

As it evolves to meet these challenges, the NSW Health Survey Program will provide a mechanism for the NSW Department of Health to meet some of its major information needs of the next decade. It will be used to improve the population's health by providing evidence of program effectiveness, information for policy development and service planning, and facilitating population health research.

2. METHODS

The following section describes the methods used for the 1997 and 1998 NSW Health Surveys. Although this report focuses on these surveys, many of the development issues and survey procedures apply also to the 1999 NSW Older Peoples' Survey and the NSW Child Health Survey. More detailed information about the development and methods of these surveys is available elsewhere.^{12,17}

Figure 1 outlines the key steps in survey development for the 1997 NSW Health Survey. The first step was the selection of mode of survey administration (see Section 2.1). Once this was decided, the NSW Health Survey Program Technical Group and its expert groups, and the NSW Health Survey Program Unit, embarked on five other key steps:

- development of the content and format of the English questionnaire (Section 2.2);
- the non English-speaking (NES) project, to ensure people who could not answer in English could participate in the surveys (Section 2.3);
- establishment of a fieldwork facility (Section 2.4);
- development of fieldwork procedures (Section 2.5);
- development of sampling methods (Section 2.6).

Although they are presented separately in this report, many of the activities involved in these steps were inter-related and their progress was reliant on the others.

2.1 Selecting a survey mode

Three main modes of survey administration are used to collect information on health: self-administered mail questionnaires, personal interview, and telephone interview.

The main advantages of self-administered mail surveys are lower costs per completed interview and possibly superior ability to obtain accurate information about socially undesirable behaviours. However, it is more difficult to control the quality of information collected, and response rates may be lower. The 1995 Welsh Health Survey used this method of delivery.

Personal interview surveys offers a number of advantages including superior ability to collect qualitative data and high response rates. This survey mode offers good control of information quality but is the most expensive in terms of interviewers costs, supervision, travel and organisation. This is the method of administration used by the Australian Bureau of Statistics National Health Surveys.

Telephone surveys have become the method of choice for many health surveys for reasons of cost (estimated to be half that for face-to-face interviews), speed, flexibility and improved quality control.^{18,19} They are particularly suitable for geographically scattered populations or urban areas where interviewer safety is a potential concern. Around 95 per cent of Australian households have a telephone connected. Telephone surveys inevitably exclude some people who rent their dwellings or are born overseas; and those who are homeless, itinerant, institutionalised, or who have disabilities that prevent them from communicating via the telephone.^{20,21} (However, most of the population subgroups excluded from telephone surveys are difficult to capture using any survey techniques.) They are used for the Behavioural Risk Factor Survey in the United States;²² the South Australian Social, Environmental and Risk Context Information System;²³ the West Australian Health Survey;²⁴ and the Queensland Regional Health Surveys.

After consideration of the features of the various administration modes, telephone interview was chosen as the preferred mode of administration for the NSW Health Survey Program.

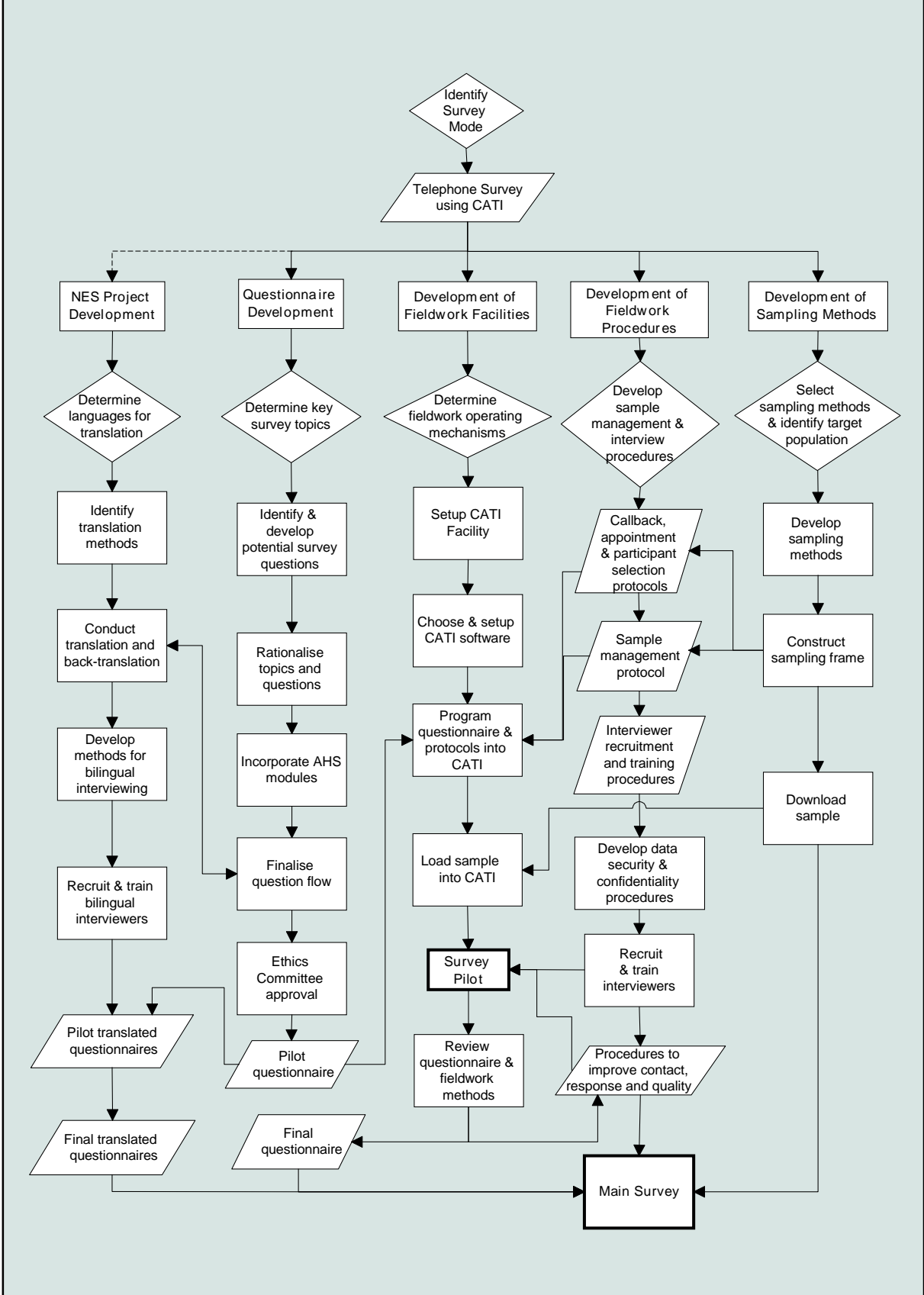
Once telephone administration had been selected, the next decision was whether to use paper copy to administer the survey over the phone, or use a computer-assisted telephone interview (CATI) system. The CATI system is an automated system for collecting information from telephone interviews. The advantages of using a CATI system for the NSW Health Surveys, as opposed to telephone administered hard-copy questionnaires, are that CATI:^{18,19,22}

