

5. 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 (everchanging) sample.
- approximately 20 GPs participate each week, 50 weeks a year.
- each GP can be selected only once per quality assurance triennium.
- the information is recorded by the GPs on structured encounter forms (on paper).

5.1 Sampling methods

- The source population includes all medical practitioners who claimed a minimum of 375 general practice A1 Medicare items in the most recently available 3-month HIC 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 Care Division of the AGDHA updates the sample frame from the HIC 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.

5.2 Recruitment methods

We approach the randomly selected GPs by letter, posted to the address provided by the AGDHA.

- During the following 10 days we use the electronic white and yellow pages to check the telephone numbers generated from the HIC data. This is necessary because many of the telephone numbers provided from the HIC data are incorrect.
- We then telephone the GPs in the order they were approached and, referring to the approach letter, ask if they will participate.
- On initial telephone contact with the practice we often find 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 a recording date approximately 3 to 4 weeks ahead.
- We send a research pack to each participant about 10 days before the planned start date.
- We make a telephone reminder to each GP in the first days of the agreed recording period – this also provides the GP with an opportunity to ask any questions they have about the recording process.

- We follow up non-returns by regular telephone calls for up to three months after the set recording time.
- Participating GPs earn up to 60 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 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.

5.3 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/Veterans' Affairs item number (where applicable) and other payment source (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 the Australian Department of Veterans' Affairs, DVA), non-English-speaking background (NESB) (patient self-report – a language other than English is the primary language at home), an 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 medications 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
 - 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.

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 data.

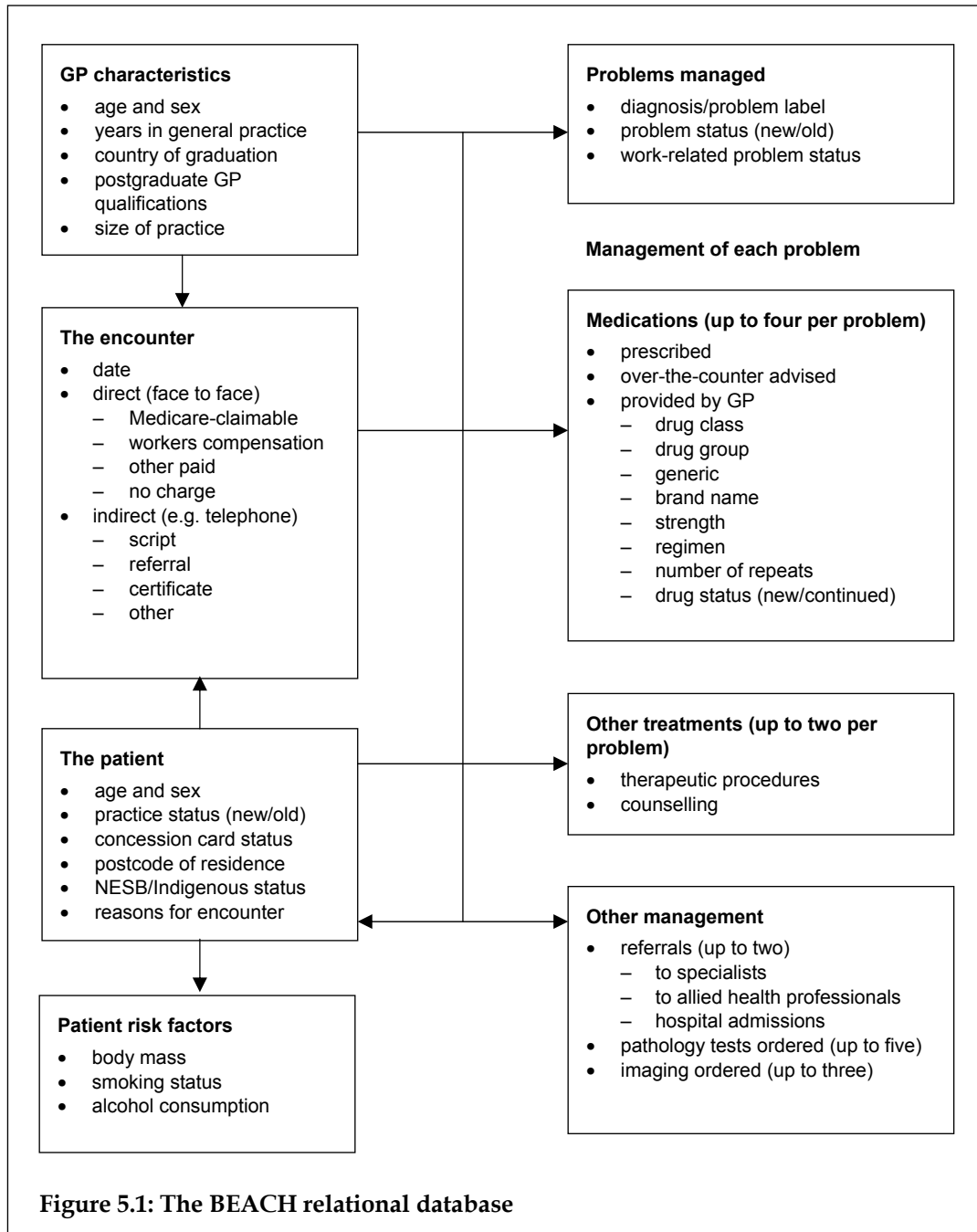
- The year-long data period is divided into 10 blocks, each of 5 weeks. Each block includes data from 100 GPs.
- Each GP's pack of 100 forms is made up of 40 forms that contain questions about patient risk factors: patient height and weight (used to calculate body mass index, BMI), alcohol intake and smoking status (patient self-report).
- The 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 in the GP recording pack is rotated, so that the 40 patient risk factor forms may appear first, second or third in the pack. Rotation of ordering of the components ensures there was no order effect on the quality of the information collected.

