2 Methods

In summary:

- each year BEACH involves a random sample of approximately 1,000 GPs
- each GP records details about 100 doctor–patient encounters of all types
- the GP sample is a rolling (ever-changing) sample
- approximately 20 GPs participate each week, 50 weeks a year
- each GP can be selected only once per quality assurance triennium
- the encounter information is recorded by the GPs on structured paper encounter forms
- each GP participant also completes a questionnaire about themselves and their practice.

2.1 Sampling methods

- The source population includes all vocationally registered GPs and all general practice registrars who claimed a minimum of 375 general practice A1 Medicare items in the most recently available 3-month Medicare data period (which equates with 1,500 A1 Medicare claims a year). This 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.
- On a quarterly basis the Primary and Ambulatory Care Division of DoHA updates the sample frame from the Medicare records, leaving out of the sample frame any GPs already randomly sampled in the current triennium, and draws a new sample from those currently in the sample frame. This ensures the timely addition of new entries to the profession, and timely exclusion of those GPs who have stopped practising.

2.2 Recruitment methods

The randomly selected GPs are approached by letter posted to the address provided by DoHA.

- Over the following 10 days the telephone numbers generated from the Medicare data are checked using the electronic white and yellow pages. This is necessary because many of the telephone numbers provided from the Medicare data are incorrect.
- The GPs are then telephoned in the order they were approached and, referring to the approach letter, asked whether they will participate.
- This initial telephone contact with the practice often indicates that the selected GP has moved elsewhere, but is still in practice. Where forward address and/or telephone number can be obtained, these GPs are followed up at their new address.
- GPs who agree to participate are set an agreed recording date several weeks ahead.
- A research pack is sent to each participant about 10 days before the planned start date.
- Each GP receives a telephone reminder in the first days of the agreed recording period—this also provides the GP with an opportunity to ask questions about the recording process.

- GPs can use a 'free-call' (1800) number to ring the research team with any questions during their recording period.
- Non-returns are followed up by regular telephone calls for up to 3 months after the set recording time.
- Participating GPs earn Clinical Audit points towards their quality assurance (QA) requirements through the RACGP. As part of this QA process, each 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 provided. In addition, GPs receive some educational material related to the identification and management of patients who smoke or consume alcohol at hazardous levels. Additional points can be earned if the participant chooses to do a follow-up audit of smoking and alcohol consumption among a sample of patients about 6 months later.

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 in Appendix 1. The GP characteristics questionnaire is provided in Appendix 2.

- **Encounter data:** date of consultation, type of consultation (direct/indirect), Medicare/DVA item numbers (where applicable) (up to three) and other payment source (where applicable) (tick boxes).
- **The patient:** date of birth, sex and postcode of residence. Tick boxes are provided for Commonwealth concession card holder, holder of a Repatriation health card (from DVA), non-English-speaking background (patient self-report—a language other than English is the primary language at home), Aboriginal person (self-identification) and Torres Strait Islander (self-identification). Space is provided for up to three patient reasons for encounter (RFEs).
- **The problems managed** at encounter (at least one and up to four). Tick boxes are provided to denote the status of each problem as new or continuing for the patient (if applicable).
- **Management of** each problem, including:
    - **medications** prescribed, supplied by the GP and advised for over-the-counter purchase including brand name, form (where required), strength, regimen, status (if new or continuing medication for this problem for this patient) and number of repeats
    - **other treatments** provided for each problem including counselling, advice and education, and procedures undertaken; and if other treatment was provided by practice nurse (tick box)
    - **new referrals** to medical specialists, allied health professionals and hospital
    - **investigations** including pathology tests, imaging and other investigations ordered at the encounter.
- **GP characteristics:** age and sex, years in general practice, number of GP sessions worked per week, number of GPs working in the practice, postcode of major practice

address, country of graduation, postgraduate general practice training and FRACGP status, after-hours care arrangements, use of computers in the practice, whether the practice is accredited, whether it is a teaching practice, work undertaken in other clinical settings, hours worked in direct patient care and hours on call per week.

2.4 Changes to data elements and reporting methods

For the first 7 years of the BEACH program (1998–99 to 2004–05), where a Medicare item number was claimable for the encounter the GP was instructed to record only one item number. Where multiple item numbers (for example, an A1 item such as 'standard surgery consultation' and a procedural item number) were claimable for an encounter the GP was instructed to record the lower of these (usually an A1 item number). For reporting purposes Medicare-claimable encounters were broken down according to the item number recorded by the GP as claimable (either through Medicare or through DVA) for the encounter.