- allows efficient collection of information directly into the system database;
- provides for accurate data entry with the use of range and logic checks to improve data quality;
- can use questionnaires with very complex branching and skip structures, for example the 17 sets of Area Health Service questions;
- automatically schedules calls and callbacks;
- monitors call attempts and outcomes, and tracks interviewer productivity;
- reduces data entry and interviewing costs.

Although more expensive to set up than paper copy telephone surveys, savings from reduced data entry, interviewing and administration costs in the longer term, and the

Figure 1

Overview of Survey Planning and Development



benefits of improved data quality and flexibility, made CATI an obvious choice for the NSW Health Survey Program.

2.2 Questionnaire development

This section focuses on the development of the questionnaire for the 1997 NSW Health Survey, which took around nine months. In the 1998 Survey, the 1997 questionnaire was slightly modified to allow the incorporation of two new topic areas: cervical cancer screening and oral health. The EQ-5D instrument for measuring health-related quality of life was also added,²⁵ and the injury questions focused on rural injury issues instead of injury incidence.

A review of the questionnaire is currently underway in light of the results from the 1997 and 1998 surveys and current national developments to improve and standardise health survey data.

2.2.1 Identifying key survey topics

Topic areas for the 1997 NSW Health Survey were developed with advice from the NSW Health Survey Program Steering Committee, the Technical Group, and its expert groups. Extensive consultation was also conducted with public health professionals and health service planners in the 17 NSW Area Health Services. The Steering Committee decided that the survey should focus on the national and state health priority areas of coronary heart disease, cancer, injury, mental health, diabetes, and asthma.

2.2.2 Identifying and developing potential survey questions

Once key topic areas were identified, the expert topic groups and the Technical Group decided whether questions should focus on risk behaviours, prevalence, disease screening, management or severity. Advice was also sought from Area Health Service representatives. Most questions were adopted or adapted from existing questionnaires. New questions were developed when no other suitable questions could be found.

In 1997, groups or individuals proposing questions for inclusion in the survey were asked to provide information about the rationale for including the questions, the prevalence of the condition of interest, and the proposed uses of the data. For the 1998 survey, a more detailed rationale was sought (Table 1).

2.2.3 Refining topics and questions

Several draft versions of the questionnaire were circulated to Area Health Service staff for comment about content. The questions, the rationale for their use, and any additional information about their ease of administration, were collated into tabular form.

After considerable deliberations, the NSW Health Survey Program Technical Group made recommendations to the Steering Committee on the topics and questions that should be included in the 1997 survey. It also made recommendations about whether the questions were 'core questions' (and should be administered yearly) or whether they should form parts of 'modules' that would be administered less frequently.

The final survey questionnaire consisted of:

- a set of core questions, including demographic details and key health indicators which would be measured annually;

- topic-specific modules (for example, asthma and diabetes) which could be included every three years to provide more detailed information on a topic;
- seventeen sets of area-specific questions developed by each Area Health Service for delivery to their residents only (see Section 2.2.4 for more information on area-specific questions).

Table 2 outlines the topics included in the 1997 and 1998 Health Surveys.

2.2.4 Incorporating Area Health Service-specific questions

So that the 1997 and 1998 NSW Health Surveys addressed local issues and priorities, each Area Health Service was asked to propose 10 additional questions (approximately three minutes of survey time). An Area-Specific Working Group was established to coordinate this aspect of questionnaire development.

The Chief Executive Officer of each Area Health Service nominated a representative to assist the NSW Health Survey Program by acting as a local point of contact; advising about local issues that might affect sampling, survey timetabling or other methodological concerns; and co-ordinating the development of area-specific questions.

The process of question development varied among Area Health Services. In 1997, members of the Area-Specific Working Group assisted Area Health Services individually with question development. In 1998, the NSW Health Survey Program unit coordinated the development of the area-specific question modules, in collaboration with Area Health Service staff. Some Area Health Services chose to repeat their 1997 questions, while others chose to collect information on newly emerging issues.

The Area Health Services chose to include a broad range of issues over the two years, including:

Table 1

Rationale for inclusion of questions in 1998 NSW Health Survey

1. What is the estimated proportion of the total sample (NSW residents 16+) who will report the behaviour–risk factor–event of interest?
2. Who will use the data and what will it be used for?
3. Will this information be useful at an Area level or should it be asked of a statewide random sample of 4,500? If useful at an Area level, please outline reasons.
4. Does it comply with State–Area information needs?
5. What other groups support the inclusion of these questions?
6. Are there other alternative data sources?
7. Has there been any work carried out on the validity, reliability and responsiveness to change over time of the questions?
8. What other information is required for interpretation of the data?
9. What are the problems or limitations associated with asking these questions?
10. What priority does each question have, if only some of the questions could be included?

