



Qualitative data processing protocols

Introduction

The qualitative data are primarily the comments included by participants in response to the final question on each survey (*Have we forgotten anything?*) and may also include responses to open-ended questions and comments written in other parts of the paper or online surveys or collected during telephone surveys. These Guidelines were revised in May 2020 and were approved by the ALSWH Study Management Committee at the May 2020 meeting.

Qualitative processing 1996-2004

The following steps were undertaken with data for the baseline ("Survey 1") surveys:

1. Potentially identifying portions of comments (eg. someone's name, addresses etc) on the paper surveys were whited out.
2. These data, without this information, were entered into Microsoft (MS) Access by ALSWH project assistants.
3. At the same time, these data were coded into the main themes and keywords that arose according to the frequency of topics made in Survey 1.
3. To check the reliability of these themes and keywords, about one in twenty surveys were checked by another operator.

Data from Surveys 2 and 3 of the 1946-51, 1921-26 and 1973-78 cohorts, and Survey 4 of the 1946-51 cohort, were entered in the following way:

1. During the editing process, potentially identifying portions of comments were whited out.
2. These data were scanned and entered by the data company (with the exception of the second survey of the 1921-26 cohort, which was entered by ALSWH staff), without the whited out material. These data were then read into MS Access. There has been no consistent notation to indicate the presence of omitted material.

Surveys 2 and 3, and Survey 4 of the 1946-51 cohort, were not coded.

In 2004 concerns were raised over the difficulties that were experienced with being unable to access data that had been whited out. Whiting out was stopped at this time.

Current qualitative processing

As of June 2005, and starting with Survey 4 of the 1921-26 cohort, data are processed as follows:

1. Paper surveys:
 - a. All qualitative comments are left as is except for names and addresses, initials, phone numbers, email addresses and staff names. Confidential information which may identify the participant or anyone connected with the participant, e.g. participant's name, doctor's name or name of husband etc., are whited out and replaced with the appropriate text enclosed in curly brackets {name}, {address}, {initials}, {phone no}, {email} or {staff name}.
 - b. Data are scanned and entered without the whited out information but with the added text placeholder. As in previous years, the data will then be read into MS Access.
2. Online surveys:
 - a. Raw qualitative data from the online surveys is imported into a SQL Server database (the backend storage used for MS Access). ALSWH Project Assistants review the data and anonymise using standard ALSWH protocols (as for 1.a. above).

The de-identified qualitative comments from the paper surveys and the online surveys are combined and the data are then forwarded to the Data Managers for distribution to researchers as required.

While these procedures are followed diligently, the volume of material means that occasionally potentially identifying details might be missed. Furthermore, there are other potentially identifying details that might need to be anonymised. Thorough anonymisation is the responsibility of the individual research team, particularly the ALSWH liaison person (see below).

Anonymising procedures

In this document, the term 'anonymise' is used to describe the process whereby qualitative data (i.e. participant comments) are altered so that participants are unlikely to be identified.

Document G

August 2020

Anonymising of all qualitative data, including those collected prior to 2005, will be the responsibility of the ALSWH liaison who is named in the Statement of Data Use. The collaborator in consultation with the liaison may nominate a member of their team who will deidentify the data, but the responsibility for ethical procedures and compliance with these guidelines remains with the ALSWH liaison.

Outputs

All outputs must be vetted by the ALSWH liaison prior to publication in any form (presentation, journal paper etc).

Data security

All electronic copies of the data must be password protected (at a minimum).

If any data are to be printed they must first be anonymised. Raw data are not to be printed. Any printed data must be held securely in keeping with NHMRC guidelines; in a locked cabinet on the premises of their institution

Suggested general guidelines for anonymising

Anonymising procedures may differ depending on the focus of particular research projects. The following guidelines are suggestions as to how data may generally be anonymised.

Dates must be removed.

All names, places and potentially identifying information can be replaced as follows:

Names can be replaced with the person's relationship to the participant or their title. For example:

{son}

{mother}

{family member}

{friend}

{family doctor}

{solicitor}

Town and suburb names can be replaced with short descriptors as per the five RRMA categories. Place names can therefore be replaced with:

{capital city}

{other metropolitan}

Document G

August 2020

{large rural}

{small rural}

{remote}

Place names can be replaced by a short descriptor in braces. For example:

Royal Newcastle Hospital replaced by {regional centre hospital}

Sydney District Courthouse replaced by {capital city courthouse}

Other less general potential identifiers have been noted, such as unique characteristics (eg. awards) and specific disabilities involving multiple family members. These types of identifiers can be anonymised by changing the characteristics involved, family make-up or other details that do not affect the nature of the analysis being conducted.

Where potential participant recognition occurs the ALSWH liaison officer must be consulted. The liaison will ensure that the participant's data are removed from the dataset.

Where a participant is very transparent and specific in their comments, in consultation with the ALSWH liaison, the analyst must decide if the data can be anonymised without losing meaning.