

A GUIDE TO MAILING QUESTIONNAIRES, DATA ENTRY & CLEANING

THE ADVICE CENTRE FOR LOCAL SURVEYS



Making patients' views count

Last updated: 25th May 2012

Contacts

The Advice Centre for local surveys
Picker Institute Europe
Buxton Court
3 West Way
Oxford
OX2 0JB

Tel: 01865 208127
Fax: 01865 208101
E-mail: advice@pickereurope.ac.uk
Website: www.nhssurveys.org

Key personnel

Chris Graham (Manager)

Esther Howell
Caroline Killpack
Jenny King
Katherine Körner
Sheena MacCormick
Alexandra Skew
Catherine Turney
Mark Waters

Guidance documents

Other guides are also available in the local surveys section of the NHS Surveys website (www.nhssurveys.org/localsurveys) to help you carry out your survey:

- Separate guides provide step-by-step instructions on how to draw a sample of patients for a specific survey (e.g. A&E Department survey, Maternity services survey, Inpatient survey)
- A guide to analysis, reporting and dissemination of survey results

Updates to guidance documents

Note that guidance documents are updated from time to time. Make sure you have the latest version of this guidance manual and other guidance documents (the date of the last update is located on the front page). All guidance documents are available on the NHS Surveys website (www.nhssurveys.org/localsurveys).

CONTENTS

1	How can this guide help me?	1
2	Implementing your survey - practicalities.....	1
3	Sending out questionnaires	4
4	Managing your survey helpline	6
5	Entering data	8
6	Data cleaning	11
	Appendix 1: Template Freephone Query Sheet	16

1 How can this guide help me?

This document provides guidance on the practicalities of implementing a postal survey, from the point at which the sample of patients or service users has been compiled¹, through to data cleaning. It is designed to help those trusts who have opted to carry out their survey mailings in-house, rather than use a survey contractor.

This document is split into sections; each provides guidance for one component of the survey process. Section 2 outlines the preparatory work that can be done before mailing commences, including printing questionnaires, organising the materials required for each mailing and setting up a FREEPOST address. The process of sending out the questionnaires and reminders is explained in section 3, whilst section 4 provides advice on managing your survey helpline during the fieldwork period. Section 5 explains how to enter the data from returned completed questionnaires, and section 6 provides advice on how to clean your survey data in accordance with the processes used for the national patient survey programme.

There is a separate guide available on the NHS Surveys website which provides advice and suggestions on data analysis, reporting and dissemination of survey findings. This report (*A guide to data analysis, reporting and dissemination of survey findings*) can be downloaded from: www.nhssurveys.org/Filestore//documents/Guide_to_analysis.pdf

2 Implementing your survey - practicalities

2.1 Printing questionnaires

The questionnaire has been tested in the format produced by the online compilation tool. We recommend that the questionnaires used by trusts should emulate this format and should be comprised of the following:

- Two columns of questions on each page.
- Questions presented with a consecutive question number, followed by the exact question wording used in the question bank, and then each of the response options presented on a separate line beneath the question, for example:

C3. Did the doctors and nurses listen to what you had to say?

- 1 ☐ Yes, definitely
- 2 ☐ Yes, to some extent
- 3 ☐ No

- Please do not arrange the response options horizontally across the page, rearrange the question options, or change the order of the questions in the questionnaire.

¹ **Please note:** there are specific guides available on the NHS Surveys website which provide step-by-step instructions on how to draw a sample of patients for a particular survey, using the same procedures as those used in the national patient survey programme. See: www.nhssurveys.org/localsurveys

- It is practical to ensure that the number of pages in a questionnaire is a multiple of four so that sheets can be printed double-sided on A3 paper and folded to make an A4 booklet, stapled in the middle. If pages are stapled at the corner, there is a greater chance that some pages will become detached and get lost. **We strongly recommend that you do not exceed 16 sides of A4:** a questionnaire that is longer than this will have an adverse effect on response rates.

Number of questionnaires

When calculating the number of questionnaires to be printed, you will need to allow for sending out duplicate questionnaires with second reminders. Printing costs can be unnecessarily high if a second print-run is required, so it is worth ensuring that the first print-run is sufficiently large to allow for contingencies. As a rule of thumb, multiply the number of patients in your sample by 1.7 to obtain the total number of questionnaires required. So, if the number of first mailing questionnaires you intend to send out is 1200, then you might want to print 1.7×1200 , or approximately 2,100 copies.