- health status and disease prevalence (for example, urinary incontinence, pain, menopause, falls, food-borne illness, and disability);
- risk behaviours (for example, sexual health practices, alcohol consumption, smoking cessation, use of prescribed and illicit drugs, physical activity, nutrition; screening practices, driver fatigue, family history of certain conditions and alcohol and driving);
- beliefs and knowledge about health problems (for example, Q fever vaccination, bowel screening, PAP tests, mental health, mosquito control, chest pain, skin cancer, and blue-green algae);
- attitudes to health issues and policies (for example, perception of personal safety, alcohol and childhood immunisation);
- health service utilisation (for example, attendance at dentists, mammography screening services; emergency services, outpatient services and support services);
- attitudes to and satisfaction with health services (for example, hospital preferences);
- sources of health information (for example, about giving up smoking and healthier diets);
- management of certain health conditions (for example, asthma);

Table 2

NSW Health Survey Topics, 1997 and 1998

Alcohol consumption (1997,1998)

Asthma (1997,1998)

Cancer screening, including mammography (1997,1998), cervical cancer screening (1998),

Colorectal cancer screening (1997,1998)

Cardiovascular risk factors including high blood pressure and high blood cholesterol (1997,1998)

Demographics (1997,1998)

Diabetes and high blood sugar (1997,1998)

General physical health status (1997,1998)

General mental health and wellbeing (1997,1998)

Height and weight (1997,1998)

Health status, one question (1997,1998), EQ-5D (1998)

Health service utilisation, access and satisfaction (1997,1998)

Hysterectomy rate (1997,1998)

Immunisation (1997,1998)

Injury Issues including rural injuries (1998), smoke alarms (1997, 1998)

Nutrition (1997,1998)

Oral health (1998)

Pain (1997)

Physical activity and inactivity (1997,1998)

Smoking behaviour (1997,1998) and smoking policy (1997)

Sun protection behaviour (1997,1998) and shade availability (1997)

- community issues (for example, carer burden, internal migration, work, sense of community and homeboundness).

2.2.5 Question flow

The questionnaires for 1997 and 1998 consisted of over 300 questions. In order to incorporate the Area Health Service-specific questions into the state-wide questionnaire so that it would flow coherently, these questions were assigned to a topic and, wherever possible, are inserted at an appropriate point. Introductory sentences were added to provide a overview of the topic to be discussed; and, where needed, connecting sentences were added to ensure coherence within topic areas. Some topic areas were considered sensitive (for example, sexual behaviour and incontinence). These questions were placed towards the end of the questionnaire, so as not to compromise the completion the survey, and were preceded by preambles emphasising their voluntary nature.

2.2.6 Ethics committee approval

Following the development of the questionnaire, sampling methods, fieldwork procedures, and data security and confidentiality measures, NSW Health Survey Program staff applied for approval to conduct the survey from the NSW Statewide Health Confidentiality and Ethics Committee. Following the incorporation of a number of recommendations about changes to survey content and protocols from the Committee, the survey was approved for administration.

2.2.7 Survey pilots

The 1997 survey was piloted in the Central Sydney and Southern Health Areas in April 1997. A total of 366 interviews were conducted. The pilot aimed to trial the CATI program and equipment; test the questionnaire; test interviewer selection, training and quality assurance procedures; and estimate the number of telephone numbers to be dialled in the main survey for each Health Area.

The 1998 pilot study was conducted with the same aims as the 1997 pilot. However, rather than piloting in one rural and one urban Health Area, 250 interviews were conducted across NSW (approximately 14 interviews per Area).

As a result of the pilot studies, changes were made to CATI programming, the questionnaire, operating procedures for interviewer selection and training, interview scheduling, and hours of operation.

2.3 Non-English Speaking Project

Around one quarter of NSW residents were born overseas and 50 per cent of these were born in non-English speaking (NES) countries.²⁶ A specific NES Project was established as part of the NSW Health Survey Program to improve the representation of people from non-English speaking backgrounds in surveys run under the program.

The main objectives of the NES project were to:

- determine the major language groups in NSW for which a translation and bi-lingual interviewing would be required;
- develop appropriate methods to ensure high quality questionnaire translations that maintained conceptual equivalence with the English language version, and used language suitable for telephone administration which would be understood by the majority of NES people in that language group who reside in NSW;

- ensure that interviews of NES people are conducted in a standardised and professional manner.

2.3.1 Identifying the languages for translation

Sufficient resources were available for translations of the questionnaire into five languages. These were selected on the basis of the number of respondents with poor English language proficiency who would be expected to be selected for interviewing in each language group. Using information from the 1991 ABS Census Profiles,²⁶ the five language groups selected for translation were Arabic, Chinese (Cantonese and Mandarin), Greek, Italian and Vietnamese.

2.3.2 Translation model

A standardised translation model was developed in collaboration with the NSW Multicultural Health Communication Centre, based on prior experience with health translations for the NSW Department of Health. The model included three major steps: pre-translation preparation, translation and verbal back-translation. Translations were conducted by translators accredited by the National Association for the Accreditation of Translators and Interpreters (NAATI). Bilingual interviewers or interpreters were used for the verbal back-translation. This helped to enhance the quality of the translation by ensuring the language was neither too formal nor too colloquial.

2.3.3 Bilingual interviewers

Two bilingual interviewers with appropriate language skills and experience in health and/or community work were selected for each language group and employed on a consultancy basis. All interviewers were NAATI accredited as translators, interpreters or language aides. All bilingual interviewers attended a compulsory intensive four-hour training session prior to interviewing.