In November 2004 four new item numbers were added to Medicare[21] to cover some selected activities conducted by a practice nurse on behalf of a medical practitioner. A nurse may see the patient in conjunction with the GP–patient consultations. In this case both the GP's professional service and the practice nurse item are claimable.

The introduction of the Medicare practice nurse items provided the research team with a challenge. In the past 'general practice activity' has been described in terms of GP–patient encounters and this was considered close to equivalent to 'general practitioner activity'. However, the introduction of the practice nurse item numbers meant that, if practice nurse activity associated with the GP–patient encounter was not included, the content of the consultation was no longer fully described.

Therefore, two changes were made to the BEACH form from 2005–06 onwards in order to capture practice nurse activity associated with the GP–patient consultations and include this activity to describe 'general practice activity in Australia':

- GPs could record multiple (up to three) Medicare item numbers.
- In the 'other treatments' section, for each problem managed, the GP was asked to tick the practice nurse box if the treatment recorded was provided by the practice nurse rather than by the GP. If the box was not ticked, the research team assumed that the GP gave the treatment.

Reporting of item numbers

In reporting about the encounters in Chapter 5, Table 5.3 and Table 5.4 count only one item number per Medicare/DVA-claimable encounter for comparability with previous years. Selection of one item number was undertaken on a priority basis: consultation item numbers override incentive item numbers, which override procedural item numbers, which override other Medicare item numbers. These results have been used when reporting changes over time. An additional table (Table 5.5) provides a breakdown of all item numbers recorded by the GPs. Chapter 13 gives a more specific description for each of the practice nurse Medicare item numbers recorded.

Reporting of other treatments

In the section on 'other treatments' in the annual results (Section 10.1), all recorded clinical and procedural treatments are included, irrespective of whether they were provided by the GP or by the practice nurse. These results are also used in the measurement of changes over time (Section 10.2).

Reporting of practice nurse activity

Chapter 13 provides a breakdown of the practice nurse Medicare items claimed, the morbidity managed with the assistance of the practice nurse, and the 'other treatments' provided by the practice nurse as recorded by the GP participants.

When viewing these results, remember that these 'practice nurse' data do not include activities undertaken by the practice nurse during the GP's BEACH recording period that were performed outside the recorded encounter. These could include Medicare-claimable activities (for example immunisations/vaccinations) provided under instruction from the GP but not at the time of the encounter recorded in BEACH, or provision of other activities not currently claimable from Medicare (for example dietary advice on a one-to-one basis, or in a group situation).

2.5 Supplementary Analysis of Nominated Data

A section at the bottom of each recording form investigates aspects of patient health or health care delivery in general practice not covered by the consultation-based data. These additional substudies are referred to as SAND, Supplementary Analysis of Nominated Data.

- The year-long data period is divided into 10 blocks, each of 5 weeks with three substudies per block. The research team aims to include data from about 100 GPs in each block.
- Each GP's pack of 100 forms is made up of 40 forms that ask for the start and finish times of the encounter and include questions about patient risk factors: patient height and weight (used to calculate body mass index, BMI), alcohol intake and smoking status (patient self-report). The methods and results of topics in the SAND substudies for alcohol consumption, smoking status and BMI are reported in Chapter 15. The start and finish times collected on these encounters is used to calculate the length of consultation. The length of consultation for Medicare-claimable encounters is reported in Section 5.1.
- The remaining 60 forms in each pack are divided into two blocks of 30. Different questions are asked of the patient in each block and these vary throughout the year.
- The order of SAND sections is rotated in the GP recording pack, so that 40 patient risk factor forms may appear first, second or third in the pad. Rotation of ordering ensures there was no order effect on the quality of the information collected.

Abstracts of results and research tools from the SAND substudies conducted in BEACH have recently been published in *Patient-based substudies from BEACH: abstracts and research tools 1999–2006,* available through the FMRC's website <www.fmrc.org.au/publications/> or the AIHW's website <www.aihw.gov.au/publications>. Abstracts and research tools for substudies conducted in 2006–07 that were not included in that report are in Chapter 16 of this report.

2.6 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
- RFEs have only an indirect relationship with problems managed as a patient may describe one RFE (such as 'repeat prescriptions') that is related to multiple problems managed, or several RFEs (such as 'runny nose' and 'cough') that relate to a single problem (URTI) managed at the encounter.
- all types of management are directly related to the problem being treated.

