Results of Field Testing
Three Modules of
Core Health Questions
for Inclusion in a Nation-Wide
Preventable Chronic Disease
and Behavioural Risk Factor
Survey Module Manual
Report jointly prepared by the WA Department of Health and the SA Department of Human Services for the Australian Bureau of Statistics.
# TABLE OF CONTENTS

1 SUMMARY 1
   1.1 Demographics/Social Characteristics 1
   1.2 Health Indicators 2
   1.3 Asthma
      1.3.1 Prevalence 2
      1.3.2 Severity 3
      1.3.3 Management 3
   1.4 Diabetes
      1.4.1 Prevalence 3
      1.4.2 To Determine Age First Told 3
      1.4.3 Gestational Diabetes 4
      1.4.4 Perception and Knowledge of Risk 4
      1.4.5 Treatment 4
      1.4.6 Complications 4
   1.5 Economic Considerations 4
   1.6 Assessing the Validity of the Prevalence Questions 4

2 INTRODUCTION 5

3 SURVEY OBJECTIVES 6

4 STUDY DESIGN 7
   4.1 Sample Selection 7
   4.1.1 Sample Table 8
   4.1.2 Letter of Introduction 8
   4.1.3 Pilot Testing 9

5 DATA COLLECTION 10
   5.1 Response Rates 10

6 RATIONALE FOR THE ANALYSIS 12
   6.1 Order Of Analysis 12

7 ESTIMATES OF RELIABILITY 13
   7.1 Reliability of Items Common to Both Versions 14
      7.1.1 Demographic Items Common to Both Versions 14
      7.1.2 SF 8 14
      7.1.3 Health Conditions 15
      7.1.4 Diabetes Questions Common to Both Versions 16
   7.2 Reliability of Items Specific to Version A 17
      7.2.1 Demographic Items Specific to Version A 17
      7.2.2 Asthma Questions Specific to Version A 18
      7.2.3 Diabetes Questions Specific to Version A 19
   7.3 Reliability of Items Specific to Version B 21
      7.3.1 Demographic Items Specific to Version B 21
      7.3.2 Asthma Questions Specific to Version B 22
      7.3.3 Diabetes Questions Specific to Version B 22

8 DIFFERENCES BETWEEN VERSIONS A AND B 25
   8.1 Demographic Item Comparisons 25
      8.1.1 Age 25
      8.1.2 Employment Status 26
      8.1.3 Hours of Work 28
      8.1.4 Marital Status 28
      8.1.5 Country of Birth 29
      8.1.6 Year of Arrival 29
      8.1.7 Education Level 30
      8.1.8 Housing 31
      8.1.9 Financial Situation 33
      8.1.10 Income 34
   8.2 Asthma Item Comparisons 35
      8.2.1 Asthma Symptoms 35
      8.2.2 Asthma Treatment 36
      8.2.3 Current Asthma Status 36
      8.2.4 Impact of Asthma on Sleep 37
      8.2.5 Medical Care - Doctor 39
      8.2.6 Emergency Department Visits 40
      8.2.7 Written Asthma Action Plan 41
      8.2.8 Relief Medication 41
      8.2.9 Prevention Medication 43
      8.2.10 Days Off Due to Asthma 45
   8.3 Diabetes Item Comparisons 46
      8.3.1 High Sugar 46
      8.3.2 Non-Gestational Diabetes 46
      8.3.3 Type of Diabetes 47
      8.3.4 Age of Diagnosis - Diabetes 48
      8.3.5 Age of Diagnosis - High Sugar Levels 48
      8.3.6 Diabetes Treatment Options 48
      8.3.7 Insulin Use 50
      8.3.8 Complications 50

9 THE INTERVIEW TIMES 53

10 DETERMINANTS OF RELIABILITY 55
   10.1 Reliability: Number of Telephones
      (Excluding Mobile Phones) 55
   10.2 Reliability: White Pages Listings 56
   10.3 Reliability: Perceived Financial Status 57
   10.4 Reliability: Education Level 58
   10.5 Reliability: Asthma and Diabetes 58
      10.5.1 Reliability, Age and Asthma 59
      10.5.2 Reliability, Education Level and Asthma 60
      10.5.3 Reliability, Age and Diabetes 61
      10.5.4 Reliability, Education and Diabetes 62

11 PREVALENCE COMPARISONS WITH THE 2001 NATIONAL HEALTH SURVEY 63
| TABLE OF CONTENTS |

12 APPENDIX 1 - BIBLIOGRAPHY  64
13 APPENDIX 2 - COMPARISONS OF DEMOGRAPHIC QUESTIONS ACROSS DATA COLLECTIONS  66
14 APPENDIX 3 - QUESTIONNAIRE A  67
   A Introduction  67
   B Demographics  68
   C SF8 and General Health  69
   D Asthma  71
   E Diabetes  73
   F Social Characteristics  76
   G Administration  79
15 APPENDIX 4 - QUESTIONNAIRE B
   A Introduction  80
   B Demographics  81
   C SF8 and General Health  82
   D Asthma  84
   E Diabetes  86
   F Social Characteristics  89
   G Administration  92
16 APPENDIX 5 - INSTRUCTIONS GIVEN TO THE INTERVIEWERS  93
17 APPENDIX 6 - COMMENTS AFTER SA PILOT  96
18 APPENDIX 7 - COMMENTS AFTER WA PILOT  98
19 APPENDIX 8 - INTERVIEWER COMMENTS POST SURVEY DEBRIEF  101
The field testing of the demographic, asthma and diabetes modules was successfully completed in November 2002. This report outlines the procedures used to conduct the field testing and the results that have been found. A summary of the main findings is presented below.

Some questions are recommended for use as they are. Some questions are not recommended for use at all and some are recommended for changes and retesting. Some questions are 'preferred' for use. The 'preferred' questions are based on economy of wording rather than one version being more reliable than another. Therefore, the use of these questions is a matter of preference rather than necessity.

1.1 DEMOGRAPHICS/SOCIAL CHARACTERISTICS

In total 19 demographic/social indicators were tested. A summary of the findings follows:

1. **Age** as determined by the use of date of birth/age last birthday/age group hierarchy (Version A) is reliable and produced the same results and was as reliable as age measured by last birthday/age group (Version B). The later version is recommended based on the validity and simplicity of the questions(s) (see Recommendation 13).

2. **Gender** as ascertained via the voice of the respondent is reliable and is recommended for use.

3. The question to assess the number of adults in the household was identical in both versions and was found to be reliable and is recommended for use.

4. The question to assess the number of children in the household was identical in both versions and was found to be reliable and is recommended for use.

5. **Post code**. The question used in both versions is reliable and recommended for use.

6. **Other language spoken at home**. The question used in both versions is reliable and recommended for use.

7. **Employment status**. Both Version A and Version B were found to be reliable but based on comparisons Version B (with modifications) is recommended for use (see recommendations 14, 15 and 16).

8. **Hours worked last week**. Both the questions used in Version A and Version B were reliable but based on comparisons Version A is recommended for use, with an added filter question (see recommendations 17 and 18).

9. The question to ascertain household composition was only asked in Version B and the response categories were found to be, on the whole, reliable. However two categories had very low reliability and some suggestion for change were made. See recommendations 9, 10 and 11 regarding use of this question.

10. Both versions to assess marital status were reliable but Version A is recommended (see recommendation 19).

11. **Country of birth**. The question used in both versions is reliable and recommended for use although the CATI TRG should access the comments found in 8.1.5 of this document regarding the UK/Scotland/Ireland/England confusion.

12. **Year of first arrival**. Both versions of the question were reliable but Version B is recommended (see recommendation 20).
13 Education. Both versions of this question had good reliability but need some further work as both have weaknesses (see recommendation 21, 22 and 23).

14 Housing. Both versions were reliable but it is recommended that a new question be developed (see recommendation 24).

15 Financial situation. Both versions proved to be reliable but Version A is recommended (see recommendation 25).

16 Household income. Both versions proved to be reliable but Version B, with alterations, is recommended (see recommendation 26 and 27).

17 The question to assess the number of phone numbers - including mobiles was identical in both versions and was found to be reliable.

18 The question to assess the number of phone numbers - excluding mobiles was identical in both versions and was found to be reliable.

19 White pages listings. This question did not have good reliability and it is not recommended for use (see recommendations 1 and 2).

1.2 HEALTH INDICATORS

It was the opinion of the interviewers that the vitality question of the SF8 instrument used a difficult wording. In order to see whether or not a more natural wording would change the responses, two sets of response categories were tested. These have also been assessed for reliability and are discussed in Section 7.1.3.

1 SF8 is a reliable instrument to use on a CATI survey although the use of individual items (eg SF1) is not recommended (see recommendation 3).

2 The questions asked identically in both questionnaires regarding prevalence of heart disease, stroke, arthritis or osteoporosis are reliable and recommended for use.

3 The questions used in both questionnaires regarding prevalence of high blood pressure (ever), high blood cholesterol (ever), depression (ever) and treatment for depression are reliable and recommended for use.

4 The use of the questions to assess the prevalence of current high blood pressure, current high cholesterol and current depression requires further clarification (see recommendations 4 and 5).

1.3 ASTHMA

1.3.1 Prevalence

1 The question used to assess the prevalence of ever having doctor diagnosed asthma was identical in both questionnaires and was found to be reliable and is recommended for use.

2 The two versions of questions used to assess symptoms of asthma were found to be reliable. Version A is preferred because of the shorter wording.

3 The two versions of questions used to assess treatment of asthma were found to be reliable. Version B is recommended but CATI TRG should note the self medication concerns associated with this question (see recommendation 28).

4 The two versions of questions used to assess current asthma status were found to be reliable. Version A is recommended (see recommendation 29).
1.3.2 Severity

1 Night awakenings - last 12 months and how often. Neither version of this question (two questions as in Version A or combined in Version B) proved to be reliable and are not recommended for use (see recommendation 30).

2 Doctor consultations - last 12 months and how often. Neither version of this question (two questions as in Version A or combined in Version B) proved to be reliable and are not recommended for use in its current form (see recommendations 31, 32 and 33).

3 Emergency attendance - last 12 months and how often. Neither version of this question (two questions as in Version A or combined in Version B) proved to be reliable and are not recommended for use (see recommendations 34, 35 and 36).

1.3.3 Management

1 Action plans. Both versions of this question proved to be reliable. Version B is preferred because of its brevity.

2 Relievers - use (last 12 months and ever). There were problems with both versions of this question and it is recommended that further development work on Version B is undertaken (see recommendations 37 and 38).

3 Relievers - how often (last week and last month). It is recommended that Version B is used (see recommendations 39 and 40).

4 Preventors - use (last 12 months and ever). Both versions were reliable. Version A is preferred because of the bounded time.

5 Preventors - how often (last week and last month). It is recommended that more work is undertaken on developing this question (see recommendation 41).

6 Days off work in last 12 months. Both versions produced low reliability scores and are not recommended for use (see recommendation 42).

1.4 DIABETES

Overall the diabetes questions require attention, especially to skip patterns and the role of high blood sugar and gestational diabetes (see recommendation 8, 12 and 44).

1.4.1 Prevalence

1 The question to determine prevalence of doctor diagnosed diabetes was identical in both versions and was found to be reliable and is recommended for use.

2 The prevalence of high sugar levels was determined using two different questions but no differences were found and either could be used.

1.4.2 To determine age first told

1 Age first told (diabetes) - no differences were found in the two versions of the question and both are recommended for use. The supplementary question did not seem to add any information.

2 The reliability of questions used to assess age first told about the high blood sugar was unable to be estimated due to small numbers.
1.4.3 Gestational Diabetes

1. The question to determine the number of women who were first told they had diabetes when they were pregnant was identical and proved to be reliable and is recommended for use.

2. The two questions asked to assess if women had been told they had diabetes other when pregnant produced similar results and either can be used although it is not recommended that type of diabetes is assessed.

1.4.4 Perception and knowledge of risk

1. The question used to assess perception of risk was identical in both questionnaires and proved to be reliable and is recommended for use.

2. The knowledge of risk factors question was asked identically in both questionnaires but proved, on the whole, to be unreliable. A reworking of the questions is recommended in 7.1.4.

1.4.5 Treatment

1. The differences in questionnaire construction regarding first treatment regime proved too different to assess and it is recommended that the question is re-worked (see recommendations 46 and 47).

2. The two versions of the questions that asked if currently on insulin produced similar results and either can be used.

3. How long on insulin. Version A is recommended although the numbers responding to this question were small (recommendation 48).

1.4.6 Complications

1. Complications. The two versions of the question produced similar responses but because of brevity version B is recommended (see recommendations 49, 50 and 51).

1.5 ECONOMIC CONSIDERATIONS

1. Days off - This question proved to have only fair reliability. The question that asked respondents to assess the number of days off was not reliable and is not recommended for use (recommendation 12).

2. Version B took significantly less time to administer than Version A. Determinants of reliability on selected questions regarding age, education level and whether or not English was spoken at home were used tested to see whether or not they influence reliability, but none did.

1.6 ASSESSING THE VALIDITY OF THE PREVALENCE QUESTIONS

1. The 2001 NHS was used as a ‘criterion’ measure against which to assess the validity of the chronic disease prevalence questions. The results demonstrate that there are problems with the use of the NHS for this purpose. Further investigation is recommended to determine suitable reference groups (recommendations 52 and 53).
Computer Assisted Telephone Interviewing (CATI) survey activity has developed in a number of Australian States and Territories during the past decade. CATI surveys are used to obtain high quality information by conducting large sample population surveys in a timely manner. In 1999 the National Public Health Information Working Group (NPHIWG) created the National CATI Technical Reference Group (TRG) to address and advance methodological and technical issues of population surveillance systems in Australia and to contribute to standard practices in CATI surveillance that would form the basis of a nationwide system.

In accordance with these aims, the members of CATI TRG have collaborated to produce modules of question sets for use in CATI surveys, nationally.

For many demographic data variables and variables relating to diabetes and asthma, the reliability of question formats, under CATI conditions, has not been formally established. The “CATI Field-Testing Survey of Demographic, Asthma and Diabetes Modules” is a formal assessment of the reliability of questions chosen for possible inclusion in proposed national question modules for these three topics. For most variables, the field-testing survey includes two alternate questions with differing wording formats that aim to obtain the same, or similar, information. These alternative question wording formats were selected following the collation of existing questions, cognitive testing and consultation with key informants. The survey is designed to establish which of the two questions should be included in the finished module. Funding was provided by the Commonwealth Department of Health and Ageing.

The survey was undertaken for the Australian Bureau of Statistics (ABS) and the CATI TRG as a joint venture between the Epidemiology & Analytical Services Branch, Western Australian Department of Health and the Centre for Population Studies in Epidemiology (CPSE), South Australian Department of Human Services.
The aim of the survey was to assess, under Computer Assisted Telephone Interviewing (CATI) conditions, the usefulness of questions that have been proposed for inclusion in the national CATI Asthma, Diabetes and Demographic question set modules. The criteria for assessment of questions are, firstly, reliability of generated estimates, and, secondly, acceptability of question wording to survey participants, parsimony and cost effectiveness of questions in terms of the length of interviewing time required per question. (In CATI surveys, response rates are affected by time on the telephone, so it is important that questions are concise in order to maximise the number of questions within a module that can be asked within a certain time limit).

The specific objectives of the field-testing are:

• to determine whether or not different question wordings produce different estimates of prevalence.

Where question wording does affect estimates, that is, there are statistically significant differences in estimates generated:

• to determine which question wording is the more reliable of two alternate question formats utilising ICC, Kappa and Weighted Kappa statistical testing.

Where question wording does not affect estimates, that is, there are no statistically significant differences in estimates generated:

• to ascertain which wording is the most acceptable to the survey participants.
• to determine the most cost effective question wording.

Initially a review of the literature related to validity and reliability, especially in regard to CATI, telephone interviewing, diabetes and asthma was undertaken. Appendix 1 includes the bibliography of relevant articles.
The questions to be tested in the field survey were determined in consultation with the CATI TRG cognitive testing and consultation with key informants. Included Demographic questions were those agreed upon at the National CATI TRG Teleconference held 10th April 2002 (see Appendix 2). Diabetes and asthma questions were sourced from the "ABS Report from Pre-testing of Proposed CATI Health Modules: Demographics, Asthma and Diabetes" conducted in July 2001. Questionnaires were developed in conjunction with the ABS. Additional questions, such as the SF8 were included in order to create questionnaires that more closely resembled those used in typical CATI health surveys. The areas covered and the number of questions asked in this survey are listed in Table 1.

Table 1  Areas Covered in Survey

<table>
<thead>
<tr>
<th>Areas covered</th>
<th>Number of items*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td>11</td>
</tr>
<tr>
<td>Diabetes</td>
<td>13</td>
</tr>
<tr>
<td>Asthma</td>
<td>11</td>
</tr>
<tr>
<td>SF8</td>
<td>8</td>
</tr>
<tr>
<td>General Health</td>
<td>2</td>
</tr>
<tr>
<td>Household data / telephone listings</td>
<td>5</td>
</tr>
</tbody>
</table>

* Items may consist of more than one actual question in one or both question formats.

The questions to be tested were incorporated into two questionnaires, "Questionnaire A" and "Questionnaire B". The questions in each of the two questionnaires aimed to collect information on the same topics, but, the question formats differed between the two questionnaires (excepting general health questions, which are the same in each questionnaire). Questionnaire A and Questionnaire B are contained in Appendix 3 and Appendix 4 respectively.

Field-testing was conducted via a process of interviewing, then re-interviewing original questionnaire respondents two to four weeks after the initial interview. The survey participants were equally divided into four groups and questionnaires were administered as shown in Table 2.

Table 2  Study Design

<table>
<thead>
<tr>
<th>Respondent Group</th>
<th>Interview Questionnaire</th>
<th>Re-interview Questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Questionnaire A</td>
<td>Questionnaire A</td>
</tr>
<tr>
<td>Group 2</td>
<td>Questionnaire B</td>
<td>Questionnaire B</td>
</tr>
<tr>
<td>Group 3</td>
<td>Questionnaire B</td>
<td>Questionnaire A</td>
</tr>
<tr>
<td>Group 4</td>
<td>Questionnaire A</td>
<td>Questionnaire B</td>
</tr>
</tbody>
</table>

4.1  SAMPLE SELECTION

Initially, a sample of 3000 households/persons was drawn. The sample included 2000 households randomly selected from the national Electronic White Pages (EWP) plus 500 persons known to have diabetes and 500 persons known to have asthma.

Of the 2000 households selected from the EWP, the number of persons selected from each state and territory was proportional to the population of that state or territory. All households with a telephone connected and the telephone number listed in the Electronic White Pages (EWP) were eligible for selection in the sample.
The sample of 500 adults known to have diabetes was selected from persons who had previously answered positively to the questions “Have you ever been told that you have diabetes?” and “Do you still have diabetes?” in previous surveys conducted by the Epidemiology & Analytical Services Branch, Western Australian Department of Health and the Centre for Population Studies in Epidemiology (CPSE), South Australian Department of Human Services (250 persons from each department). Similarly, the sample of 500 adults known to have asthma was selected from persons who had replied positively to the questions “Have you ever been told you have asthma?” and “Do you still have asthma?” in previous surveys.

Within each household selected from the EWP, only one person per household, the person aged 18 years or over who was last to have a birthday, was interviewed. There was no replacement for noncontactable persons. For adults known to have diabetes or asthma, the required participant was sought by name on contact with the household.

4.1.1 Sample Table

The sample sizes were determined by the known prevalence of diabetes and asthma. The numbers were calculated to give between a 3% and 5% error margin and a 95% to 99% confidence limit for these prevalence estimates. The demographic sample was determined using the 5% estimate as there were a variety of prevalence estimates making up the ‘demographic’ area. The totals are more than adequate to give reliable estimates from the samples and to provide enough power for a cross over design analysis. Table 3 presents the minimum number of respondents expected in each cell.

<table>
<thead>
<tr>
<th>Test-retest</th>
<th>Dr diagnosed Diabetics</th>
<th>Dr diagnosed Asthmatics</th>
<th>Demographics</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A:A</td>
<td>62</td>
<td>62</td>
<td>183</td>
<td>307</td>
</tr>
<tr>
<td>A:B</td>
<td>62</td>
<td>62</td>
<td>183</td>
<td>307</td>
</tr>
<tr>
<td>B:A</td>
<td>62</td>
<td>62</td>
<td>183</td>
<td>307</td>
</tr>
<tr>
<td>B:B</td>
<td>62</td>
<td>62</td>
<td>183</td>
<td>307</td>
</tr>
<tr>
<td>Totals</td>
<td>248</td>
<td>248</td>
<td>732</td>
<td>1228</td>
</tr>
</tbody>
</table>

Table 4 presents the actual number of interviews achieved in each cell.

4.1.2 Letter of Introduction

Normally a letter introducing the health survey would be sent to the household of each randomly selected telephone number. However, the lack of complete addresses for many of the sample meant that there was a high probability that many letters would be returned. In order to standardise the study design, it was decided not to send any letters. The potential for biasing the response rates was also avoided.
4.1.3 Pilot Testing

Before the conduct of the main survey, the questionnaire was pilot tested in South Australia (SA) on September 2, 2002 (n=50) using paper and pencil recording of responses. After changes were made on the basis of the first pilot, it was piloted a second time in Western Australia (WA) on September 9, 2002. This time, the pilot was run as a CATI survey. The aims of the pilot tests were to establish if there are any problems with the questionnaires. A representative from ABS attended both debriefing sessions.

Many of the questions had been used in other studies and so were to some extent tried and tested. However, it was important that the pilot interviewers gave the investigators their full and detailed observations on the questionnaires. In particular interviewers were asked to comment if they thought:

• the wording of any questions was confusing?
• any of the questions were too difficult to answer over the telephone?
• the questionnaire was too long?
• any of the questions were too sensitive / embarrassing?
• any of the questions cause respondents to be angry or frustrated?
• older/younger people experience more difficulty with the questionnaire than others?

The instructions given to the pilot interviewers is contained in Appendix 5. The detailed comments recorded from each pilot are included in Appendix 6 (SA pilot) and Appendix 7 (WA pilot). The questionnaires were amended, in consultation with the ABS, as a result of these findings.
Although interviewing was planned to be undertaken in both WA and SA an unforeseen incident within the SA Department of Human Services resulted in all data collection being undertaken by the WA contracted agency, the University of Western Australia Survey Research Centre. The investigators were very confident that this change in plans would not affect the outcome of the study. This premise was based on firstly, the successful completion of the WANTS study where differences between interviews in each state proved to be minimal and secondly, the results of the pilot testing for this project where consistent comments and observations were presented.

The first stage survey interviews for WA commenced on September 10, 2002. Second stage interviews were staggered to fit with a two week interval between first and second interviews. The final second stage interviews were conducted in the first week of October 2002. The first stage survey interviews conducted in WA on behalf of SA were unable to begin before the end of October 2002 and the second stage interviews were similarly staggered to fit with a two week interval between first and second interviews. The final interviews were conducted in the third week of November. Telephone calls were made between 9:30 am and 9:00 p.m. in the state of the participant’s residence, seven days a week. Professional interviewers conducted the interviews and were supervised by the Manager of the CATI centre at UWA SRC and Western Australian Department of Health Epidemiology & Analytical Services Branch personnel. Disposition codes and data were supplied to Epidemiology & Analytical Services Branch staff regularly, to ensure careful monitoring of survey activities.

On contacting the household, the interviewer initially identified themselves and explained they were conducting a community health survey and no mention was made at this time of any future possible contact.

The CATI III (Computer Assisted Telephone Interview) system was used to conduct the interviews. This system allows immediate entry of data from the interviewer’s questionnaire screen to the computer database. The main advantages of this system are the precise ordering and timing of callbacks and correct sequencing of questions as specific answers are given. The CATI system enforces a range of checks on each response with most questions having a set of predetermined response categories. When open ended responses were required; these were transcribed exactly by the interviewer.

At least seven callbacks were made to the telephone number selected. Different times of the day or evening were scheduled for each callback. If a person could not be interviewed immediately they were rescheduled for interview at a time suitable to them. Replacement interviews for persons who could not be contacted or interviewed were not permitted.

The second interview was conducted approximately two weeks after the initial interview. Recruitment to take part in the second interview was negotiated with the respondent at the beginning of the second interview, with 94% or more agreeing in each group. No respondents had any prior warning that they would be recalled and re-interviewed before that time.

5.1 RESPONSE RATES

Tables 5 to 8 show the response rates for the first and second interviews. There were differences in response rate by State with New South Wales and Victoria having the lowest response rates. The tables 5 to 8 give overall response rates only.

1 The WANTS study was a collaborative study between South Australia, Western Australia and the Northern Territory. South Australia conducted the CATI survey on behalf of the other jurisdictions. WA went on to survey an additional 7500 people to supplement their sample size. The results for data collected by SA on behalf of WA and those collected in WA were almost identical in distribution of responses.
The initial response rates are lower than normal for WA and SA but may be typical for other States with larger populations and more survey demands. The SA and WA response rates for the Asthma and Diabetes respondents ranged from 80% to 95% with most being over 90%. The second contact response rates are excellent and lend confidence to the reliability estimates.

### Table 5  Response Rates - AA Group

<table>
<thead>
<tr>
<th></th>
<th>First Contact</th>
<th>Second Contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>1035</td>
<td>415</td>
</tr>
<tr>
<td>Non Contact</td>
<td>452</td>
<td>0</td>
</tr>
<tr>
<td>Eligible Contact</td>
<td>583</td>
<td>415</td>
</tr>
<tr>
<td>Unable to participate</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Refused</td>
<td>144</td>
<td>19</td>
</tr>
<tr>
<td>Complete</td>
<td>415</td>
<td>396</td>
</tr>
<tr>
<td>Response Rate</td>
<td>415/583</td>
<td>396/415</td>
</tr>
</tbody>
</table>

### Table 6  Response Rates - AB Group

<table>
<thead>
<tr>
<th></th>
<th>First Contact</th>
<th>Second Contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>1035</td>
<td>429</td>
</tr>
<tr>
<td>Non Contact</td>
<td>456</td>
<td>0</td>
</tr>
<tr>
<td>Eligible Contact</td>
<td>579</td>
<td>429</td>
</tr>
<tr>
<td>Unable to participate</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Refused</td>
<td>127</td>
<td>10</td>
</tr>
<tr>
<td>Complete</td>
<td>429</td>
<td>419</td>
</tr>
<tr>
<td>Response Rate</td>
<td>429/579</td>
<td>419/429</td>
</tr>
</tbody>
</table>

### Table 7  Response Rates - BA Group

<table>
<thead>
<tr>
<th></th>
<th>First Contact</th>
<th>Second Contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>1037</td>
<td>389</td>
</tr>
<tr>
<td>Non Contact</td>
<td>476</td>
<td>0</td>
</tr>
<tr>
<td>Eligible Contact</td>
<td>561</td>
<td>389</td>
</tr>
<tr>
<td>Unable to participate</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Total Refusal</td>
<td>147</td>
<td>18</td>
</tr>
<tr>
<td>Complete</td>
<td>389</td>
<td>371</td>
</tr>
<tr>
<td>Response Rate</td>
<td>389/561</td>
<td>371/389</td>
</tr>
</tbody>
</table>

### Table 8  Response Rates - BB Group

<table>
<thead>
<tr>
<th></th>
<th>First Contact</th>
<th>Second Contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>1039</td>
<td>421</td>
</tr>
<tr>
<td>Non Contact</td>
<td>422</td>
<td>0</td>
</tr>
<tr>
<td>Eligible Contact</td>
<td>617</td>
<td>397</td>
</tr>
<tr>
<td>Unable to participate</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Total Refusal</td>
<td>175</td>
<td>25</td>
</tr>
<tr>
<td>Complete</td>
<td>421</td>
<td>397</td>
</tr>
<tr>
<td>Response Rate</td>
<td>421/617</td>
<td>371/389</td>
</tr>
</tbody>
</table>
RATIONALE FOR THE ANALYSIS

In outline, the brief for this analysis was to use data from a mixed repeated measures design to examine the reliability and validity of two versions of the health survey interview. The differences between the versions were many and are identified in detail within the report. In general, the differences involved: wording of questions, wording of responses, open versus closed-ended response sets, and use of different filter questions.

Two versions of the interview (A and B) were delivered to a sample of 1655 respondents on two occasions. Four delivery groups were identified by random allocation and the delivery of the two versions was mixed to create four groups who received the versions in the following order: AA (n=415), AB (n=429), BA (n=389), BB (n=422).

The original analysis plan (see Appendix 1) was designed to take account of the following principles:

   - Reliability,
   - Validity,
   - Maximisation of response.

6.1 ORDER OF ANALYSIS

Reliability was examined first, following the principle that an unreliable indicator cannot be valid. For reliability estimation, standard indicators of reliability were calculated for each version based upon statistical estimates using the AA and BB groups. In addition, order effects were examined by detection of systematic variation between the first and second interviews. These results are presented in Section 7.