2.3.4 Refining questionnaire translations

Bilingual interviewers conducted several practice interviews with members of their translation team. Following these trials, they piloted the questionnaire among a small group of community members with poor English language proficiency. Translators then revised the translations according to recommendations from the bilingual interviewers. This process helped to further identify sections of the questionnaire that might be poorly understood by community members, helped the interviewers to gain familiarity with the hard copy tool, and further ensured consistency of delivery of the questions between interviewers

2.3.5 Bilingual data collection

A standard protocol was used to establish the language spoken in households where NES people were encountered by English-speaking interviewers (see Section 2.5.1.4). Once an English-speaking interviewer had identified a specific NES person for interview, the call was assigned to a bilingual interviewer. Batches of interviews were assigned to the relevant interviewer, and were completed on paper copy. Completed questionnaires were returned to the Health Survey Program Unit along with information on the outcome of calls that were incomplete. Data entry of completed interviews and call outcomes was undertaken by Health Survey Program Unit staff.

2.4 Development of fieldwork facilities

There were three options for the conduct of fieldwork for the NSW Health Survey Program, as follows:

- to outsource the survey fieldwork to a market research firm, which would then be responsible for all aspects of data collection, from the development of the survey to the analysis and reporting;
- to outsource some of the survey fieldwork (such as the survey interviewing) to a market research firm, with the NSW Health Survey Program Unit being responsible for the rest;
- for the NSW Health Survey Program Unit to carry out all fieldwork. This option would mean establishing a facility where telephone interviewing could be conducted 'in-house'.

Based on the experience of members of the Survey Program Team, and discussions with individuals who had used in-house methods and outsourced surveys to market research firms, it was decided that the NSW Health Survey Program would conduct all fieldwork in-house. This maximised control of the major factors influencing the outcomes of the survey administration process, and thus the quality of the resulting data. The establishment of an in-house facility also gave the NSW Health Survey Program Unit flexibility to respond to urgent information needs related to acute public health issues, such as disease outbreaks or disasters.

2.4.1 The NSW Health Survey Program CATI Facility

In late 1996, work began to identify suitable premises for a facility where CATI telephone interviewing would be conducted for the NSW Health Survey Program. Space was found at the NSW Department of Health premises at Surry Hills. These premises were modified to provide secure accommodation for 19 CATI telephone interviewing stations, a room for conducting NES interviews, an office for NSW Health Survey Program Unit staff, a training room and staff amenities. An isolated computer network was set up to ensure a stable platform for the CATI software and enhanced data security.

In early 2000, the CATI facility and Health Survey Program staff moved to join the rest of the Epidemiology and Surveillance Branch in North Sydney. The new premises have improved security and provide better access to the NSW Department of Health computer network and administrative support services.

2.4.2 CATI software

The NSW Health Survey Program has used three CATI software packages since 1997: Sawtooth Ci3 CATI software (DOS version);²⁷ Surveycraft CATI software (DOS version);²⁸ and, most recently, Sawtooth Ci3 CATI software (Windows version).²⁹ Each package offers different advantages in terms of management of the sample, reporting features, costs, and technical support, but none so far meets all the needs of the NSW Health Survey Program.

The NSW Health Survey Program experience indicates that the most important characteristics of CATI software are:

- advanced call management functions to effectively and efficiently schedule and manage appointments, and to minimise missed callbacks;
- ease of managing the sample database;

- on-line help and technical support;
- modifiable and easy to use reporting function;
- ease of data export;
- ability to search the sample data base and view a complete call history;
- ease of learning for interviewers, supervisors and data manipulators.

2.4.3 Programming questionnaires and protocols into CATI

One member of the NSW Health Survey Program Unit was responsible for questionnaire programming, while another did program checking. Programming commenced when the survey instruments were in the final stages of development, leading to a period of intensive work so as not to delay the commencement of fieldwork.

The programmer wrote and simultaneously tested the code and then checked the completed code again against the final questionnaire. Before piloting, survey supervisors tested questionnaire skips and layout on-screen. Once fieldwork procedures were finalised, these were programmed into CATI to automate interviewing, sample management, and quality control. During training and piloting, the interviewing team checked questionnaire skips and layout, question flow and adherence to fieldwork protocols. Following changes made as a result of piloting, a further process of re-testing was undertaken by the programmer, supervisors and interviewing team before the surveys were put into the field.

2.5 Fieldwork procedures

2.5.1 Sample management and interviewing procedures

2.5.1.1 Callback and contact procedures

To maximise opportunities for contacting households and potential respondents, the protocols were developed for the CATI systems used for the 1997 and 1998 NSW Health Surveys to sequence calls and callbacks on different days of the week and at different times of the day and evening, and prioritise appointments to ensure that they were kept and contact was made on time.

In 1997, up to 10 call-backs were made to make contact with each number. Once contact had been made, a further five attempts were undertaken to interview the selected respondent. Analysis of the call data from the 1997 survey showed a very poor conversion rate to an interview or other final disposition after seven attempts to contact. Therefore in 1998, the protocol was changed to seven attempts to call back, and a further four attempts to interview once contact with the household was made.

Appointments were offered at a time suitable for the respondent. If a call was made but the respondent was not home, another appointment time was scheduled. Messages were left on answering machines where possible, but only when previous contact with the household had been made.

2.5.1.2 Dialling out

Once the target sample size (1,000) was achieved for each Health Area, no new numbers (numbers with no prior attempt to contact) for that Area were dialled. The call protocol, including up to 10 callbacks (seven in 1998) to make contact, and up to a further five (four in 1998) to contact the selected respondent, was completed for each number that

had already been dialled at least once. This ‘dialling out’ process ensured that respondents who required many phone calls for contact and/or interview (typically young adults who are rarely at home) were not differentially excluded from the sample.

2.5.1.3 Participant selection

A single respondent aged 16 years or over was randomly selected from each household. Upon contact, the CATI system led interviewers through a series of questions to ascertain the number of people aged 16 years or more living there, and generated a random number ranging between one and that number. Interviewers then asked to speak to the selected person. For example, if there were four eligible household members, the CATI system randomly generated a number between one and four. If the computer generated the number ‘2’, the scripting said ‘May I please speak to the second oldest person in your household’.