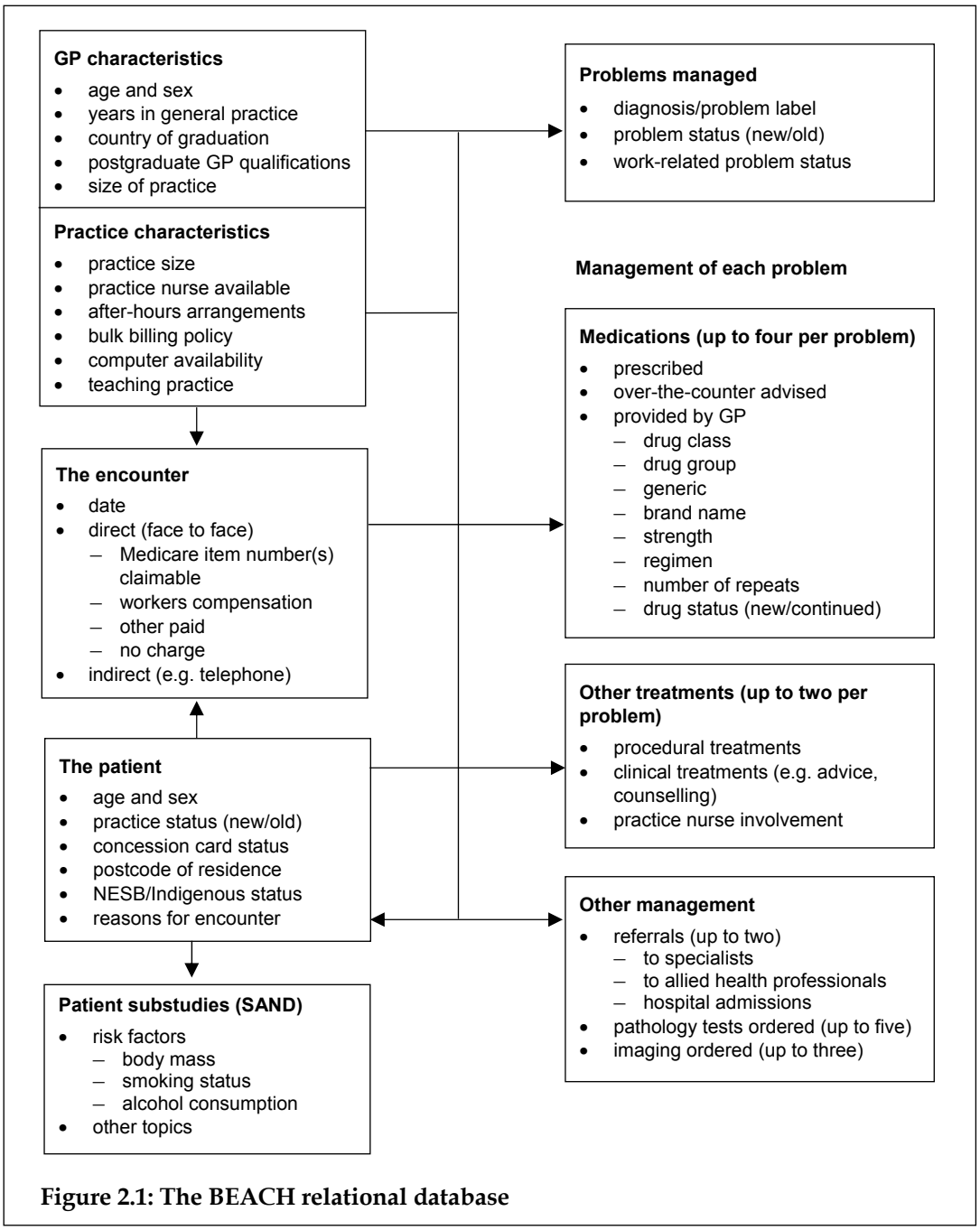

Figure 2.1: The BEACH relational database

Note: NESB—non-English-speaking background.

2.7 Statistical methods

The analysis of the 2006–07 BEACH data was conducted with SAS version 9.1[22] and the encounter is the primary unit of inference. Proportions (%) are used only when describing the distribution of an event that can arise only once at a consultation (for example age, sex) or to describe the distribution of events within a class of events (for example problem *A* as a percentage of total problems). Rates per 100 encounters are used when an event can occur more than once at the consultation (for example 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 results present the number of observations *(n)*, the rate per 100 encounters and the 95% confidence interval.