2.2 Covering letters

You will need headed paper from your trust for covering letters for the first and third mailing. The standard covering letter for the particular survey you would like to carry out is available in Microsoft Word format on the NHS Surveys website for you to download and add your own trust's details (see: www.nhssurveys.org/localsurveys). A reminder letter is used for the second mailing (also available on the NHS Surveys website). If you are using a survey contractor to carry out the survey work, it is preferable that the paper does not include a telephone number for the trust, as patients should call the contractor's FREEPHONE line, rather than the trust.

2.3 Mailing labels

Up to three mailing labels are needed for each patient. One set of labels will be used for the first mailing, one for the first reminder and one for the second reminder. The total number of mailing labels used will therefore depend on whether the patient is included in the second and third mailings.

We recommend using the mail merge feature in a word processing package to create the mailing labels from the database of patient names and addresses. **It is essential that the patient record number is on each address label**, as this has to be matched with the number on the front of the questionnaire. The label should not include any other information except the patients' name, address and postcode details, and the patient record number.

Note on the patient record number (PRN)

The patient record number (PRN) is a unique number allocated to all patients at the start of the survey that allows their responses to be kept separate from their name and address, but allows matching up of the response data with the sample data. It also allows patients to identify themselves if they contact the trust or contractor without needing to provide name and address information. The PRN should be centrally placed and large enough to be visible to all patients. The Royal National Institute of the Blind recommends the PRN be printed in size 14 font and located inside the box on the lower half of the front page of the questionnaire.

If patients delete the PRN from the cover page and then return the questionnaire, we recommend that responses should be added to the bottom of the file, and **not** matched to patients with similar demographics.

2.4 Mail out envelopes

It is important that the envelope(s) which you use to mail out your survey materials to patients does not show any indication of the NHS Trust. Some patients may not have told family or friends that they have gone to hospital/received health services and, under data protection regulations, it is important that this information remains confidential to the patient. We would therefore recommend that the return address used on any mail out envelope(s) does not indicate Trust name or address. It is, however, important that you record questionnaires which are returned undelivered as this affects response rate. We would therefore advise that, where necessary (i.e. for Trusts conducting the survey in-house), a PO Box address is set up for envelopes which are returned undelivered.

Please note that the above does not apply to the address on the reply paid envelope (which can be a hospital address) as we assume that the patient is responsible for opening his or her own mail.

For further details on how to set up a PO Box please go to:

<http://www2.royalmail.com/delivery/inbound-mail/po-box>

2.5 Materials required for each mailing

First mailing

For the first mailing, you will need each of the following items:

- Printed questionnaires
- Large envelopes for mailing questionnaires to patients
- Labels for addressing envelopes
- Labels for sender address on reverse of envelopes
- FREEPOST envelopes for return of questionnaires
- Covering letters using the trust's letterhead
- Multi-language helpline sheet.

Second mailing (first reminder)

First reminders are sent to all patients who do not respond to the first mailing (except, of course, those who withdraw). Usually you will need to send first reminders to around 55-75% of the original patient sample. The following items are needed:

- First reminder letters
- Envelopes
- Labels for addressing envelopes
- Labels for sender address on reverse of envelopes.

Third mailing (second reminder)

The second reminder should replicate the first mailing, and you will need to send this to around 45-65% of the original sample, depending on the number of responses to the previous two mailings.

The following items are needed:

- Printed questionnaires
- Large envelopes for mailing questionnaires to patients
- Labels for addressing envelopes
- Labels for sender address on reverse of envelopes
- FREEPOST envelopes for returning questionnaires
- Reminder letters using the trust's letterhead
- Multi-language helpline sheet.

2.6 Setting up a FREEPOST address

A FREEPOST address allows patients to return completed questionnaires at no cost to themselves. After you have paid for the licence, you will only pay for the responses you receive. The FREEPOST address can be printed on the envelopes you send out with the questionnaires. Printed envelopes must comply with Royal Mail guidelines. Details of how to apply for a FREEPOST licence can be found at the Royal Mail website: <http://www.royalmail.com>

Alternatively, you can call your local Sales Centre on 0845 7950 950.