A comprehensive assessment of validity is not possible with this set of data, since there is no 'gold standard' whereby the versions can be assessed. However, any systematic difference between versions of the interview is an indicator of problems with validity. A comparison of the first interviews for each version allowed for such differences to be detected using all four groups. These results are presented in Section 8.

Recommendations on the different versions are made where necessary, but are largely based upon an assessment of both reliability and validity, since a reliable indicator may not be valid, but a valid indicator must be reliable. The recommendations are inserted in the text at an appropriate point, but are largely contained in Section 8. No recommendation is made where the versions are equivalent.

A further factor which should influence the design of a survey instrument is to maximise response and to ensure the quality of information collected. This can be assessed by examination of the refused and don't know responses for items which are differently worded. These considerations have been incorporated in Section 8 where relevant. In addition, the frequency of 'other' responses has been examined where appropriate.

It was originally intended to conduct a cost utility analysis based upon data for the time of each interview (used as a cost indicator). The data could not be used for this purpose since it was not collected for individual items. A descriptive analysis of the time for each version has been conducted and is presented in Section 9.

Where there is low reliability, it is important to examine factors which may contribute to the measurement error. One important source of measurement error is caused by systematic associations between response errors and other independent variables. A preliminary analysis of determinants of reliability is reported in Section 10.

The final section of the report contains a validation of the prevalence estimates from the two versions of the interview. Data was obtained from the ABS National Health Survey and compared with the prevalence estimates from this survey. The results are reported in Section 11.
The analysis was conducted using SPSS Version 11.0 and PEPI Version 4.0. Many items were common to both A and B versions of the questionnaire. In order to maximise the precision of the reliability coefficients the estimates of reliability for these items were based on the total test-retest data for the whole sample (n=1583). Data from each common question were merged to produce aggregated variables. The variables were labeled by whether or not they were asked at first contact, or second contact.

The reliability was estimated using the following statistics with appropriate confidence intervals:

1. Kappa (K) for nominal categorical variables
2. Weighted Kappa ($K_w$) for ordinal variables
3. Intra-class correlation (ICC) coefficient for interval and ratio scale variables

McNemar's related samples chi squared test was used to detect bias for the Kappa statistics and the F statistic was used to detect bias for the ICC.

Bias was defined as a statistically significant difference between the first and second interview ($P_{	ext{crit}} = .05$). The existence of bias is indicative of an order effect between the first and second interview.

The generally accepted values of Landis and Koch\(^4\) for the interpretation of these statistics have been modified slightly. The following values are suggested:

- Less than 0.40, poor reliability;
- 0.40-0.59, fair reliability;
- 0.60-0.79, good reliability;
- 0.80-1.00, excellent reliability.

The ICC which is based upon Analysis of Variance requires normally distributed variables or, at a minimum, a symmetrical distribution. A number of items in both versions did not meet either assumption and weighted Kappa was calculated in these circumstances.

For ordinal variables, both Kappa and weighted Kappa has been estimated and reported. The estimate of weighted Kappa should be used in these circumstances.

Where reliability falls in the poor and fair range, the lower 95% confidence limit ($\text{LCL}_\text{95\%}$) can be used to determine whether or not the reliability is statistically significantly less than 0.4\(^2\). These circumstances indicate that an item should be carefully considered before being used.

Recommendations will be restricted to questions that require further development work or are not recommended. Questions that are working and reliable have been already summarised (Section 1).


7.1 RELIABILITY OF ITEMS COMMON TO BOTH VERSIONS

7.1.1 Demographic Items Common to Both Versions

The first set of common items were associated with demographics. Both questionnaires used identical questions about number in the household, postcode, language other than English spoken at home, country of birth, perceived financial status and items about telephones listings. Gender was coded by the interviewer for both interviews. Copies of the questionnaires can be found in Appendices 3 and 4. Table 9 presents the reliability estimates around these items.

Most of the estimates indicate excellent or good reliability. The number of listings in the White Pages shows only fair reliability with weighted Kappa.

Examination of the two related items on telephone ownership shows that the reliability of both total number of telephones including mobiles that can be used to speak to someone in the household has excellent reliability and total number of telephones excluding mobiles that can be used to speak to someone in the household has good reliability.

Given that the telephone items are used as the basis of weighting for probability of selection, these varying reliabilities are a cause for concern. They do suggest that the number of listings in the White Pages should not be used for weighting.

Table 9 Common Variables - Demographic Characteristics

<table>
<thead>
<tr>
<th>Demographic Characteristics</th>
<th>Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>Weighted Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>ICC</th>
<th>Lower CI</th>
<th>Upper CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>0.995</td>
<td>0.99</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. in household 18+</td>
<td>0.86</td>
<td>0.84</td>
<td>0.88</td>
<td>0.86</td>
<td>0.83</td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. in household &lt;18</td>
<td>0.95</td>
<td>0.93</td>
<td>0.97</td>
<td>0.97</td>
<td>0.96</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postcode*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9644</td>
<td>0.9591</td>
</tr>
<tr>
<td>Country of birth</td>
<td>0.96</td>
<td>0.95</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other language</td>
<td>0.91</td>
<td>0.82</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phone numbers excl mobiles</td>
<td>0.69</td>
<td>0.63</td>
<td>0.75</td>
<td>0.68</td>
<td>0.61</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phone numbers incl mobiles</td>
<td>0.76</td>
<td>0.73</td>
<td>0.79</td>
<td>0.82</td>
<td>0.80</td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White Page Listings</td>
<td>0.54</td>
<td>0.48</td>
<td>0.60</td>
<td>0.54</td>
<td>0.47</td>
<td>0.61</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 9 Note: Postcode as given by the respondent.

**Recommendations 1 and 2:**

1. The question that asks about the number of times telephone numbers are listed in the White Pages for weighting purposes should be assessed.

2. As mobile telephones are not used to survey people at this time, the number of telephones excluding mobiles is the only question that has a reasonable reliability. It may be worth looking for other alternatives for questions in this area.

7.1.2 SF 8

The next set of common items across all versions is the SF 8. While Kappa and weighted Kappa was derived for each item, it is important to note that the developers of the scale specifically state that individual items should not be used in isolation. Table 10 presents the results.
With the exception of the question about vitality, all other items showed statistically significant bias (p<.05, McNemar's test). Bias Adjusted Kappa (BAK) was calculated for all these items. The bias for these items is likely to be a consequence of a ceiling effect because of the generally high scores on these items. The individual items had weighted Kappa indicating fair reliability only, but the summary component scores showed excellent reliability.

It is worth noting that the first item which asks respondents to rate their health generally and is often referred to as the SF1, has only fair reliability and this may indicate that this item is not a good population based indicator of health status.

Table 10 Common Variables - SF8 Items and Component Scores

<table>
<thead>
<tr>
<th>Variable</th>
<th>Kappa</th>
<th>BAK</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>Weighted Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>ICC</th>
<th>Lower CI</th>
<th>Upper CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Overall health</td>
<td>0.40</td>
<td>0.40</td>
<td>0.37</td>
<td>0.43</td>
<td>0.56</td>
<td>0.53</td>
<td>0.59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Physical function</td>
<td>0.40</td>
<td>0.39</td>
<td>0.36</td>
<td>0.42</td>
<td>0.55</td>
<td>0.52</td>
<td>0.58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Role physical</td>
<td>0.39</td>
<td>0.39</td>
<td>0.36</td>
<td>0.42</td>
<td>0.54</td>
<td>0.51</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Bodily pain</td>
<td>0.36</td>
<td>0.36</td>
<td>0.33</td>
<td>0.39</td>
<td>0.55</td>
<td>0.52</td>
<td>0.58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Vitality (Very Much)</td>
<td>0.40</td>
<td>na</td>
<td>0.35</td>
<td>0.45</td>
<td>0.51</td>
<td>0.46</td>
<td>0.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Vitality (A Lot)</td>
<td>0.38</td>
<td>na</td>
<td>0.33</td>
<td>0.43</td>
<td>0.50</td>
<td>0.45</td>
<td>0.55</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Social function</td>
<td>0.36</td>
<td>0.36</td>
<td>0.01</td>
<td>0.71</td>
<td>0.48</td>
<td>0.44</td>
<td>0.52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Mental health</td>
<td>0.37</td>
<td>0.37</td>
<td>0.34</td>
<td>0.40</td>
<td>0.51</td>
<td>0.48</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Role emotional</td>
<td>0.37</td>
<td>na</td>
<td>0.33</td>
<td>0.41</td>
<td>0.48</td>
<td>0.44</td>
<td>0.52</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Physical component score (PCS) | 0.8698 | 0.8564 | 0.8821
Mental component score (MCS)   | 0.8140 | 0.7948 | 0.8315

a Note: Significant indication of bias, McNemar test, p < 0.05. Bias Adjusted Kappas (BAK) shown on the table.
b Note: Vitality questions were presented in two forms and the reliability for these is presented separately.

For the vitality question, there were two forms of wording for the alternatives which were randomised within both versions of the interview. The first alternative wording of 'very much' was replaced with 'a lot'. To examine the effect of this variation in wording, the reliability (K) was estimated for both versions and the equivalence of the two versions was tested using a $\chi^2$ test of heterogeneity. The results of the analysis show that the versions are equivalent ($\chi^2 = 0.02$, p = 0.89).

**Recommendation 3:**
3. The SF8 in its entirety is a reliable instrument that can be used to measure health status in population health surveys.

### 7.1.3 Health Conditions

The final set of items that were common to both questionnaires are about health conditions, including the prevalence of asthma (ever) and diabetes (ever). Table 11 presents the prevalence estimates and the reliability estimates. These prevalence estimates are biased by the non random selection of some of the sample and only the reliability is being assessed here.
Table 11  Common Variables - Health Conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>1st Contact (%)</th>
<th>2nd Contact (%)</th>
<th>Don't Know</th>
<th>Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart disease</td>
<td>12.1</td>
<td>12.4</td>
<td>0.83</td>
<td>0.78</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>Stroke</td>
<td>3.5</td>
<td>3.1</td>
<td>0.80</td>
<td>0.72</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>Arthritis</td>
<td>27.6</td>
<td>26.9</td>
<td>0.77</td>
<td>0.74</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>7.6</td>
<td>7.2</td>
<td>0.74</td>
<td>0.68</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>None of the above</td>
<td>62.3</td>
<td>63.4</td>
<td>0.86</td>
<td>0.84</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>High BP, ever</td>
<td>37.2</td>
<td>37.8</td>
<td>0.86</td>
<td>0.84</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>High BP, current</td>
<td>17.9</td>
<td>16.1</td>
<td>1.8 - 2.2</td>
<td>0.54</td>
<td>0.47</td>
<td>0.61</td>
</tr>
<tr>
<td>High cholesterol, ever</td>
<td>28.1</td>
<td>28.1</td>
<td>0.7 - 0.5</td>
<td>0.82</td>
<td>0.79</td>
<td>0.85</td>
</tr>
<tr>
<td>High cholesterol, current</td>
<td>9.4</td>
<td>7.9</td>
<td>4.8 - 4.9</td>
<td>0.48</td>
<td>0.41</td>
<td>0.56</td>
</tr>
<tr>
<td>Depression, ever</td>
<td>21.4</td>
<td>21.0</td>
<td>0.83</td>
<td>0.79</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>Depression, current</td>
<td>7.5</td>
<td>7.9</td>
<td>0.56</td>
<td>0.46</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>Depression, treatment</td>
<td>7.3</td>
<td>7.0</td>
<td>0.82</td>
<td>0.75</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>Asthma, ever</td>
<td>36.7</td>
<td>37.1</td>
<td>0.92</td>
<td>0.90</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>Diabetes</td>
<td>26.5</td>
<td>27.2</td>
<td>0.96</td>
<td>0.94</td>
<td>0.97</td>
<td></td>
</tr>
</tbody>
</table>

a Note: All prevalences are estimated from the sample for each contact (1655 1st contact and 1583 2nd contact)
b Note: There were 72 missing values (4.4%) on the second contact - combined data.

Most of the Kappas indicate excellent reliability or are the top end of the range of good reliability. The three exceptions where reliability is only fair are the estimates of current high blood pressure, current high cholesterol and current depression. It may be that people are on medication and so don't know if they are still classified as having the condition. Note that the reliability for being treated for depression is excellent. It may be that more current treatment information would provide a more reliable estimate of current disease status.

There were no statistically significant differences between the two prevalence estimates on any of the items.

**Recommendations 4 and 5:**

4. More questions seeking information about medication use are developed for high blood pressure, high cholesterol and depression.

5. Reliability is tested again using the added information on current treatment.

7.1.4 Diabetes Questions Common to Both Versions

Other than the question asking whether the respondent had been diagnosed with diabetes by a doctor, there were three diabetes items that were common across both versions. The first question asked whether or not the diabetes was diagnosed during pregnancy, the second asked if respondents considered themselves at risk of getting diabetes (asked only of those not diagnosed with diabetes). The third question asked all respondents to identify risk factors for diabetes. This was an unprompted question and had multiple response options for coding.

The reliability estimates for the risk factors were generally low and in the fair range. Only one risk factor, being overweight, was in the good reliability range. Almost all the items showed an indication of bias and a Biased Adjusted Kappa (BAK) was calculated. Table 12 presents the prevalence at first and second contact and reliability estimates.
While the prevalence question about gestational diabetes had a Kappa in the excellent range, all the risk factors had very low reliability. Only being overweight or physically inactive had a reasonable reliability. However, respondents were asked for the risk factors of diabetes unprompted. It may be that asking the question and prompting would improve reliability. Another possibility would be to ask respondents to indicate factors from a list. This is discussed more fully in Section 7.

Table 12 Common Variables - Perception of Risk Factors Associated with Diabetes

<table>
<thead>
<tr>
<th>Diabetes Risk Factors</th>
<th>Prevalence 1st Contact (%)</th>
<th>Prevalence 2nd Contact (%)</th>
<th>Don't Know (%)</th>
<th>Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pregnant when first diabetic</td>
<td>1.1</td>
<td>0.9</td>
<td>0.96</td>
<td>0.92</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Do you consider yourself at high risk</td>
<td>14.8</td>
<td>14.6</td>
<td>5.1 - 6.8</td>
<td>0.61</td>
<td>0.55</td>
<td>0.66</td>
</tr>
<tr>
<td>Getting older</td>
<td>4.1</td>
<td>4.4</td>
<td>0.40</td>
<td>0.29</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>Being overweight</td>
<td>29.7</td>
<td>34.2</td>
<td>0.64</td>
<td>0.60</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Being physically inactive</td>
<td>23.3</td>
<td>25.8</td>
<td>0.56</td>
<td>0.51</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>Having a stressed life</td>
<td>5.6</td>
<td>5.0</td>
<td>0.48</td>
<td>0.38</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>Having a family history</td>
<td>19.0</td>
<td>21.5</td>
<td>0.44</td>
<td>0.38</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>Being of Aboriginal origin</td>
<td>0.1</td>
<td>0.5</td>
<td>0.20</td>
<td>-0.14</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>Eating too much sugar</td>
<td>12.2</td>
<td>14.2</td>
<td>0.46</td>
<td>0.40</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>Being pregnant</td>
<td>1.1</td>
<td>0.9</td>
<td>0.48</td>
<td>0.27</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Poor diet/wrong diet/junk food</td>
<td>38.4</td>
<td>43.8</td>
<td>0.48</td>
<td>0.44</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>Smoking</td>
<td>2.2</td>
<td>2.5</td>
<td>0.41</td>
<td>0.20</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>Genetic</td>
<td>12.1</td>
<td>11.6</td>
<td>0.35</td>
<td>0.29</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>2.8</td>
<td>0.7</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Don't know</td>
<td>17.7</td>
<td>14.2</td>
<td>0.54</td>
<td>0.49</td>
<td>0.60</td>
<td></td>
</tr>
</tbody>
</table>

Note: Significant indication of bias, McNemar test, p < 0.05. Bias Adjusted Kappas were all the same as Kappa.

Recommendations 6 and 7:

6. Further work is done to develop a question to assess diabetes risk factor knowledge.
7. In the interim, the existing question assessing knowledge of risk factors for diabetes is asked as a multiple response question.

7.2 RELIABILITY OF ITEMS SPECIFIC TO VERSION A

A number of items in Version A were different in some respect to items asking the same information in Version B (see Section 7 for more detail). The reliability of these items was tested separately, using the AA and BB groups.

7.2.1 Demographic Items Specific to Version A

Table 13 presents the reliability estimates for the demographic variables that were specific to Version A.

As with the demographic variables in common across all forms, the reliability estimates are good or excellent. The exception was the question asking unemployed people how long they had been unemployed. This question had few respondents and little consistency. Although all who said that they were unemployed in either contact also answered the question about the length of time unemployed, there were only five cases that said they were unemployed on both occasions. It is not possible to assess the reliability on so few cases.
Table 13 Version A Specific - Demographic Characteristics

<table>
<thead>
<tr>
<th>Demographic Characteristics</th>
<th>Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>Weighted Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>ICC</th>
<th>Lower CI</th>
<th>Upper CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9997</td>
<td>0.9997</td>
<td>0.9998</td>
</tr>
<tr>
<td>Employment status</td>
<td>0.77</td>
<td>0.72</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time looking for employment’</td>
<td>na</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hours worked per week</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.7865</td>
<td>0.7075</td>
<td>0.8442</td>
</tr>
<tr>
<td>Marital status</td>
<td>0.97</td>
<td>0.95</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year of arrival in Australia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9980</td>
<td>0.9970</td>
<td>0.9987</td>
</tr>
<tr>
<td>Education</td>
<td>0.69</td>
<td>0.64</td>
<td>0.74</td>
<td>0.78</td>
<td>0.73</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Housing</td>
<td>0.91</td>
<td>0.87</td>
<td>0.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived financial status</td>
<td>0.61</td>
<td>0.54</td>
<td>0.67</td>
<td>0.66</td>
<td>0.60</td>
<td>0.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>0.71</td>
<td>0.66</td>
<td>0.77</td>
<td>0.84</td>
<td>0.80</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Note: Kappa not estimated due to small number of responses (n<5)
b Note: Financial Status categories Don’t Know and Refused coded to missing for weighted Kappa estimate of the ordered categories.

Section 7 discusses some of the issues related to the lower reliability of education, perceived financial status and income in more detail.

7.2.2 Asthma Questions Specific to Version A

Table 14 presents the prevalence estimates for first and second contact and the estimated reliability for each item.

Most of the items were in the good reliability range with one, asthma treatment within the last year showing excellent reliability. The items asking for respondents to remember over time generally showed only fair reliability. Exceptions were how often they used reliever and prevention medication over the last week. Both these items had reliability estimates within the good range based on weighted Kappa.

Table 14 Version A Specific - Asthma Items

<table>
<thead>
<tr>
<th></th>
<th>Prevalence 1st Contact (%)</th>
<th>Prevalence 2nd Contact (%)</th>
<th>Don’t Know</th>
<th>Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>Weighted Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asthma symptoms past year</td>
<td>27.7</td>
<td>26.0</td>
<td></td>
<td>0.74</td>
<td>0.61</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asthma treatment past year</td>
<td>28.0</td>
<td>28.0</td>
<td></td>
<td>0.88</td>
<td>0.78</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Still get asthma</td>
<td>27.0</td>
<td>27.2</td>
<td>1.0 - 0.7</td>
<td>0.70</td>
<td>0.57</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Woke at night</td>
<td>12.3</td>
<td>12.0</td>
<td>0.0 - 0.2</td>
<td>0.60</td>
<td>0.45</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Woken how often</td>
<td></td>
<td></td>
<td></td>
<td>0.47</td>
<td>0.25</td>
<td>0.69</td>
<td>0.53</td>
<td>0.32</td>
<td>0.74</td>
</tr>
<tr>
<td>Doctor consult</td>
<td>19.0</td>
<td>18.3</td>
<td>0.2 - 0.0</td>
<td>0.67</td>
<td>0.53</td>
<td>0.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consult how often</td>
<td></td>
<td></td>
<td></td>
<td>0.44</td>
<td>0.27</td>
<td>0.61</td>
<td>0.53</td>
<td>0.37</td>
<td>0.69</td>
</tr>
<tr>
<td>Attend ED</td>
<td>1.2</td>
<td>1.9</td>
<td></td>
<td>0.76</td>
<td>0.49</td>
<td>1.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ED how often’</td>
<td></td>
<td></td>
<td></td>
<td>na</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asthma plan</td>
<td>7.7</td>
<td>5.5</td>
<td>0.2 - 0.0</td>
<td>0.61</td>
<td>0.44</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asthma relief</td>
<td>25.5</td>
<td>25.5</td>
<td></td>
<td>0.68</td>
<td>0.47</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relief last week</td>
<td></td>
<td></td>
<td></td>
<td>0.47</td>
<td>0.34</td>
<td>0.60</td>
<td>0.60</td>
<td>0.47</td>
<td>0.73</td>
</tr>
<tr>
<td>Asthma prevention</td>
<td>20.5</td>
<td>19.0</td>
<td></td>
<td>0.55</td>
<td>0.39</td>
<td>0.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevention last week</td>
<td></td>
<td></td>
<td></td>
<td>0.54</td>
<td>0.37</td>
<td>0.71</td>
<td>0.67</td>
<td>0.51</td>
<td>0.83</td>
</tr>
<tr>
<td>Days off work</td>
<td>3.1</td>
<td>4.1</td>
<td>0.2 - 0.0</td>
<td>0.35</td>
<td>0.10</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of days’</td>
<td>5.5</td>
<td>8.8</td>
<td></td>
<td>na</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Note: Cell size of 5 or less
Four items had a Kappa or weighted Kappa with a LCL$_{95\%}$ less than 0.4. These are, how many nights the respondent woke due to asthma, how many times they consulted their doctor, whether or not they used a prevention medication last week and whether or not they took any days off work. These items should be examined with care. Section 8 will discuss the issues around these items in more detail.

### 7.2.3 Diabetes Questions Specific to Version A

Version A asked questions of those diagnosed with diabetes (that was not gestational only), of those who were told that they had high blood sugar only and of both groups. The results are presented in Tables 15 to 17 by these classifications.

The first group of questions applied only to those with diabetes that is not gestational (Table 15).

#### Table 15 Version A Specific - Diabetes Items

<table>
<thead>
<tr>
<th>Asked of Diabetes only</th>
<th>1st Contact (%)</th>
<th>2nd Contact (%)</th>
<th>Don't Know Refused</th>
<th>Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>ICC</th>
<th>Lower CI</th>
<th>Upper CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I</td>
<td>0.5</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type II</td>
<td>22.2</td>
<td>23.4</td>
<td>3.1 - 1.9</td>
<td>0.51</td>
<td>0.24</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age of diagnosis$^a$</td>
<td>57.9</td>
<td>57.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9865</td>
<td>0.9801</td>
<td>0.9909</td>
</tr>
<tr>
<td>Have insulin injections</td>
<td>4.6</td>
<td>4.8</td>
<td></td>
<td>0.97</td>
<td>0.91</td>
<td>1.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length of time injecting</td>
<td>4.6</td>
<td>4.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9919</td>
<td>0.9789</td>
<td>0.9969</td>
</tr>
</tbody>
</table>

a Note: Kappa for Type of diabetes question and don't know percentages given with Type II prevalence.

b Note: Mean age at first diagnosis of diabetes - there were two cases who didn't remember at first contact and these were not included in the mean. These two cases did give an age at the second contact.

The striking thing about Table 15 is the fact that the reliability of knowing the type of diabetes is low (LCL$_{95\%}$ <0.4), but the age of diagnosis, being an insulin injector and the time injecting all have excellent reliability.

An examination of the data revealed that there had been a question flow problem and that the question about what type of diabetes the respondent had was asked only of women who had diabetes other than gestational diabetes. These issues are discussed further in Section 8.

Note that in the first interview, 3.1% said that they didn't know what type they had which fell to 1.9% in the second interview. It may be that they were prompted by the interview question to find out what type of diabetes they had.

The second group of questions (Table 16) applied only to those who had been told that they had high blood sugar by a doctor.

#### Table 16 Version A Specific - High Blood Sugar Items

<table>
<thead>
<tr>
<th>Asked of High Blood Sugar Respondents Only</th>
<th>1st Contact (%)</th>
<th>2nd Contact (%)</th>
<th>Don't Know Refused</th>
<th>Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>ICC</th>
<th>Lower CI</th>
<th>Upper CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Blood Sugar Levels</td>
<td>3.9</td>
<td>4.6</td>
<td></td>
<td>0.74</td>
<td>0.56</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age of diagnosis$^a$</td>
<td>57.9</td>
<td>55.2</td>
<td>3.1 - 1.9</td>
<td>0.51</td>
<td>0.24</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Note: Mean age at first diagnosis of high blood sugar levels - N=9 cases
The reliability of the prevalence of high blood sugar is acceptable, but the age of diagnosis shows only fair reliability \((\text{LCL}_{95\%} < 0.4)\). There were only 3 responses to this question at first interview, and 4 responses at second interview, resulting in very poor precision for this estimate.

The third group of questions was asked of people with either diabetes or high blood sugar (Table 17).

The gestational items show complete consistency. With the exception of tablets, the reliability for initial treatment items is poor \((\text{LCL}_{95\%} < 0.4)\). This may be due to the fact that the question asks about treatment at first diagnosis. For those with diabetes this is 7 years ago on average (Range: 0 - 37 years). It is quite likely that people may not remember initial recommendations over that period of time, but would remember what treatment they are currently on.

The questions about complications of diabetes had variable reliability, ranging from excellent to poor. The excellent reliability relates to limb amputation and renal disease, which suggests that reliability is related to the severity of the complication.

**Table 17**  Version A Specific - Diabetes or High Blood Sugar (Combined Wording) Items

<table>
<thead>
<tr>
<th>Diabetes or High Blood Sugar Items</th>
<th>Prevalence 1st Contact (%)</th>
<th>Prevalence 2nd Contact (%)</th>
<th>Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gestational, high blood sugar or diabetes (Yes)*</td>
<td>0.5</td>
<td>0.5</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-gestational, high Blood Sugar (Yes)*</td>
<td>0.5</td>
<td>0.5</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other than gestational, diagnosis within last 12 mos*</td>
<td>0.0</td>
<td>0.0</td>
<td>na</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment - Dietary change</td>
<td>6.7</td>
<td>8.2</td>
<td>0.50</td>
<td>0.32</td>
<td>0.68</td>
</tr>
<tr>
<td>Treatment - Insulin*</td>
<td>0.0</td>
<td>0.0</td>
<td>na</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment - Tablets</td>
<td>14.5</td>
<td>13.7</td>
<td>0.76</td>
<td>0.64</td>
<td>0.88</td>
</tr>
<tr>
<td>Treatment - Lose Weight</td>
<td>12.8</td>
<td>11.3</td>
<td>0.33</td>
<td>0.15</td>
<td>0.50</td>
</tr>
<tr>
<td>Treatment - Exercise</td>
<td>15.7</td>
<td>15.7</td>
<td>0.31</td>
<td>0.14</td>
<td>0.48</td>
</tr>
<tr>
<td>Treatment - Monitor blood glucose daily*</td>
<td>11.6</td>
<td>8.9</td>
<td>0.36</td>
<td>0.19</td>
<td>0.54</td>
</tr>
<tr>
<td>Treatment - Nothing*</td>
<td>0.5</td>
<td>0.7</td>
<td>0.66</td>
<td>0.04</td>
<td>1.00</td>
</tr>
<tr>
<td>Treatment - Other*</td>
<td>0.0</td>
<td>0.0</td>
<td>na</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment - Don't Know*</td>
<td>0.0</td>
<td>0.0</td>
<td>na</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complication - Leg symptoms ever</td>
<td>10.6</td>
<td>11.3</td>
<td>0.70</td>
<td>0.57</td>
<td>0.84</td>
</tr>
<tr>
<td>Complication - Problems with vision</td>
<td>3.9</td>
<td>4.1</td>
<td>0.46</td>
<td>0.29</td>
<td>0.63</td>
</tr>
<tr>
<td>Complication - Hypoglycaemic attack</td>
<td>5.1</td>
<td>4.8</td>
<td>0.70</td>
<td>0.52</td>
<td>0.87</td>
</tr>
<tr>
<td>Complication - Foot ulcer*</td>
<td>0.7</td>
<td>0.7</td>
<td>0.32</td>
<td>0.00</td>
<td>0.61</td>
</tr>
<tr>
<td>Complication - Limb amputation*</td>
<td>0.2</td>
<td>0.2</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Complication - Renal disease</td>
<td>1.9</td>
<td>2.2</td>
<td>0.81</td>
<td>0.78</td>
<td>0.84</td>
</tr>
<tr>
<td>Complication - None</td>
<td>21.9</td>
<td>22.4</td>
<td>0.63</td>
<td>0.47</td>
<td>0.80</td>
</tr>
</tbody>
</table>

* Note: Less than 5 cases - no further analysis conducted
* Note: No one selected this response category

Section 7 contains further discussion of the issues relating to these items.

**Recommendation 8:**

8. Generally, the low reliability of the initial treatment and complications items suggests that the diabetes questions should be reviewed.
7.3  RELIABILITY OF ITEMS SPECIFIC TO VERSION B

7.3.1 Demographic Items Specific to Version B

Table 18 presents the reliability estimates for the demographic variables that were specific to Version B.