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 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.[23] 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 simple random sample 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.[24]

Changes over time

SAS version 9.1[22] was used for all analysis of 2006–07 data (as was the case in 2005–06). All data from previous years (1999–00 to 2004–05) were originally analysed using SAS version 6.12[25] (with additional programming to adjust for the cluster sample study design). This year the research team re-calculated all previous data originally analysed with SAS V6.12, using SAS V9.1. This has resulted in slightly tighter confidence intervals and minor variations in point estimates (of up to 0.1) when compared with the data published in earlier annual reports for the 1998–2004 data years.

In measuring changes over time, the research team compared the 2006–07 results with those from 1998–99 wherever possible. However, as in any long-term research program, changes occur over the years. For example, in response to requests from the DoHA (then the Department of Health and Aged Care), more detailed coding systems for pharmaceuticals, pathology and imaging test orders were developed, and these were applied from year 3 (2000–01) onwards. Where this has occurred, change was measured from 2000–01 because earlier years are not comparable.

Where the BEACH 2006–07 results demonstrate a significant change over time, the team calculated the estimated national change across total GP Medicare services from 1998–99 (or where appropriate 2000–01) to 2006–07.

Some concepts have been grouped for comparability over the 9 years of the study. Where concepts have been grouped the change has been footnoted in the table. Due to this grouping some figures may be different from those previously published.

Extrapolated national estimates

In past years BEACH estimates have been extrapolated to the total number of unreferred general practice attendances in Australia (that is, A1 and A2 items combined) as reported by Medicare. However, most of the more recent additions to Medicare item numbers claimable through general practice are not classed as A1 or A2. Therefore an increasing proportion of general practice Medicare claims were not being counted when the extrapolation was limited to A1 and A2 items of service. Table 2.1 demonstrates the proportion of total GP Medicare claims that are accounted for in these other types of general practice Medicare items. Please refer to Section 2.10 for discussion of limitations regarding extrapolations.

Table 2.1 provides the breakdown of Medicare groups that were used to calculate the total GP Medicare item claims. These data were drawn from Medicare Benefits Schedule (MBS) statistics reports[26] and an estimate (based on BEACH data) was applied to the Antenatal attendance Medicare claims, of the proportion likely to be claimed by GPs.

The total GP Medicare claims rounded to the nearest 100,000 were used to calculate the extrapolations in each of the changes over time sections. The numbers used to extrapolate were 103,500,000 for 1998–99, 101,200,000 for 2000–01 (where applicable), and 102,800,000 for 2006–07.

- The national estimates were calculated by multiplying the encounter rate for 1998–99 (or 2000–01 where appropriate) by the estimated total number of general practice services claimed through Medicare in that year (see Table 2.1) to give the estimated annual number of events in 1998–99 (or 2000–01). The same was done for 2006–07. The difference between the two estimates (rounded to the nearest 10,000) gives the estimated national change in the rate of encounters for that event.
- This is expressed as the estimated increase or decrease over the study period (between 1998–99 or 2000–01 and 2006–07), in the number of general practice contacts for that event (for example an increase or decrease in the number of contacts where a problem was managed or management provided) occurring in Australia.

Table 2.1: Number of general practice Medicare items claimed in Australia, 1998–99, 2000–01 and 2006–07

Medicare group descriptor		1998–99	2000–01	2006–07
A1	General practice attendances	90,800,767	89,814,608	90,678,610
A2	Other non-referred attendances	11,180,126	9,972,657	4,283,879
Total A1+A2 items claimed		**101,980,893**	**99,787,265**	**94,962,489**
A5	Prolonged attendances	8,311	9,169	9,581
A6	Group therapy	27,040	24,894	16,890
A7	Acupuncture	901,414	736,691	592,291
A14	Health assessments	0	97,513	376,107
A15	(Subgroup 1 only): GP management plans, team care, care plans and case conferences	0	29,783	1,596,717
A17	Domiciliary medication management review	0	0	54,555
A18	GP attendance practice incentive payments (PIP)	0	0	252,275
A19	Other non-referred attendances associated with practice incentive payments (PIP)	0	0	4,907

(continued)

Table 2.1 (continued): Number of general practice Medicare items claimed in Australia, 1998–99, 2000–01 and 2006–07