2.7 Setting up a FREEPHONE line

The covering letter to patients should include a telephone number for patients to call if they have any questions or concerns about the survey. If you run the survey in-house, and you have not already got a FREEPHONE number (such as a PALS Freephone number), you might want to set up a FREEPHONE line for this purpose so that patients who have a query about the survey are not charged for calling. Instead the trust pays to receive the call. Alternatively, many approved contractors offer this service.

There are a number of companies that can set up a FREEPHONE line for you (i.e. numbers generally starting with 0800).

Please see section 4 for information about managing the survey helpline during the survey period.

3 Sending out questionnaires

3.1 First mailing

The envelope sent to each patient at the first mailing should include the following:

- 1) A questionnaire **numbered with the patient record number**. The number must match the number on the address label, and the number on the list of patient details.
- 2) The first covering letter.
- 3) The multi-language helpline sheet (recommended).¹
- 4) A large envelope for returning the questionnaire, labelled with the FREEPOST address on it.

These items should be packed into an envelope that has a return address on the outside. This should be the contact at the NHS trust, or the approved contractor. Remember that the return address should not indicate the name or address of the Trust.

Please note: We recommend that all mailings are sent as second class mail. It is essential that the appropriate postage rate is paid. The postage may exceed the standard letter rate.

¹ This document can be found on the NHS Surveys website alongside the cover letters (see: www.nhssurveys.org/localsurveys). The languages covered by this document are: Arabic, Bengali, Chinese (Cantonese), Farsi /Persian, Gujarati, Hindi, Kurdish, Chinese (Mandarin), Punjabi, Tamil, Thai, Turkish, Urdu, French, Italian, Polish, Portuguese, Russian, Somali, Spanish.

If a **survey contractor** is carrying out most of the work for you, but is not operating under a contract which allows the transfer of patient names and addresses, they should send pre-packed questionnaires to your trust for mailing out. The envelopes should be clearly marked with the patient record numbers so that trust staff can match these with their patient list and put on appropriate patient address labels. This procedure ensures that patient names and addresses are not shared outside the trust but allows questionnaires to be matched to the correct patients.

3.2 Booking in questionnaires

When questionnaires are received, match up the patient record numbers against the list of patients in your sample file¹, so that you can record (in the *outcome* column) which patients have returned questionnaires and will not therefore need to be sent reminders. We recommend you keep paper copies (or scanned copies) of all of the pages of the questionnaires, including the front page, of any questionnaires that are returned to you until three months after the survey.

If a **survey contractor** carries out the work, questionnaires will be returned directly to them, so they will be able to record these returns against the list of patient record numbers. You should inform the contractor of any questionnaires that were returned to the trust undelivered, and of any patients who inform the trust that they do not wish to be included in the survey, or if any patient has died during or prior to the survey. The contractor can then record these details in their own patient list, and ensure that reminders are not sent out to those patients.

3.3 Sending out reminders

For results to be representative, it is essential to get a good response rate. To achieve this, we strongly recommend that you send out two reminders to non-responders. Depending on the time that has elapsed since you first checked your patient list for deaths, it might be necessary to send your list back to the Demographic Batch Service (DBS) for a further check before you send out reminders.

For further information about DBS, please refer to the separate guidance, available in the 'local surveys' section of the NHS Surveys website, on how to draw your survey sample for a particular survey (i.e. '*A Guide to conducting a local [survey type name]*').

First reminders

The first reminder should be sent to patients who have not responded after **one to two weeks**. The first reminder should reach the participant while they are still in possession of the first questionnaire. We recommend approximately **ten days** between the mailing day of the first questionnaire and the mailing day of the first reminder.

The standard first reminder is available in Microsoft Word format on the NHS Surveys website (in the 'local surveys' section) for you to download. It can be printed on A5 paper.

¹ See separate guidance, available in the 'local surveys' section of the NHS Surveys website, on how to draw your survey sample and create your sample file for a particular survey (i.e. '*A Guide to conducting a local [survey type name]*').

Second reminders

Second reminders should be sent out approximately **two to three weeks** after the first reminder to patients who have not yet responded.

The envelopes should include the following:

- 1) A questionnaire numbered with the patient record number. The number must match (or correspond to) the number on the address label and the number on the list of patient details.
- 2) The second reminder covering letter.
- 3) A multi-language helpline sheet (if used in first mailing).
- 4) A large envelope, labelled with the FREEPOST address on it.