<table>
<thead>
<tr>
<th>Demographic Characteristics</th>
<th>Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>Weighted Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>ICC</th>
<th>Lower CI</th>
<th>Upper CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.9997</td>
<td>0.9997</td>
<td>0.9998</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment status</td>
<td>0.86</td>
<td>0.82</td>
<td>0.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hours worked per week</td>
<td>0.97</td>
<td>0.95</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital status</td>
<td>0.95</td>
<td>0.90</td>
<td>0.90</td>
<td>0.97</td>
<td>0.91</td>
<td>0.97</td>
<td></td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>Year of arrival in Australia</td>
<td>0.94</td>
<td>0.90</td>
<td>0.94</td>
<td>0.93</td>
<td>0.91</td>
<td>0.97</td>
<td></td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>Living alone</td>
<td>0.94</td>
<td>0.90</td>
<td>0.96</td>
<td>0.93</td>
<td>0.88</td>
<td>0.94</td>
<td></td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>Living with mother</td>
<td>0.94</td>
<td>0.90</td>
<td>1.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living with father</td>
<td>0.94</td>
<td>0.91</td>
<td>1.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living with partner</td>
<td>0.94</td>
<td>0.91</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living with children</td>
<td>0.94</td>
<td>0.91</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living with stepchildren</td>
<td>0.94</td>
<td>0.90</td>
<td>1.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living with siblings</td>
<td>0.94</td>
<td>0.91</td>
<td>1.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living with other relative</td>
<td>0.94</td>
<td>0.91</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living with non family</td>
<td>0.94</td>
<td>0.91</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refused</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>0.55</td>
<td>0.49</td>
<td>0.61</td>
<td>0.81</td>
<td>0.76</td>
<td>0.86</td>
<td></td>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td>Housing</td>
<td>0.93</td>
<td>0.85</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived financial status*</td>
<td>0.61</td>
<td>0.54</td>
<td>0.67</td>
<td>0.70</td>
<td>0.64</td>
<td>0.76</td>
<td></td>
<td>0.76</td>
<td>0.76</td>
</tr>
<tr>
<td>Income</td>
<td>0.80</td>
<td>0.76</td>
<td>0.85</td>
<td>0.91</td>
<td>0.88</td>
<td>0.94</td>
<td></td>
<td>0.94</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Note: Financial Status categories Don’t Know and Refused coded to missing for weighted Kappa estimate of the ordered categories.

The reliability of most of the items on Table 18 is in the good or excellent range. The only exceptions were the items about 'living with other relative' and 'living with stepchildren' (LCL < 0.4). While all the others show good reliability it is not obvious what the purpose of the question is in relation to health.

The low reliability of 'hours worked per week' and perceived financial status are discussed further in Section 8.

**Recommendations 9, 10 and 11:**
9. Unless there is a good reason to keep the response category 'living with other relative', it should be deleted from the available selections.
10. The 'living with other relative' and 'non family' could be captured in a more general 'other' category.
7.3.2 Asthma Questions Specific to Version B

Table 19 presents the prevalence estimates for first and second contact and the estimated reliability for each item.

As with Version A, the unweighted Kappas ranged from fair to good, with only one item in the excellent range, 'taken treatment for asthma in the past year'.

Table 19 Version B Specific - Asthma Items

<table>
<thead>
<tr>
<th>Item</th>
<th>Prevalence 1st Contact (%)</th>
<th>Prevalence 2nd Contact (%)</th>
<th>Don’t Know</th>
<th>Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>Weighted Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asthma symptoms past year</td>
<td>29.1</td>
<td>26.5</td>
<td>0.73</td>
<td>0.58</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asthma treatment past year</td>
<td>27.0</td>
<td>26.5</td>
<td>0.88</td>
<td>0.78</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Still have asthma</td>
<td>28.2</td>
<td>26.1</td>
<td>0.62</td>
<td>0.47</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Woken how often</td>
<td></td>
<td></td>
<td>0.40</td>
<td>0.29</td>
<td>0.52</td>
<td>0.61</td>
<td>0.50</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>Consult Dr how often</td>
<td></td>
<td></td>
<td>0.37</td>
<td>0.25</td>
<td>0.49</td>
<td>0.54</td>
<td>0.43</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Attend ED how often</td>
<td></td>
<td></td>
<td>0.05</td>
<td>-0.35</td>
<td>0.45</td>
<td>0.56</td>
<td>0.22</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>Asthma plan</td>
<td>9.2</td>
<td>7.3</td>
<td>0.0 - 0.2</td>
<td>0.63</td>
<td>0.48</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asthma relief</td>
<td>26.3</td>
<td>25.1</td>
<td>0.2 - 0.2</td>
<td>0.54</td>
<td>0.28</td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relief last month</td>
<td></td>
<td></td>
<td>0.55</td>
<td>0.43</td>
<td>0.67</td>
<td>0.72</td>
<td>0.62</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>Asthma prevention</td>
<td>16.6</td>
<td>16.6</td>
<td>0.55</td>
<td>0.40</td>
<td>0.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevention last month</td>
<td></td>
<td></td>
<td>0.48</td>
<td>0.30</td>
<td>0.66</td>
<td>0.65</td>
<td>0.48</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>Days off work</td>
<td>2.4</td>
<td>2.4</td>
<td>0.45</td>
<td>0.16</td>
<td>0.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Prevalence estimates will be elevated due to sampling scheme.

The weighted Kappas, however, were generally in the good range. The items requiring memory over a period of time had lower reliability than other types of items. These items are discussed further in Section 7.

7.3.3 Diabetes Questions Specific to Version B

Diabetes questions specific to Version B asked some questions of those diagnosed with diabetes and some of those who were told that they had high blood sugar. The results are presented by classification. The first group of questions applied only to those with diabetes. The results are presented on Table 20.
### Table 20  Version B Specific - Diabetes Items

<table>
<thead>
<tr>
<th>Diabetes Questions Only</th>
<th>1st Contact (%)</th>
<th>2nd Contact (%)</th>
<th>Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>ICC</th>
<th>Lower CI</th>
<th>Upper CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gestational plus high blood sugar or diabetes</td>
<td>1.9</td>
<td>1.7</td>
<td>0.89</td>
<td>0.68</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type †</td>
<td>1.7</td>
<td>1.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type ‡</td>
<td>10.7</td>
<td>9.7</td>
<td>0.79</td>
<td>0.60</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age of diagnosis ‡</td>
<td>54.9</td>
<td>55.6</td>
<td>0.9708</td>
<td>0.9552</td>
<td>0.9801</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currently injecting</td>
<td>3.3</td>
<td>3.1</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yrs after diagnosis injecting</td>
<td>10.4</td>
<td>8.6</td>
<td>0.7163</td>
<td>-0.6512</td>
<td>0.9512</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment - Dietary change</td>
<td>18.7</td>
<td>16.1</td>
<td>0.40</td>
<td>0.21</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment - Insulin †</td>
<td>2.4</td>
<td>1.9</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Tablets</td>
<td>8.8</td>
<td>8.5</td>
<td>0.85</td>
<td>0.74</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment - Lose Weight</td>
<td>5.7</td>
<td>6.6</td>
<td>0.33</td>
<td>0.13</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Exercise</td>
<td>8.1</td>
<td>9.0</td>
<td>0.21</td>
<td>0.01</td>
<td>0.40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment - Monitor blood glucose daily ‡</td>
<td>7.8</td>
<td>8.8</td>
<td>0.23</td>
<td>0.04</td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment - Nothing ‡</td>
<td>1.2</td>
<td>2.6</td>
<td>0.47</td>
<td>0.11</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment - Other ‡</td>
<td>0.7</td>
<td>0.2</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment - Don't Know ‡</td>
<td>0.2</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complication - Leg symptoms ever</td>
<td>10.7</td>
<td>10.4</td>
<td>0.66</td>
<td>0.51</td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complication Problems with vision</td>
<td>4</td>
<td>5.0</td>
<td>0.57</td>
<td>0.37</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complication Hypoglycaemic attack</td>
<td>6.2</td>
<td>5.2</td>
<td>0.56</td>
<td>0.38</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complication - Foot ulcer ‡</td>
<td>0.7</td>
<td>0.5</td>
<td>0.80</td>
<td>0.40</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complication - Limb amputation ‡</td>
<td>0.0</td>
<td>0.2</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complication - Renal disease</td>
<td>0.7</td>
<td>0.2</td>
<td>0.49</td>
<td>0.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complication None</td>
<td>12.3</td>
<td>11.6</td>
<td>0.60</td>
<td>0.44</td>
<td>0.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Days off</td>
<td>1.4</td>
<td>1.2</td>
<td>0.79</td>
<td>0.51</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How many days off</td>
<td>4.0</td>
<td>15.2</td>
<td>0.33</td>
<td>0.09</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Note: Kappa for Type of diabetes question and don't know percentages given with Type II prevalence.
b Note: In each version, 0.2% of those asked did not know the type of diabetes that they had.
c Note: No one chose this response category

Some of these items had good or excellent reliability: gestational high blood sugar or diabetes, type of diabetes, age of diagnosis and whether respondents were currently injecting insulin. The years after diagnosis that injecting began has very poor reliability (LCL < 0.4).

The reliability is low for the type of treatment, with only insulin and tablets having acceptable reliability. Similarly, reliability for type of complication is low. Only leg symptoms and no complications have a lower confidence limit greater than 0.4. This is discussed further in Section 8.

The second group of questions applied only to those who had been told that they had high blood sugar by a doctor. Table 21 presents the results.
Table 21  Version B Specific - High Blood Sugar Items

<table>
<thead>
<tr>
<th>High Blood Sugar Questions Only</th>
<th>Prevalence 1st Contact (%)</th>
<th>Prevalence 2nd Contact (%)</th>
<th>Kappa</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>ICC</th>
<th>Lower CI</th>
<th>Upper CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Blood Sugar Levels</td>
<td>5.2</td>
<td>6.4</td>
<td>0.80</td>
<td>0.67</td>
<td>0.94</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age of diagnosis</td>
<td>44.8</td>
<td>48.2</td>
<td>0.89</td>
<td>0.68</td>
<td>1.00</td>
<td>0.7466</td>
<td>0.0581</td>
<td>0.9318</td>
</tr>
<tr>
<td>Treatment - Dietary change</td>
<td>1.9</td>
<td>1.7</td>
<td>0.89</td>
<td>0.68</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Insulin</td>
<td>1.4</td>
<td>3.3</td>
<td>0.63</td>
<td>0.27</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Tablets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment - Lose Weight</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Exercise</td>
<td>0.7</td>
<td>1.9</td>
<td>0.60</td>
<td>0.13</td>
<td>1.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment - Monitor blood glucose daily</td>
<td>0.5</td>
<td>2.1</td>
<td>-0.12</td>
<td>-0.32</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment - Nothing</td>
<td>0.0</td>
<td>0.7</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment - Other</td>
<td>2.8</td>
<td>1.7</td>
<td>0.53</td>
<td>0.18</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment - Don't Know</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Note: On first contact 0.2% of those asked didn’t know whether or not they had been diagnosed with high blood sugar.
b Note: Mean age at first diagnosis of high blood sugar levels - N=7 cases
c Note: No or very few responses

The reliability for these items was better than for diabetes items, but cell sizes are very small, leading to wide confidence intervals and the lower confidence limit was less than 0.4 in many cases. The monitoring of blood glucose treatment category had a negative reliability. As with the other items, these results suggest that a review of diabetes items is required. This is discussed further in Section 8.

**Recommendation 12:**

12 Generally, the variability of the Kappas suggests that the diabetes and high blood sugar questions may need to be reviewed.
The number of questions varies between the versions with Version A having more questions overall. Version A has 20 demographic questions and Version B has 15 demographic questions; Version A has 16 asthma questions and Version B has 14 asthma questions; Version A has 18 diabetes questions and Version B has 15 diabetes questions. Because of this difference, the word items will be used rather than questions when describing differences between Versions for similar areas. The two versions of the questionnaire differ in a number of ways:

1. Some response categories are not read out to the respondents in Version B that are read out in Version A (e.g. Highest level of education completed).
2. Version A contains filter or screening questions before frequency questions whereas Version B does not filter before frequency items (e.g. Doctor visited for an asthma attack).
3. Different question wording for the same response categories (e.g. Asthma symptoms in the past twelve months).
4. Different question wording and different response categories about the same information (e.g. Woken during the night).
5. Different time intervals for dichotomous responses (e.g. use of reliever medication).
6. Frequency questions with different time frames for recall (e.g. reliever medication 1 week (A) and 1 month (B)).
7. Version A deals with high blood sugar differently from Version B (e.g. treatment question).
8. Extra questions are included in Version A to assess gestational diabetes.
9. Version A has more information on the response categories for Types of Diabetes.
10. Response categories ordered differently (e.g. complications of diabetes).
11. Version B has two extra questions about time off work.

As respondents are normally only interviewed once, all comparisons are made on the first contact interviews to determine any systematic difference in responses. Statistical tests used to compare the versions included Fisher's exact test, likelihood ratio $\chi^2$, t-test and F test as appropriate.

### 8.1 DEMOGRAPHIC ITEM COMPARISONS

These items were split between the beginning and the end of the questionnaires. Age, gender, household composition, and postcode came at the beginning. Employment, hours of work, further household composition (Version B only), marital status, country of birth, year of arrival, language use, education, housing, financial status, income, telephone ownership, and White Pages listings came at the end. Many of these items are identical between the two versions and these will not be discussed further in this section. Items which differed between the two versions are compared below using the first contact interviews. Information on reliability from Section 7 is incorporated in the discussion.

#### 8.1.1 Age

The two versions used different methods to collect age information. Version A requested date of birth, with 'age last birthday' requested if birth date was not provided. Version B requested 'age last birthday' only. Both versions followed with identical questions requesting age group in ten-year categories (except 18-24 and 75+) if no other age information was provided. For Version A, age was calculated by subtracting the birth date from the interview date and a comparison was made between the mean ages for each version using an independent samples t-test.
There was no significant difference between the two versions for age ($t = 0.96$, d.f. = 1614, $p = 0.34$). Ten (1.2%) respondents to Version A chose to give ‘age last birthday’ rather than date of birth. No respondents declined to provide an age and chose to provide age group only. Some instances of data entry errors were detected, with years of birth greater than 2002 being recorded. Although these were few (approximately 1%) the validity and simplicity of the Version B form suggests that it is preferable, particularly since both forms have excellent reliability (ICC = 0.9997 for both forms, see Tables 13 and 18, Section 6).

**Recommendation 13:**

13. Information about age is collected as in Version B.

### 8.1.2 Employment Status

The two versions used different introductory wording with Version A introducing a one week time frame. Version B was not bounded by time. The two versions also requested information for the unemployed in a different manner. In Version A, those who were unemployed were asked a supplementary question about the length of unemployment, whereas in Version B this information was incorporated in the alternatives in the employment question.

**Version A:** “Which of these descriptions fits you best? Last week, were you:

- Self-employed,
- Employed for wages, salary or payment in kind,
- Unemployed,
- Engaged in home duties,
- Student,
- Retired,
- Unable to work,
- Other (specify),
- Refused.”

Supplementary: “Have you been unemployed for more than one year?”

**Version B:** “Which of these best describes your current employment status?

Are you:

- Self-employed,
- Employed for wages, salary or payment in kind,
- Unemployed for less than one year,
- Unemployed for more than one year,
- Engaged in home duties,
- A student,
- Retired,
- Unable to work,
- Other (specify),
- Refused.”

By recoding the unemployed in Version A based upon the supplementary response, it was possible to produce comparable categories for comparison of the two forms. If the respondent answered yes to being unemployed for more than one year, they were mapped to category 4 of Version B, otherwise they were mapped to response category 3 of Version B. All other response categories were worded identically. Table 22 presents the recoded responses.
The two versions produced significantly different responses (likelihood ratio $\chi^2 = 31.2$, d.f. = 8, $p < 0.001$). There were no refusals to either form. Although there was a difference in the frequency of responses related to unemployment for less than one year (A: 0.4%, B: 1.7%) it is minor and not significant being based upon few responses ($n=18$).

There was little difference in the responses related to unemployment for more than one year. A much bigger and statistically significant difference is related to the frequency of 'other' responses ($\chi^2 = 9.2$, d.f. = 1, $p < 0.01$), with many more of these to Version A than to Version B (A: 2.6%, B: 0.1%).

An examination of the specified responses shows that many of the 'other' responses, particularly to Version A, contained appropriate responses given the wording of the questions. These responses largely fell into four categories, which were: on leave (holiday, maternity, sickness) ($n=9$), voluntary work ($n=6$), pensioner (aged, disability) ($n=5$) and casual workers ($n=3$). It is clear that many of these categories are transitory or intermittent and have been provoked by the one week time frame in Version A. If the primary reason for collecting this information is to assess socio-economic status, such responses are largely inappropriate. This difference may also explain the different reliability coefficients for the two forms (Kappa, A: 0.77, B: 0.86, Tables 5 and 10, Section 7).

It is recommended that the wording for Version B be used and the 'retired' alternative be reworded to include pensioners. This alternative should also include those who are retired but do voluntary work if it is to be used as a socio-economic indicator. There may be a case for making a separate category for this small group since voluntary work in retirement may have an influence on health status. However, a separate question on voluntary work would be preferred since employed people also engage in voluntary activities.

**Recommendations14, 15 and 16:**

14. Questions about employment status that are being used to assess socio-economic status be worded as in Version B without a timeframe.

15. The response category of Retired be expanded to include pensioners and those whose only work is voluntary.

16. A separate question for those doing volunteer work be considered for addition.
8.1.3 Hours of Work

This open-ended item was worded differently in the two versions, and was inserted between the employment and unemployment questions in Version A. In Version A the time frame of the previous week was specified.

**Version A:** “How many hours did you work last week in total?”

**Version B:** “How many hours do you work per week?”

The responses to the two forms were compared using a t-test. There was no significant difference between the two means (A: 36.1 hours, B: 37.3 hours, t = 1.00, d.f. = 716.9, p = 0.33, equal variance not assumed).

There is a significant difference in the variance of the two versions (F = 11.55, p = 0.001), with Version A having a larger variance (s.d. = 18.7) than Version B (s.d. = 15.8). It may be that limiting the responses to a specific time frame allows transitory factors to have a greater influence on the responses. This increased variance may be a result of those on leave, or doing voluntary work including their working hours. It is notable that the two versions differ in reliability (see Tables 5 and 10), with ICC for Version A of 0.79 and for Version B of 0.62. Version A should be the preferred wording. However this variable should be limited to those who are employed for payment. Perhaps a filter question could be used for this purpose, or responses limited to those who were employed for payment and who worked last week.

---

**Recommendations 17 and 18:**

17. When asking a question about the number of hours respondents work, the more specific wording of Version A is preferable.

18. Filtering be used to avoid inappropriate responses.

---

8.1.4 Marital Status

The two versions used different question wording and provided different alternatives for marital status. The major difference was the limitation in Version B to ‘formal’ status and the inclusion of de facto relationships in Version A.

**Version A:** “Which of the following best describes your current marital status? Are you:

- Married
- Living with a partner (de facto)
- Widowed
- Divorced
- Separated
- Never married
- Not stated.”

**Version B:** “What is your formal, current marital status? Are you:

- Married
- Widowed
- Separated but not divorced
- Divorced
- Never married
- Don’t know
- Refused.”
The different alternatives make it difficult to formally compare the two versions with a statistical test. The frequency of the responses is shown in Table 23.

The proportion of ‘married’, ‘widowed’ and ‘separated’ respondents is similar for both forms. As would be expected, the major difference lies with the ‘living with a partner (de facto)’ category. It appears that respondents to Version B who might have given this response (based upon Version A) have mostly opted for ‘never married’ and, to a lesser extent, ‘divorced’.

Table 23  Comparison of Version A and Version B Marital Status Item

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Version A</th>
<th></th>
<th>Version B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>Married</td>
<td>483</td>
<td>57.23%</td>
<td>458</td>
<td>56.47%</td>
</tr>
<tr>
<td>Living with a partner (de facto)</td>
<td>44</td>
<td>5.21%</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Widowed</td>
<td>109</td>
<td>12.91%</td>
<td>92</td>
<td>11.34%</td>
</tr>
<tr>
<td>Divorced</td>
<td>62</td>
<td>7.35%</td>
<td>84</td>
<td>10.36%</td>
</tr>
<tr>
<td>Separated</td>
<td>26</td>
<td>3.08%</td>
<td>84</td>
<td>10.36%</td>
</tr>
<tr>
<td>Never married</td>
<td>119</td>
<td>14.1%</td>
<td>153</td>
<td>18.87%</td>
</tr>
<tr>
<td>Not stated/refused</td>
<td>1</td>
<td>0.12%</td>
<td>2</td>
<td>0.25%</td>
</tr>
<tr>
<td>Total</td>
<td>844</td>
<td>100.0%</td>
<td>811</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

It is difficult to say which of the two forms is most valid since the information is not really comparable. It might, however, be argued that current relationship is more relevant to health status than a formal status based upon legal definition. If this is so, then Version A seems to provide more relevant information and should be preferred over Version B. Both versions have excellent reliability (Kappa = 0.97).

Recommendation 19:

19. If the purpose of the marital status question is to assess the effect on health, a question asking for the current relationship be asked as in Version A.

8.1.5 Country of Birth

This item was identical in both versions, and so a formal comparison between them is not appropriate. However it is worth noting that some unreliability related to the categories ‘England’ and ‘Scotland’. A number of ‘other’ responses specified country of birth as ‘Ireland’, ‘Wales’, ‘N. Ireland’, ‘Britain’, etc. It may be worth considering more accurate categories of ‘United Kingdom’ and ‘Ireland’ for this item.

8.1.6 Year of Arrival

For those not born in Australia, a supplementary, open-ended question collected information on the year of arrival. The wording differed between the two versions, with Version A being more specific in defining “arrival”.

Version A: “In which year did you first arrive in Australia to live here for one year or more?”

Version B: “In which year did you first settle in Australia?”
In order to compare the responses to these two forms, the number of years living in Australia was calculated by subtracting the year of arrival from 2002. There was no significant difference in mean years since arrival between the two forms (A: 34.0 yrs, B: 31.8 yrs, t = 1.39, d.f. = 371.1, p = 0.16, equal variance not assumed). There were no refusals or don’t know responses. Since this variable has an excellent reliability coefficient (ICC = 0.9941 (0.9928-0.9952)), the choice between the two forms of wording is a matter of style. Perhaps the shorter Version B form should be preferred on these grounds.

**Recommendation 20:**

20. If economies are needed on wording of questions, Version B wording could be used to ask the question about year of arrival in Australia.

8.1.7 Education Level

The difference between the two versions for this item was that the alternatives were read out for Version A, but not for Version B. In the latter case the responses were interviewer coded to the same ten categories.

**Version A:** “What is the highest level of education you have completed? Have you completed …

University, CAE or some other tertiary institute degree including post-university i.e. postgraduate Diploma, Masters, PhD,
TAFE or Trade Certificate or Diploma,
Completed high school,
Some high school,
Completed primary school,
Some primary school,
Never attended school,
Other (specify),
Don’t know,
Refused.”

**Version B:** “What is the highest level of education you have completed?”

The responses to each version are shown in Table 24.

**Table 24 Comparison of Version A and Version B Education Item**

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Version A</th>
<th></th>
<th>Version B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>University CAE or some other tertiary institute</td>
<td>144</td>
<td>17.06%</td>
<td>145</td>
<td>17.88%</td>
</tr>
<tr>
<td>TAFE or Trade Certificate or Diploma</td>
<td>230</td>
<td>27.25%</td>
<td>124</td>
<td>15.29%</td>
</tr>
<tr>
<td>Completed high school</td>
<td>175</td>
<td>20.73%</td>
<td>153</td>
<td>18.87%</td>
</tr>
<tr>
<td>Some high school</td>
<td>227</td>
<td>26.9%</td>
<td>281</td>
<td>34.65%</td>
</tr>
<tr>
<td>Completed primary school</td>
<td>52</td>
<td>6.16%</td>
<td>43</td>
<td>5.3%</td>
</tr>
<tr>
<td>Some primary school</td>
<td>7</td>
<td>0.83%</td>
<td>3</td>
<td>0.37%</td>
</tr>
<tr>
<td>Never attended school</td>
<td>1</td>
<td>0.12%</td>
<td>1</td>
<td>0.12%</td>
</tr>
<tr>
<td>Other (specify)</td>
<td>8</td>
<td>0.95%</td>
<td>60</td>
<td>7.4%</td>
</tr>
<tr>
<td>Don’t know</td>
<td>0</td>
<td>0.0%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>Refused</td>
<td>0</td>
<td>0.0%</td>
<td>1</td>
<td>0.12%</td>
</tr>
<tr>
<td>Total</td>
<td>844</td>
<td>100.0%</td>
<td>811</td>
<td>100.0%</td>
</tr>
</tbody>
</table>
There is a statistically significant difference between the two forms (likelihood ratio $\chi^2 = 87.7$, d.f. = 8, $p < 0.001$). This difference is largely confined to the two categories 'TAFE or Trade Certificate or Diploma' and 'Other'. Much of the difference relates to those who were educated overseas, with 12 of the 'other' cases and 18 of the 'TAFE or Trade Certificate or Diploma' in Version B coming from those born overseas. Since this pattern is not apparent in the responses to Version A, it appears that respondents themselves convert their overseas education to Australian equivalents. An examination of the specific responses in the 'other' category, indicate some other problems for both the interviewers and respondents in coding these responses in Version B, for example, a number of respondents had gained qualifications during military service.

It is difficult to know which of the two forms is most desirable. Version A produces the more complete information, but reading ten alternatives to a single question during an interview is undesirable and disliked by respondents (See Interviewer comments Appendix 7). The two issues were the time it took to read the responses, which must be read even if the respondent has already answered, and the fact that the highest education category is read out first, which seemed to cause some discomfort for respondents who had little schooling. However, the large number of 'other' responses reduces the reliability of Version B as an ordinal variable. As the reliability for Version B ($K = 0.81$) is similar to that for Version A ($K = 0.78$), either the Version A closed ended question could be used with the response categories reversed, or the open ended Version B question could be used if the interviewers are given better coding instructions for responses.

Recommendations 21, 22 and 23:

1. Both versions of the question be retested.
2. Version A response categories be reversed to read from lowest to highest.
3. Version B response categories are annotated well enough for interviewers to code the responses to the question, particularly for overseas qualifications.

8.1.8 Housing

This item differed between the two versions, with different wording for the introduction, and different alternatives.

Version A: "I would like to ask you some questions about your housing arrangements. Are you …

- Paying rent or board,
- Paying off this dwelling,
- Outright owner/fully owned,
- Living rent free,
- Purchasing under a rent/buy scheme,
- Occupying under a life-tenure scheme,
- Other (specify),
- Don't know,
- Refused."

Version B: "Is this dwelling …

- Owned or being purchased by the occupants,
- Rented from a government organization or any other public agency,
- Rented privately,
- Other (specify),
- Don't know,
- Refused."
Because of the different, non-comparable categories it is not possible to make a formal comparison of the two forms using statistical analysis. The frequency distributions of the responses to the two versions are shown in Tables 25 and 26 below.

Table 25  Version A - Housing Item

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paying rent or board</td>
<td>162</td>
<td>19.2</td>
</tr>
<tr>
<td>Paying off this dwelling</td>
<td>210</td>
<td>24.9</td>
</tr>
<tr>
<td>Outright owner/fully owned</td>
<td>416</td>
<td>49.4</td>
</tr>
<tr>
<td>Living rent free</td>
<td>37</td>
<td>4.4</td>
</tr>
<tr>
<td>Occupying under a life-tenure scheme</td>
<td>6</td>
<td>0.7</td>
</tr>
<tr>
<td>Other - [specify]</td>
<td>10</td>
<td>1.2</td>
</tr>
<tr>
<td>Refused</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>Total</td>
<td>842</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 26  Version B - Housing Item

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owned or being purchased by the occupants</td>
<td>634</td>
<td>78.2</td>
</tr>
<tr>
<td>Rented from a government organisation or any other public agent</td>
<td>72</td>
<td>8.9</td>
</tr>
<tr>
<td>Rented privately</td>
<td>93</td>
<td>11.5</td>
</tr>
<tr>
<td>Other - [specify]</td>
<td>7</td>
<td>0.9</td>
</tr>
<tr>
<td>Refused</td>
<td>5</td>
<td>0.6</td>
</tr>
<tr>
<td>Total</td>
<td>811</td>
<td>100</td>
</tr>
</tbody>
</table>

There are very few responses to the 'Living rent free' (4.4%), ‘Purchasing under a rent/buy scheme’ (0%), and ‘Occupying under a life-tenure scheme’ (0.7%) alternatives in Version A, which suggests that these alternatives could be included in the 'Other' category. The alternatives in Version B have much greater frequency, but it is not clear why there should be a distinction between renting privately and renting from the government. Perhaps this distinction is designed to identify those with subsidised housing.