The results of topics in the SAND substudies for alcohol consumption, smoking status and BMI are included in this report. Abstracts of results for other substudies are available through the website of the Family Medicine Research Centre (of which the AGPSCC is a part), at www.fmrc.org.au/publications/SAND_abstracts.htm.

5.4 The BEACH relational database

The BEACH relational database is described diagrammatically in Figure 5.1. Note that:

- all variables can be directly related to GP, patient characteristics and to the encounter
- RFEs have only an indirect relationship with problems managed
- all types of management are directly related to the problem being treated.



5.5 Statistical methods

The analysis of the BEACH database is conducted with SAS versions 6.12⁴⁴ and 8.2²¹ 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 (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 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 results present the number of observations (n), the 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.⁴⁵ 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 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.⁴⁶ For comparability with previous years, we have continued to use SAS 6.12 for the tables in Chapter 2—Summary of changes from 1998–99 to 2004–05 (red margin), Chapter 4—Annual results BEACH 2004–05 (blue margin) and in Appendix 5 (comparative data from each of five years 2000–01 to 2004–05, available from <www.aihw.gov.au/publications/index.cfm>).

Changes over time

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 formulae. SAS version 8.2²¹ now includes procedures that calculate the robust standard error to adjust for the intra-cluster correlation of the cluster sample.

In Chapter 3—Selected topics—changes over time (red margin), we used SAS V8.2 regression procedures that adjust the standard error to allow for the design effect of the cluster sample.²¹ The adjusted standard error gives more conservative tests of statistical significance than would be calculated if the data were analysed using simple random sample methods.

Changes over time in medications prescribed/supplied or advised were examined for specific problems of interest. Linear regression was performed to detect:

- a change over time in the medication management for the problem of interest, *or*
- whether a change in medication rate was explained by a commensurate change in management rate of the problem(s) for which the medication is prescribed.

Outcomes are expressed as rates per 100 encounters for medications and problems managed. When examining changes in medication rates within specified morbidities (e.g. arthritis), rates are expressed per 100 specified problems. All analyses were weighted for the GP's age, sex and activity level.

Extrapolated national estimates

- Where we detected a significant change over time, we calculated the estimated annual rate of change.
- We extrapolated the national estimates by multiplying the encounter rate for 1998–99 by the number of unreferred attendances (A1 and A2 items) claimed through Medicare in that year to give the estimated number of encounters for that event in 1998–99. The same

was done for 2004–05. The difference between the two estimates was averaged over six years to give the estimated annual rate of change in encounters.

- This is expressed as the mean annual increase (or decrease) over the study period, in the number of general practice encounters for that problem or medication occurring in Australia each year.

5.6 Classification of data

The patient RFEs, problems managed, other (non-pharmacological) treatments, referrals, and pathology and imaging tests ordered are coded using ICPC-2 PLUS.⁴⁷ 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).⁴⁸

The ICPC 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 Classifications⁴⁹ and has been declared the national standard in Australia for reporting of health data from general practice and patient self-reported health information.⁵⁰

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

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 5.2). 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 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 (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, 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 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 been 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.⁴⁷

Classification of pharmaceuticals

Pharmaceuticals that are prescribed, provided by the GP and those 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 which, when combined with the CAPS code, give an opportunity to derive 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

When reporting pharmaceutical data we have 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 it 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 then 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 2004–05 annual results for pharmaceutical data, we have used the CAPS database in the tables reporting the ‘most frequent medications’ (Tables 4.26 to 4.28 inclusive).
- When reporting the annual results for pharmaceutical in terms of the ATC hierarchy (Table 4.25), we have reported using ATC Levels 1, 3, and 5. 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 described above.
- In measuring changes in medications over time (in Chapter 2–red margin), we have chosen 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.