Interviewers made up to five call-backs (four in 1998) in order to contact the selected person. They were not permitted to substitute an alternative household member. Using this method, the probability of selection for each respondent was inversely proportional to the number of eligible people in the household. A special callback protocol was used where language difficulties were encountered (see Section 2.5.1.4).

2.5.1.4 Identifying the language spoken in NES households

The language spoken in households was not known prior to the first call attempt. In order to identify a language when a communication problem was experienced, English-language interviewers, prompted by the CATI system, followed the following protocol:

- the interviewer attempted to determine if there was an English speaker in the household. If so, they asked to speak to them;
- if no person proficient in English could be located, the interviewer asked the respondent what language they spoke and, if unable to be understood, what was their country of birth;
- if the interviewer was unable to identify the language on the first call, a survey supervisor or specially trained English language interviewer called the number a second time;
- once the language had been ascertained, the interviewer entered the relevant disposition code into the CATI system;
- if the language was still not identified after the second call, or the household spoke a language for which there was no translation, the details were recorded so the number could be removed from the sample.

Interviews to be conducted in the translated languages were batched and given to the bilingual interviewers for completion, as outlined in Section 2.3.5.

2.5.2 Data security and confidentiality

The NSW Health Survey Code of Conduct and Confidentiality Agreement was developed to address specific data security and confidentiality issues relating to conducting health survey interviews using CATI and paper copy (for non-English interviews), and storage and access to survey data. All interviewers were required to sign the Code of Conduct and Confidentiality Agreement.

The Code of Conduct ensures that the NSW Health Survey Program complies with codes and regulations regarding protecting the privacy of personal health information (including the NSW Health Information Privacy Code of Practice), and current industry codes of behaviour (including the Code of Professional Behaviour adopted by the Market Research Society of Australia, and the Code of Ethics for Interpreters and Translators).

2.5.3 Interviewer recruitment and training

Interviewers for the 1997 and 1998 NSW Health Surveys were recruited as casual staff through advertisements in newspapers and university bulletin boards. The selection procedure comprised assessment of the applicant's curriculum vitae, followed by personal interview conducted by a three-person selection panel. During the selection interview, applicants conducted a trial telephone interview, allowing the panel to assess telephone manner and voice quality. All recommended applicants were subject to a criminal record check before they were employed.

All interviewers were issued with a training manual and attended two, four-hour, training sessions which included both theory and practice interviews. A further two hours of practice sessions were completed prior to commencing interviewing when needed.

Interviewer training covered:

- the background of the NSW Health Survey Program;
- objectives of the survey;
- how to use the CATI system;
- procedures to be followed, including number of call-backs
- methods for improving response rate (for example, making appointments to ring back at a more convenient time);
- ways to identify the language spoken by NES households and to determine whether the selected respondent spoke English well enough to complete the interview in English;
- the intent and meaning of each question;
- how responses should be recorded or coded;
- referral organisations for respondents requiring more information on specific issues;
- what to do if any problems are encountered;
- the NSW Health Survey Code of Conduct and Confidentiality Agreement.

Throughout fieldwork for the surveys, regular interviewer briefing meetings were held to discuss issues that arose during the survey process and to provide additional training opportunities if needed.

2.5.4 Procedures to improve contact rates

Initially, interviewing was conducted during the following hours to maximise opportunities for contact with potential respondents:

- Monday to Thursday, 10.00 a.m. to 9.00 p.m.
- Friday and Saturday, 10.00 a.m. to 6.00 p.m.
- Sunday, 11.00 a.m. to 7.00 p.m.

Following the 1997 Health Survey, data were analysed to identify the best time of day and best days to contact respondents. As expected, Monday and Tuesday evenings were the best times to either establish initial contact or complete interviews. Fridays were the least productive for both contact and interview completion. Thursdays, Saturdays and Sundays had variable results depending on the time of year, the weather, and other factors such as special events and benefits payment days. These results were

used to develop and refine call and call-back protocols for the CATI system, and guide operating hours for the CATI facility.

2.5.5 Procedures to improve response rates

Interviewer training and careful scripting of the introduction to the survey were key to maximising response rates. The scripted CATI text enabled interviewers to reiterate the importance of the study and further explain its purpose. If respondents said they were too busy, interviewers offered them appointments for a more suitable time. Where respondents were unsure as to the authenticity of the survey, interviewers provided them with a Freecall (1800) telephone number for verification.

Several other strategies were also used to increase response:

- A Web site (www.health.nsw.gov.au/public-health/survey/surveyfaq.html) was set up so that potential respondents could find out more about the survey.
- Introductory letters were sent to all selected households with telephone numbers included in the electronic residential listing. Introductory letters have been shown to increase response rate by up to 10 per cent.³⁰ The letters explained the purpose of the survey, indicated when the household would be contacted, and stressed the importance of the study and the confidentiality of results. The Freecall number was provided for further information or queries. A 'frequently asked questions' sheet was included with the letter; and, in metropolitan areas only, an abbreviated version of the letter, translated into the five non-English survey languages.
- Letters of verification were provided on request.
- A series of media releases, including the Freecall number and Web site address, were prepared and distributed to major and local newspapers, television and radio stations-including the major ethnic media-every 6–8 weeks to increase public awareness about the survey. A column was also incorporated into Health Columns, a newsletter sent to NES communities by the Multicultural Health Communication Service.
- An information sheet about the survey, including answers to frequently asked questions, was distributed to relevant Area Health Service staff, including public affairs officers, switchboard and reception and other 'front line' staff, and staff of the Health Care Interpreter Services.

2.5.6 Survey quality procedures

During the conduct of fieldwork for the surveys, supervisors monitored interviewers on a regular basis to ensure that they conducted interviews appropriately and accurately. Additionally, the CATI system produced a series of reports that assisted in reviewing interviewer performance. These detailed interviewer statistics such as the numbers of completed interviews and appointments made per 100 contacts, and the average time taken to complete interviews.