Some information is being lost in Version B by combining those who own their home with those who are paying it off. The frequency distributions suggest that four categories drawn from both versions, together with an 'other' category would maximize the information for this item. Since both forms have excellent reliability (Kappa, A = 0.91, B = 0.90) such a combination can be expected to have excellent reliability.

Recommendation 24:

24. The response categories to the questions about housing in Version A and Version B should be combined to form fewer more general categories. Possible alternatives could be:

- Rented from government or public authority
- Rented privately,
- Paying off this dwelling,
- Outright owner/fully owned,
- Other (specify).
8.1.9 Financial Situation

This item collects information on respondent's perception of their financial situation. Both forms contain the same alternatives, but with different introductory questions. In particular, Version A includes a one year time frame whereas Version B is not time specific.

**Version A:** “Thinking of your household's financial situation over the last 12 months, which one of the following statements best describes your household's financial situation?

- I am/we are spending more money than I/we get,
- I/we have just enough money to get us through to the next pay day,
- There's some money left over each week but I/we just spend it,
- I/we can save a bit every now and then,
- I/we can save a lot,
- Don't know,
- Refused.”

**Version B:** “Which best describes your family's money situation?

- I am/we are spending more money than I/we get,
- I/we have just enough money to get us through to the next pay day,
- There's some money left over each week but I/we just spend it,
- I/we can save a bit every now and then,
- I/we can save a lot,
- Don't know,
- Refused.”

Table 27 presents the results.

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Version A</th>
<th></th>
<th>Version B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>I am/we are spending more money than I/we get,</td>
<td>41</td>
<td>4.99%</td>
<td>36</td>
<td>4.44%</td>
</tr>
<tr>
<td>I/we have just enough money to get us through to the next pay day,</td>
<td>217</td>
<td>26.4%</td>
<td>207</td>
<td>25.52%</td>
</tr>
<tr>
<td>There's some money left over each week but I/we just spend it,</td>
<td>86</td>
<td>10.46%</td>
<td>40</td>
<td>4.93%</td>
</tr>
<tr>
<td>I/we can save a bit every now and then,</td>
<td>383</td>
<td>46.59%</td>
<td>393</td>
<td>48.46%</td>
</tr>
<tr>
<td>I/we can save a lot,</td>
<td>95</td>
<td>11.56%</td>
<td>108</td>
<td>13.32%</td>
</tr>
<tr>
<td>Don't know,</td>
<td>0</td>
<td>0.0%</td>
<td>11</td>
<td>1.36%</td>
</tr>
<tr>
<td>Refused</td>
<td>0</td>
<td>0.0%</td>
<td>16</td>
<td>1.97%</td>
</tr>
<tr>
<td>Total</td>
<td>822</td>
<td>100.0%</td>
<td>811</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Rather surprisingly, a comparison of the responses to each form shows that there is a significant difference between the two (likelihood ratio $\chi^2 = 56.1$, d.f. = 6, $p < 0.001$). Much of the difference lies in the 'There's some money left over each week but I/we just spend it' and in the 'Don't know' and 'Refused' alternatives. Why this should be so is unclear, but it does suggest that Version A with no 'Don't know' and 'Refused' responses is providing more information. Perhaps Version B responses are being influenced by recent, short-term circumstances, since no time frame is given. Both versions have similar good reliability ($K_{aa} = 0.66$, $B = 0.70$).

**Recommendation 25:**

25. When asking about perceived financial situation, the wording of Version A is preferred.
8.1.10 Income

The two versions differ in both the wording of the introductory questions and the alternatives for the income item. As with other items, Version A specifies a time frame (one year) and gives more explanation in the question wording. It also gives a greater number of alternative categories at a lower income level than Version B. Because the categories are not comparable the versions will be compared by examination of the frequency distributions.

Version A: “I would now like to ask you about your household’s income. We are interested in how income relates to health, lifestyle and access to health services. Before tax is taken out, which of the following ranges best describes your household’s income, from all sources, over the last 12 months?

Less than $10,000,
$10,001-$20,000,
$20,001-$40,000,
$40,001-$60,000,
$60,001-$80,000,
Over $80,000,
Don’t know/not sure,
Refused.”

Version B: “I would now like to ask you about your household’s income. We are interested in how income relates to health, lifestyle and access to health services. Before tax is taken out, which of the following ranges best describes your household’s income, from all sources, over the last 12 months?

Less than $25,000,
$25,001-$50,000,
$50,001-$100,000,
More than $100,000,
Don’t know/not sure,
Refused.”

Tables 28 and 29 present the results.

Table 28  Version A - Income Item

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than $10,000</td>
<td>66</td>
<td>7.8%</td>
</tr>
<tr>
<td>$10,001 - $20,000</td>
<td>185</td>
<td>21.9%</td>
</tr>
<tr>
<td>$20,001 - $40,000</td>
<td>178</td>
<td>21.1%</td>
</tr>
<tr>
<td>$40,001 - $60,000</td>
<td>122</td>
<td>14.5%</td>
</tr>
<tr>
<td>$60,001 - $80,000</td>
<td>89</td>
<td>10.5%</td>
</tr>
<tr>
<td>Over $80,000</td>
<td>93</td>
<td>11%</td>
</tr>
<tr>
<td>Don’t know/not sure</td>
<td>85</td>
<td>10.1%</td>
</tr>
<tr>
<td>Refused</td>
<td>26</td>
<td>3.1%</td>
</tr>
<tr>
<td>Total</td>
<td>844</td>
<td>100.0%</td>
</tr>
</tbody>
</table>
Table 29  Version B - Income Item

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than $25,000</td>
<td>281</td>
<td>34.6%</td>
</tr>
<tr>
<td>$25,001 - $50,000</td>
<td>212</td>
<td>26.1%</td>
</tr>
<tr>
<td>$50,001 - $100,000</td>
<td>167</td>
<td>20.6%</td>
</tr>
<tr>
<td>More than $100,000</td>
<td>70</td>
<td>8.6%</td>
</tr>
<tr>
<td>Don't know</td>
<td>41</td>
<td>5.1%</td>
</tr>
<tr>
<td>Refused</td>
<td>40</td>
<td>4.9%</td>
</tr>
<tr>
<td>Total</td>
<td>811</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

There is a significant difference in the number of 'don't know' and refused responses between the two forms (A: 13.2%, B: 10.0%) (Likelihood ratio $\chi^2 = 18.0$, d.f. = 2, p < 0.0001), which would recommend Version B. This conclusion is supported by the greater reliability of Version B ($K_A = 0.84$, $B = 0.91$). However, the most frequent category in Version B is also the lowest ('Less than $25,000') which suggests that the minimum specified value is too high.

**Recommendations 26 and 27:**

26. Fewer categories as in Version B will be likely to produce better reliability  
27. In assessing income, the lower limit response category might be set at 'less than $20,000' and a '$20,001-$30,000' category be introduced.

### 8.2 ASTHMA ITEM COMPARISONS

Using only the first contact interviews, Versions A and B were compared to determine any difference in responses. The comparison examined the response patterns and the frequency of refusals and don't know responses.

#### 8.2.1 Asthma Symptoms

The second item about asthma examined symptoms of asthma over the preceding 12 months. The two versions differed in the wording of this item, with Version B describing the symptoms. The same response alternatives were offered.

**Version A:** “Excluding times when you have had coughs, colds or other infections, in the last 12 months have you had symptoms of asthma?”

**Version B:** “Symptoms of asthma include coughing, wheezing, shortness of breath and chest tightness when you don't have a cold or respiratory infection. During the past 12 months did you have any symptoms of asthma?”

Table 30 presents the responses. The two versions produced very similar estimates and there was no significant difference between the two versions (Fisher's exact test, $p=0.71$). There were no refusals or don't know responses. The two versions were equally reliable ($K_A=0.74$, $B=0.73$, Tables 19 and 24 respectively). Given that the questions are equivalent, the shorter Version A should be preferred.
### 8.2.2 Asthma Treatment

The third item examined treatment of asthma during the preceding 12 months. The two versions differed in the wording of this item, with Version B specifying the treatment.

**Version A:** “In the last 12 months, have you taken treatment for asthma?”

**Version B:** “During the past 12 months did you take asthma medication that was prescribed or given to you by a doctor? This includes using an inhaler, puffer or nebuliser.”

There was no significant difference between the two versions (Fisher’s exact test, \( p=0.23 \)). There were no refusals or don’t know responses (Table 31).

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Version A</th>
<th>Version B</th>
<th>A &amp; B Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
</tr>
<tr>
<td>Yes</td>
<td>236</td>
<td>76.6%</td>
<td>216</td>
</tr>
<tr>
<td>No</td>
<td>72</td>
<td>23.4%</td>
<td>83</td>
</tr>
<tr>
<td>Total</td>
<td>308</td>
<td>100.0%</td>
<td>299</td>
</tr>
</tbody>
</table>

Given that the items had equal reliability estimates (\( K, A=0.88, B=0.88 \)), the shorter Version A might be preferred. However, the higher proportion responding to this wording might suggest that self medication is being included. If this is undesirable, the wording from Version B should be adopted.

**Recommendation 28:**

28. The wording of Version A is preferred for determination of asthma treatment during the past year, but this might specify that only doctor prescribed medication should be included.

### 8.2.3 Current Asthma Status

The fourth item examined current asthma status. The two versions differed slightly in the wording of this item.

**Version A:** “Do you still get asthma?”

**Version B:** “Do you still have asthma?”

Table 32 presents the results. There was no significant difference between the two versions (Likelihood ratio \( \chi^2, p=0.40 \)). There were slightly more refused/don’t know responses for Version B (2.3% vs. 1.3%). In addition, Version A is slightly more reliable (\( K, A=0.70, B=0.62 \)), so Version A is preferred.
Table 32  Comparison of Version A and Version B - Still Have/Get Asthma

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Version A</th>
<th></th>
<th>Version B</th>
<th></th>
<th>A &amp; B Combined</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>Yes</td>
<td>235</td>
<td>76.3%</td>
<td>216</td>
<td>72.2%</td>
<td>451</td>
<td>74.3%</td>
</tr>
<tr>
<td>No</td>
<td>69</td>
<td>22.4%</td>
<td>76</td>
<td>25.4%</td>
<td>145</td>
<td>23.9%</td>
</tr>
<tr>
<td>Don't Know/Refused</td>
<td>4</td>
<td>1.3%</td>
<td>7</td>
<td>2.3%</td>
<td>11</td>
<td>1.8%</td>
</tr>
<tr>
<td>Total</td>
<td>308</td>
<td>100.0%</td>
<td>299</td>
<td>100.0%</td>
<td>607</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Recommendation 29:
29. The wording of Version A is preferred when assessing whether or not respondents get current asthma.

8.2.4 Impact of Asthma on Sleep

The fifth item examined the effect of asthma symptoms on sleep. The two versions differed with Version A including a filter question and a reduced number of alternatives for the frequency question which followed.

Version A: “In the last 12 months, have you ever woken up during the night because of your asthma?” (Yes/No/Don’t know or refused.)
“In the last 12 months, on average about how often did this occur? Was it:
- 7 nights a week,
- 4-6 nights a week,
- 2-3 nights a week,
- Once a week,
- Less than once a week.”

Version B: “How often do you awaken during the night with asthma?
- Nightly,
- Most nights,
- Twice weekly,
- Weekly,
- Monthly,
- Less often than monthly,
- Only certain times of the year,
- Never,
- Refused.”

The responses were recoded in an attempt to produce comparable categories. If the respondent answered No to the filter question in Version A, then that was considered equivalent to the response category 5 ‘Never’ in Version B. The categories from 1 to 4 in each version’s question mapped readily to each other. If the respondent said, don’t know or refused to the filter question in Version A, then it was considered equivalent to response category 8 in Version B. The remaining three response categories in Version B, 5, 6 and 7 had to map to a single response category in Version A, 5.

Recoding the responses to produce comparable categories showed a significant difference between the two forms (Likelihood ratio \( \chi^2 \), \( p<0.001 \)), but this may be a consequence of the lack of congruity between the two versions. The frequency distributions of both versions are shown below in Tables 33, 34 and 35.
Table 33  Version A: Woke Up Because Of Asthma

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>105</td>
<td>42.3%</td>
</tr>
<tr>
<td>No</td>
<td>141</td>
<td>56.9%</td>
</tr>
<tr>
<td>Don’t Know/Refused</td>
<td>2</td>
<td>0.8%</td>
</tr>
<tr>
<td>Total</td>
<td>248</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Table 34  Version A: On Average About How Often Did This Occur?”

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 nights a week</td>
<td>7</td>
<td>6.7%</td>
</tr>
<tr>
<td>4-6 nights a week</td>
<td>6</td>
<td>5.7%</td>
</tr>
<tr>
<td>2-3 nights a week</td>
<td>22</td>
<td>21.0%</td>
</tr>
<tr>
<td>Once a week</td>
<td>17</td>
<td>16.2%</td>
</tr>
<tr>
<td>Less than once a week</td>
<td>53</td>
<td>50.5%</td>
</tr>
<tr>
<td>Total</td>
<td>105</td>
<td></td>
</tr>
</tbody>
</table>

Table 35  Version B: “How Often Do You Awaken During the Night With Asthma?”

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nightly</td>
<td>8</td>
<td>3.4%</td>
</tr>
<tr>
<td>Most nights</td>
<td>8</td>
<td>3.4%</td>
</tr>
<tr>
<td>Twice weekly</td>
<td>12</td>
<td>5.1%</td>
</tr>
<tr>
<td>Weekly</td>
<td>12</td>
<td>5.1%</td>
</tr>
<tr>
<td>Monthly</td>
<td>13</td>
<td>5.6%</td>
</tr>
<tr>
<td>Less often than monthly</td>
<td>15</td>
<td>6.4%</td>
</tr>
<tr>
<td>Only certain times of the year</td>
<td>98</td>
<td>41.9%</td>
</tr>
<tr>
<td>Never</td>
<td>67</td>
<td>28.6%</td>
</tr>
<tr>
<td>Refused</td>
<td>1</td>
<td>0.4%</td>
</tr>
<tr>
<td>Total</td>
<td>234</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

It is particularly striking that 98 respondents to Version B (42%) chose the “Only certain times of the year” alternative. This suggests that the implied assumption of Version A, that the effect of asthma on sleep is a stable experience, is incorrect. It appears that the more straightforward single question in Version B is providing more information in a more efficient manner than Version A.

There were very few don’t know/refused responses to either version. The reliability of these questions was in the fair to good range with Version B having the better weighted Kappa ($K_w$, A=0.53, B=.061 - See Tables 18 and 23).

**Recommendation 30:**

30. The limited reliability of these items suggests that further development is required to establish both the time interval and the response categories that will result in reliable estimates.
8.2.5 Medical Care - Doctor

The sixth item examined seeking medical care from a doctor because of asthma. The two versions differed with Version A including a filter question and a slightly reduced number of alternatives for the frequency question which followed.

**Version A:** “In the last 12 months, have you consulted a doctor about your asthma? (Yes/No/Don't know or refused.)

In the last 12 months, how often did you consult a doctor about your asthma?

Would you say it was:
- Once a month or more frequently,
- Once every two or three months,
- Two to three times a year,
- About once,
- Don't know/can't remember."

**Version B:** “How often do you see your doctor about your asthma?

Would you say it was:
- Once a month or more frequently,
- Once every two or three months,
- Two to three times a year,
- About once a year,
- Less often than once a year,
- Don't know/can't remember/refused."

Because the two versions have comparable sets of alternatives it is possible to directly compare the two versions. The filter question answer of no was mapped to the response category ‘Less often than once a year’ and all the other response categories mapped directly.

**Table 36 Comparison of Version A and Version B - See Your Doctor**

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Version A</th>
<th>Version B</th>
<th>Both A &amp; B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
</tr>
<tr>
<td>Less often than once a year</td>
<td>76</td>
<td>30.8%</td>
<td>59</td>
</tr>
<tr>
<td>Once a month or more frequently</td>
<td>13</td>
<td>5.3%</td>
<td>25</td>
</tr>
<tr>
<td>Once every two to three months</td>
<td>37</td>
<td>15.0%</td>
<td>38</td>
</tr>
<tr>
<td>Two to three times a year</td>
<td>71</td>
<td>28.7%</td>
<td>65</td>
</tr>
<tr>
<td>About once a year</td>
<td>50</td>
<td>20.2%</td>
<td>46</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>247</strong></td>
<td><strong>100.0%</strong></td>
<td><strong>233</strong></td>
</tr>
</tbody>
</table>

Table 36 shows that there is no significant difference between the two (Likelihood ratio $\chi^2 = 6.0$, d.f. = 4, $p=0.20$). The reliability of the two frequency questions was similar ($K_w, A=0.53, B=0.54$) but not high, whereas the screening question in Version A had good reliability ($K=0.67$). Unless there is a very good reason to collect the frequency data, it is recommended that the single introductory question in Version A be used.
Recommendation 31, 32 and 33:

31. A single question to determine doctors visits is preferred.
32. Given the low reliability estimates of the frequency for doctor visits, it is recommended that this question not be used.
33. If frequency is needed, then develop response categories that give more reliable results.

8.2.6 Emergency Department Visits

The seventh item examined seeking medical care from an emergency department because of asthma. As with the previous item, the two versions differed, with Version A including a filter question. The frequency question which followed was open-ended. Version B contained only the open-ended frequency question.

Version A: “In the last 12 months, have you visited a hospital emergency department because of your asthma? (Yes/No/Don't know or refused.)

How many times in the last 12 months have you visited a hospital emergency department because of an asthma attack you had?”

Version B: “How many times in the last 12 months have you visited a hospital emergency department for an attack of asthma?”

By recoding the “no” responses of the filter question in Version A to zero, it is possible to make comparable response sets for the two versions. The response frequencies are very similar for the two versions as shown in Table 37. There is no significant difference between the two (Likelihood ratio $\chi^2 = 1.52$, d.f. = 3, p=0.68). There were no “don't know/refused” responses to either form. The reliability of Version A was not able to be calculated due to the very small cell size and the reliability for Version B was not high at $K=0.53$. As with the doctor visits the screening question had good reliability, $K=0.76$.

Table 37 Comparison of Version A and Version B - Visited an Emergency Department

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Version A</th>
<th>Version B</th>
<th>A &amp; B Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
</tr>
<tr>
<td>Never</td>
<td>234</td>
<td>94.4%</td>
<td>220</td>
</tr>
<tr>
<td>Once</td>
<td>9</td>
<td>3.6%</td>
<td>9</td>
</tr>
<tr>
<td>Twice</td>
<td>4</td>
<td>1.6%</td>
<td>5</td>
</tr>
<tr>
<td>Four Times</td>
<td>1</td>
<td>0.4%</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>248</td>
<td>100.0%</td>
<td>234</td>
</tr>
</tbody>
</table>

Recommendations 34, 35 and 36:

34. A single question to determine emergency department visits is preferred.
35. Given the low reliability estimates of the frequency for ED visits, it is recommended that this question not be used.
36. If frequency is needed, then develop response categories that give more reliable results.
8.2.7 Written Asthma Action Plan

The eighth item examined the provision of a written asthma action plan. The two versions contained the same response alternatives but were worded differently. In Version A the action plan is described before the question, whereas in Version B the description follows.

**Version A:** “A written asthma action plan contains instructions written by a medical professional on how to manage your asthma. Do you have a written asthma action plan from a medical professional?”

**Version B:** “Do you have a written asthma action plan from a medical professional? That is, written instructions written on how to manage your asthma.”

Table 38 shows that the two versions produce very similar responses (Likelihood ratio $\chi^2 = 2.96$, d.f. = 2, $p=0.23$). There were two don’t know responses, both occurred with Version A. Both versions have similar, good reliability ($K,A=0.61$, $b=0.63$). Both versions appear to be equivalent so perhaps the shorter Version B is to be preferred.

| Response Category | Version A | | Version B | | A & B Combined |
|-------------------|-----------|---------------------------|-----------|---------------------------|
|                   | Number    | Percent                  | Number    | Percent                  | Number    | Percent                  |
| Yes               | 65        | 26.2%                    | 67        | 28.6%                    | 132       | 27.4%                    |
| No                | 181       | 73.0%                    | 167       | 71.4%                    | 348       | 72.2%                    |
| Don’t Know/Refused| 2         | 0.8%                     | 0         | 0.0%                     | 2         | 0.4%                     |
| Total             | 248       | 100.0%                   | 234       | 100.0%                   | 482       | 100.0%                   |

8.2.8 Relief Medication

The ninth item examines the use of medication for the relief of symptoms. Both versions begin with a yes/no question about use, but with different wording. Version A limits the responses to the previous 12 months, whereas Version B is not bounded by a timeframe. Version B describes and provides examples of the relevant medication.

**Version A:** “In the last 12 months, have you used any medication for the relief of your asthma?”

**Version B:** “Do you use reliever medication (puffer or nebuliser) such as Ventolin, Asthmol, Bricanyl?”

Because of the differing time frame one should be cautious about the interpretation of the responses to the two versions.

Nevertheless, as can be seen from Table 39, the two sets of responses are remarkably similar, and not statistically significantly different (Likelihood ratio $\chi^2 = 1.73$, d.f. = 2, $p=0.42$). The reliability of Version A however, was not as good as Version B ($K,A=0.34$, $b=0.55$), but both are low.

| Response Category | Version A | | Version B | | A & B Combined |
|-------------------|-----------|---------------------------|-----------|---------------------------|
|                   | Number    | Percent                  | Number    | Percent                  | Number    | Percent                  |
| Yes               | 222       | 89.5%                    | 205       | 87.6%                    | 427       | 88.6%                    |
| No                | 26        | 10.5%                    | 28        | 12.0%                    | 54        | 11.2%                    |
| Don’t Know/Refused| 0         | 0.0%                     | 1         | 0.4%                     | 1         | 0.2%                     |
| Total             | 248       | 100.0%                   | 234       | 100.0%                   | 482       | 100.0%                   |
Recommendations 37 and 38:
37. More development is needed to raise the reliability of this item.
38. Version B wording be used until a better version has been developed (however, see next recommendation).

The second part of this item requested information about the frequency of relief medication use. The two versions differ in terms of the time frame, one week in Version A and one month in Version B.

**Version A:** “In the last week, how often have you used reliever medication for your asthma? Was it:

- Every day,
- 4-6 days a week,
- 1-3 days a week,
- Less than once a week,
- Not at all,
- Don’t know/refused.”

**Version B:** “In the last month, how often have you used reliever medication?

- Every day,
- Most days,
- About half the days,
- Less than half the days,
- Not at all
- Don’t know/refused.”

Comparison of these responses must be conducted carefully. Only the first, and perhaps the second categories are strictly comparable. The frequency distributions of the two versions are shown below in Tables 40 and 41.

**Table 40  Version A - Use of Reliever Medication**

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Every day</td>
<td>77</td>
<td>34.7%</td>
</tr>
<tr>
<td>4-6 days a week</td>
<td>7</td>
<td>3.2%</td>
</tr>
<tr>
<td>1-3 days a week</td>
<td>43</td>
<td>19.4%</td>
</tr>
<tr>
<td>Less than once a week</td>
<td>17</td>
<td>7.7%</td>
</tr>
<tr>
<td>Not at all</td>
<td>77</td>
<td>34.7%</td>
</tr>
<tr>
<td>Don’t know/refused</td>
<td>1</td>
<td>0.5%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>222</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

**Table 41  Version B - Use of Reliever Medication**

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Every day</td>
<td>75</td>
<td>36.6%</td>
</tr>
<tr>
<td>Most days</td>
<td>19</td>
<td>9.3%</td>
</tr>
<tr>
<td>About half the days</td>
<td>13</td>
<td>6.3%</td>
</tr>
<tr>
<td>Less than half the days</td>
<td>67</td>
<td>32.7%</td>
</tr>
<tr>
<td>Not at all</td>
<td>31</td>
<td>15.1%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>205</td>
<td>100.0%</td>
</tr>
</tbody>
</table>
Both versions have produced very similar estimates of everyday use (A: 34.7%, B: 36.6%), but they produce quite different estimates of frequent use. 22.6% are using weekly (but not every day) in Version A, whereas 15.6 are using most or half the days of the month. Perhaps the most striking difference lies in the respondents for whom there is no information for frequency because the time frame is not appropriate for their pattern of use (A: 34.7%, B: 15.1%). Since there are very few “don’t know/refused” responses it appears that respondents have no difficulty in recollection over the longer time frame. It appears that Version B produces more information on usage patterns. The reliability of Version B is higher than Version A (K, A=0.60, B=0.72). Almost all those who answered ‘No’ to the first question in Version B answered ‘not at all’ to the second question (31 and 28 responses respectively).

**Recommendations 39 and 40:**

39. The wording of Version B should be used to determine patterns of relief medication.

40. The screening question should be discarded as it appears to be redundant.

8.2.9 Prevention Medication

The tenth item examines the use of medication for the prevention of symptoms. As with the relief medication, both versions begin with a yes/no question about use, but with different wording. In a similar manner, Version A limits the responses to the previous 12 months, whereas Version B is not bounded by a timeframe. Version B describes and provides examples of the relevant medication.

**Version A:** “In the last 12 months, have you used preventer medication for your asthma?”

**Version B:** “Do you use preventer medication such as Becotide, Becloforte, Aldecin, Pulmicort, Flixotide, Intaflorte, Chromogen or Tilade?”

Table 42 presents the results. As above, one should be cautious about the interpretation of the responses to the two versions because of the differing time frame. Unlike the responses to the reliever medication there is a significant difference between the two forms (Likelihood ratio $\chi^2 = 6.86$, d.f. = 2, $p=0.03$). It is interesting that the wording in Version A has produced a much higher proportion of positive responses. This may be a consequence of the greater precision of the version which defines the time frame. Overall it would appear that the wording in Version A should be preferred. One response to Version A was a ‘don’t know’. Both versions had similar fair reliability (K, A=0.55, B=0.55).

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Version A</th>
<th>Version B</th>
<th>A &amp; B Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
</tr>
<tr>
<td>Yes</td>
<td>174</td>
<td>70.2%</td>
<td>141</td>
</tr>
<tr>
<td>No</td>
<td>73</td>
<td>29.4%</td>
<td>93</td>
</tr>
<tr>
<td>Don’t Know/Refused</td>
<td>1</td>
<td>40.0%</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>248</td>
<td>100.0%</td>
<td>234</td>
</tr>
</tbody>
</table>

As with the reliever medication, the second part of this item requested information about the frequency of preventive medication use, and the two versions differ in terms of the time frame, one week in Version A and one month in Version B.
Version A: “In the last week, how often have you used preventer medication for your asthma? Was it:
Every day,
4-6 days a week,
1-3 days a week,
Less than once a week,
Not at all,
Don't know/refused.”

Version B: “In the last month, how often have you used preventer medication?
Every day,
Most days,
About half the days,
Less than half the days,
Not at all
Don't know/refused.”

The frequency distributions of the two versions are shown in Tables 43 and 44. As with relief medication, there are a similar proportion who respond ‘every day’ to both versions (A: 56.9%, B: 63.1%). More frequent users (weekly but not every day, or most and half the days) are again rather different (A: 10.9%, B: 6.4%). Once again, the most striking difference is the number who have responded to the ‘not at all’ category. As one might expect, this is much higher for the shorter time frame. The reliability is very similar for both versions (K, A=0.67, B=0.65).

Table 43  Version A - Use of Prevention Medication

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Every day</td>
<td>99</td>
<td>56.9%</td>
</tr>
<tr>
<td>4-6 days a week</td>
<td>5</td>
<td>2.9%</td>
</tr>
<tr>
<td>1-3 days a week</td>
<td>14</td>
<td>8.0%</td>
</tr>
<tr>
<td>Less than once a week</td>
<td>5</td>
<td>2.9%</td>
</tr>
<tr>
<td>Not at all</td>
<td>51</td>
<td>29.3%</td>
</tr>
<tr>
<td>Total</td>
<td>174</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Table 44  Version B - Use of Prevention Medication

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Every day</td>
<td>89</td>
<td>63.1%</td>
</tr>
<tr>
<td>Most days</td>
<td>2</td>
<td>1.4%</td>
</tr>
<tr>
<td>About half the days</td>
<td>7</td>
<td>5.0%</td>
</tr>
<tr>
<td>Less than half the days</td>
<td>21</td>
<td>14.9%</td>
</tr>
<tr>
<td>Not at all</td>
<td>22</td>
<td>15.6%</td>
</tr>
<tr>
<td>Total</td>
<td>141</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

There are response categories in both versions and for both types of medication which contribute very little information (e.g. A: ‘4-6 days a week’, B: ‘most days’). This suggests that neither of the alternative responses represent the typical patterns of medication use. Alternative wording may bring better precision and reliability.

Recommendation 41:
41. This question be reformed with a one month time frame and more precise categories.
8.2.10 Days Off Due to Asthma

The last item in the asthma module requested information about days of incapacity. Version A contained one question which was a yes/no format concerning days of incapacity in the past year. In Version B the same question was followed by an open-ended frequency question.