5.7 Patient risk factor methods

Patient risk factors are investigated for a subsample of patients using the SAND methods (see Section 5.3). The patient risk factors measured include self-reported height and weight (for calculation of body mass index, BMI), alcohol consumption and smoking status.

Body mass index

The BMI for an individual is calculated by dividing weight (kilograms) by height (metres) squared. A person with a BMI less than 20 is considered underweight, 20–24 is normal, 25–29 overweight, and more than 30 is considered to be obese.

The GPs were instructed to ask the patients (or their carer in the case of children):

- What is your height in centimetres?
- What is your weight in kilograms?

Metric conversion tables (feet and inches; stones and pounds) were provided to the GP.

The standard BMI calculation described above is not appropriate in the case of children. Cole et al. have developed a method which calculates the age–sex-specific BMI cut-off levels for overweight and obesity specific to children.⁵¹ This method, based on international data from developed Western cultures, is applicable in the Australian setting.

The BEACH data on BMI are presented separately for adults (aged 18 and over) and children. The standard BMI cut-offs have been applied for the adult population, and the method described by Cole et al. Has been used for defining overweight and obesity in children (aged 2–17 years).⁵¹ There are three categories defined for childhood BMI: underweight/normal, overweight and obese.

Smoking

As part of the current study, the GPs were instructed to ask the patients (18 years and over):

- What best describes your smoking status?
Smoke daily
Occasional smoker
Previous smoker
Never smoked

Respondents were limited to adults aged 18 years and over because there are ethical concerns about approaching the younger patient group to ask for information on smoking and alcohol consumption for survey purposes. In addition, the reliability of this information from patients aged 14–17 years may be compromised if a parent is present at the consultation.

Alcohol consumption

To measure alcohol consumption, BEACH uses three items from the WHO Alcohol Use Disorders Identification Test (AUDIT),⁵² with scoring for an Australian setting.⁵³ Together, these three questions assess 'at-risk' alcohol consumption. The scores for each question range from zero to four. A total (sum of all three questions) score of five or more for males or four or more for females suggests that the person's drinking level is placing him or her at risk.⁵³

GPs were instructed to ask the patient (18 years and over):

- How often do you have a drink containing alcohol?
Never
Monthly or less
Once a week/fortnight
2–3 times a week
4+ times a week
- How many standard drinks do you have on a typical day when you are drinking?

- How often do you have 6 or more standard drinks on one occasion?
Never
Less than monthly
Monthly
Weekly
Daily or almost daily

A standard drinks chart was provided to each GP to help the patient identify the number of standard drinks consumed.

The wording of the responses to the first and third questions was changed from 2001–02 onwards to reflect exactly the AUDIT instrument from which the responses are derived. This update, along with a data entry change enabling more specific entry for the second question, slightly increased the rates of at-risk drinking. The data collected from 2001–02 onwards are a more accurate reflection of the alcohol consumption of general practice patients and these are the years compared in this report.

5.8 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.

5.9 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 adopted to ensure maximum reliability of coding and data entry and the statistical techniques applied have been described above.

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.⁵⁴ Other studies have reported agreement between GP-reported patient RFEs and problems managed and those recalled by the patient,⁵⁵ the reliability of secondary coding of RFEs⁵⁶ and problems managed,⁵⁷ and the validity of ICPC as a tool for classifying the data.⁵⁸

Limitations regarding the reliability and validity of practitioner-recorded morbidity have been discussed elsewhere and should always be borne in mind. However, these apply equally to data drawn from medical records (whether paper-based or electronic) and to active data collection methods.^{59,60} Further, irrespective of the differences between individual GPs in their labelling of 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.⁶¹

5.10 Methodological issues

How many individual GPs have participated in BEACH to date?

Over the first seven years of the BEACH program, 697,400 encounters have been recorded by 6,974 GPs. Since GPs may be sampled from the HIC data once in each QA triennium, we are often asked the extent to which GPs have participated more than once over the seven years.

This year we investigated the extent of 'double ups' and found that the 6,974 participants represented 5,929 individuals. GPs who had participated twice since March 1998 number 970. A further 37 GPs had participated three times. This means that we have to date sampled more than one-third of the GPs (approximately 17,500 in any one year) who have qualified for inclusion in the original sample frame (for definition see Section 5.1).

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.¹² The cluster effect is dealt with through SAS 8.2 (see Section 5.5).