The standard second reminder letter is available in Microsoft Word format on the NHS Surveys website (in the 'local surveys' section) for you to download and add your trust's details.

Remember that you should check your trust's own records for deaths before sending out reminders.

If using a survey contractor, when reminders are due to be sent out, the contractor should send the pre-packed envelopes bearing the patient record numbers of the non-responders. Again, the envelopes should be clearly marked with the patient record number so that those carrying out the mailing can match these with their patient list and put on appropriate address labels.

4 Managing your survey helpline

4.1 Introduction

As stated in section 2.7, the covering letter to patients should include a telephone number for patients to call if they have any questions or complaints about the survey. All staff who are likely to take calls should be properly briefed about the details of the survey, and be aware of the questions or complaints they are likely to receive.

Where appropriate, ask patients who call to tell you their patient record number, which should be on the address label of the envelope they received, and on the questionnaire itself. You can then use this number to identify people who do not want to receive any further reminders.

4.2 Offering interpretation services

The patients who respond to your survey should be representative of all of the patients who use the trust, so it is important that groups with limited understanding of English are not excluded. The questionnaire has been written in as plain language as possible to facilitate optimum understanding by all respondents.

It is good practice to offer help or interpretation services to those who might require them. You can do this by subscribing to a specialist interpreting service, most of whom offer telephone interpretation on a pay-as-you-go basis. This normally involves a three-way conversation between you (or your helpline operator), the patient and the interpreter. Your trust may already have arrangements with such a service.

A multi-language leaflet¹ template is available on the NHS Surveys website, and this can be included with your first and third mailings. You can use this leaflet by inserting the appropriate helpline number for the survey. The leaflet gives directions in the 20 most common non-English languages spoken in England. We recommend you also make provision for those with learning difficulties.

4.3 Making a record of the calls

It is important to keep a record of all calls received and the reasons patients called, as this can help to make improvements to future surveys and can provide useful additional information on patients' concerns. A standard form should be produced for completion by those taking the calls (a template form is shown in Appendix 1, which you may wish to use). The relevant details of each call can be recorded so that survey organisers can monitor any problems and remove patients who wish to be excluded from the mailing list.

4.4 Dealing with patients' queries

Below are some questions and comments commonly asked by patients, and some advice on how they can be managed:

I have had two or more admissions/attendances - which one should I refer to?

Patients should be advised to refer to their **most recent** hospital admission (for inpatients) or attendance (at A&E/Outpatients). Usually, this is the admission or attendance covered by your sampling period but, for the few patients who have been re-admitted or reattended since you drew the sample; it is simpler to tell them to refer to their most recent stay or visit. It will not make the results invalid if a few of the patients refer to a more recent episode than the others.

I have a specific comment, complaint or question about my care or treatment. Who can I contact at the trust?

Patients can be referred to the trust's PALS, the complaints manager or patient services manager. Survey contractors should be given the contact details of the PALS office or an appropriate member of trust staff to whom calls can be referred.

I do not wish to participate in this survey.

A few patients might call to say that they do not want to be involved in the survey, and fewer still may object to being sent the questionnaire in the first place. Staff should apologise to the patient and reiterate the statement in the covering letter - that taking part in the survey is voluntary, and that the patient's care will not be affected in any way if they do not respond. It might be helpful to point out the purpose of the survey, and to emphasise the potential value of the patient's responses. If the patient is willing to tell the staff member the identification number (patient record number) written on their survey, it might also be possible to prevent any further reminders being sent to that patient. It is also advisable to ask the patient to ignore any future reminders that they might receive. If the patient is unwilling to provide their patient record number they should be

¹ The languages covered by this document are: Arabic, Bengali, Chinese (Cantonese), Farsi /Persian, Gujarati, Hindi, Kurdish, Chinese (Mandarin), Punjabi, Tamil, Thai, Turkish, Urdu, French, Italian, Polish, Portuguese, Russian, Somali, Spanish.

advised that they will remain on the mailing list and receive further correspondence. These patients should be logged as outcome = four (opt out).

The person to whom the questionnaire is addressed is unable to understand the questionnaire.

Relatives or carers may call to pass on this information. In some cases, they may offer to complete the questionnaire for the patient, but this is only advisable if there is a good chance that the responses are a true reflection of the patients' views.

The person to whom the questionnaire is addressed has died.