**Version A:** “In the last 12 months, have you had any days off work, school or home duties due to your asthma?”

**Version B:** “In the last 12 months, have you had any days off work, school or home duties due to your asthma? How many days would you estimate?”

Table 45 presents the results.

### Table 45 Comparison of Version A with Version B - Time Off Because of Asthma

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Version A</th>
<th></th>
<th>Version B</th>
<th></th>
<th>A &amp; B Combined</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>Yes</td>
<td>29</td>
<td>11.7%</td>
<td>27</td>
<td>11.5%</td>
<td>56</td>
<td>11.6%</td>
</tr>
<tr>
<td>No</td>
<td>218</td>
<td>87.9%</td>
<td>207</td>
<td>88.5%</td>
<td>425</td>
<td>88.2%</td>
</tr>
<tr>
<td>Don’t Know/Refused</td>
<td>1</td>
<td>0.4%</td>
<td>0</td>
<td>0.0%</td>
<td>1</td>
<td>0.2%</td>
</tr>
<tr>
<td>Total</td>
<td>248</td>
<td>100.0%</td>
<td>234</td>
<td>100.0%</td>
<td>482</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

There was no significant difference between the two versions (likelihood ratio $\chi^2 = 1.34$, d.f. = 2, $p=0.51$). The reliability of each version is low ($K_A=0.35$, $B=0.45$).

The frequency follow-up in Version B produced a wide range of values as shown in Table 46.

### Table 46 Version B - Days Off Because of Asthma

<table>
<thead>
<tr>
<th>Number of Days</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>14.8</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>11.1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>11.1</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>11.1</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>11.1</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>3.7</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>7.4</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>11.1</td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td>7.4</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td>3.7</td>
</tr>
<tr>
<td>28</td>
<td>1</td>
<td>3.7</td>
</tr>
<tr>
<td>56</td>
<td>1</td>
<td>3.7</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

**Recommendation 42:**

42. Given the poor reliability of these questions, it is recommended that they be discarded.
8.3 DIABETES ITEM COMPARISONS

8.3.1 High Sugar

The two versions differed in the wording used to establish a history of high blood sugar.

**Version A:** “Have you ever been told by a doctor that you have high sugar levels in your blood or urine?”

**Version B:** “Have you ever been told by a doctor that you have high blood sugar or a touch of sugar?”

Table 47 presents the comparison of the two. There was no significant difference in the response to each version (likelihood ratio $\chi^2 = 3.4$, d.f. = 2, $p = 0.19$).

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Version A</th>
<th>Version B</th>
<th>A &amp; B Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
</tr>
<tr>
<td>Yes</td>
<td>31</td>
<td>5.0%</td>
<td>43</td>
</tr>
<tr>
<td>No</td>
<td>582</td>
<td>94.5%</td>
<td>557</td>
</tr>
<tr>
<td>Don't Know</td>
<td>3</td>
<td>0.5%</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>616</td>
<td>100.0%</td>
<td>601</td>
</tr>
</tbody>
</table>

Since both versions have identical reliability ($K = 0.74$), there is little to recommend one version over another.

8.3.2 Non-Gestational Diabetes

The next item to differ between the two versions was concerned with distinguishing between gestational and non-gestational diabetes among those with a history of the former. The two versions contained different wording for this item. In addition, Version A asked respondents to differentiate between diabetes and high blood sugar.

**Version A:** “Other than when you were pregnant, have you ever been told that you have diabetes or high sugar levels?”

Supplementary: “Were you told you had diabetes or high sugar levels? 
Diabetes, 
High sugar levels.”

**Version B:** “Have you ever been told that you had diabetes or high blood sugar, other than when you were pregnant?”

There was no significant difference between the two versions, suggesting that they are equivalent (Fisher’s exact test, $p = 1.00$). Table 48 presents the results.

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Version A</th>
<th>Version B</th>
<th>A &amp; B Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
</tr>
<tr>
<td>Yes</td>
<td>6</td>
<td>25.0%</td>
<td>12</td>
</tr>
<tr>
<td>No</td>
<td>18</td>
<td>75.0%</td>
<td>30</td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td>100.0%</td>
<td>42</td>
</tr>
</tbody>
</table>

Both versions have excellent reliability ($K = 1.00$ and $B = 0.89$)
As the supplementary question was only asked in Version A it is not possible to compare the two versions. The question was only relevant to six respondents, three of whom opted for each alternative. It is recommended that this question not be included as it produces so little information.

8.3.3 Type of Diabetes

The next item aimed to identify the type of diabetes for respondents who had answered positively that they had a history of diabetes. The two versions were different, with alternative responses being read out in Version A. Version B was open ended and coded by the interviewers.

Version A: “Other than the diabetes you had when you were first pregnant, what type of diabetes were you told you had?  
Type I - insulin dependent, juvenile onset,  
Type II - non-insulin dependent, mature onset,  
Don’t know.”

Version B: “Other than the gestational diabetes, what type of diabetes were you told you had?”  
Type I - insulin dependent, juvenile onset,  
Type II - non-insulin dependent, mature onset,  
Don’t know,  
Other (specify).

During the analysis it became apparent that there was a problem with the filtering for Version B. The filtering was designed to identify gestational diabetes/high sugar levels, so that this was only asked of female respondents. As a result, male respondents were inadvertently excluded from this question. Thus this analysis is limited to female respondents only. Table 49 presents the results.

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Version A</th>
<th>Version B</th>
<th>A &amp; B Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
</tr>
<tr>
<td>Type I</td>
<td>3</td>
<td>2.7%</td>
<td>11</td>
</tr>
<tr>
<td>Type II</td>
<td>96</td>
<td>86.5%</td>
<td>85</td>
</tr>
<tr>
<td>Don’t Know</td>
<td>12</td>
<td>10.8%</td>
<td>11</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>0.0%</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>111</td>
<td>100.0%</td>
<td>109</td>
</tr>
</tbody>
</table>

The frequency of female responses to the two versions was significantly different (likelihood ratio $\chi^2 = 8.3$, d.f. = 3, $p = 0.04$), with more respondents originally classified as Type I in Version B (10.1%) than Version A (2.7%) and corresponding differences in the proportions for Type II. It is likely that a similar difference would have been observed for males, since current insulin use does not differ significantly between male and female respondents (Fisher's exact test, $p = 0.34$). The two versions also differ in reliability, with Version B having much better reliability ($K, A=0.51, B=0.79$). Version B wording is to be preferred.

The problem with the filtering is undesirable. In order to avoid such problems it is important that the filtering arrangements be kept simple and that the order of questions be modified to facilitate this. A better sequence would be achieved if this question was third in this module, before any filtering occurred, for both versions.


Recommendations 43 and 44:
43. An unprompted question such as used in Version B is preferred.
44. The sequencing of the diabetes prevalence and type needs to be reviewed. It is suggested that asking type after the prevalence question would simplify the flow.

8.3.4 Age of Diagnosis - Diabetes

This item used different wording in the two versions, with different instructions for respondents.

**Version A:** “The following questions relate to diabetes only. At what age were you first told you had diabetes?”

**Version B:** “Other than the gestational diabetes, how old were you when you were first told you had diabetes?”

There was no difference in the means for each version (A: 61.1 yrs, B: 69.9 yrs, t = -0.94, d.f. = 403, p = 0.35, equal variance assumed). As both versions have excellent reliability (K, A=0.99, B=0.97), it can be concluded that the different versions are equivalent.

8.3.5 Age of Diagnosis - High Sugar Levels

This item used different wording in the two versions, with different instructions for respondents.

**Version A:** “The following questions relate only to the high sugar levels you had when you were not pregnant. At what age were you first told you had high sugar levels?”

**Version B:** “Other than the gestational diabetes, how old were you when you were first told you had high blood sugar?”

There were too few responses to this question, particularly in Version A, to make a valid comparison between the two. Version A had 3 respondents and Version B had 24 respondents. This difference does, however, suggest that the two versions are producing different responses and maybe it should not be used.

Recommendation 45:
45. Discard the question asking for time of diagnosis of high blood sugar other than when pregnant.

8.3.6 Diabetes Treatment Options

This item differed between the two versions, with Version A being applied to those with a history of either diabetes or high blood sugar. Version B asked the same question separately for these two groups but the response categories remained the same.

**Version A:** “Other than the diabetes or high sugar levels you had when you were pregnant, when you were first told you had diabetes or high sugar levels what treatments did the doctor put you on?

- **Dietary change,**
- **Insulin,**
- **Tablets,**
- **Lose weight,**
- **Exercise,**
- **Monitor blood glucose daily,**
- **Nothing,**
- **Don't know/refused,**
- **Other.”**
**Version B:** “When you were first told you had diabetes what treatment did the doctor put you on? and “When you were first told you had high blood sugar what treatment did the doctor put you on?

- Dietary change,
- Insulin,
- Tablets,
- Weight,
- Exercise,
- Monitor blood glucose daily,
- Nothing,
- Don’t know/refused,
- Other.”

It is difficult to combine this data in a way which creates comparable categories. Table 50 shows the positive response frequencies for each of the questions, and compares the Version A question with the diabetes question in Version B using Fisher’s exact test.

The two versions appear to be collecting very different data. Although there is some similarity between Version A and Version B diabetes, there are important differences. Version B shows that the treatment recommendations for diabetes are very different to those for high blood sugar and this suggests that it would be very unwise to combine the two as in Version A.

**Table 50  Comparison of Version A and Version B - Treatment**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Version A Combined</th>
<th>Version B Diabetes</th>
<th>Version B High Blood Sugar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dietary change</td>
<td>78.0</td>
<td>73.6</td>
<td>39.4</td>
</tr>
<tr>
<td>Insulin***</td>
<td>1.7</td>
<td>8.2</td>
<td>0.0</td>
</tr>
<tr>
<td>Tablets</td>
<td>44.4</td>
<td>39.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Lose weight**</td>
<td>41.9</td>
<td>29.5</td>
<td>12.1</td>
</tr>
<tr>
<td>Exercise*</td>
<td>51.5</td>
<td>40.5</td>
<td>18.2</td>
</tr>
<tr>
<td>Monitor blood glucose</td>
<td>42.7</td>
<td>35.5</td>
<td>3.0</td>
</tr>
<tr>
<td>Nothing</td>
<td>1.7</td>
<td>4.1</td>
<td>57.6</td>
</tr>
<tr>
<td>Don’t know/refused</td>
<td>0.0</td>
<td>0.5</td>
<td>0.0</td>
</tr>
<tr>
<td>Other</td>
<td>4.1</td>
<td>1.4</td>
<td>0.0</td>
</tr>
</tbody>
</table>

A combined significantly different from B diabetes; *p<0.05  **p<0.01  ***p<0.001

The reliability of this data is poor, with the possible exception of insulin and tablets for Version A (combined) and Version B (diabetes). It is recommended that this question only request data for these alternatives plus an ‘other’ category of those with diabetes only.

**Recommendations 46 and 47:**

46. Restrict treatment questions to diabetes respondents only.
47. Include fewer categories. A grouping might be: Tablets, Insulin, Other.
8.3.7 Insulin Use

This item used two questions in each version of the interview, with different wording for both.

**Version A:** “Do you currently have insulin injections?"
- Yes,
- No.”

“For how long have you been having insulin injections for your diabetes?”

**Version B:** “Do you currently use insulin injections?”
- Yes,
- No.”

“How long after diagnosis did you start the insulin?”

Table 51 presents the results.

### Table 51 Comparison of Version A With Version B - Current Insulin Use

<table>
<thead>
<tr>
<th>Insulin Use</th>
<th>Version A</th>
<th>Version B</th>
<th>A &amp; B Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
</tr>
<tr>
<td>Yes</td>
<td>32</td>
<td>14.9%</td>
<td>33</td>
</tr>
<tr>
<td>No</td>
<td>183</td>
<td>85.1%</td>
<td>187</td>
</tr>
<tr>
<td>Total</td>
<td>215</td>
<td>100.0%</td>
<td>220</td>
</tr>
</tbody>
</table>

There was no significant difference between the two versions for current insulin use (Fisher's exact test, \( p = 1.00 \)). They both have excellent reliability (K, A=0.97, B=1.00). This suggests the versions are equivalent.

There were too few responses to the follow-up questions to make a formal comparison between the two versions (Table 52). However the much larger number of non-responses to Version B suggests that Version A should be preferred for this question, particularly since it has greater reliability (ICC, A=0.99, B=0.72).

### Table 52 Version A Compared With Version B - Time Using Insulin

<table>
<thead>
<tr>
<th>Time Using Insulin</th>
<th>Number</th>
<th>Mean Time</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version A</td>
<td>14</td>
<td>27</td>
<td>6.18</td>
</tr>
<tr>
<td>Version B</td>
<td>3</td>
<td>44</td>
<td>17.44</td>
</tr>
</tbody>
</table>

**Recommendations 48:**

48. Use version A to determine the length of time of using insulin.

8.3.8 Complications

The two versions differed in the way in which they collected data about complications as a result of diabetes. Version A used three questions whereas Version B used a single question with more alternatives. The categories across the questions were identical.
Version A: “The next questions are about other health conditions you may have had. Have you ever experienced frequent tingling, pins and needles, burning or pain, or numbness in your legs and feet?
Yes, No,
Don’t know/refused.”

“Do you have any vision problems because of your diabetes or high blood sugar?
Yes, No,
Don’t know/refused.”

“Have you ever had any of the following?
Hypoglycaemic attack, that is low blood sugar,
A foot ulcer,
A limb amputation,
Kidney failure/disease,
None.”

Version B: “Have you ever had any of the following?
Frequent tingling, pins and needles, burning or pain, or numbness in your legs and feet,
Hypoglycaemic attack (low blood sugar),
Vision problems because of your diabetes,
A foot ulcer,
A limb amputation,
Kidney failure/disease,
None.”

Table 53 presents the results.

Table 53  Comparison of Version A with Version B - Complications of Diabetes

<table>
<thead>
<tr>
<th>Complications</th>
<th>Version A</th>
<th>Version B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>Numbness in your legs and feet</td>
<td>100</td>
<td>41.5%</td>
</tr>
<tr>
<td>Hypoglycaemic attack (low blood sugar)</td>
<td>38</td>
<td>18.7%</td>
</tr>
<tr>
<td>Vision problems because of your diabetes</td>
<td>45</td>
<td>15.8%</td>
</tr>
<tr>
<td>A foot ulcer</td>
<td>5</td>
<td>2.1%</td>
</tr>
<tr>
<td>A limb amputation</td>
<td>1</td>
<td>0.4%</td>
</tr>
<tr>
<td>Kidney failure/disease</td>
<td>11</td>
<td>4.6%</td>
</tr>
<tr>
<td>Total who replied to the questions</td>
<td>241</td>
<td></td>
</tr>
</tbody>
</table>

The frequency of responses was similar between the two versions, but there was a significant difference between the two versions for vision problems associated with diabetes (likelihood ratio $\chi^2=13.7$, d.f.=2, $p=0.001$). This difference is largely a consequence of the 4.1% of 'don't know' responses to Version A, with no such responses to Version B.

There does not appear to be any advantage in using three questions when one has produced the same information with fewer don't know responses, thus it is recommended that Version B be used if this item is retained.
Since the reliability for these items is very low (See Tables 17 and 20, Section A), it is recommended that these items be examined further. It is also implied by the context of the item that these conditions are related to diabetes, but this is not stated, other than for vision problems (where there were a number of ‘don't know’ responses). If this is the intent behind the item it should be stated in order to reduce confusion, but this may produce more ‘don't know’ responses.

**Recommendations 49, 50 and 51:**

49. Response categories for the complications of diabetes section be examined.

50. The wording of the question be revised to make it clear that the complications are related to diabetes.

51. Combining the complications into a multiple response question as in Version B is recommended as a basis for further development.
The time for each interview (minutes) was recorded. An examination of the distribution showed that it was negatively skewed. A natural log transformation produced a symmetrical distribution, but this was not normal, due to high kurtosis. Further analysis was conducted using the natural logarithm of the interview time. A comparison of interview times shows that there was a significant difference between the time to interview using Version A and Version B (Table 54). For the first interview the average time (geometric mean) for A was 11.9 minutes and for B it was 10.5 minutes, a difference of 1.4 minutes (8.5%) \((t=7.45, \, df=1517.0, \, P<.001, \, \text{equal variance not assumed})\). For the second interview the average time for A was 8.9 minutes and for B it was 7.8 minutes, an 8.1% difference \((t=7.47, \, df=1506, \, p<.001, \, \text{equal variance assumed})\).

<table>
<thead>
<tr>
<th>Number Interviewed</th>
<th>Arithmetic Mean Time</th>
<th>Std. Dev.</th>
<th>Geometric mean time</th>
<th>Geometric SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Contact</td>
<td>Second Contact</td>
<td>First Contact</td>
<td>Second Contact</td>
<td></td>
</tr>
<tr>
<td>815</td>
<td>734</td>
<td>768</td>
<td>774</td>
<td></td>
</tr>
<tr>
<td>12.50</td>
<td>9.44</td>
<td>11.22</td>
<td>8.31</td>
<td></td>
</tr>
<tr>
<td>4.34</td>
<td>3.70</td>
<td>4.57</td>
<td>3.30</td>
<td></td>
</tr>
<tr>
<td>11.88</td>
<td>8.91</td>
<td>10.48</td>
<td>7.8</td>
<td></td>
</tr>
<tr>
<td>1.36</td>
<td>1.40</td>
<td>1.43</td>
<td>1.39</td>
<td></td>
</tr>
</tbody>
</table>

As this might imply, there is also a significant difference between the time for the first and second interviews. For Version A, and the AA group, the mean difference between first and second was 1.35 minutes (paired \(t = 13.44, \, \text{d.f.} = 380, \, p<0.001\)). For Version B, and the BB group, the mean difference between first and second was 1.31 minutes (paired \(t = 13.76, \, \text{d.f.} = 372, \, p<0.001\)).

After controlling for version of interview, factors which influence the time of interview are age and a variety of health measures. Gender is not associated with interview time. Age and the two summary SF8 measures were regressed on interview time (natural log) for the first interview and were found to be roughly equally predictive of the score (Table 55).

### Regression Results for Factors Influencing Interview Time

<table>
<thead>
<tr>
<th>Regression Results for Factors Influencing Interview Time</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td>(Constant)</td>
<td>2.894</td>
<td>.081</td>
<td>35.859</td>
<td>.000</td>
</tr>
<tr>
<td>Version (B vs. A)</td>
<td>-.114</td>
<td>.016</td>
<td>-.169</td>
<td>-.098</td>
</tr>
<tr>
<td>Age</td>
<td>3.34E-03</td>
<td>.000</td>
<td>.174</td>
<td>6.825</td>
</tr>
<tr>
<td>Gender</td>
<td>-9.335E-03</td>
<td>.017</td>
<td>-.013</td>
<td>-5.60</td>
</tr>
<tr>
<td>SF8 PCS score</td>
<td>-4.875E-03</td>
<td>.001</td>
<td>-.154</td>
<td>-5.998</td>
</tr>
<tr>
<td>SF8 MCS score</td>
<td>-5.048E-03</td>
<td>.001</td>
<td>-.140</td>
<td>-5.760</td>
</tr>
</tbody>
</table>

A further model was developed to detect any other factors which may be important in predicting long interview times. These factors may be taken into consideration in the design of further interviews. Table 56 shows the results.
Table 56  Model 2 Predicting Interview Time

<table>
<thead>
<tr>
<th>Regression Results for Factors Influencing Interview Time</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td>(Constant)</td>
<td>2.738</td>
<td>.076</td>
<td></td>
<td>.000</td>
</tr>
<tr>
<td>Version (B vs. A)</td>
<td>-.116</td>
<td>.016</td>
<td>-.171</td>
<td>.000</td>
</tr>
<tr>
<td>Age</td>
<td>5.239E-03</td>
<td>.001</td>
<td>.272</td>
<td>.000</td>
</tr>
<tr>
<td>SF8 PCS score</td>
<td>-4.999E-03</td>
<td>.001</td>
<td>-.158</td>
<td>.000</td>
</tr>
<tr>
<td>SF8 MCS score</td>
<td>-4.367E-03</td>
<td>.001</td>
<td>-.121</td>
<td>.000</td>
</tr>
<tr>
<td>Unemployed</td>
<td>-.4567E-02</td>
<td>.049</td>
<td>-.022</td>
<td>.356</td>
</tr>
<tr>
<td>Home duties</td>
<td>-4.271E-04</td>
<td>.025</td>
<td>.000</td>
<td>.986</td>
</tr>
<tr>
<td>Student</td>
<td>9.099E-02</td>
<td>.047</td>
<td>.048</td>
<td>.052</td>
</tr>
<tr>
<td>Retired</td>
<td>-.103</td>
<td>.025</td>
<td>-.138</td>
<td>.000</td>
</tr>
<tr>
<td>TAFE</td>
<td>4.830E-02</td>
<td>.026</td>
<td>.058</td>
<td>.061</td>
</tr>
<tr>
<td>High school</td>
<td>-9.621E-03</td>
<td>.026</td>
<td>-.011</td>
<td>.712</td>
</tr>
<tr>
<td>Less than high school</td>
<td>3.693E-02</td>
<td>.023</td>
<td>.054</td>
<td>.115</td>
</tr>
<tr>
<td>English not spoken at home</td>
<td>.211</td>
<td>.032</td>
<td>.156</td>
<td>.000</td>
</tr>
</tbody>
</table>

The incorporation of education, employment status and English spoken at home added little to the model. Only being retired and not speaking English at home were significantly associated with the length of time to complete the interview after controlling for age and the summary SF8 scores. Those who were retired took less time compared with the employed, and those who speak another language at home also took longer.

One further analysis examined the influence of specific health conditions on interview length. The results are shown in Table 57.

Table 57  Model 3 Predicting Interview Time

<table>
<thead>
<tr>
<th>Regression Results for Factors Influencing Interview Time</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td>(Constant)</td>
<td>2.645</td>
<td>.078</td>
<td></td>
<td>.000</td>
</tr>
<tr>
<td>Version (B vs. A)</td>
<td>-.115</td>
<td>.016</td>
<td>-.169</td>
<td>.000</td>
</tr>
<tr>
<td>Age</td>
<td>2.428E-03</td>
<td>.001</td>
<td>.126</td>
<td>.000</td>
</tr>
<tr>
<td>SF8 PCS score</td>
<td>-3.207E-03</td>
<td>.001</td>
<td>-.102</td>
<td>.000</td>
</tr>
<tr>
<td>SF8 MCS score</td>
<td>-3.620E-03</td>
<td>.001</td>
<td>-.100</td>
<td>.000</td>
</tr>
<tr>
<td>Asthma</td>
<td>.119</td>
<td>.017</td>
<td>.169</td>
<td>7.087</td>
</tr>
<tr>
<td>Diabetes</td>
<td>.114</td>
<td>.020</td>
<td>.149</td>
<td>5.762</td>
</tr>
<tr>
<td>Depression</td>
<td>5.069E-02</td>
<td>.020</td>
<td>.062</td>
<td>2.502</td>
</tr>
<tr>
<td>High cholesterol</td>
<td>-7.973E-03</td>
<td>.020</td>
<td>-.011</td>
<td>.408</td>
</tr>
<tr>
<td>High blood pressure</td>
<td>3.019E-02</td>
<td>.018</td>
<td>.043</td>
<td>1.644</td>
</tr>
<tr>
<td>Heart disease</td>
<td>2.913E-02</td>
<td>.026</td>
<td>.028</td>
<td>1.102</td>
</tr>
<tr>
<td>Stroke</td>
<td>.106</td>
<td>.044</td>
<td>.058</td>
<td>2.404</td>
</tr>
<tr>
<td>Arthritis</td>
<td>-1.892E-02</td>
<td>.020</td>
<td>-.025</td>
<td>.936</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>2.777E-02</td>
<td>.031</td>
<td>.022</td>
<td>.892</td>
</tr>
</tbody>
</table>

As would be expected, given the asthma and diabetes modules, both of these conditions are associated with longer times. The only other conditions which are associated are depression (ever) and stroke, both of which are associated with increased times.
For most variables the reliability is very good. As a result there is little variance to allow for an examination of variables which might influence reliability. Thus this section will examine the influence of selected demographic factors on reliability of key variables where the reliability is limited.

The key determinants which will be examined are age, having a non-English language background, and education. These variables have been chosen as those which are most likely to have an influence on reliability.

Of the initial demographic variables which are common to both questionnaires, only those concerning telephone ownership have reduced reliability. Thus the reliability ($K_c$) of the number of telephones in the household (excluding mobile phones) and the number of White Pages listings will be explored. These variables were chosen because of their importance as weighting variables to control for selection bias of CATI surveys.

The Heterogeneity $\chi^2$ was used to test the null hypothesis that the reliability is homogeneous across the groups defined by the independent variable. If the Heterogeneity test is negative, that is, if the reliability estimates are statistically homogenous, then it can be concluded that the independent variable is not influencing reliability.

### 10.1 RELIABILITY: NUMBER OF TELEPHONES (EXCLUDING MOBILE PHONES)

Table 58 presents the reliability estimates across age groups. The test does not suggest that the null hypothesis should be rejected, and the test has substantial power given the sample size in each age group. This means that age does not appear to be influencing reliability. Nevertheless it should be noted that $K_c$ falls substantially above age 55, and becomes poor above age 75.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>18 - 24.9</td>
<td>0.72</td>
<td>110</td>
</tr>
<tr>
<td>25 - 34.9</td>
<td>0.72</td>
<td>192</td>
</tr>
<tr>
<td>35 - 44.9</td>
<td>0.73</td>
<td>282</td>
</tr>
<tr>
<td>45 - 54.9</td>
<td>0.71</td>
<td>262</td>
</tr>
<tr>
<td>55 - 64.9</td>
<td>0.63</td>
<td>280</td>
</tr>
<tr>
<td>65 - 74.9</td>
<td>0.56</td>
<td>255</td>
</tr>
<tr>
<td>75 +</td>
<td>0.32</td>
<td>171</td>
</tr>
</tbody>
</table>

Overall $K_c = 0.68$ (precision-based), Heterogeneity $\chi^2 = 6.57, p = 0.362$

Table 59 presents the reliability estimates for different education levels. The significant value of $\chi^2$ indicates that $K_c$ is not homogenous across the levels of education. This is consistent with the reduction beyond the first two levels. Education has a significant influence on reliability, particularly above high school level.

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tertiary</td>
<td>0.82</td>
<td>276</td>
</tr>
<tr>
<td>TAFE, trade Cert., diploma</td>
<td>0.73</td>
<td>338</td>
</tr>
<tr>
<td>Completed high school</td>
<td>0.56</td>
<td>312</td>
</tr>
<tr>
<td>Some high school</td>
<td>0.59</td>
<td>487</td>
</tr>
<tr>
<td>Less than high school</td>
<td>0.53</td>
<td>163</td>
</tr>
</tbody>
</table>

Overall $K_c = 0.71$ (precision-based), Heterogeneity $\chi^2 = 11.67, p = 0.020$
Table 60 compares the reliability of groups who do and do not speak English at home. Although reliability is lower for those who speak a language other than English at home, this difference is not significant.

**Table 60  Number of Telephones (Excluding Mobile Phones) by Non-English Spoken at Home**

<table>
<thead>
<tr>
<th>English Spoken at Home</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.68</td>
<td>1473</td>
</tr>
<tr>
<td>No</td>
<td>0.55</td>
<td>107</td>
</tr>
</tbody>
</table>

Overall $K = 0.68$ (precision-based), Heterogeneity $\chi^2 = 0.54$, $p = 0.463$

### 10.2 RELIABILITY: WHITE PAGES LISTINGS

Comparisons across groups have also been made with regard to the number of times telephone numbers appear in the White Page. Table 61 presents the different reliability estimates across age groups. The poor reliability for this variable is homogeneous across age categories and age does not influence the estimates.

**Table 61  Number of White Pages Listings by Age**

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>18 - 24.9</td>
<td>0.59</td>
<td>90</td>
</tr>
<tr>
<td>25 - 34.9</td>
<td>0.57</td>
<td>176</td>
</tr>
<tr>
<td>35 - 44.9</td>
<td>0.44</td>
<td>271</td>
</tr>
<tr>
<td>45 - 54.9</td>
<td>0.52</td>
<td>244</td>
</tr>
<tr>
<td>55 - 64.9</td>
<td>0.54</td>
<td>271</td>
</tr>
<tr>
<td>65 - 74.9</td>
<td>0.65</td>
<td>246</td>
</tr>
<tr>
<td>75 +</td>
<td>0.57</td>
<td>162</td>
</tr>
</tbody>
</table>

Overall $K = 0.55$ (precision-based), Heterogeneity $\chi^2 = 3.88$, $p = 0.692$

Table 62 presents the reliability estimates across education levels. As with age, there is no evidence to suggest that the poor reliability for this variable is a consequence of education as a confounder.