GP participation rates

The response rate of GPs in the seventh year of BEACH was 28.1% of those we could contact—somewhat of an improvement since the previous year when it was 23.7%. The 2004–05 result is comparable with the 28.9% in the fifth BEACH year (2002–03), 32.3% in the fourth year, and the 29.8% in the third year. In the first two years of BEACH, response rates were far higher, at 39.1% in the second year and 38.4% in the first year (1998–99).

But what is the denominator?

One of the difficulties in reliably reporting response rate is the changing size of the denominator. The sample frame includes all non-specialists who have claimed more than 375 A1 Medicare items of service in the previous quarter. This means that the sample frame includes:

- current registrars, who in the past, were not required to undertake QA until the triennium after graduation but who are now required to undertake QA activities in the triennium in which they complete training. The annual intake of registrars to the training program for general practice is now close to 600 per year.
- overseas trained doctors employed in areas of workforce shortage, the number of whom is increasing. It is expected there will be an additional 725 such doctors working in Australia by 2007.⁶² Until 2004 these doctors were not required to do QA but they were counted in the denominator. Now they are required to do QA.

There is no differentiation between recognised GPs and those other medical practitioners who can claim Medicare A1 service items through the MedicarePlus initiatives.⁶² As the pool of overseas trained doctors and other medical practitioners who are paid A1 items of service increases,⁶² the denominator used to calculate the response rate grows. Unfortunately there is no way we can identify the size of this effect.

How many can we contact?

In recent years we have expressed increasing concern over the (in)accuracy of the contact details provided by the HIC 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, and this is not always successful as GPs don't usually have a work telephone number in their own name. In spite of these inaccuracies we have, in all previous years, still established contact with a minimum of 90% of the GPs for whom details were provided in our HIC sample. This year we managed to contact only 85.7%. The proportion of all sampled GPs who were found to have died, moved to an untraceable location, or to have retired doubled from 4.0% in 2003–04 to 8.3% this year. As the aim is to represent active, practising GPs, the exclusion of these GPs from the sample is a valid and necessary action.

What about the young GPs?

In all previous years we have had an under-representation of GPs aged less than 35 years. We corrected for this under-representation in the final BEACH data set each year using post stratification weighting. This year we do not have an under-representation of young GPs. In the past we have hypothesised that the under-representation of young GPs reflected the lack of requirement for GP registrars to undertake QA activities during training or during the QA triennium on completion of training. This hypothesis appears to have been correct – recent changes have meant that the registrars now have to complete QA during the triennium in which they complete their training – and this year was the first since BEACH commenced in which GPs aged less than 35 years were not under-represented in the participating sample.

Electronic BEACH data collection

The BEACH program is currently a paper-based data collection program. Many people have suggested that with the increased GP uptake of electronic prescribing systems or full clinical systems (electronic health records – EHRs), national data could soon be drawn passively, directly from the GPs' computers. Although an attractive proposition, there are still many barriers to its implementation.

- To obtain a national random sample of practising GPs, each GP must have an equal chance of selection. Until all GPs are using EHRs, this would not be the case. Further, with the recognised variance between GPs⁶³ it is likely that those who do not have EHRs differ from those who do. Sampling from only those GPs with EHRs would therefore give a biased national result.
- Many GPs currently use electronic prescribing systems rather than full EHRs, or use their EHRs for prescribing only (see Chapter 4, blue margin). The extent to which data are entered at encounters that do not involve a prescription is not known. Where GPs do not record the problem managed unless a prescription is provided, measurement of changes in prescribing behaviour over time becomes impossible. For example, if GPs significantly decrease the prescribing of antibiotics for URTI, and in parallel record problems only where a medication is prescribed, the recorded rate of antibiotic prescriptions for URTI will either not change or may increase. Further, this report has demonstrated that drug prescription is only one of many management techniques used by GPs. The measurement of GP clinical activity should not be confined to the measurement of prescribing behaviour any more than it should be limited to activities claimed only through the MBS.
- The structure of electronic clinical systems varies, as do the coding and classification systems used in each. Drawing reliable and representative data from electronic clinical systems will require the introduction of a standardised minimum data set and use of standard coding and classification systems in all electronic clinical systems.
- Issues of privacy and confidentiality also need to be resolved.

5.11 Other BEACH applications

The AGPSCC has recently completed the data collection phase of a study measuring the experience gained by GP registrars during each stage of their training. The BEACH methods have been applied in this study which is being conducted in collaboration with Monash University and the Victoria Metropolitan Alliance. The results may help to better define the areas in which registrars should receive training and may identify areas in which they are not gaining experience.

Another parallel BEACH study is being conducted in Victoria Community Health Centres, for the Victoria Department of Human Services. There is currently limited information available 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.