Survey supervisors re-contacted a random 10 per cent of all respondents to check that a household had actually been contacted, that the correct household member had been selected and interviewed, that accurate information had been recorded, and to obtain the respondent's rating of the interviewer's manner.

2.6 Sampling methods

2.6.1 Selecting a sampling method

During the development of the 1997 NSW Health Survey, two basic options for the telephone number sampling frame were considered by the Technical Group: electronic telephone listings (based on the White Pages directory), or random digit dialling methods. The electronic telephone listings offer ease of use and lower costs. This sampling frame permits all potential respondents to receive an introductory letter, and yields higher contact and response rates than random digit dialling methods.^{18,19,31} Random digit dialling methods are more complex and expensive, but capture both listed and unlisted numbers.

Information obtained from Telstra in 1997 indicated that up to 20 per cent of households in some areas of Sydney had unlisted telephone numbers. As a result, the NSW Health Survey Program Technical Group decided to adopt a random digit dialling method for the 1997 NSW Health Survey, to ensure that this substantial proportion of the population was not excluded from the sampling frame.

Following the completion of the 1997 NSW Health Survey, data were analysed to estimate the proportion of households with unlisted numbers, and to examine potential differences in estimates calculated using data from all participants and from those participants that had listed telephone numbers only. In rural Area Health Services, the proportion of households with unlisted telephone numbers ranged from 5–12 per cent. This proportion ranged from 20–27 per cent in inner and outer metropolitan Sydney Area Health Services. For certain key demographic and risk behaviour questions, there were significant differences in the estimates obtained using the whole sample and the sample with listed numbers only, but only for respondents aged 16–64 years, and not those aged 65 years or more. Results from similar studies show inconsistent findings.^{31,32,33}

These findings informed the sampling strategy for subsequent surveys. The 1998 NSW Health Survey (and the current 2001 NSW Child Health Survey) used random digit dialling methods. The 1999 NSW Older People's Health Survey, which interviewed people 65 years and older, used the electronic telephone listings as its sampling frame.

2.6.2 Target and source populations for the 1997 and 1998 NSW Health Surveys

The target population for both surveys was residents of NSW aged 16 years and over. The source population was NSW residents aged 16 years and over living in households with private telephones.

The target sample for each survey comprised 1,000 people aged 16 years and over from each of the 17 NSW Health Areas (total sample 17,000 people). A two-stage stratified sample design was used, with:

- simple random sampling of household telephone numbers within each NSW Health Area;
- simple random sampling of household residents aged 16 years or more within each household.

2.6.3 Constructing household sampling frames

In 1997 and 1998, household sampling frames were developed using five steps:

- mapping NSW telephone number prefixes to statistical local areas;
- assigning number prefixes to NSW Health Areas;
- generating lists of all potentially active numbers;
- removing continuous blocks of unlisted numbers, and known business and fax numbers;
- sorting the resulting lists in random order.

Sampling frames for Health Area strata were constructed by mapping NSW telephone exchange areas to statistical local areas (SLAs). In 1997, Spatial Decisions Incorporated,³⁴ a commercial subsidiary of Telstra, was contracted to produce these files, using electronic exchange boundary files and MapInfo mapping software.³⁵ The result was a list of exchanges which included the SLAs covered by that exchange, and the proportion of the exchange area falling into each SLA that was used for both the 1997 and 1998 NSW Health Surveys.

The NSW telephone exchanges were then assigned to NSW Health Areas (which in most instances are aggregates of SLAs). Telephone exchanges that supplied two or more Health Areas were assigned to the Health Area containing 55 per cent or more of the exchange area. Where no Health Area met this criterion, the exchange was assigned to that Health Area judged, on the basis of population estimates,³⁶ to contain 50 per cent or more of the population served by the exchange. A similar approach was adopted where exchange areas crossed state boundaries.

In 1997 and 1998, 'Exchange Reference Files' (ERF) for 1997 and 1998 were supplied by Telstra and used to assign 3- or 4-digit telephone number prefixes to exchanges and then to Health Areas.³⁷ The 1998 ERF were not complete, and were enhanced using the most recent version of the electronic residential listings to identify newer prefixes.³⁷ Based on potentially active number ranges given in the 1997 ERF, and the enhanced 1998 ERF, lists of all potentially active numbers for each Health Area were then generated.

Following discussions with Telstra to confirm that unlisted numbers were assigned using the same processes as all other residential numbers, and not in specific blocks, the following procedure was used to remove blocks of unassigned numbers: the lists of numbers were matched with electronic residential listings,^{38,39} and continuous blocks of 100 or more (1997) or 50 or more (1998) numbers that did not include at least one listed number were deleted.

To remove known business numbers, the lists were also matched with the electronic business listings,^{38,39} and any numbers that were included in the business listings, but not the residential listings, were deleted from the sample. Any numbers specified as fax numbers in the electronic residential or business listings were also deleted.

The resulting lists were sorted in random order. The final sampling frames for household selection comprised lists of potentially active telephone numbers for each NSW Health Area, sorted in random order. Samples were drawn from these randomly ordered phone number lists for each Area Health Service to allow for 1,000 completed interviews of respondents aged 16 years or older.

Note: After 1998, ERFs were no longer supported by Telstra and new methods were developed for assigning telephone number prefixes to NSW Area Health Services.

Telephone number prefixes were mapped to SLAs by geocoding address information for telephone numbers listed in the electronic telephone number listings,³⁷ using MapInfo mapping software,³⁴ and then assigned to NSW Health Areas. Where prefixes supplied two or more Health Areas, they were assigned to the Health Area containing 50 per cent or more of the listed numbers with that prefix. A similar approach was adopted where prefixes crossed state boundaries. Additionally, the size of continuous blocks of unlisted numbers that were deleted from the sample was reduced from 50 to 30.