**Table 62  Number of White Pages Listings by Education Level**

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tertiary</td>
<td>0.49</td>
<td>255</td>
</tr>
<tr>
<td>TAFE, trade Cert., diploma</td>
<td>0.64</td>
<td>320</td>
</tr>
<tr>
<td>Completed high school</td>
<td>0.53</td>
<td>294</td>
</tr>
<tr>
<td>Some high school</td>
<td>0.46</td>
<td>455</td>
</tr>
<tr>
<td>Less than high school</td>
<td>0.62</td>
<td>157</td>
</tr>
</tbody>
</table>

Overall $K = 0.54$ (precision-based), Heterogeneity $\chi^2 = 5.05$, $p = 0.282$.

Table 63 presents the reliability estimates of those who speak English at home compared with those who do not. As with the other variables examined, there is no evidence of heterogeneity of reliability for this variable, suggesting there is no influence on the reliability estimates.
Table 63  Number of White Pages Listings by Non-English Spoken at Home

<table>
<thead>
<tr>
<th>English Spoken at Home</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.54</td>
<td>1388</td>
</tr>
<tr>
<td>No</td>
<td>0.63</td>
<td>94</td>
</tr>
</tbody>
</table>

Overall $K_a = 0.54$ (precision-based), Heterogeneity $\chi^2 = 0.42, p = 0.517$.

Of the later demographic variables which differ between versions, perceived financial status, education and income have the lowest reliability. These variables will be examined for influence by age, and non-English speaking background. The influence of education will be examined for perceived financial status and income.

10.3 RELIABILITY: PERCEIVED FINANCIAL STATUS

The question on perceived financial status was tested for heterogeneity and found significant heterogeneity across age groups (See Table 64). In particular, reliability falls substantially above age 55 and is low for those under 25. Perhaps this is a consequence of very varied financial circumstances of students and those who are retired.

Table 64  Perceived Financial Status by Age

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>18 - 24.9</td>
<td>0.61</td>
<td>110</td>
</tr>
<tr>
<td>25 - 34.9</td>
<td>0.76</td>
<td>188</td>
</tr>
<tr>
<td>35 - 44.9</td>
<td>0.78</td>
<td>279</td>
</tr>
<tr>
<td>45 - 54.9</td>
<td>0.70</td>
<td>259</td>
</tr>
<tr>
<td>55 - 64.9</td>
<td>0.58</td>
<td>274</td>
</tr>
<tr>
<td>65 - 74.9</td>
<td>0.54</td>
<td>250</td>
</tr>
<tr>
<td>75 +</td>
<td>0.47</td>
<td>163</td>
</tr>
</tbody>
</table>

Overall $K_a = 0.67$ (precision-based), Heterogeneity $\chi^2 = 44.51, p < 0.00001$.

Table 65 shows the influence of education on perceived financial status. As for age, there is significant heterogeneity in financial status across educational levels. This is largely because of the higher reliability in those with the most education. Below tertiary level there is much less heterogeneity.

Table 65  Perceived Financial Status by Education Level

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tertiary</td>
<td>0.74</td>
<td>275</td>
</tr>
<tr>
<td>TAFE, trade Cert., diploma</td>
<td>0.63</td>
<td>329</td>
</tr>
<tr>
<td>Completed high school</td>
<td>0.63</td>
<td>305</td>
</tr>
<tr>
<td>Some high school</td>
<td>0.63</td>
<td>478</td>
</tr>
<tr>
<td>Less than high school</td>
<td>0.60</td>
<td>162</td>
</tr>
</tbody>
</table>

Overall $K_a = 0.66$ (precision-based), Heterogeneity $\chi^2 = 9.94, p = 0.041$. 
Table 66 shows that speaking a language other than English at home has no influence on reliability.

**Table 66  Perceived Financial Status by Non-English Spoken at Home**

<table>
<thead>
<tr>
<th>Other Language Spoken at Home</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>0.65</td>
<td>1444</td>
</tr>
<tr>
<td>Yes</td>
<td>0.63</td>
<td>106</td>
</tr>
</tbody>
</table>

Overall $K_w = 0.65$ (precision-based), Heterogeneity $\chi^2 = 0.09$, $p = 0.587$.

10.4 RELIABILITY: EDUCATION LEVEL

Next the effect of age on education level reliability was tested. Table 67 presents the reliability estimates across age groups.

**Table 67  Education Level by Age**

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>18 - 24.9</td>
<td>0.79</td>
<td>102</td>
</tr>
<tr>
<td>25 - 34.9</td>
<td>0.80</td>
<td>182</td>
</tr>
<tr>
<td>35 - 44.9</td>
<td>0.82</td>
<td>257</td>
</tr>
<tr>
<td>45 - 54.9</td>
<td>0.77</td>
<td>239</td>
</tr>
<tr>
<td>55 - 64.9</td>
<td>0.76</td>
<td>241</td>
</tr>
<tr>
<td>65 - 74.9</td>
<td>0.64</td>
<td>213</td>
</tr>
<tr>
<td>75 +</td>
<td>0.70</td>
<td>145</td>
</tr>
</tbody>
</table>

Overall $K_w = 0.77$ (precision-based), Heterogeneity $\chi^2 = 18.25$, $p < 0.006$.

The reliability of education level varies with age. It tends to be lower at older ages, which is presumably a consequence of the longer period of time since education and may also be a consequence of changes in the education system over that longer period.

The next test was to determine whether or not speaking a language other than English at home affected the education reliability estimates. Table 68 presents the results that show there is no influence on reliability estimates.

**Table 68  Education Level by Non-English Spoken at Home**

<table>
<thead>
<tr>
<th>Other Language Spoken at Home</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>0.76</td>
<td>1310</td>
</tr>
<tr>
<td>Yes</td>
<td>0.84</td>
<td>96</td>
</tr>
</tbody>
</table>

Overall $K_w = 0.77$ (precision-based), Heterogeneity $\chi^2 = 2.95$, $p = 0.086$.

10.5 RELIABILITY: ASTHMA AND DIABETES

The final tests that were conducted examined the effects of age, education and language spoken at home on the most unreliable asthma and diabetes items. The choice of items was difficult because it could only be applied to questions which were sufficiently similar in both versions to allow the data to be combined. This provided sufficient sample size for a stratified analysis of heterogeneity to be conducted. In view of the different wording between the versions, this analysis was limited to cases where the same version was presented in the test-retest (Groups AA and BB). Four asthma items and two diabetes items were identified.
For asthma these were:
- Asthma symptoms in the past year,
- Still have/get asthma,
- Use of asthma relief medication,
- Use of asthma preventive medication.

For diabetes, these were:
- Initial treatment of dietary change,
- History of hypoglycaemic attack.

In view of the lack of association between non-English speaking status and reliability of demographic items, this further analysis was stratified by age group and education level.

10.5.1 Reliability, Age and Asthma

None of the items examined showed that reliability was influenced significantly by age. However, the power of some of these statistical tests was not great and it might be prudent to conduct these analyses again on bigger sample sizes. Tables 69 to 72 present the results.

| Table 69 | Reliability Estimates - Symptoms of Asthma in Previous 12 Months |
| --- | --- | --- |
| Age Group | Weighted Kappa | N |
| 18 - 24.9 | 0.50 | 24 |
| 25 - 34.9 | 0.81 | 34 |
| 35 - 44.9 | 0.81 | 54 |
| 45 - 54.9 | 0.62 | 50 |
| 55 - 64.9 | 0.82 | 56 |
| 65 - 74.9 | 0.62 | 38 |
| 75+ | 0.86 | 23 |

Overall $K_w = 0.77$ (precision-based), Heterogeneity $\chi^2 = 5.23$, $p < 0.514$.

| Table 70 | Reliability Estimates - Still Have/Get Asthma |
| --- | --- | --- |
| Age Group | Weighted Kappa | N' |
| 18 - 24.9 | 0.68 | 23 |
| 25 - 34.9 | 0.93 | 34 |
| 35 - 44.9 | 0.63 | 53 |
| 45 - 54.9 | 0.63 | 47 |
| 55 - 64.9 | 0.74 | 55 |
| 65 - 74.9 | 0.80 | 37 |
| 75+ | 0.35 | 20 |

a Note: There were 8 don’t know responses - these were not included in the analysis.

Overall $K_w = 0.80$ (precision-based), Heterogeneity $\chi^2 = 10.88$, $p < 0.092$. 
Table 71  Reliability Estimates - Use of Reliever Medication

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Weighted Kappa</th>
<th>N'</th>
</tr>
</thead>
<tbody>
<tr>
<td>18 - 24.9</td>
<td>1.00</td>
<td>15</td>
</tr>
<tr>
<td>25 - 34.9</td>
<td>0.77</td>
<td>21</td>
</tr>
<tr>
<td>35 - 44.9</td>
<td>0.88</td>
<td>42</td>
</tr>
<tr>
<td>45 - 54.9</td>
<td>0.64</td>
<td>44</td>
</tr>
<tr>
<td>55 - 64.9</td>
<td>0.45</td>
<td>47</td>
</tr>
<tr>
<td>65 - 74.9</td>
<td>0.63</td>
<td>34</td>
</tr>
<tr>
<td>75 +</td>
<td>0.55</td>
<td>17</td>
</tr>
</tbody>
</table>

a Note: There were 3 don't know responses - these were not included in the analysis.

Overall $K_w = 0.65$ (precision-based), Heterogeneity $\chi^2 = 2.97, p < 0.812$.

Table 72  Reliability Estimates - Use of Prevention Medication

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>18 - 24.9</td>
<td>0.86</td>
<td>15</td>
</tr>
<tr>
<td>25 - 34.9</td>
<td>0.67</td>
<td>21</td>
</tr>
<tr>
<td>35 - 44.9</td>
<td>0.47</td>
<td>42</td>
</tr>
<tr>
<td>45 - 54.9</td>
<td>0.57</td>
<td>44</td>
</tr>
<tr>
<td>55 - 64.9</td>
<td>0.49</td>
<td>47</td>
</tr>
<tr>
<td>65 - 74.9</td>
<td>0.65</td>
<td>34</td>
</tr>
<tr>
<td>75 +</td>
<td>0.32</td>
<td>19</td>
</tr>
</tbody>
</table>

Overall $K_w = 0.60$ (precision-based), Heterogeneity $\chi^2 = 7.53, p < 0.275$.

10.5.2  Reliability, Education Level and Asthma

As with age, education did not significantly affect reliability estimates of the chosen asthma items. Tables 73 to 76 present the results.

Table 73  Reliability Estimates - Symptoms of Asthma in the Previous 12 Months

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tertiary</td>
<td>0.73</td>
<td>47</td>
</tr>
<tr>
<td>TAFE, trade Cert., diploma</td>
<td>0.75</td>
<td>56</td>
</tr>
<tr>
<td>Completed high school</td>
<td>0.74</td>
<td>58</td>
</tr>
<tr>
<td>Didn't complete high school</td>
<td>0.71</td>
<td>121</td>
</tr>
</tbody>
</table>

Overall $K_w = 0.73$ (precision-based), Heterogeneity $\chi^2 = 0.08, p < 0.994$. 
Table 74  Reliability Estimates - Still Have/Get Asthma

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Weighted Kappa</th>
<th>N'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tertiary</td>
<td>0.81</td>
<td>44</td>
</tr>
<tr>
<td>TAFE, trade Cert., diploma</td>
<td>0.71</td>
<td>51</td>
</tr>
<tr>
<td>Completed high school</td>
<td>0.57</td>
<td>57</td>
</tr>
<tr>
<td>Didn’t complete high school</td>
<td>0.74</td>
<td>120</td>
</tr>
</tbody>
</table>

a Note: There were 8 don’t know responses - these were not included in the analysis.

Overall $K_w = 0.73$ (precision-based), Heterogeneity $\chi^2 = 1.58$, $p < 0.664$.

Table 75  Reliability Estimates - Use of Reliever Medication

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Weighted Kappa</th>
<th>N'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tertiary</td>
<td>0.87</td>
<td>36</td>
</tr>
<tr>
<td>TAFE, trade Cert., diploma</td>
<td>0.80</td>
<td>39</td>
</tr>
<tr>
<td>Completed high school</td>
<td>0.63</td>
<td>51</td>
</tr>
<tr>
<td>Didn’t complete high school</td>
<td>0.39</td>
<td>97</td>
</tr>
</tbody>
</table>

a Note: There were 3 don’t know responses - these were not included in the analysis.

Overall $K_w = 0.73$ (precision-based), Heterogeneity $\chi^2 = 5.84$, $p < 0.120$.

Table 76  Reliability Estimates - Use of Prevention Medication

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tertiary</td>
<td>0.71</td>
<td>36</td>
</tr>
<tr>
<td>TAFE, trade Cert., diploma</td>
<td>0.52</td>
<td>39</td>
</tr>
<tr>
<td>Completed high school</td>
<td>0.62</td>
<td>52</td>
</tr>
<tr>
<td>Didn’t complete high school</td>
<td>0.44</td>
<td>98</td>
</tr>
</tbody>
</table>

Overall $K_w = 0.58$ (precision-based), Heterogeneity $\chi^2 = 4.64$, $p < 0.200$.

10.5.3 Reliability, Age and Diabetes

As with the Asthma items, age did not show any statistically significant influence on reliability for the selected diabetes items. As there were too few cases in the younger age groups for these items, ages 18 up to 44 were combined.

Tables 77 and 78 present the results.

Table 77 Reliability Estimates - Dietary Changes as Treatment for Diabetes

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 45</td>
<td>0.56</td>
<td>24</td>
</tr>
<tr>
<td>45 - 54.9</td>
<td>0.31</td>
<td>31</td>
</tr>
<tr>
<td>55 - 64.9</td>
<td>0.35</td>
<td>53</td>
</tr>
<tr>
<td>65 - 74.9</td>
<td>0.55</td>
<td>61</td>
</tr>
<tr>
<td>75 +</td>
<td>0.36</td>
<td>46</td>
</tr>
</tbody>
</table>

Overall $K_w = 0.45$ (precision-based), Heterogeneity $\chi^2 = 2.37$, $p < 0.668$. 
Table 78  Reliability Estimates - Hypoglycaemic Attack

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 45</td>
<td>0.47</td>
<td>24</td>
</tr>
<tr>
<td>45 - 54.9</td>
<td>0.76</td>
<td>31</td>
</tr>
<tr>
<td>55 - 64.9</td>
<td>0.55</td>
<td>53</td>
</tr>
<tr>
<td>65 - 74.9</td>
<td>0.61</td>
<td>61</td>
</tr>
<tr>
<td>75 +</td>
<td>0.79</td>
<td>46</td>
</tr>
</tbody>
</table>

Overall $K_w = 0.64$ (precision-based), Heterogeneity $\chi^2 = 2.52$, $p < 0.641$.

10.5.4 Reliability, Education and Diabetes

Similarly, education did not influence reliability on the two selected diabetes items. Tables 79 and 80 present the results.

Table 79  Reliability Estimates - Dietary Changes as Treatment for Diabetes

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tertiary</td>
<td>0.49</td>
<td>22</td>
</tr>
<tr>
<td>TAFE, trade Cert., diploma</td>
<td>0.30</td>
<td>47</td>
</tr>
<tr>
<td>Completed high school</td>
<td>0.52</td>
<td>26</td>
</tr>
<tr>
<td>Didn’t complete high school</td>
<td>0.49</td>
<td>122</td>
</tr>
</tbody>
</table>

Overall $K_w = 0.46$ (precision-based), Heterogeneity $\chi^2 = 1.40$, $p < 0.706$.

Table 80  Reliability Estimates - Hypoglycaemic Attack

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Weighted Kappa</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tertiary</td>
<td>0.86</td>
<td>22</td>
</tr>
<tr>
<td>TAFE, trade Cert., diploma</td>
<td>0.69</td>
<td>47</td>
</tr>
<tr>
<td>Completed high school</td>
<td>0.33</td>
<td>26</td>
</tr>
<tr>
<td>Didn’t complete high school</td>
<td>0.62</td>
<td>122</td>
</tr>
</tbody>
</table>

Overall $K_w = 0.67$ (precision-based), Heterogeneity $\chi^2 = 4.50$, $p < 0.213$. 
The National Health Survey, conducted in 2001, was used as the basis for comparing the prevalence of the selected health conditions. The ABS supplied a table of disease prevalence estimates by ten year age group. These were used to compare with the field testing prevalence estimates.

For this comparison, only randomly selected respondents were used in the analysis. Adults who had been selected because they were diagnosed as having diabetes or asthma were not included.

As the numbers were small for many age groups, indirect standardisation was used to calculate the expected numbers of cases which would have occurred in the survey if it had been applied to the same population as the reference survey (the NHS 2001 Health Survey), correcting for any age differences between the two surveys. These expected numbers were then compared with observed values and the $\chi^2$ test was used to test the hypothesis that the ratio of observed to expected numbers was unity (d.f. = 1). Table 81 presents the results.

<table>
<thead>
<tr>
<th>Condition</th>
<th>First Contact</th>
<th></th>
<th></th>
<th></th>
<th>Second Contact</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed</td>
<td>Expected</td>
<td>Ratio</td>
<td>Chi Square</td>
<td>Observed</td>
<td>Expected</td>
<td>Ratio</td>
</tr>
<tr>
<td>Heart Condition</td>
<td>68</td>
<td>20.2</td>
<td>3.4</td>
<td>113.52</td>
<td>64</td>
<td>18.7</td>
<td>3.4</td>
</tr>
<tr>
<td>Stroke</td>
<td>14</td>
<td>7.7</td>
<td>1.8</td>
<td>5.10</td>
<td>*</td>
<td>13</td>
<td>7.2</td>
</tr>
<tr>
<td>High Blood Pressure (current)</td>
<td>112</td>
<td>138.8</td>
<td>0.8</td>
<td>5.17</td>
<td>*</td>
<td>90</td>
<td>131.2</td>
</tr>
<tr>
<td>Arthritis</td>
<td>185</td>
<td>182.4</td>
<td>1.0</td>
<td>0.04</td>
<td>ns</td>
<td>166</td>
<td>172.1</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>54</td>
<td>22.3</td>
<td>2.4</td>
<td>45.24</td>
<td>***</td>
<td>51</td>
<td>20.8</td>
</tr>
<tr>
<td>High Blood Cholesterol (current)</td>
<td>56</td>
<td>80.7</td>
<td>0.7</td>
<td>7.56</td>
<td>**</td>
<td>51</td>
<td>76.7</td>
</tr>
<tr>
<td>Asthma (current)</td>
<td>100</td>
<td>91.7</td>
<td>1.1</td>
<td>0.75</td>
<td>ns</td>
<td>100</td>
<td>86.4</td>
</tr>
<tr>
<td>Diabetes</td>
<td>62</td>
<td>39.8</td>
<td>1.6</td>
<td>12.35</td>
<td>***</td>
<td>64</td>
<td>37.6</td>
</tr>
<tr>
<td>Feeling Depressed (current)</td>
<td>58</td>
<td>5.8</td>
<td>10.1</td>
<td>473.47</td>
<td>****</td>
<td>58</td>
<td>5.4</td>
</tr>
</tbody>
</table>

*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

For many conditions there is little or no difference between the prevalence estimates produced by this survey and the National Health Survey, 2001 (high blood pressure, arthritis, high blood cholesterol, and asthma). For other conditions, such as stroke and diabetes, there are bigger differences which may be a consequence of selection biases which operate differently for the two surveys. For heart disease, osteoporosis and depression, the difference is very large. Such a big difference suggests that the survey instruments are producing substantial measurement bias. This would need further examination.

**Recommendations 52 and 53:**

52. Further examination should be conducted to determine what factors may be influencing the differences in prevalence between the 2001 NHS and the CATI survey

53. Other standards for comparison with CATI surveys be investigated


11. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33::159-74.


## APPENDIX 2 - COMPARISONS OF DEMOGRAPHIC QUESTIONS ACROSS DATA COLLECTIONS

Source of demographic question formats determined for inclusion in the CATI field-testing survey, as agreed at the CATI TRG Teleconference 10th April 2002.

<table>
<thead>
<tr>
<th>Question Topic</th>
<th>Questionnaire A</th>
<th>Questionnaire B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>NSW Continuous Health Survey</td>
<td>CATI SNAP(S) Workshop</td>
</tr>
<tr>
<td>Marital status</td>
<td>ABS Cognitive Testing</td>
<td>NSW Continuous Health Survey</td>
</tr>
<tr>
<td>Birth country</td>
<td>ABS NHS 2001</td>
<td>ABS NHS 2001</td>
</tr>
<tr>
<td>Year of arrival in Australia</td>
<td>ABS Cognitive Testing</td>
<td>CATI TRG</td>
</tr>
<tr>
<td>Languages spoken</td>
<td>ABS Cognitive Testing</td>
<td>NSW Continuous Health Survey</td>
</tr>
<tr>
<td>Education</td>
<td>NSW Continuous Health Survey with CATI (SNAP(S) options for part 2 of question (highest qualification). Options for part 2 read out by interviewer.</td>
<td>NSW Continuous Health Survey with CATI (SNAP(S) options for part 2 of question (highest qualification). Options for part 2 not read out by interviewer.</td>
</tr>
<tr>
<td>Work status</td>
<td>ABS Cognitive Testing</td>
<td>CATI SNAP(S) Workshop with unemployed category divided into 1 year or more than 1 year.</td>
</tr>
<tr>
<td>Household size/structure</td>
<td>ABS Cognitive Testing</td>
<td>ABS Cognitive Testing</td>
</tr>
<tr>
<td>Housing tenure</td>
<td>NSW Continuous Health Survey</td>
<td>CATI TRG</td>
</tr>
<tr>
<td>Income adequacy</td>
<td>ABS Cognitive Testing</td>
<td>CATI TRG</td>
</tr>
</tbody>
</table>
APPENDIX 3 - QUESTIONNAIRE A

CATI Question Validation Survey - Questionnaire A
September 2002

A INTRODUCTION

A1 Record sample source of respondent.
1 Known diabetic [    ]
2 Known asthmatic [    ]
3 EWP random selection [    ]
If A.1 = 1 or 2, go to Intro A4

Intro A1

Good ....... My name is ....... from the South Australian Dept of Human Services / Western Australian Department of Health. We are conducting a survey about the health of Australians.

Intro A2

We recently sent you a letter telling you about the survey. Did you receive the letter?
(Single Response)
1 Yes [    ]
2 No [    ]
3 Don't know [    ]

Intro A3

Could I please speak with the person in the household, aged 18 and over, who was the last to have a birthday.

Repeat Intro A1 [and intro A2 if A1 = 2]
Go to Intro A6

Intro A4

Good.........My name is .........and I am calling on behalf of the South Australian Department of Human Services / Western Australian Department of Health

Please may I speak to...state name of required person.

When speaking to required household member, read Intro A5.

Intro A5

Hello, my name is .......and I am calling on behalf of the South Australian Dept of Human Services / Western Australian Department of Health. As you took part in a previous health survey and said (at that time) you would be willing to answer some more questions, please may I ask you to do that now (if convenient) or may I call you back. This survey should take no more than ten minutes.

Intro A6

I can assure you that information given will remain confidential. The answers from all people interviewed will be gathered together and presented in a report. No individual answers will be passed on.
**B DEMOGRAPHICS**

As some of the next questions relate to certain groups of people only I would like to ask you some general questions first of all.

**B.1 to 0 CATI SNAP(S) Workshop - December 2001**

**B.1** Could you please tell me your date of birth?
*(Single response)*
1. Date of birth \__/__/____
2. Not stated
   If B.1 = 1, go to B.4

**B.2** What was your age last birthday?
*(Single Response)*
1. Enter age
2. Not stated
   If B.2 = 1, go to B.5

**B.3** Which age group are you in? Would it be
*(Read Options. Single Response)*
1. 18 to 24
2. 25 to 34
3. 35 to 44
4. 45 to 54
5. 55 to 64
6. 65 to 74
7. 75 years or older
8. Refused Terminate interview

**B.4** Voice (ask if unsure)
1. Male
2. Female

**B.5** Including yourself how many people aged 18 years and over usually live in this household?
*(Single Response. Interviewer note: prompt respondent to ensure that respondent has included themselves. Enter number of people 18 years and over)*
1. Enter number
2. Not stated

**B.6**, **B.7**, **B.8**

**B.6 - WANTS - 2000**

B6 How many children (including babies) under 18 years live in your household?
*(Single Response. Interviewer note: enter number of babies and children under 18.)*
1. Enter number
2. None
3. Not stated

B7 What is your postcode?
*(Single Response)*
1. Enter postcode
2. Not stated
   If B.7 = 1, go to C.

B8 What is your suburb, town or community?
*(Single Response.)*
1. Enter town /suburb
2. Not stated
C  SF8 AND GENERAL HEALTH

Now we would like to ask you about your health.

C1  Overall, how would you rate your health during the past 4 weeks?
(Read options. Single response)
1  Excellent [    ]
2  Very good [    ]
3  Good [    ]
4  Fair [    ]
5  Poor [    ]
6  Very poor [    ]

C2  During the past 4 weeks, how much did physical health problems limit your usual physical activities (such as walking or climbing stairs)?
(Read Options. Single Response)
1  Not at all [    ]
2  Very little [    ]
3  Somewhat [    ]
4  Quite a lot [    ]
5  Could not do physical activities [    ]

C3  During the past 4 weeks, how much difficulty did you have doing your daily work, both at home and away from home, because of your physical health?
(Read Options. Single Response)
1  None at all [    ]
2  A little bit [    ]
3  Some [    ]
4  Quite a lot [    ]
5  Could not do daily work [    ]

C4  How much bodily pain have you had during the past 4 weeks?
(Read Options. Single Response)
1  None [    ]
2  Very mild [    ]
3  Mild [    ]
4  Moderate [    ]
5  Severe [    ]
6  Very severe [    ]

C5  During the past 4 weeks, how much energy did you have?
(Read Options. Single Response)
1  Very much [    ]
2  Quite a lot [    ]
3  Some [    ]
4  A little [    ]
5  None [    ]

C6  During the past 4 weeks, how much did your physical health or emotional problems limit your usual social activities with family or friends?
(Read Options. Single Response)
1  Not at all [    ]
2  Very little [    ]
3  Somewhat [    ]
4  Quite a lot [    ]
5  Could not do social activities [    ]

C7  During the past 4 weeks, how much have you been bothered by emotional problems (such as feeling anxious, depressed or irritable)?
(Read Options. Single Response)
1  Not at all [    ]
2  Slightly [    ]
3  Moderately [    ]
4  Quite a lot [    ]
5  Extremely [    ]

C8  During the past 4 weeks, how much did personal or emotional problems keep you from doing your usual work, school or other daily activities?
(Read Options. Single Response)
1  Not at all [    ]
2  Very little [    ]
3  Somewhat [    ]
4  Quite a lot [    ]
5  Could not do daily activities [    ]

C9  Have you ever been told by a doctor that you have any of the following conditions?
(Read Options. Multiple response).
1  Heart disease [    ]
2  Stroke [    ]
3  Arthritis [    ]
4  Osteoporosis [    ]
5  None of these [    ]
6  Refused [    ]
C10 Have you ever been told by a doctor that you have high blood pressure?
(Single response. Interviewer note: include never measured response into no category)
1 Yes [ ]
2 No [ ]
3 Don't know [ ]
If C.10 > 1, go to C.12

C11 Do you still have high blood pressure?
(Single response)
1 Yes [ ]
2 No [ ]
3 Don't know [ ]

C12 Have you ever been told by a doctor that you have high cholesterol?
(Single response. Interviewer note: include never measured response into no category)
1 Yes [ ]
2 No [ ]
3 Don't know [ ]
If C.12 > 1, go to C.14

C13 Do you still have high cholesterol?
(Single response)
1 Yes [ ]
2 No [ ]
3 Don't know [ ]

C14 Have you ever been told by a doctor that you are depressed?
(Single response. Interviewer note: Only medically diagnosed depression equals "Yes", not depression diagnosed by a counsellor).
1 Yes [ ]
2 No [ ]
3 Don't know [ ]
4 Refused [ ]
If C.14 > 1, go to D.

C15 Are you still depressed?
(Single response. Interviewer note: record answers such as "somewhat" or "a little" as Yes)
1 Yes [ ]
2 No [ ]
3 Don't know [ ]
4 Refused [ ]

C16 Are you currently receiving treatment for depression?
(Single response. Interviewer note - includes telephone treatment)
1 Yes [ ]
2 No [ ]
3 Refused [ ]

Changing the subject, slightly...
D ASTHMA

Now some questions about asthma.

D1- Agreed TRG & ABS question

Have you ever been told by a doctor that you have asthma?
(Single Response)
1 Yes [    ]
2 No [    ]
3 Don't know / refused [    ]
If Q D.1 = 2, go to E.

D2 & D.3- ABS recommended questions

Excluding times when you have had coughs, colds or other infections, in the last 12 months, have you had symptoms of asthma?
(Single Response)
1 Yes [    ]
2 No [    ]
3 Not stated / refused [    ]

D3 In the last 12 months, have you taken treatment for asthma?
(Single Response)
1 Yes [    ]
2 No [    ]
3 Not stated / refused [    ]

D4 Do you still get asthma?
(Single Response)
1 Yes [    ]
2 No [    ]
3 Don't know / refused [    ]
If D.2 > 1 and D.3 > 1, go to E

D.5 & D.6 ABS recommended question formats.