Even with the use of a deceased patients tracing service, it may not be possible to identify all deceased patients, particularly those who have died most recently. It is very important that staff who take the calls are aware of this possibility and are prepared to respond sensitively to such calls. These patients should be logged as outcome = three (patient deceased).

I would like to take part but English is not my first language.

If a patient's spoken English is better than their written English, they may be willing to have someone, such as a family member, fill in a form on their behalf. It is also strongly recommended that you offer access to a telephone interpretation service, where the questionnaire can be filled in over the telephone. A multi-language sheet template in the 20 most commonly spoken non-English languages is available on our website, and trusts or contractors can make use of this by inserting the appropriate number for their helpline and/or translation service (see section 4.2).

5 Entering data

5.1 Using the pre-designed 'data entry file'

The data from your survey can be entered into the pre-designed Excel file (the '*data entry spreadsheet*'). This is available in the relevant survey folder under the 'Local surveys' section of the NHS Surveys website. The data entry file contains a series of worksheets that take the response data entered in the first sheet ('Data') and calculates the number and percentage of responses to each question. An advantage of using this pre-designed file is that the data entry sheet is validated so that out-of-range responses cannot be entered.

At the bottom of the Excel screen there are labelled tabs for each of the worksheets within the workbook, which correspond to each section of the questionnaire. These sheets automatically provide a summary report of the data entered (frequency counts and percentages on each question). The first of these tabs is labelled "Data". Click on this tab to show the data entry window.

You should transfer your sample file (i.e. the file which includes the patient record number, and various data fields such as patients' age and sex) into the first 'Data' worksheet. Your file should have a row for each patient included in your sample. Response data from completed returned questionnaires can then be entered in this file (using the unique patient record number enables you to enter the data from a completed questionnaire into the correct row).

This file should not contain any patient names or address information. Only this 'anonymised' file should be used to record patients' responses.

As you will only have responses from about 50% of the patients (that is, those who have returned a completed questionnaire, and who will therefore have an outcome code “1”), some rows will not have any data in the ‘response’ section of the file.

Please note: The pre-designed data entry spreadsheet includes all questions that are in the question bank tool. You will therefore need to delete those questions that you have not included in your survey, both from the ‘Data’ sheet and from the relevant worksheet.

5.2 Entering and coding data

Data should be entered using the following guidelines:

- Each row records one patient's responses to the survey (use one column of data for each item of patient information or response).
- For each question, the small number next to the box ticked by the patient should be entered as the response.
- If a response is missing for any reason, the data field should be left blank, or coded as a full stop (.)¹
- If two boxes are ticked (where only one should be ticked), the response should be left blank or coded as a full stop (.)
- Where a respondent has crossed out a response, this should not be entered in the data. Where a respondent has crossed out a response and instead ticked a second response option, this second choice should be entered into the data.
- Where a respondent has given their response inconsistently with the formatting of the questionnaire but where their intended response is nonetheless unambiguous on inspection of the completed questionnaire, then the respondent's intended response should be entered. For example, where a patient has written their date of birth in the boxes for the year of birth question, but written their year of birth in at the side of this, then the respondent's year of birth should be entered.
- For the year of birth question, unrealistic responses should still be entered except following point three above. For example, if a respondent enters ‘2012’ in the year of birth box, this should still be entered unless the respondent has unambiguously indicated their actual year of birth to the side.
- All responses should be entered into the dataset, regardless of whether or not the respondent was meant to respond to the question (e.g. where patients answer questions that they have been directed to skip past, these responses should still be entered).²
- For most questions, each column corresponds to one survey question. However, there are some exceptions to this rule. For multiple response questions that give the instruction “Tick all that apply”, each response option is treated as a separate question (see example below).

¹ If you want to use the pre-designed data input file to display frequencies on the other pages of the workbook, you will need to fill in the blank cells with a full stop (.) Data may be missing for a number of reasons: the patient may have skipped a question or a set of questions by following instructions or a patient may have not answered for some other reason. However, all missing data should be left blank or coded as a full stop (.), regardless of the reason for the omission.

² See section 6 for advice on how to clean your data once all responses have been entered and checked.