Medicare group descriptor		1998–99	2000–01	2006–07
A20	GP mental health care	0	0	338,078
A22	GP after-hours	0	0	4,056,368
A23	Non-referred after-hours	0	0	261,977
A27	Pregnancy support counselling	0	0	700
T4	(Item 16500 only): Antenatal attendance (estimated GP portion)	540,000	483,000	244,000
Total other items claimed (estimated GP claims)		**1,476,765**	**1,381,050**	**7,794,865**
Per cent of total item numbers claimed		**1.4%**	**1.4%**	**7.6%**
Total general practice items claimed (estimate)		**103,457,658**	**101,168,315**	**102,766,935**

2.8 Classification of data

The following data elements are classified according to the International Classification of Primary Care—Version 2 (ICPC-2), a product of the World Organization of Family Doctors (Wonca).[14]

- patient reasons for encounter (RFEs)
- problems managed
- clinical treatments (for example counselling, advice)
- procedural treatments
- referrals
- investigations ordered (including pathology, imaging and other investigations).

The ICPC-2 is used in more than 45 countries as the standard for data classification in primary care. It has recently been accepted by the World Health Organization (WHO) in the WHO Family of Classifications[27] and has been declared the national standard in Australia for reporting of health data from general practice and patient self-reported health information.[28]

The ICPC-2 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 RFEs or problems managed.

Components 2 to 6 cover the process of care and are common throughout all chapters. The processes of care, including referrals, other (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 (for example 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, requires 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 FMRC (then the Family Medicine Research Unit) developed an extended vocabulary of terms classified according to the ICPC, now called ICPC-2 PLUS.[29] This is an interface terminology, developed by the FMRC from all the terms used by GPs in studies such as the Australian Morbidity and Treatment Survey 1990–91,[30] the Morbidity and Therapeutic Index 1992–1998 (a clinical audit tool that was available to GPs) and BEACH 1998–2007, that together have included close to 1.5 million encounter records. These terms are classified according to ICPC-2 to ensure international standards for reporting. Readers interested in seeing how coding works can download the ICPC-2 PLUS Demonstrator at <www.fmrc.org.au/icpc2plus/demonstrator.htm>.

	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—Version 2 (ICPC–2)

Presentation of data classified in ICPC-2

When the free-text data are received from the GPs, trained secondary coders (who are undergraduate health information management students) code the data in more specific terms using ICPC-2 PLUS. Reporting, however, is almost always at the level of the ICPC-2 classification (for example acute otitis media/myringitis—ICPC-2 code H71). However, there are some exceptions where data are grouped either above the ICPC-2 level or across the ICPC-2 level. These grouped codes for morbidity, pathology and imaging data are defined in Appendix 4 and for chronic morbidity in Appendix 5 (see <www.aihw.gov.au/publications/index.cfm/subject/19>).

Reporting morbidity with groups of ICPC-2 codes

- When recording problems managed, the GP may not always be very specific. For example, in recording the management of 'diabetes', they may simply record the problem as 'diabetes'. In ICPC-2, 'Diabetes unspecified' is classified as non-insulin dependent diabetes (code T90). There is another code for insulin dependent diabetes (T89). In some cases the GP may simply have failed to tell us that the patient had 'insulin dependent diabetes'. The research team therefore feels that for national data reporting, it

is more reliable to group the two codes T90 and T89 and label this 'Diabetes—all*'—the asterisk indicating that multiple ICPC-2 codes (as in this example) or ICPC-2 PLUS codes (see below) are included.

Reporting morbidity with groups of ICPC-2 PLUS codes

- In other cases a concept can be classified within (but be only part of) multiple ICPC-2 codes. For example, 'osteoarthritis' is classified in ICPC-2 in multiple broader codes according to site, for example L92—shoulder syndrome (includes bursitis, frozen shoulder, osteoarthritis of shoulder, rotator cuff syndrome). When reporting 'osteoarthritis' in this publication, all the more specific osteoarthritis ICPC-2 PLUS terms are grouped within all the appropriate ICPC-2 codes. This group is labelled 'Osteoarthritis*', the asterisk again indicating multiple codes, but in this case they are PLUS codes rather than ICPC-2 codes.

Reporting pathology and imaging test orders

- All the pathology and imaging tested are coded very specifically in ICPC-2 PLUS but the ICPC-2 classifies pathology and imaging tests very broadly (for example a test of cardiac enzymes is classified in K34—Blood test associated with the cardiovascular system; a CT scan of the lumbar spine is classified as L41—Diagnostic radiology/imaging of the musculoskeletal system). In Australia the MBS classifies pathology and imaging tests in groups that are relatively well recognised. The team therefore re-grouped all pathology and imaging ICPC-2 PLUS codes into MBS standard groups. This allows comparison of data between data sources. These groups are marked with an asterisk and included in Appendix 4.