D5 In the last 12 months, have you ever woken up during the night because of your asthma?
(Single Response)
1 Yes [    ]
2 No [    ]
3 Don't know / refused [    ]
If D.5 > 1, go to D.7

D7 In the last 12 months, on average about how often did this occur? Was it...
(Read Options. Single Response)
1 7 nights a week [    ]
2 4-6 nights a week [    ]
3 2-3 nights a week [    ]
4 Once a week [    ]
5 Less than once a week [    ]

D.7 & D.8 - ABS recommended question formats.

D7 In the last 12 months, have you consulted a doctor about your asthma?
(Single Response)
1 Yes [    ]
2 No [    ]
3 Don't know / refused [    ]
If D.7 > 1, go to D.9

D8 In the last 12 months, how often did you consult a doctor about your asthma? Would you say it was
(Read Options. Single Response)
1 Once a month or more frequently [    ]
2 Once every two to three months [    ]
3 Two to three times a year [    ]
4 About once [    ]
5 Don't know / Can't remember [    ]
**APPENDIX 3 - QUESTIONNAIRE A**

**D.9 & D.10 - ABS recommended question formats.**

**D9** In the last 12 months have you visited a hospital emergency department because of your asthma?

(Read options. Single Response)

1. Yes
2. No
3. Don't know / refused

If D.9 > 1, go to D.11

**D10** How many times in the last 12 months have you visited a hospital emergency department because of an asthma attack you had?

(Read options. Single Response)

1. Number of visits
2. None
3. Don't know / refused

**D11** A written asthma action plan contains instructions written by a medical professional on how to manage your asthma. Do you have a written asthma action plan from a medical professional?

(Read options. Single Response)

1. Yes
2. No
3. Don't know / refused

If Q D.14 > 1, go to D.16

**D.12 & D.13 - ABS questions.**

**D12** In the last 12 months, have you used any medication for the relief of your asthma?

(Read options. Single Response)

1. Yes
2. No
3. Don't know / refused

If D.12 > 1, go to D.14

**D13** In the last week, how often have you used reliever medication for your asthma? Was it...

(Read options. Single Response)

1. Every day
2. 4-6 days a week
3. 1-3 days a week
4. Less than once a week
5. Not at all
6. Don't know / refused

**D.14 & D.15 - ABS questions**

**D14** In the last 12 months, have you used preventer medication for your asthma?

(Read options. Single Response)

1. Yes
2. No
3. Don't know / refused

If Q D.14 > 1, go to D.16

**D15** In the last week, how often have you used preventer medication for your asthma? Was it...

(Read Options. Single Response)

1. Every day
2. 4-6 days a week
3. 1-3 days a week
4. Less than once a week
5. Not at all
6. Don't know / refused

**D.16 - ABS recommended question format.**

**D16** In the last 12 months, have you had any days off work, school or home duties due to your asthma?

(Read options. Single Response)

1. Yes
2. No
3. Don't know / refused

If D.12 > 1, go to D.14
**E DIABETES**

The next questions are about diabetes and high sugar levels.

**E.1- Agreed ABS & TRG question**

**E1 Have you ever been told by a doctor that you have diabetes?**  
*(Single Response.)*  
1 Yes \[    \]  
2 No \[    \]  
3 Don’t know / Refused \[    \]  
*If E.1 = 1, go to E.4*

**E.2 - ABS recommended question**

**E2 Have you ever been told by a doctor that you have high sugar levels in your blood or urine?**  
*(Single Response.)*  
1 Yes \[    \]  
2 No \[    \]  
3 Don’t know / refused \[    \]  
*If E.1 = 1 and B.4 = 1, go to E.8  
If E.2 = 1 and B.4 = 1, go to E.9  
If E.2 > 1, go to F.*

**E.3 Do you consider yourself at high risk of getting diabetes?**  
*(Single Response.)*  
1 Yes \[    \]  
2 No \[    \]  
3 Don’t know \[    \]  
4 Refused \[    \]  
*If E.1 = 1 and E.5 > 1, go to E.8  
If E.1 > 1 and E.5 > 1, go to E.9*

**E.4 What do you think increases a person’s risk of getting Diabetes?**  
*(Multiple Response. Interviewer note: prompt “anything else” after first response.)*  
1 Getting older \[    \]  
2 Being overweight \[    \]  
3 Being physically inactive \[    \]  
4 Having a stressed life \[    \]  
5 Having a family history \[    \]  
6 Genetics \[    \]  
7 Being of Aboriginal origin \[    \]  
8 Eating too much sugar \[    \]  
9 Poor diet / wrong diet / junk food \[    \]  
10 Being pregnant \[    \]  
11 Smoking \[    \]  
12 Other (specify) \[    \]  
*13 None of these \[    \]  
14 Don’t know \[    \]  
15 Refused \[    \]  
*If E.1 = 1 and B.4 = 1, go to E.8  
If E.2 = 1 and B.4 = 1, go to E.9  
If E.2 > 1, go to F.*

**E.5- ABS recommended question**

**E5 Were you pregnant when you were first told you had [diabetes / high sugar levels]?**  
*(Single Response.)*  
1 Yes \[    \]  
2 No \[    \]  
3 Don’t know / Refused \[    \]  
*If E.1 = 1 and E.5 > 1, go to E.8  
If E.1 > 1 and E.5 > 1, go to E.9*

**E.6 - ABS recommended question**

**E6 Other than when you were pregnant, have you ever been told that you have diabetes or high sugar levels?**  
*(Single Response.)*  
1 Yes \[    \]  
2 No \[    \]  
3 Don’t know / Refused \[    \]  
*If E.6 > 1, go to F.*
E7  Were you told you had diabetes or high sugar levels?
    (Single Response.)
    1  Diabetes [ ]
    2  High sugar levels [ ]
    If E.7 = 1, go to E.8
    If E.7 = 2, go to E.11

E8  [Other than the diabetes you had when you were pregnant] what type of diabetes were you told you had?
    (Single Response.)
    1  Type 1 - Insulin dependent,
        Juvenile onset [ ]
    2  Type 2 - Non-insulin dependent,
        Mature onset [ ]
    3  Don't know [ ]
    Go to E.10

E9  At what age were you first told you had diabetes / high sugar levels?
    (Single Response.)
    1  Less than one year of age [ ]
    2  Age in years [ ]
    3  Don't know / refused [ ]
    If B.1 = 2, go to E.12, other wise go to E.13

E10 [The following questions relate to Type 1 / Type 2 / other diabetes only]. At what age were you first told you had [Type 1 / Type 2 / other] diabetes?
    (Single Response.)
    1  Less than one year of age [ ]
    2  Age in years [ ]
    3  Don't know / refused [ ]
    If B.1 = 2, go to E.12, other wise go to E.13

E11 The following questions relate only to the high sugar levels you had when you were not pregnant. At what age were you first told you had high sugar levels?
    (Single Response)
    1  Less than one year of age [ ]
    2  Age in years [ ]
    3  Don't know / refused [ ]
    If B.1 = 2, go to E.12, other wise go to E.13

E12 [Other than the [diabetes/high sugar levels] you had while you were pregnant], were you first told you had [diabetes / high sugar levels] in the last 12 months?
    (Single response)
    1  Yes [ ]
    2  No [ ]
    3  Refused [ ]

E13 [Other than the [diabetes/high sugar levels] you had while you were pregnant], when you were first told you had [diabetes / high sugar levels], what treatments did the doctor put you on?
    (Read options. Multiple Response.)
    1  Dietary change [ ]
    2  Insulin Go to E.16 [ ]
    3  Tablets [ ]
    4  Lose weight [ ]
    5  Exercise [ ]
    6  Monitor blood glucose daily [ ]
    7  Nothing [ ]
    8  Don't know / refused [ ]
    9  Other [ ]
    If E.1 > 1, go to E.16

E14 Do you currently have insulin injections?
    (Single Response)
    1  Yes [ ]
    2  No [ ]
    If E.14 > 1, go to E.16
E15  For how long have you been having insulin injections for your diabetes?
(Single Response.)
1  Months (specify number) [  ]
2  Years (specify number) [  ]
3  Don't know / refused [  ]

E.16, E.17 & E.18 ABS recommended questions - conditions in C9 excluded from E.18, and, order of options amended.

E16  The next questions are about other health conditions you may have had. Have you ever experienced frequent tingling, pins and needles, burning or pain, or numbness in your legs or feet?
(Single Response)
1  Yes [  ]
2  No [  ]
3  Don't know / refused [  ]

E17  Do you have any vision problems because of your [diabetes / high sugar levels]?
(Single Response. Interviewer note: record yes only if respondent currently has vision problems which they know to be related to diabetes)
1  Yes [  ]
2  No [  ]
3  Don't know / refused [  ]

E18  Have you ever had any of the following?
(Read options. Multiple response)
1  Hypoglycaemic attack, that is, low blood sugar [  ]
2  A foot ulcer [  ]
3  A limb amputation [  ]
4  Kidney failure / disease [  ]
5  None [  ]
F SOCIAL CHARACTERISTICS

Now, some general questions to complete the survey.

F1 Which of these descriptions fits you best?
Last week, were you…
(Read Options. Single Response)
1 Self employed [ ]
2 Employed for wages, salary or payment in kind [ ]
3 Unemployed [ ]
4 Engaged in home duties [ ]
5 A student [ ]
6 Retired [ ]
7 Unable to work [ ]
8 Other (Specify) [ ]
9 Refused [ ]
If F1 = 1 or 2, go to F2
If F1 = 3 Go to F3
Otherwise Go to F4

F2 How many hours did you work last week in total?
(Single Response)
1 Enter number of hours per week [ ___ ]
2 Don’t know [ ]
Go to F4

F3 Have you been unemployed for more than one year?
(Single Response)
1 More than one year [ ]
2 One year or less [ ]

F4 Which of the following best describes your current marital status? Are you…
(Read Options. Single Response. Interviewer note: defacto equals Living with partner)
1 Married [ ]
2 Living with partner [ ]
3 Widowed [ ]
4 Divorced [ ]
5 Separated [ ]
6 Never Married [ ]
7 Not stated [ ]

F5 In which country were you born?
(Single Response)
1 Australia(1101) [ ]
2 England(2102) [ ]
3 New Zealand(1201) [ ]
4 Italy(3104) [ ]
5 Vietnam(5105) [ ]
6 Scotland(2105) [ ]
7 Greece(3207) [ ]
8 Germany(2304) [ ]
9 Philippines(5204) [ ]
10 Netherlands(2308) [ ]
11 Other (specify) [ ]

F6 In which year did you first arrive in Australia to live here for one year or more?
(Single Response)
1 Enter year [ ____ ]
2 Don’t know [ ]
F.7 - ABS Cognitive Testing 2001

F7 Do you speak a language, other than English, at home? (Single Response)
1 Yes
2 No
3 Refused
If F.7 > 1, go to F10

F.8 - ABS Cognitive Testing 2001

F8 Which language or languages do you speak at home? (Single Response. Interviewer note: record English if this is volunteered by respondent. Where respondents speaks many dialects of a language, (such as Mandarin Chinese) record only the language.
1 Specify
   1a__________
   1b__________
   1c__________
2 Refused
(Code language to ASCL)

F.9 - ABS Cognitive Testing 2001

F9 Which language do you use most often? (Single Response)
1 Specify
2 Refused
(Code language to ASCL)

F.10 CATI SNAP(S) Workshop -Dec 2001

F10 What is the highest level of education you have completed? Have you completed...
(Read Options. Single Response)
1 University, CAE or some other tertiary institute degree, including post-university (ie Post Graduate Diploma, Master, PhD)
2 TAFE or trade certificate or diploma
3 Completed high school
4 Some high school
5 Completed primary school
6 Some primary school
7 Never attended school
8 Other (specify) [__________]
9 Don't know
10 Refused

F.11 - NSW Continuous Health Survey - February 2002

F11 I would like to ask some questions about your housing arrangements. Are you: (Read options. Single Response)
1 Paying rent or board
2 Paying off this dwelling
3 Outright owner / fully owned
4 Living rent free
5 Purchasing under a rent / buy scheme
6 Occupying under a life tenure scheme
7 Other (specify)
8 Don't know
9 Refused
F12  Thinking of your [household's] financial situation over the last 12 months, which one of the following statements best describes your [household's] financial situation?  

*Read options. Single Response*

1. I am/we are spending more money than [I/we] get [ ]
2. [I/we] have just enough money to get [us] through to the next pay day [ ]
3. There's some money left over each week but [I/we] just spend it [ ]
4. [I/we] can save a bit now and then [ ]
5. [I/we] can save a lot [ ]
6. Don't know [ ]
7. Refused [ ]

F13  I would now like to ask you about your household's income. We are interested in how income relates to health, lifestyle and access to health services. Before tax is taken out, which of the following ranges best describes your household's income, from all sources, over the last 12 months?  

*Read options. Single response*

1. Less than $10,000 [ ]
2. $10,001 - $20,000 [ ]
3. $20,001 - $40,000 [ ]
4. $40,001 - $60,000 [ ]
5. $60,001 - $80,000 [ ]
6. Over $80,000 [ ]
7. Don't know / not sure [ ]
8. Refused [ ]
G ADMINISTRATION

G1 How many residential telephone numbers, excluding mobile phones, can be used to speak to someone in this household?
(Single Response. Interviewer note: do not include Internet or fax numbers)
1 Enter number [    ]
2 Don't know/refused [    ]

G2 How many residential telephone numbers, including mobile phones, can be used to speak to someone in this household?
(Single Response. Interviewer note: do not include Internet or fax numbers)
1 Enter number [    ]
2 Don't know/refused [    ]

G3 How many times [do these/does this] number(s) appear in the White Pages?
(Single Response. Interviewer note: do not include Internet or fax numbers. Total number of entries includes numbers that are listed more than once.)
1 Enter number [    ]
2 Don't know/refused [    ]

G4 In a survey like this, issues often arise which require further explanation. If we need to could we phone you at a later date to help clarify some issues?
(Single Response)
1 Yes (specify - record first name only) [    ]
2 No [    ]

G5 Date of interview [    ]

G6 Day of week interview undertaken [    ]

G7 Time of day interview undertaken [    ]

That concludes the survey. On behalf of the South Australian Dept of Human Services/Western Australian Department of Health, thank you very much for taking part in this survey.
A INTRODUCTION

A1 Record sample source of respondent.
1 Known diabetic [ ]
2 Known asthmatic [ ]
3 EWP random selection [ ]
If A1 = 1 or 2, go to Intro A4.

Intro A1

Good ....... My name is ....... from the South Australian Dept of Human Services / Western Australian Department of Health. We are conducting a survey about the health of Australians.

Intro A2

A2 We recently sent you a letter telling you about the survey. Did you receive the letter?
(Single Response)
1 Yes [ ]
2 No [ ]
3 Don't know [ ]

Intro A3

[If A2 > 1 read “The survey will gather information about the general health of Australians, and, about certain health conditions. To make sure that we gather information on both males and females of all age groups, we are wanting to speak with the person in the household, aged 18 and over, who was the last to have a birthday.”]

Intro A4

Good..........My name is ...........and I am calling on behalf of the South Australian Department of Human Services / Western Australian Department of Health.

Please may I speak to...state name of required person.

Intro A5

Hello, my name is .......and I am calling on behalf of the South Australian Dept of Human Services / Western Australian Department of Health. As you took part in a previous health survey and said (at that time) you would be willing to answer some more questions, please may I ask you to do that now (if convenient) or may I call you back. This survey should take no more than ten minutes.

Intro A6

I can assure you that information given will remain confidential. The answers from all people interviewed will be gathered together and presented in a report. No individual answers will be passed on.

Intro A7

Could I please speak with the person in the household, aged 18 and over, who was the last to have a birthday.

Repeat Intro A1 [and intro A2 if A1 = 2]
Go to Intro A6.
B DEMOGRAPHICS

As some of the next questions relate to certain groups of people only I would like to ask you some general questions first of all.

B1 & B2 - currently used CATI questions.

B1 What was your age last birthday?  
(Single Response.)  
1 Enter age in years [ ]  
2 Not stated [ ]  
If Q B1 = 1, go to B3

B2 Which age group are you in? Would it be  
(Read Options. Single Response)  
1 18 to 24 [ ]  
2 25 to 34 [ ]  
3 35 to 44 [ ]  
4 45 to 54 [ ]  
5 55 to 64 [ ]  
6 65 to 74 [ ]  
7 75 years or older [ ]  
8 Refused Terminate interview [ ]

B3 Voice (ask if unsure)  
1 Male [ ]  
2 Female [ ]

B4 Including yourself how many people aged 18 years and over usually live in this household?  
(Single Response. Interviewer note: prompt respondent to ensure that respondent has included themselves. Enter number of people 18 years and over)  
1 Enter number [ ]  
2 Not stated [ ]

B5 How many children (including babies) under 18 years live in your household?  
(Single Response. Interviewer note: enter number of babies and children under 18.)  
1 Enter number [ ]  
2 None [ ]  
3 Not stated [ ]

B6 What is your postcode?  
(Single Response)  
1 Enter postcode [ ]  
2 Not stated [ ]  
If B6 = 1, go to C.

B7 What is your suburb, town or community?  
(Single Response.)  
1 Enter town /suburb [ ]  
2 Not stated [ ]
C SF8 AND GENERAL HEALTH

Now we would like to ask you about your health.

C1 Overall, how would you rate your health during the past 4 weeks?
(Read options. Single response)
1 Excellent [ ]
2 Very good [ ]
3 Good [ ]
4 Fair [ ]
5 Poor [ ]
6 Very poor [ ]

C2 During the past 4 weeks, how much did physical health problems limit your usual physical activities (such as walking or climbing stairs)?
(Read options. Single response)
1 Not at all [ ]
2 Slightly [ ]
3 Somewhat [ ]
4 Quite a lot [ ]
5 Could not do physical activities [ ]

C3 During the past 4 weeks, how much difficulty did you have doing your daily work, both at home and away from home, because of your physical health?
(Read options. Single response)
1 None at all [ ]
2 A little bit [ ]
3 Some [ ]
4 Quite a lot [ ]
5 Could not do daily work [ ]

C4 How much bodily pain have you had during the past 4 weeks?
(Read options. Single response)
1 None [ ]
2 Very mild [ ]
3 Mild [ ]
4 Moderate [ ]
5 Severe [ ]
6 Very severe [ ]

C5 During the past 4 weeks, how much energy did you have?
(Read options. Single response)
1 Very much [ ]
2 Quite a lot [ ]
3 Some [ ]
4 A little [ ]
5 None [ ]

C6 During the past 4 weeks, how much did your physical health or emotional problems limit your usual social activities with family or friends?
(Read options. Single response)
1 Not at all [ ]
2 Very little [ ]
3 Somewhat [ ]
4 Quite a lot [ ]
5 Could not do social activities [ ]

C7 During the past 4 weeks, how much have you been bothered by emotional problems (such as feeling anxious, depressed or irritable)?
(Read options. Single response)
1 Not at all [ ]
2 Slightly [ ]
3 Moderately [ ]
4 Quite a lot [ ]
5 Extremely [ ]

C8 During the past 4 weeks, how much did personal or emotional problems keep you from doing your usual work, school or other daily activities?
(Read options. Single response)
1 Not at all [ ]
2 Very little [ ]
3 Somewhat [ ]
4 Quite a lot [ ]
5 Could not do daily activities [ ]

C9 Have you ever been told by a doctor that you have any of the following conditions?
(Read Options. Multiple response).
1 Heart disease [ ]
2 Stroke [ ]
3 Arthritis [ ]
4 Osteoporosis [ ]
5 None of these [ ]
6 Refused [ ]

C10 Have you ever been told by a doctor that you have high blood pressure?
(Single response. Interviewer note: include never measured response into no category)
1 Yes [ ]
2 No [ ]
3 Don't know [ ]

If C10 > 1, go to C12
C11  Do you still have high blood pressure?
     (Single response)
     1  Yes  [  ]
     2  No   [  ]
     3  Don't know  [  ]

C12  Have you ever been told by a doctor that you have high cholesterol?
     (Single response. Interviewer note: include never measured response into no category)
     1  Yes  [  ]
     2  No   [  ]
     3  Don't know  [  ]
     If C12 > 1, go to C14

C13  Do you still have high cholesterol?
     (Single response)
     1  Yes  [  ]
     2  No   [  ]
     3  Don't know  [  ]

Changing the subject, slightly...

C14  Have you ever been told by a doctor that you are depressed?
     (Single response. Interviewer note: Only medically diagnosed depression equals "Yes", not depression diagnosed by a counsellor)
     1  Yes  [  ]
     2  No   [  ]
     3  Don't know  [  ]
     4  Refused  [  ]
     If C14 > 1, go to D

C15  Are you still depressed?
     (Single response. Interviewer note: record answers such as "somewhat" or "a little" as Yes)
     1  Yes  [  ]
     2  No   [  ]
     3  Don't know  [  ]
     4  Refused  [  ]

C16  Are you currently receiving treatment for depression?
     (Single response. Interviewer note - includes telephone treatment)
     1  Yes  [  ]
     2  No   [  ]
     3  Refused  [  ]
D ASTHMA

Now some questions about asthma.

D1- agreed TRG & ABS Question

D1 Have you ever been told by a doctor that you have asthma?
(Single Response)
1 Yes [   ]
2 No [   ]
3 Don't know [   ]
If D1 = 2, go to E.

D2 & D3 - currently used CATI questions

D2 Symptoms of asthma include cough, wheezing, shortness of breath and chest tightness when you don't have a cold or respiratory infection.

During the past 12 months, did you have any symptoms of asthma?
(Single Response)
1 Yes [   ]
2 No [   ]
3 Don't know/refused [   ]

D3 During the past 12 months, did you take asthma medication that was prescribed or given to you by a doctor? This includes using an inhaler, puffer or nebuliser.
(Single Response)
1 Yes [   ]
2 No [   ]
3 Don't know/refused [   ]

D4 - TRG question

D4 Do you still have asthma?
(Single Response)
1 Yes [   ]
2 No [   ]
3 Don't know/refused [   ]
If D2 > 1 and D3 > 1, go to E.

D5- TRG question

D5 How often do you awaken during the night with asthma?
(Read options. Single Response)
1 Nightly [   ]
2 Most nights [   ]
3 Twice weekly [   ]
4 Weekly [   ]
5 Monthly [   ]
6 Less often than monthly [   ]
7 Only certain times of the year [   ]
8 Never [   ]
9 Refused [   ]

D6- TRG question

D6 How often do you see your doctor about your asthma? Would you say...
(Read options. Single Response)
1 Once a month or more frequently [   ]
2 Once every two to three months [   ]
3 Two to three times a year [   ]
4 About once a year [   ]
5 Less often than once a year [   ]
6 Don't know/Can't remember/refused [   ]

D7- TRG question

D7 How many times in the last 12 months have you visited a hospital emergency department for an attack of asthma?
(Single Response)
1 Number of visits [   ]
2 None [   ]
3 Not stated/refused [   ]

D8 Do you have a written asthma action plan from a medical professional? That is, written instructions on how to manage your asthma.
(Single Response)
1 Yes [   ]
2 No [   ]
3 Don't know/refused [   ]
**D9 & D10 - TRG question**

**D9** Do you use a reliever medication (puffer or nebuliser) such as Ventolin, Asmol or Bricanyl?  
*(Single Response. - Interviewer note If respondent claims to use a reliever or preventer that is not included in the medications which are read out, record "Yes")*

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Don't know / refused</td>
<td>[ ]</td>
<td></td>
</tr>
</tbody>
</table>

*If D9 > 1, go to D11*

**D10** In the last month, how often have you used reliever medication?  
*(Read options. Single Response)*

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Every day</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Most days</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>About half the days</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Less than half the days</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Not at all</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Don't know / refused</td>
<td>[ ]</td>
<td></td>
</tr>
</tbody>
</table>

*If D13 = 2, go to E.*

**D11 & D12 - TRG question**

**D11** Do you use preventer medication such as Becotide, Becloforte, Aldecin, Pulmicort, Flixotide, Intal forte, Cromogen or Tilade?  
*(Single Response. - Interviewer note If respondent claims to use a reliever or preventer that is not included in the medications which are read out, record "Yes")*

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Don't know / refused</td>
<td>[ ]</td>
<td></td>
</tr>
</tbody>
</table>

*If D11 > 1. Go to D13*

**D12** In the last month, how often have you used preventer medication?  
*(Read options. Single Response)*

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Every day</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Most days</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>About half the days</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Less than half the days</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Not at all</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Don't know / refused</td>
<td>[ ]</td>
<td></td>
</tr>
</tbody>
</table>

**D13 - TRG question**

**D13** In the last twelve months have you had any days lost from work, school or home duties from asthma?  
*(Single Response)*

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>[ ]</td>
<td></td>
</tr>
</tbody>
</table>

*If D13 = 2, go to E.*

**D14** How many days would you estimate?  
*(Single Response.)*

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter number of days</td>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Don't know</td>
<td>[ ]</td>
<td></td>
</tr>
</tbody>
</table>
E DIABETES
The next questions are about diabetes and high sugar levels.

E1 Agreed ABS & TRG question

E1 Have you ever been told by a doctor that you have diabetes?
(Single Response.)
1 Yes [ ]
2 No [ ]
3 Don’t know/Refused [ ]
If E1 = 1, go to E4

E2 Have you ever been told by a doctor that you have high blood sugar or a touch of sugar?
(Single Response.)
1 Yes [ ]
2 No [ ]
3 Don’t know/refused [ ]

E3 Do you consider yourself at high risk of getting diabetes?
(Single Response.)
1 Yes [ ]
2 No [ ]
3 Don’t know [ ]
4 Refused [ ]

E4 What do you think increases a person’s risk of getting Diabetes?
(Multiple Response. Interviewer note: prompt “anything else” after first response).
1 Getting older [ ]
2 Being overweight [ ]
3 Being physically inactive [ ]
4 Having a stressful life [ ]
5 Having a family history [ ]
6 Genetics [ ]
7 Being of Aboriginal origin [ ]
8 Eating too much sugar [ ]
9 Poor diet/wrong diet/junk food [ ]
10 Being pregnant [ ]
11 Smoking [ ]
12 Other (specify) __________________________

13 None of these [ ]
14 Don’t know [ ]
15 Refused [ ]

If E1 = 1 and B4 = 1, go to E8
If E2 = 1 and B4 = 1, go to E9
If E2 > 1, go to F.

E5 Were you pregnant when you were first told you had diabetes or high blood sugar?
(Single Response.)
1 Yes [ ]
2 No [ ]
3 Don’t know/refused [ ]
If E1 = 1 and E5 > 1, go to E7
If E2 = 1 and E5 > 1, go to E10

E6 Have you ever been told that you have diabetes or high blood sugar other than when you were pregnant?
(Single Response.)
1 Yes [ ]
2 No [ ]
If E2 = 1 and E6 > 1, go to E11
If E1 = 1 and E6 > 1, go to E9
If E2 = 1 and E6 = 1, go to E10
APPENDIX 4 - QUESTIONNAIRE B

E7 - TRG recommended question

E7 When you were first told you had high blood sugar, what treatment did the doctor put you on?
(Read out. Multiple Response)
1 Dietary change [ ]
2 Insulin [ ]
3 Tablets [ ]
4 Lose weight [ ]
5 Exercise [ ]
6 Monitor blood glucose daily [ ]
7 Nothing [ ]
8 Don't know [ ]
9 Other [ ]
Go to F.

E8, E9, E10 & E11 - TRG recommended questions asked separately of those with diabetes / high blood sugar.

E8 When you were first told you had diabetes, how old were you when you first were told you had diabetes?
(Single Response)
1 Age in Years [ ]
2 Don't know [ ]

E9 When you were first told you had diabetes, what treatment did the doctor put you on?
(Read out. Multiple Response)
1 Dietary change [ ]
2 Insulin [ ]
3 Tablets [ ]
4 Lose weight [ ]
5 Exercise [ ]
6 Monitor blood glucose daily [ ]
7 Nothing [ ]
8 Don't know [ ]
9 Other [ ]
Go to E12

E10 When you were first told you had high blood sugar, how old were you when you first were told you had high blood sugar?
(Single Response)
1 Age in Years [ ]
2 Don't know [ ]

E12 Do you currently use insulin injections?
(Single Response)
1 Yes [ ]
2 No [ ]
If E9 = 2, go to E14
If E12 > 1, go to E14

E13 How long after diagnosis did you start the insulin?
(Single Response. Interviewer note: Prompt if necessary. Tick response 1 or 4, or record number of months or years)
1 Immediately [ ]
2 Months [ ]
3 Years [ ]
4 Don't know [ ]
**E14- TRG recommended question. Conditions included in C9 excluded.**

**E14** Have you ever had any of the following?  
*(Read options. Multiple response. Interviewer note: record vision problem only if the respondent knows that the problem is related to diabetes)*  
1. Frequent tingling, pins and needles, burning or pain or numbness in your legs or feet  
2. Hypoglycaemic attack, (low blood sugar)  
3. Vision problems because of your diabetes  
4. A foot ulcer  
5. A limb amputation  
6. Kidney failure / disease  
7. None of the above

**E15 & E16 For possible inclusion**

**E15** In the last 12 months have you had any days lost from work, school or home duties or usual activities due to diabetes?  
*(Single Response)*  
1. Yes  
2. No  
*If E15 = 2, Go to F.*

**E16** How many days would you estimate?  
*(Single Response. Enter 999 if not stated)*  
1. Enter days  
2. Don’t know
F SOCIAL CHARACTERISTICS

Now, some general questions to complete the survey.