These samples were then downloaded into the CATI system.

2.7 Data cleaning and coding

Data cleaning processes for the 1997 and 1998 NSW Health Surveys were complex, particularly in 1997, when the CATI system stored the survey data in multiple databases. In that year, the data cleaning process began with the download and concatenation or merging of 51 data sets from the CATI system. In 1998, the CATI system was configured to minimise this problem.

Data cleaning processes included checking for duplicate records, reassigning incorrect outcome codes, distinguishing completed interviews from those that are incomplete, checking data consistency for each variable, recoding variables, coding and categorising text data, creating formats and labels for variables, previewing data and, finally, preparing coding manuals.

Occupational data were coded using the ABS Australian Standard Code of Occupations (ASCO) and cause of injury information was coded using external cause, intent, and activity codes based on the National Minimum Data Set for Injury Surveillance (NMDS-IS). New coding frames were developed to code complex text data for patient satisfaction questions.

A new standardised format for survey data was developed in 1999. All previous surveys were recoded to match this data format, and a new coding manual prepared.

2.8 Data analysis

2.8.1 Software

Call and interview data were manipulated and analysed using SAS version 8.1 software.⁴⁰

2.8.2 Weighting

Two levels of weights were generated, sampling weights and post-stratification weights.

Sampling weights adjust for differences in the probabilities of selection among respondents. Persons with more than one telephone number for their household had a higher probability for selection than those with only one phone number. Since only one eligible respondent was selected to be interviewed, persons living in households with more eligible persons were less likely to be selected than persons living in households with fewer eligible persons.

Sampling weights were calculated as follows:

i) Calculate a probability of selection variable (*pselect*) for each respondent:

$$pselect = \frac{\text{number of eligible persons in household}}{\text{number of household telephone connections}}$$

ii) Calculate a sampling weight (*sampwgt*) for each respondent, scaled so the total weighted number of respondents is the same as the total population of NSW, and the weighted number of respondents for each Area Health Service equals the total population of that Area Health service:

$$sampwgt = pselect \times \frac{\text{AHS sample}}{\text{sum of pselect for AHS}} \times \frac{\text{AHS population}}{\text{NSW population}}$$

Note: Information on the number of telephone connections (not for dedicated fax or Internet use) in the household was collected in 1998 only. Because of a CATI programming problem in 1997, the number of eligible persons in the household was missing for some respondents. These were assigned the median value for that variable (2).

Poststratification weights adjust for exclusion from the survey, or differential non-response among certain population groups.

Because different segments of the population are more likely than others to live in households with a telephone and to take part in a survey, certain groups of people were over- or under-represented in the data file.

The variables used to poststratify were gender (male-female) and age (16–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, 74–79, 80+ years). Australian Bureau of Statistics mid-year population estimates for 1997 and 1998 for each Area Health Service were used. These proportions were adjusted to exclude people resident in institutions (nursing homes, public, private and psychiatric hospitals, hostels for the disabled, corrective institutions for children, prisons, corrective and detention centres for adults).

The poststratification weights (*wgt*) were calculated as follows:

$$wgt = \frac{\text{proportion of AHS sample in age - sex - cell (weighted by sampwgt)}}{\text{proportion of AHS population in age - sex - cell}} \times sampwgt$$

2.8.3 Adjustment for survey design

Following the release SAS version 8.1 in 2000,⁴⁰ the SURVEYMEANS procedure was used to calculate standard errors that take into account the stratified sample design. This procedure uses the Taylor expansion method to estimate sampling errors of estimators based on the stratified random sample.^{41,42} The method obtains a linear approximation for the estimator and then uses the variance estimate for this approximation to estimate the variance of the estimate itself. Variance estimates for each strata are then pooled to compute the overall variance estimate.

2.8.4 Response Rates

The response rates were calculated using the following formula:

$$\frac{\text{Completed Interviews}}{\text{Completed interviews} + \text{Household Refusals} + \text{Personal Refusals} + \text{Terminations}}$$

This calculation excludes: not connected; business; unable to contact; persons unable to be interviewed in English, Arabic, Chinese, Greek, Italian, or Vietnamese; persons with a disability that prevented them from being interviewed; and selected respondents who were unavailable for the duration of the survey.

2.9 Reporting

In 1998, data from the 1997 NSW Health Survey were analysed and prepared as an electronic report that was placed on the NSW Health Web site in that same year. The report presented data on self-reported health status; risk factors; health service use; and key health indicators; by age, sex and Area Health Service, and by other demographic variables as appropriate. The data was also made available throughout NSW Health, on-line, through the Health Outcomes Information Statistical Toolkit (HOIST).

In 2000–2001, combined data from the 1997 and 1998 NSW Health Surveys were analysed and prepared as an electronic report which is available at www.health.nsw.gov.au/public-health/nswhs/hsindex.htm. This report includes more extensive analysis than the earlier report, and has a wider selection of graphs and downloadable files for the user to choose from. In addition to age, sex and Area Health Service, indicators are examined by region and remoteness, country of birth, language spoken at home and language of interview, highest level of education, socioeconomic status, employment, indigenous status, and self-rated health status.

The combined dataset from the 1997 and 1998 NSW Health Surveys is available throughout NSW Health, on-line through the Health Outcomes Information Statistical Toolkit (HOIST).