For all grouped morbidity, pathology and imaging codes, a full list of inclusions is provided in Appendix 4 <www.aihw.gov.au/publications/index.cfm/subject/19>.

Classification of pharmaceuticals

Pharmaceuticals that are prescribed, provided by the GP or advised for over-the-counter purchase 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.
- Strength and regimen are independent fields that, when combined with the CAPS code, give an opportunity to derive the prescribed daily dose for any prescribed medication or group of medications.
- CAPS is mapped to the Anatomical Therapeutic Chemical (ATC)[31] classification, which is the Australian standard for classifying medications at the generic level.

The ATC has a hierarchical structure with five levels. For example:

- Level 1: C—Cardiovascular system
- Level 2: C10—Serum lipid reducing agents
- Level 3: C10A—Cholesterol and triglyceride reducers
- Level 4:C10AA—HMG CoA reductase inhibitors
- Level 5: C10AA01—Simvastatin (the generic drug).

Use of the medication classifications in reporting

For pharmaceutical data there is the choice of reporting in terms of the CAPS coding scheme or the ATC. They each have advantages in different circumstances.

In the CAPS system, a new drug enters at the product and generic level, and is immediately allocated a generic code. Therefore, the CAPS classification uses a bottom–up approach.

In the ATC, a new generic may initially enter the classification at any level (1 to 5), not necessarily always at the generic level. Reclassification to lower ATC levels may occur later. Therefore, the ATC uses a top–down approach.

When analysing medications across time, a generic medication that is initially classified to a higher ATC level will not be identifiable in that data period and may result in under-enumeration of that drug during earlier data collection periods.

- When reporting the 2006–07 annual results for pharmaceutical data, the CAPS database is used in tables of the 'most frequent medications' (tables 9.2 to 9.4 inclusive).
- When reporting the annual results for pharmaceuticals in terms of the ATC hierarchy (Table 9.1), ATC Levels 1, 3, and 5 were used. The reader should be aware that the results reported at the generic level (Level 5) may differ slightly from those reported in the 'most frequent medication' tables for the reasons described above.
- In measuring changes in medications over time (in Section 9.2), the team chose to report at Level 2 of the ATC (which is more stable over time than Level 3), and in CAPS for the generic-level drugs.

2.9 Quality assurance

All morbidity and therapeutic data elements were secondarily coded by staff entering key words or word fragments and selecting the required term or label from a pick list. This was then automatically coded and classified by the computer. A QA 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 logical data checks are conducted through SAS on a regular basis.

2.10 Methodological issues

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 analysis and reporting of recorded data are described in Section 2.7. 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 GPs.[32] Other studies have reported the degree to which GP-reported patient RFEs and problems managed accurately reflect those recalled by the patient[33] and the reliability of secondary coding of RFEs[34] and

problems managed.[30] The validity of ICPC as a tool with which to classify the data has also been investigated in earlier work.[35]

However, the question of the extent to which the GP-recorded data are 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[36] and Barsky[37] suggest that a firm and clear diagnosis is not apparent in about half of GPs' consultations, and others suggest the proportion may be even greater.[38] 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.[39,40,40] As a result, it is often necessary for a practitioner to record a problem in terms of symptoms, signs, patient concerns, or the service that is requested, such as immunisation. For this reason, this report refers to patient 'problems' rather than 'diagnoses'.

A number of studies have demonstrated wide variance in the way a GP perceives the patient's RFE and the manner in which the GP 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.[36] Two other factors that might affect GPs' descriptions of patient RFEs have been identified: while individuals may select the same stimuli, some label each stimulus separately whereas others cluster them under one label; individuals differ in the number of stimuli they select (selective perception).[41]

The extent to which therapeutic decisions may influence the diagnostic label selected has also been discussed. Howie[42] and Anderson[39] 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 antidepressants.[39] This should be remembered when considering the changes in general practice described in this report.

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.[43] 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 geography, age, sex or class differences in the practice populations.[44] Differences in the way male and female GPs label problems also appear to be independent of such influences.[45]

These problems are inherent in the nature of general practice. Knotterus 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.[46] 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'.[39] 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, whether paper or electronic, as they do to active data collection methods.[47,48,48] 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.[49]

Cluster sampling

The statistical techniques applied in BEACH recognise that the sampling is based on GPs and that for each GP there is a cluster of encounters. Each cluster may have its own characteristics, being influenced by the characteristics of the GP. Although ideally the sample should be a random sample of GP–patient encounters, such a sampling method is impractical in the Australian health care system. The reader should, however, be aware that the larger the GP sample and the smaller the cluster, the better. The sample size of 100,000 encounters from a random sample of 1,000 GPs has been demonstrated to be the most suitable balance between cost and statistical power and validity.[50] The cluster effect is dealt with through SAS version 9.1 (see Section 2.7).