F1 Which of these best describes your current employment status? Are you...
(Read Options. Single Response)
1 Self employed [ ]
2 Employed for wages, salary or payment in kind [ ]
3 Unemployed for less than one year [ ]
4 Unemployed for more than one year [ ]
5 Engaged In home duties [ ]
6 A student [ ]
7 Retired [ ]
8 Unable to work [ ]
9 Other (specify) [ ]
10 Refused [ ]
If F1 = 1 or 2 Go to F2,
Otherwise Go to F3

F2 How many hours do you work per week?
1 Specify number of hours [ ]
2 Not stated [ ]

F3 Besides yourself, who lives in your household?
(Multiple response. Read out options if necessary)
1 No one (lives alone) [ ]
2 Mother [ ]
3 Father [ ]
4 Partner [ ]
5 Stepmother [ ]
6 Stepfather [ ]
7 Grandparents [ ]
8 Sons/daughters [ ]
9 Stepsons/stepdaughters [ ]
10 Brothers/sisters [ ]
11 Step-brothers/sisters [ ]
12 Other relatives [ ]
13 Non-family members [ ]
14 Other (specify) [ ]
15 Don’t know [ ]
16 Refused [ ]

F4 What is your formal current marital status?
Are you...
(Read Options. Single Response)
1 Married [ ]
2 Widowed [ ]
3 Separated but not divorced [ ]
4 Divorced [ ]
5 Never Married [ ]
6 Don’t know [ ]
7 Refused [ ]
**F5** - NSW Continuous Health Survey - February 2002

**In which country were you born?**

*Single Response*

1. Australia (1101) [ ]
2. England (2102) [ ]
3. New Zealand (1201) [ ]
4. Italy (3104) [ ]
5. Vietnam (5105) [ ]
6. Scotland (2105) [ ]
7. Greece (3207) [ ]
8. Germany (2304) [ ]
9. Philippines (5204) [ ]
10. Netherlands (2308) [ ]
11. Other (specify) [ ____________ ]

*(Option 11 to be coded to the SACC)*

*If F5 = 1, go to F7*

**F6** - CATI TRG November 2000

**In which year did you first settle in Australia?**

*Single Response*

1. Enter year [ _ _ _ _ ]
2. Don't know [ ]

**F7** - NSW Continuous Health Survey - February 2002

**Do you speak a language other than English at home?**

*Single Response*

1. Yes [ ]
2. No [ ]
3. Refused [ ]

*If F7 > 1, go to F9*

**F8** - NSW Continuous Health Survey - February 2002

**What language do you usually speak at home?**

*Single Response. Interviewer note: record English if this is volunteered by respondent. Where respondents speak many dialects of a language, (such as Mandarin Chinese) record only the language.*

1. Specify [ ____________ ]
2. Refused [ ]

*(Code language to ASCL)*

**F9** - CATI SNAP(S) Workshop - December 2001: Categories NOT read out.

**What is the highest level of education you have completed?**

1. University, CAE or some other tertiary institute degree, including post-university (ie Post Graduate Diploma, Master, PhD). [ ]
2. TAFE or trade certificate or diploma [ ]
3. Completed high school [ ]
4. Some high school [ ]
5. Completed primary school [ ]
6. Some primary school [ ]
7. Never attended school [ ]
8. Other (specify) [ ____________ ]
9. Don't know [ ]
10. Refused [ ]

**F10** - CATI TRG November 2000

**Is this dwelling?**

*Read options. Single response*

1. Owned or being purchased by the occupants [ ]
2. Rented from a government organisation or any other public agency [ ]
3. Rented privately [ ]
4. Other (specify) [ ____________ ]
5. Don't know [ ]
6. Refused [ ]
F11 - Which best describes your family's money situation? (Read options. Single response)
1. [I am/we are] spending more money than [I/we] get
2. [I/we] have just enough money to get [us] through to the next pay day
3. There's some money left over each week but [I/we] just spend it
4. [I/we] can save a bit now and then
5. [I/we] can save a lot
6. Don't know
7. Refused

F12 - Which of the following categories does your gross annual household income from all sources fall into? That is the total income from all members of your household before tax is deducted. Would it be: (Read options. Single response)
1. Less than $25,000
2. $25,001 - $50,000
3. $50,001 - 100,000
4. Don't know
5. Refused
G ADMINISTRATION

G1 How many residential telephone numbers, excluding mobile phones, can be used to speak to someone in this household?  
(Single Response. Interviewer note: do not include Internet or fax numbers)  
1 Enter number [ ]  
2 Don't know [ ]

G2 How many residential telephone numbers, including mobile phones, can be used to speak to someone in this household?  
(Single Response. Interviewer note: do not include Internet or fax numbers)  
1 Enter number [ ]  
2 Don't know [ ]

G3 How many times [do these/does this] number(s) appear in the White Pages?  
(Single Response. Interviewer note: do not include Internet or fax numbers. Total number of entries includes numbers that are listed more than once.)  
1 Enter number [ ]  
2 Don't know [ ]

G4 In a survey like this, issues often arise which require further explanation. If we need to could we phone you at a later date to help clarify some issues?  
(Single Response)  
1 Yes (specify - record first name only) [ ]  
2 No [ ]

G5 Date of interview _/__/____

G6 Day of week interview ______

G7 Time of day interview undertaken ______

That concludes the survey. On behalf of the South Australian Department of Human Services/Western Australian Health Department, thank you very much for taking part in this survey.
APPENDIX 5 - INSTRUCTIONS GIVEN TO THE INTERVIEWERS

General Rules For Telephone Interviewing

**Talk clearly and at a moderate pace.** The interviewer should constantly aim to be in control of the pacing and flow of the interview. By the interview stage you will have heard and become familiar with each question. Each respondent, however, is hearing every question for the first time and won't be able to comprehend and digest the question if you speak quickly.

Studies have shown that the reading pace is one of the most critical elements of the interview.

Take care to enunciate each word and allow the respondent time to listen and comprehend the whole question. If you race through the interview the respondent may do so as well. One clear indication that you are asking the questions too quickly is when the respondent continues to ask you to repeat questions.

(Should a respondent want to go beyond the specifics of the questionnaire to discuss his/her problems do not attempt to become a "telephone therapist" but try politely to get the interview back on course!)

**Do not sacrifice quality for quantity.** Although it is important you aim to complete as many interviews as possible, we would prefer you to conduct each interview properly and complete fewer, than race through each interview and complete many.

**Play a neutral role.** Ideally, a questionnaire item should mean exactly the same thing to every respondent. Although this is a somewhat unrealistic goal it is important that the interviewer does not affect a respondent's perception of a question nor the answer given. Therefore, never express your own personal opinions, even if the respondent asks, as this can bias the respondent. Also, interviewers should aim to have a pleasant, confident manner without being either too enthusiastic or too detached. Never sound bored or impatient if someone is taking rather a long time over their responses since if the interviewer sounds bored this may well be reflected in the quality of responses.

**Be familiar with the questionnaire to avoid stumbling over words or phrases.** Take time to read the questionnaire making sure that you understand the questions and possible answers.

**Follow the wording of the questions exactly.** The questions displayed on your computer screen should be read exactly as they appear. Read the lines as naturally as possible in a conversational type manner but following the exact wording set down in the questionnaire. A great deal of effort goes in to constructing the questions and even a slight change in the wording of a questionnaire may lead a respondent to answer "yes" instead of "no".

**Record responses exactly.** If a question has several pre-coded response categories as well as an 'Other' category and a respondent's answer doesn't really fit any of the pre-coded categories, the response should be typed into the 'Other (specify)' category. Do not choose a category that is 'sort of, but not quite' what the respondent said, in order to avoid having to type all of their response.

If a question is open-ended, then record the responses exactly and do not attempt to summarise in your own words.

If a respondent does not wish to answer a particular question this should be recorded as "Refused to answer".
16.1 Ask every question it is not unusual for a respondent to provide in one answer the answer to another question that comes later in the questionnaire. In such an instance, you may be tempted to skip the later questions under the assumption that the answer has already been provided. This is not acceptable practice. The answer received in the context of one question may not be the same as when the specific question is asked directly.

When you receive information that seems to answer an upcoming question, record the information under the question where it is received. Later, when the related questions occur, acknowledge to the respondent that you remember what was said earlier. For example you may say: “We’ve already talked about this topic a bit, but let me ask…..”, then ask the question as it is worded in the questionnaire.

The only circumstances under which questions may be omitted is when skip instructions are given. In this case the asking of one question depends upon the answer to a previous questions.

16.2 Never tamper with the order of the questions. The order in which questions are asked has been carefully planned. They may proceed from general to more specific questions. They may be in order so that the answer to one question cannot influence the answers to others. They will usually be arranged so that the questions flow logically from one area of inquiry to another. Therefore, to change the order in which questions are presented may adversely influence the accuracy of the answers you obtain.

16.3 If the respondent does not understand the question repeat questions that are misinterpreted or misunderstood by the respondent. There are 2 common situations where you may need to repeat a question.

A respondent’s reply may indicate that he/she has not understood the question (e.g. an irrelevant or unfeasible answer). Second, a respondent may simply greet a question with a prolonged silence (another way of saying, “What are you asking me?”). In both cases, the best procedure is simply to say “Let me ask this question again” AND REPEAT THE QUESTION AS WRITTEN.

If the respondent asks you to interpret or define a word/phrase in the question reassure the respondent with a neutral remark (such as, “There are no right or wrong answers. I am interested in your response”). NB: A silence may indicate that the respondent is thinking about his/her answer, and in these cases you should allow time for the respondent to formulate an answer.

16.4 Dealing with “I don’t know” responses during the interview, just as in everyday conversation, you will have to distinguish the legitimate “I don’t know” response from the reply that signifies that the respondent wants time to think about the question. Answering “I don’t know” might be the respondent’s way of stalling for time and in these instances it is best to wait expectantly. It will become easier as the interview progresses as you will get a better feel for the respondent’s mannerisms.
Some respondents may try to avoid answering questions they consider sensitive by answering “I don’t know”. Here you should try and reassure the respondent that all answers are important and that any information provided will be helpful. You may also like to refer to the confidentiality of the data. In cases where the respondent persists in providing an “I don’t know” answer even after reassurance and probing, accept the response.

16.5 **AVOID SUGGESTING ANSWERS TO THE RESPONDENT THAT MERELY NEED TO BE ‘VERIFIED’ BY THE RESPONDENT. THIS MAY OCCUR WHEN YOU ALREADY HAVE PARTIAL INFORMATION ABOUT UPCOMING QUESTIONS AND HAVE ARRIVED AT A REASONABLE ANSWER BY PIECING BITS OF INFORMATION TOGETHER.**

It is also inappropriate to anticipate a respondent’s answer based on their manner. When you suggest an answer, the respondent may accept your implied answer because they assume that it is either correct or desired.

16.6 **WHAT TO DO WHEN A RESPONDENT REFUSES TO ANSWER A QUESTION** SOMETIMES THE QUESTIONING ENTERS A SENSITIVE AREA, WHICH THE RESPONDENT IS RELUCTANT TO DISCUSS. IF YOU REMAIN MATTER-OF-FACT AND PURELY PROFESSIONAL IN YOUR MANNER, YOU WILL HELP TO OVERCOME ANY SUCH RELUCTANCE. YOU MIGHT REMIND THE RESPONDENT THAT ALL RESPONSES TO THE STUDY ARE CONFIDENTIAL. IF THE RESPONDENT STILL REFUSES TO ANSWER A PARTICULAR QUESTION THEN NOTE THIS AND CONTINUE WITH THE INTERVIEW.

**Respect confidentiality.** We have an obligation to keep all interviews confidential. Please do not tell anyone the substance of any interview or part of an interview no matter how fascinating/funny it was.

**Probing.** Good questions allow respondents to express themselves freely and answer the question easily and accurately. Occasionally, though, for some questions there is a need to help respondents to say what they want to say (not what you want to hear). Probing should be done in a neutral manner, appearing natural and conversational to the respondent. For example, if a respondent states that something just isn't appealing to them, you may like to ask: "Is there any particular reason why ... doesn't appeal to you?"

**Be honest about the length of the interview.** If respondents ask about the length of the interview, answer honestly. Don't lie about this to try to get them to agree to participate. The length of interviews vary depending on factors like how chatty the respondent is, and whether all sections of the questionnaire are relevant to them so tell them the usual range of times e.g. "the interview usually takes between x and y minutes, depending on your answers"
CATI QUESTION FIELD-TESTING SURVEY, 2002.

DETAILED COMMENTS ON SA PILOT

Timing of Interviews

<table>
<thead>
<tr>
<th>QUESTIONNAIRE A</th>
<th>QUESTIONNAIRE B</th>
</tr>
</thead>
<tbody>
<tr>
<td>General public - 7 to 11 minutes</td>
<td>General public - 5 to 12 minutes</td>
</tr>
<tr>
<td>Known diabetics - 10 to 15 minutes</td>
<td>Known diabetics - 10 to 18 minutes</td>
</tr>
<tr>
<td>Known asthmatics - 8 to 16 minutes</td>
<td>Known asthmatics - 5 to 12 minutes</td>
</tr>
</tbody>
</table>

Agreed Changes to the Questionnaire

Questionnaire A and Questionnaire B

C14 Have you ever been told by a doctor that you are depressed?  
Introduction "Changing the subject, slightly…" added pre-question.  
"Refused" response category added.

C15 Are you still depressed?  
"Refused" response category added.  
Interviewer note: record answers such as "somewhat" and "a little" as Yes added.

D1 Have you ever been told by a doctor that you have asthma?  
"Don't know" response category added.

D4 Do you still have asthma?  
Skip "If D2 (During the past 12 months, did you have any symptoms of asthma?) > 1 and D3(During the past 12 months, did you take asthma medication…?) > 1, go to next section" added after question D4.

E1 What do you think increases a person's risk of getting Diabetes?  
Responses "Poor diet / wrong diet. junk food" and "smoking" added to response categories.

Questionnaire A

D5 In the last 12 months, have you ever woken up during the night because of your asthma  
Skip pattern amended.

D13 In the last week, how often have you used reliever medication for your asthma? was it…  
"Not at all" response category added.

D15 In the last week, how often have you used preventer medication for your asthma? Was it…  
Response category 2 changed to "4-6 days a week (to make categories exclusive)."

D16 In the last 12 months have you had any days off work, school or home duties due to your asthma?  
Response categories changed to "not-bold" (that is, "not read out").

E9, E10, E11, At what age were you first told you had diabetes / high sugar levels? (three versions of question).  
"first told" emphasised in question.  
Response category 1 changed to "Less than one year of age"
E12  Were you first told you had diabetes/high sugar levels in the last 12 months?  
"first told" emphasised in question.

E13  When you were first told you had diabetes/high sugar levels, what treatments did the doctor put you on?  
"first told" emphasised in question.  
"Read options" included in interviewer instructions.  
Skip after E13 and skip for response 2, both changed to "…go to E16"  

E14  Do you currently have insulin injections?  
Skip after E14 changed to "…go to E16"  

E17  Do you have any vision problems because of your diabetes / high sugar levels?  
Interviewer note added "record yes only if respondent currently has vision problems which they know to be related to diabetes"

E18  Have you ever had any of the following? (complications of diabetes).  
"Read options" added to interviewer instruction.

Section F Social characteristics.  
Introduction now bold (read out).

F4  Which of the following best describes your current marital status? Are you…  
Bullet numbering corrected to 1-7.

Questionnaire B  

C14, C16 Have you ever been told by a doctor that you are depressed?, Are you currently receiving treatment for depression?  
Response categories changed to "not-bold" (that is, "not read out").  

E10, E11 Bullet numbering corrected.  

E13  How long after diagnosis did you start the insulin?  
Response categories changed to "not-bold" (that is, "not read out").  

E14  Have you ever had any of the following? (complications of diabetes)  
Interviewer note added "Record vision problem only if the respondent knows that the problem is related to diabetes"

F3  Besides yourself, who lives in the household?  
Response category "Stepsons / stepdaughters" added.  
Response category "No one (0-15 years) remains omitted (per Margo Eyeson Annan)"

F9  What is the highest level of education you have completed? Have you completed…  
"Have you completed…" removed from question (responses not read out).  

G2  Bullet numbering corrected.

Other Comments Discussed  

SF8 questions - some respondents found these "confusing" - no action.

Questions relating to languages spoken at home - discussion whether "English" would be volunteered by respondents, recorded by interviewers. Decision - no change to questionnaire.
APPENDIX 7 - COMMENTS AFTER WA PILOT

CATI QUESTION FIELD-TESTING SURVEY.

WESTERN AUSTRALIA - PILOT SURVEY DEBRIEF, 6TH SEPTEMBER 2002.

Pilot sample

<table>
<thead>
<tr>
<th>QUESTIONNAIRE A</th>
<th>QUESTIONNAIRE B</th>
</tr>
</thead>
<tbody>
<tr>
<td>General public - 21</td>
<td>General public - 21</td>
</tr>
<tr>
<td>Known diabetics - 2</td>
<td>Known diabetics - 2</td>
</tr>
<tr>
<td>Known asthmatics - 2</td>
<td>Known asthmatics - 2</td>
</tr>
</tbody>
</table>

Timing

<table>
<thead>
<tr>
<th>QUESTIONNAIRE A</th>
<th>QUESTIONNAIRE B</th>
</tr>
</thead>
<tbody>
<tr>
<td>General public - 5 to 7 minutes</td>
<td>General public - 5 to 6 minutes</td>
</tr>
<tr>
<td>Known diabetics - 10 minutes</td>
<td>Known diabetics - 7 to 9 minutes</td>
</tr>
<tr>
<td>Known asthmatics - 7 minutes</td>
<td>Known asthmatics - 7 to 8 minutes</td>
</tr>
<tr>
<td>Known diabetic with asthma - 15 minutes</td>
<td></td>
</tr>
</tbody>
</table>

No persons with diabetes of pregnancy were interviewed during the pilot survey.

Timing of interviews in the survey will be automated.

INTERVIEWER COMMENTS AND QUESTIONNAIRE CHANGES.

Interviewer allocation of questionnaires

The interviewers felt that the questionnaire wording would be more accurately delivered if the interviewers worked with only Questionnaire A or Questionnaire B.

However, as the interviewers perceptions of which questions worked best (for particular topics) is part of the assessment process, a system where interviewers work with only one questionnaire for 50% of the survey, then, only with the other questionnaire for the latter half of the survey will be proposed.

SUMMARY OF QUESTIONNAIRE CHANGES

Questionnaire A and Questionnaire B

Introduction - Intro A4 and Intro A5 replaced with WA introductions.

Interviewers felt the WA version of the introductory paragraphs Introductory paragraphs for "known" diabetics / asthmatics worked best and created good rapport with respondents. This will be incorporated into the questionnaire.

Questionnaire A, B5 and Questionnaire B, B4 Including yourself how many people aged 18 years and over usually live in this household?

Prompt added to ensure respondent has included him/herself in the number of persons 18 years and over.
C14 Have you ever been told by a doctor that you are depressed?  
Interviewer note: (Only medically diagnosed depression equals "Yes", not depression diagnosed by a counsellor) added.

Changes to Questionnaire A

Order of diabetes questions E1 to E4 changed as follows:
- E2 changed to E1
- E3 changed to E2
- E4 changed to E3
- E1 changed to E4

Skips for these four questions changed as follows:
- After E1: If E1 = 1, go to E4.
- After E4: If E1 = 1 and B4 = 1 (sex is male), go to E8
  - If E2 = 1 and B4 = 1 (sex is male), go to E9
  - If E2 > 1, go to F

E4 (formerly E1) What do you think increases a person's risk of getting diabetes? Response category "genetic" added.  
Interviewer note: prompt "anything else" after first response added.

F8 Which language or languages do you speak at home?  
Interviewer note record English if this is volunteered by respondent. Where respondents speaks many dialects of a language, (such as Mandarin Chinese) record only the language) added.

Questionnaire B

D9 - and D11 Do you use a reliever medication (puffer or nebuliser) such as…?, Do you use a preventer medication such as…?  
Interviewer note If respondent claims to use a reliever or preventer that is not included in the medications which are read out, record "Yes")

Interviewer comments - not acted upon between pilot and actual survey.

Interviewers perceived that Questionnaire A flowed well on CATI, in comparison to Questionnaire B.

Asthma questions - wording in Questionnaire A preferred by interviewers.

Questionnaire A and Questionnaire B.

Intro for general public worked fine.

SF8 question wording and response categories were confusing for respondents, especially C5 (During the past 4 weeks, how much energy did you have?) response option 1 "very much".

QA E2 (formerly E3) Have you ever been told by a doctor that you have high sugar levels in your blood or urine?  
Interviewers preferred wording to QB E3 Have you ever been told by a doctor that you have high blood sugar or a touch of sugar?

QB, F10 (Is this dwelling…) 'v' QA, F11 (I would like to ask some questions about your housing arrangements. Are you…) Interviewers preferred Questionnaire A wording and responses (excepting responses 5 (purchasing under a rent/buy scheme) & 6 (occupying under a life tenure scheme).
Questionnaire A

B1 Could you please tell me your date of birth?
Most people gave DOB, 1 person was offended by the question which he perceived as invasive, 1 person had a language difficulty with the question. Interviewers found asking age only on Questionnaire B easier.

D12 - D15 (Asthma questions)
Some respondents did not know whether their medication was a "reliever" or a "preventer" and wanted examples of these medications. The interviewer felt that by not prompting or assisting the respondent that the respondent was put in a situation of sounding stupid for not knowing the answer to the question.

F10 What is the highest level of education you have completed? Response category 1 (University, CAE or some other tertiary institute degree, including post-university (i.e. Post Graduate Diploma, Master, PhD) was felt to be "too wordy". Also, respondents wanted to volunteer their response prior to hearing all of the response categories.

F11 I would like to ask some questions about your housing arrangements. Are you…
Response categories 5 (purchasing under a rent/buy scheme) & 6 (occupying under a life tenure scheme) were not understood by respondents. Also, respondents wanted to volunteer their response after hearing the first four categories.

F13 I would now like to ask you about your household's income. We are interested in how income relates to health, lifestyle and access to health services. Before tax is taken out, which of the following ranges best describes your household's income, from all sources, over the last 12 months?
Introductory sentence made asking question about income easier for interviewer and more acceptable to respondent.

Questionnaire B

D2 Question on asthma symptoms
Interviewer suggests replace "cough" with "coughing".

D5 How often do you awaken during the night with asthma?
Wording "do you awaken" does not flow well.

D5 How often do you awaken during the night with asthma? and D6 How often do you see your doctor about your asthma?
Interviewer suggests reverse order of response categories.

F4 What is your formal current marital status? Are you…
Interviewers found it strange to ask "formal" marital status. One newly widowed lady had to think hard about her answer. It appeared to the interviewer that up until this time, she had considered herself still to be a married person.

F7 Do you speak a language other than English at home?
Having "at home" at the end of the question was problematic as respondents began to report languages learned at school.

F9 What is the highest level of education you have completed?

18.1 QUESTION WORKED WELL AS "NON-READ OUT, BUT INTERVIEWER FELT THAT THE QUESTION WORDING OR A PROMPT NEEDED TO ENSURE THAT PERSONS WHO HAD A DIPLOMA OR TRADE CERTIFICATE CONSIDERED THESE IN THEIR ANSWER.
A full debriefing lasting over 2 ½ hours was hosted by the UWA Survey Research Centre (SRC) upon completion of the field testing. Representatives from the Australian Bureau of Statistics (ABS), the SA Department of Human Services and the WA Department of Health were present. The Director of the UWA SRC and the team of interviewers who worked on the project provided detailed feedback. A brief summary of the main points raised follows.

General Observations

Interviewers were assigned to one or other version of the questionnaire. This was done because of the similarity of the questions on the two versions. Experienced interviewers begin to memorise the questions and the slight wording changes were difficult for them to remember. In the interests of keeping the versions properly delivered, the interviewers were split into two teams, the A team and the B team. Two interviewers, who were found to be able to deal with the subtle changes without problems did interviews with both versions. They did not notice any real difference in the delivery of the two versions. They were considered to be equivalent from that point of view.

All interviewers felt that prompts were an essential part of the interviewing process. Prompting could simply be a pause after an answer, or just an 'is there anything else' question. The interviewers were unanimous in their belief that subtle prompts improved the quality of the data collected.

They believed that the instructions to read questions and offer no clarification also impeded their ability to get good quality and accurate information.

It may be that a more considered set of instructions needs to be developed in consultation with experience interviewers.

Section A - The Introduction:

Some of the A Team felt that the Introduction was too long, this was not an issue on for the B Team.

Section B - Key Demographic Information

Both teams believed that people did not include themselves in their answer to the number of adults in the household. They all felt that a further prompt along the lines of 'and does that include yourself?' would be helpful.

Both teams also believed that older people who lived alone were reluctant to say so and suggested that maybe these questions would be better asked at the end.

Section C - Chronic Conditions

Both teams said that the questions on current high blood pressure (C10) and current high cholesterol (C12) confused people. People recalled having high blood pressure but were less sure about whether or not they still had it. Intervening questions about treatment/medication use would be helpful.

Section D - Asthma

The B Team had difficulties with questions D5 and D6, which asked how often people woke up at night and how often they visited their doctor because of asthma. They felt that the categories were difficult for people. The A team (D6 and D7) did not have this difficulty.
The interviewers suggested that maybe asking about the pattern of asthma attack before asking about the frequency would help with the response.

Both teams suggested that including specialist along with your doctor would help things on the question about consulting a doctor because of asthma.

Both teams had noted that the question asking how often the doctor was consulted had left out a category - namely once every 1 to 2 months (D6 on Version B and D8 on Version A).

Both teams felt that the questions on reliever and preventor medications caused confusion. They thought that the names on Version B assisted (D9 and D11) but that the frequency categories on Version A were better (D13 and D15).

The A team questioned the 1 week time interval on D13 and felt the one month interval as asked on Version B would be better (D10).

Team A thought that D11, the question about the asthma action plan was too long. The B team thought their equivalent but shorter D8 was better.

**Section E - Diabetes**

Generally, all interviewers said that the flow of the questions on diabetes could be improved. They thought that questions separating diabetes from high blood sugar led to better quality responses.

Both teams agreed that E4, the unprompted question asking about risk factors for Diabetes need to be revised. The teams felt that either prompts or reading out the options would work better.

Team B found that the question asking about the first treatment that people with diabetes and high blood sugar were put on caused problems (E9 and E11). The interviewers felt that maybe reading out the options would help. Team A, whose question only asked about treatment rather than first treatment, had no problem with the question.

Both teams said that their respondents had a problem with the complication 'kidney disease' (E18 on Version A and E14 on Version B). They said that some extra definition would be helpful.

**Section F - Demographics**

Both teams felt that F1, the question on current employment status could use some work. Their opinion was that a great deal of information was being missed by forcing one response, and thought that multiple response would work better.

Both teams thought that F2, the question on the number of hours worked needed an average time added, or some qualifying statement so that those on holidays in the past week would not be lost.

F3, the question on who currently lives in the household, on Version B frequently elicited names rather than relationships.

F4 on Version B, the question about marital status needs to have de facto or living with a partner added as people defined themselves that way.

The questions about language spoken at home F7 and F8 on Version A were considered to work better than Version B.
Both teams had some problems with E10 (Version A) and E9 (Version B), the question on education. The interviewers felt that reading out the options was better but that there needed to be better instructions about how to convert overseas qualifications. They also thought that the highest level being read out first was off putting to respondents who may not have much education. They would prefer to start at the bottom and work up. All interviewers thought that reducing the options may improve the quality of the responses.

F10 on Version B, the question about the ownership of the dwelling, caused Team B some problems. The occupant may not mean the respondent and this category should be clarified.

In Version A F11, the term 'dwelling' caused problems to Team A who would prefer to substitute house.

Both teams felt that the question on perceived financial status could be further developed (F12 Version A and F11 Version B). Two suggestions offered were 1) add an option for being able to save regularly and 2) take out each week from response category 3.

Both teams thought that the introduction to the income question on Version A was really good (F13). However they also thought that some work needed to be done on the ranges of income. Generally the smaller set of ranges worked better. They also thought that an added question for older people might work and suggested 'well, do you know how much you get a week?'.

**Section G - Administration**

Both teams thought that an introduction to this section was required. The thought that household might be a difficult term for people.

They also reported that G3 the question about White Pages listings was confusing for people but did not make any suggestions for change.

**Timing**

The two interviewers who did both versions were not aware of any difference in the time taken to interview using Version A or Version B. However, in the pilot, Version A took longer and the interviewers only did a couple of both types, so they were giving impressions rather than factual information.