References

1. NSW Department of Health. *Performance Agreement between Director General of NSW Department of Health and Area Health Service*. Sydney: Performance Management Division, NSW Department of Health, July 1997.
2. NSW Department of Health. *Health Outcome Performance Indicators (HOPIs): Monitoring Health Improvement*. Sydney: Health Improvement Branch, NSW Department of Health, May 1997.
3. Chey T, Smith WT, Jalaludin B, Salkeld G, Hanratty SJ, Capon AG. *Surveillance of health status using telephone methodology: a feasibility study*. Sydney: Western Sydney Area Health Service, 1992.
4. Roberts C, Chey T, Capon AG. *Surveillance of health status: results of a telephone survey in the Wentworth Health Area*. Penrith: Western Sydney Area Health Service, 1992.
5. Williamson M, Jorm L, Cardona M, Chey T. *Blacktown Health Survey 1993*. Sydney: Western Sydney Area Health Service, 1994.
6. Astbury L, Jorm L. *Auburn and Holroyd Health Surveys 1995*. Sydney: Western Sydney Area Health Service, 1997.
7. Kolbe A, Gilchrist K, Stubbs N, Clark E. *Behavioural Risk Factor Study—Culcairn, Henty, Holbrook*. Albury: South West Region Department of Health, 1992.
8. Gilchrist K, Kolbe A. *Behavioural Risk Factor Survey—Griffith 1993*. Albury: South West Centre for Public Health, 1993.
9. Gilchrist K, Kolbe A. *Behavioural Risk Factor Survey—Balranald 1993*. Albury: South West Centre for Public Health, 1993.
10. Graham-Clarke P, Howell S, Bauman A, Nathan S. *NSW Health Promotion Survey 1994—Technical Report*. Sydney: National Centre for Health Promotion and NSW Department of Health, 1995.
11. Jorm L, Puech M. *Strategy for population health surveillance in NSW*. Sydney: Epidemiology and Surveillance Branch, NSW Department of Health, 1997.
12. Public Health Division. *New South Wales Older People's Health Survey 1999*. Sydney: NSW Department of Health, 2000.
13. Puech MC, McAnulty JM, Lesjak M, Shaw N, Heron L, Watson JM. A statewide outbreak of Cryptosporidiosis in New South Wales associated with swimming at public pools. *Epidemiol and Infections* 2001 (in press).
14. Vicky Sheppeard, Stephen Corbett, Geoffrey Morgan. A statewide 'outbreak' of asthma in NSW, February 1999. *NSW Public Health Bulletin* 2000; 11(11):187–89.
15. The Sydney Water Incident. July–September 1998. *NSW Public Health Bulletin* 1998; 9(8 & 9): 91–94.
16. National Public Health Partnership. *National Public Health Information Development Plan 1999*. Canberra: National Public Health Partnership, 1999.
17. Quaine J, Jorm L, Williamson M. The Development of the NSW Child Health Survey. *NSW Public Health Bulletin* 2000; 11(5): 75–77.
18. Frey J. *Survey Research by Telephone (2nd Edition)*. London: Sage Library of Social Research, 1989.
19. Groves RM et al. (editors). *Telephone Survey Methodology*. New York: John Wiley and Sons, 1988.
20. Australian Bureau of Statistics. *Household telephone connections in Australia*. Canberra: Australian Bureau of Statistics, 1986. Catalogue no. 4110.0.
21. Smith W, Mitchell P, Attebo K, Leeder S. Selection bias from sampling frames: telephone directory and electoral roll compared with door-to-door population census: the Blue Mountains Eye Study. *Aust NZ J Public Health* 1997; 21: 127–133.

22. Centers for Disease Control and Prevention. *Behavioral Risk Factor Surveillance System User's Guide*. Atlanta, USA: Department of Health and Human Services, Centers for Disease Control and Prevention, 1998.
23. South Australian Social, Environmental and Risk Context Information System at www.dhs.sa.gov.au/pehs/cpse/SERCIS.html.
24. Daly A, Milligan R, Unwin E, Thomson N. *The 1995 Western Australian Health Survey: Overview*. Perth: Health Authorities of Western Australia, February 1996.
25. Kind P, Dolan P, Gudex C, Williams A. Variations in population health status: results from a United Kingdom national questionnaire survey. *BMJ* 1998; 316: 736–41.
26. Australian Bureau of Statistics. *Population Profile 1991—Volume 7: Demographic & Social indicators*. Canberra: Australian Bureau of Statistics, 1994.
27. Ci3 CATI [software]. Evanston, Illinois: Sawtooth Technologies, 1997.
28. SurveyCraft [software]. Melbourne: SurveyCraft, 1998.
29. WinCati [software]. Evanston, Illinois: Sawtooth Technologies, 1999.
30. Smith W, Chey T, Jalaludin B, Salkeld G, Capon T. Increasing response rates in telephone surveys: a randomised trial. *Journal of Public Health Medicine* 1995;17(1): 33–8.
31. Wilson DH, Starr GJ, Taylor AW, Dal Grande E. Random digit dialling and electronic White Pages samples compared: demographic profiles and health estimates. *Aust N Z J Public Health* 1999; 23(6): 627–33.
32. Bennet D, Steel D. *An evaluation of a large scale CATI survey using random digit dialling*. Conference Proceedings of the Statistical Society of Australia 14 Biennial Conference. 1998.
33. Kigawa M, Nishikawa H. Differences between listed and unlisted telephone subscribers among telephone survey respondents. *Japanese Journal of Public Health* 1998; 45(4): 309–19.
34. Spatial Decisions Inc, Melbourne (note: company no longer exists).
35. MapInfo [software]. Troy, NY: MapInfo Corporation, 1997.
36. Australian Bureau of Statistics. Population estimates [electronic], HOIST. Sydney: Epidemiology and Surveillance Branch, NSW Department of Health.
37. Exchange Reference File [software]. Melbourne: Telstra, 1997.
38. Australia on Disc [software]. Sydney: Australia on Disc, 1998 and 1999.
39. Oz on Disc [software]. Sydney: Oz on Disc, 1997 (note: company no longer exists).
40. SAS version 8.1 [software]. SAS Institute Inc, 2000.
41. Fuller WA. Regression Analysis for sample Survey. *Sankhy* 1975, 37, Series C, Part 3, 117–132.
42. Woodruff RS. A simple method for approximating the variance of a complicated estimate. *Journal of American Statistical Association* 1971; 66: 411–414.