GP participation

How many individual GPs have participated in BEACH to date?

Over the 9 years of the BEACH program, 892,300 encounters have been recorded by 8,923 GPs. Since GPs may be sampled from the Medicare data once in each QA triennium, the research team are often asked about the extent to which GPs have participated more than once over the 9 years.

The team investigated the extent of 'double ups' and found that the 8,923 participants in the first 9 years of BEACH represented 6,949 individuals. This means that by March 2007 about 40% of GPs and registrars (approximately 17,500 in any one year) who qualify for inclusion in the original sample frame (for definition see Section 2.1) have participated to date.

Response rates

The response rate of GPs in the ninth year of BEACH was 22.9% of those who could be contacted.

Response rates have fluctuated over the 9 years of BEACH, being highest in the first year (1998–99) at 39.1%. Fluctuations appear related to the QA cycle. In each QA triennium the best response rate occurs in the first year, followed by the third year, and last is the second year of the triennium. GPs are keen to earn their points in the first year, and some are keen to 'catch up' needed points in the third year. In the middle year it seems there is far less interest. It will be interesting to see if the response rate picks up for the 2007–08 data year as it is the third year of the cycle.

Nevertheless, even with a response rate of 22.9%, after post-stratification weights for GP activity and the under-representation of young GPs were applied to the raw data, the age–sex distribution of the patients at encounters this year again demonstrated excellent precision when compared with Medicare data (see Chapter 3). Those concerned about the BEACH response rate should remember that commercial data sources in Australia fail to publicly report response rates, so comparisons cannot be made.

How many GPs can be contacted?

In recent years the research team has expressed increasing concern over the (in)accuracy of the contact details provided by Medicare Australia for sampled GPs. About 15–20% of addresses provided are no longer current and approximately 90% of telephone numbers are incorrect when the sample is received. A considerable amount of time is invested by the recruitment team in locating practitioners. This is not always successful as GPs do not usually have a work telephone number in their own name. In spite of these inaccuracies the recruitment team has, in all previous years, established contact with a minimum of 90% of the GPs for whom details were provided in our Medicare sample. This year the team managed to contact only 88.7%. The proportion of all sampled GPs who were found to have retired, died or moved to an untraceable location was 7.6% this year. As the aim is to represent active, practising GPs, the exclusion of these GPs from the denominator when calculating response rates is a valid and necessary action.

What about the young GPs?

In all years except 2004–05, GPs aged less than 35 years have been under-represented. This under-representation is corrected in the final BEACH data set each year using post-stratification weighting.

For 2006–07, the team investigated the proportion of these young GPs who were not traceable when contacted at the practice address provided from Medicare Australia records by DoHA. We found that 27.0% of those drawn in the sample could not be traced, for they had left the practice to move on through their training. This compares with a non-contactable rate of 9.8% for GPs aged 35 years or more. The team believes that this has a significant impact on the chances of successfully recruiting GPs in this youngest age group. The only way to overcome this problem is to ensure that registrars leave a forwarding address at all practices during training.

It would seem, therefore, that the reason for the under-representation of young GPs in BEACH is that they move through the training program and are no longer contactable by the time they are randomly selected and we attempt to recruit them to the program. Any national general practice study relying on samples being drawn from Medicare data for recognised GPs and registrars would be faced with similar problems. All such studies should check the final participating sample against the sample frame and use post-stratification weighting to adjust for any under-representation of this age group.

Limitations of extrapolations

National estimates

The extrapolation to total estimated encounters occurring nationally in any one year is only an estimate. It is likely to provide:

- an underestimate of the true 'GP workload' of a condition/treatment because the extrapolations are made to GP Medicare items claimed, not to the total number of GP encounters per year (which include indirect encounters and those paid by other sources than Medicare, including DVA, state governments, work cover, employers)
- an overestimate of the management rate of a group of conditions (for example 'cardiovascular disease') because there is a chance that more than one problem of this type will be managed at a single encounter. In the extrapolations two cardiovascular problems managed at the same encounter will be counted as two encounters.

Further, the base numbers used in the extrapolations are rounded to the nearest 100,000 and the extrapolations are rounded to the nearest 10,000.

