from the specialist, the GP often having to draw up general diagnostic hypotheses related to probability, severity and consequences (Knottnerus 1991). Anderson suggests that morbidity statistics from family practice should therefore be seen as 'a reflection of the physician's diagnostic opinions about the problems that patients bring to them rather than an unarguable statement of the problems managed' (Anderso, 1980). In any case, doctors base their actions on problems as they perceive them.

While these findings regarding limitations in the reliability and validity of practitioner-recorded morbidity should be borne in mind, they apply equally to data drawn from medical records as to active data collection methods (Britt et al. 1996c;Gehlbach 1979). There is as yet no more reliable method of gaining detailed data about morbidity and its management in general practice. Further, irrespective of the differences between individual GPs in their labelling of the problems, morbidity data collected by GPs in active data collection methods have been shown to provide a reliable overview of the morbidity managed in general practice (Britt et al. 1998).