



Qualitative processing protocols

Introduction

This document was compiled to clarify and extend upon current protocols that are employed in processing qualitative data collected by the Australian Longitudinal Study on Women's Health (ALSWH). The qualitative data are primarily the comments included by participants in response to the final question on each survey (*Have we forgotten anything?*) and also include responses to open-ended questions that may be asked in the future. These guidelines were referred to the ALSWH Steering Committee and the ALSWH Data Management Group and (DMG) were first revised in accordance with decisions reached during the Steering Committee meeting 5/5/2005 and DMG meeting 15/6/2005. The guidelines were then revised by the Qualitative Research Group which formed in 2008 in response to increased use of the qualitative data. The Steering Committee approved the revised guidelines 18/02/2009 and University of Newcastle HREC approved the changes 09/07/2009.

These guidelines were again reviewed in May 2014 by the Qualitative Research Group. The Steering Committee endorsed these changes in August 2014.

Qualitative processing 1996-2004

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

1. Potentially identifying comments (eg. someone's name, addresses etc) were whited out.
2. The data, without this information, were entered into Access by ALSWH project assistants
3. At the same time, the 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, and Survey 4 of the 1946-51 cohort, were entered in the following way:

1. During the editing process, potentially identifying comments were whited out.

Document G

May 2017

2. The data were scanned and entered by the data company (with the exception of Old 2, which was entered by ALSWH staff), without the whited out material. The data were then read into 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 were being processed as follows:

1. All qualitative comments were left as is except for names and addresses. Changes of names and addresses were whited out and replaced with {change name} and {change address}. Where women include their own name and address for other purposes, these were whited out. Names are replaced with {name}, and addresses with {address}.
2. Data will be scanned and entered by the data company without whited out names and addresses. As in previous years, the data will then be read into Access.
3. In terms of online surveys, raw qualitative data is downloaded from the online surveys and imported into an MS Access database. ALSWH Project Assistants review the data and anonymise using standard ALSWH protocols. The data are then forwarded to the Data Managers for distribution to researchers as required.

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.

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.

Document G

May 2017

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.

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}

{large rural}

{small rural}

{remote}

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

Document G

May 2017

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.