However, these are used to measure differences between 1998–99 (or 2000–01) and 2006–07 and these limitations apply equally in all years. The extrapolation therefore still provides an indication of the size of the effect of measured change nationally.

Using SAND to estimate prevalence of disease in the attending population

Many SAND substudies ask an opening question to ascertain if the patient present at the encounter has a named condition or to measure the prevalence of a number of diseases among the respondents. Using a qualified medical practitioner to record morbidity in conjunction with patient self-report may provide a more accurate classification of patients' major health problems than self-report alone.[51,52] In the SAND substudies, the patient rather than the content of the encounter is the subject of interest. This overcomes the problem of trying to estimate prevalence of disease among the attending patients, where the disease of interest was not managed at the encounter.

However, in the SAND substudies patients who attend more often have a greater chance of being sampled than those who attend less frequently, so these raw results cannot be used to estimate prevalence of a disease in the total population of attenders. Further, up to 20% of the population currently do not visit a GP in 1 year[3] and these non-attenders cannot be sampled in SAND.

It can be stated that, based on SAND prevalence estimates, a GP would see, on average, 'x number' of patients who have this morbidity in any average GP working week, regardless of whether the GP manages that morbidity at that time.

Further, SAND prevalence estimates of morbidities covered in the National Health Priority Areas have recently been used in combination with age–sex-specific attendance rates (from Medicare statistics) to gain estimates of the prevalence of selected morbidity in the general practice patient population. It was assumed that the 20% of the population that did not visit a GP that year (and therefore had no chance to be selected) did not have the disease in question under ongoing medical management, and extrapolated to estimates of total population prevalence. This method provided prevalence estimates that are somewhat higher than those from the National Health Survey, which relies on self-report of a random sample of the population for diabetes, depression, anxiety, hypertension, hyperlipidaemia and ischaemic heart disease.[53]

2.11 Other BEACH applications

The BEACH methodology can be applied in various health settings. In the past the AGPSCC has used the methodology to conduct a variety of studies in collaboration with other organisations. Examples of past studies are described below.

In 2004 a study was conducted in collaboration with Monash University and the Victorian Metropolitan Alliance. The BEACH methods were used to measure the experience gained by GP registrars during each stage of their training. The results will help to better define the areas in which registrars should receive training and identify areas in which they are not gaining experience.

Another registrar study was conducted in 2003 as a consultancy for North Coast GP Training Ltd and the Institute of General Practice Education. This study looked at the clinical activities of registrars compared with those of their supervisors, to assess their education program in terms of actual practice.

A study in the Victoria Community Health Centres was done in 2004 in collaboration with the Victorian Department of Human Services. The project aimed to provide information about the clinical role of Community Health Service GPs and the characteristics of the patients they see, and how these may differ from the 'average' GP in Australia. The department will use the results to help them plan future health services.

From 2002–04 the BEACH methods were used in the Alternative Pathway Program to assess the educational needs of each GP enrolled in the program. The Alternative Pathway Program was conducted by the National Consortium for Education in Primary Medical Care. The results for each GP were used in identifying specific educational needs and in planning an educational program for the individual practitioner.

In 2002–03 the AGPSCC conducted a longitudinal, matched, controlled trial of active computerised data collection compared with paper-based data collection in the western, north-western and south-western areas of Sydney. Software was developed that reflected the data elements collected in BEACH; the software did not interact with any clinical system being used by GPs. This study demonstrated that active GP computerised data collection in structured, stand-alone software does not provide a reliable and valid measure of GP activity and could not be adopted at this stage as an acceptable alternative to the paper-based data collection methods currently being used.

Due to the fact that BEACH collects data nationally it is possible to analyse data at a level specific to local areas. For example, reports have been published comparing general practice in the different states and territories of Australia and investigating the differences between metropolitan and rural general practice. The research team is also developing Statistical Evaluation Areas (SEAs) that are aimed to provide localised data for divisions of general practice.

Studies have been conducted for the Townsville and Inner South East Melbourne divisions of general practice. These studies were conducted in 1999 and involved oversampling the GPs in each division to provide sufficient samples for statistical analysis of general practice activity within the divisions.

A study investigating changes over time in Victorian general practice was conducted in 1998. The Victorian Morbidity and Treatment Survey used the same methodology as BEACH to measure changes in general practice activity from 1990–91 to 1998. The 1990–91 data were collected in the Australian Morbidity and Treatment Survey (AMTS).[30]