Example: how to code a multiple response question**K10. Do you have any of the following long-standing conditions? (Tick all that apply)**

- 1 ☒ Deafness or severe hearing impairment
- 2 ☐ Blindness or partially sighted
- 3 ☐ A long-standing physical condition
- 4 ☐ A learning disability
- 5 ☒ A mental health condition
- 6 ☐ I have a long-standing illness, such as cancer, HIV, diabetes, chronic heart disease, or epilepsy
- 7 ☐ No, I do not have a long-standing condition

Responses to each part of this question are coded: **1 if the box is ticked**
0 if the box is not ticked

Please note: if a respondent does not answer any part of a multiple response question, (i.e. does not tick any of the response options) then it should be left blank or coded as a full stop (.)

This question takes up seven columns in the data file, labelled as follows:

Column headings	K10_1	K10_2	K10_3	K10_4	K10_5	K10_6	K10_7
Codings for this example	1	0	0	0	1	0	0

We recommend that responses from patients who are missing their unique Patient Record Numbers should be added to the bottom of the file, and not matched to patients with similar demographics.

5.3 Checking the data for errors

We suggest the data is thoroughly checked for errors resulting from problems with data entry or similar. Ensuring high data quality is paramount and errors resulting from data entry problems can and should be corrected by checking against the appropriate completed questionnaire. We recommend that at least 5% of all data entered is checked for any errors by cross-checking against the original questionnaires.

Once the data has been entered, no responses should be removed or changed in any way except where responses are known to have been entered incorrectly or where inspection of the questionnaire indicates that the patient's intended response has not been captured.

The following checks on the data should be carried out:

Duplicate questionnaires:

Data sets should be examined for duplicate questionnaires, which should then be excluded (e.g. where a respondent has sent the original questionnaire and the questionnaire with the final reminder). We recommend retaining the most complete version of the questionnaire.

Range checks:

Data should be checked to ensure that question responses are within the defined response range. Check your data covers the full range of values specified for each variable (there may be exceptional cases where the data ranges do not reach the maximum in the range for a question, but the reasons for that should be apparent when the particular circumstances of the question are taken into account). The data should not exceed expected boundaries (e.g. are any dates in the future or too far in the past to be feasible? Are age distributions as would be expected for this population?).

Matching sample and response information:

To check that sample information had been correctly matched with response information the data drawn from patient files and their responses to the survey can be compared on fields where there is duplication. For example, does the year-of-birth in the sample information match the patient's responses on their year of birth? A few mis-matches are to be expected, but if there are more than around 3% per data set, this may indicate that the sample information has not been properly matched up with the response information and this should be checked and rectified. The same checks can be carried out for gender.

Before you start to carry out any data cleaning (as outlined in the following section), we strongly recommend you **save your 'raw' or 'uncleaned' data file**. You may need to return to this file if you make any mistakes whilst cleaning your data.

6 Data cleaning

6.1 Introduction

This section provides a description and specification of the processes that are used by the Co-ordination Centre to clean the data submitted by contractors and trusts as part of the national patient survey programme. By following the guidance contained in this section, it should be possible for trusts (and contractors on behalf of trusts) to recreate this cleaning process.

The Co-ordination Centre for the national acute patient survey programme uses the term 'data cleaning' to refer to all editing processes undertaken upon survey data once the survey has been completed and the data have been entered and collated. The aim in cleaning the data is to ensure an optimal balance between data quality and completeness. Thus, it is sought to remove responses that are known to be erroneous or inappropriate but to do this in a relatively permissive way to enable as many responses as possible to contribute to the overall survey results.

6.2 Filter questions

Some of the questions included in the survey are only relevant to a subset of respondents, and in these cases filter instructions are included in the questionnaire to route respondents past questions that are not applicable to them. For example, in the A&E survey, people who tick “no” to the question “Were you in any pain while you were in the A&E Department?” are instructed to skip the further questions on pain.

It is necessary to clean the data to remove responses where filter instructions have been incorrectly followed. In such cases, participants’ responses to questions that were not relevant to them are deleted from the dataset. Responses are **only** deleted where respondents have answered filtered questions despite ticking an earlier response on a routing question¹ instructing them to skip these questions (e.g. a respondent ticking “No” to the question in the example above but then answering the three questions about pain). Responses to filtered questions are **NOT** removed where the response to the routing question is missing because it is not certain that these questions should not have been answered.

A worked example of the cleaning process for removing unexpected responses to filtered questions is included in the *‘data cleaning manual’* for the most recent national survey, available on the NHS Surveys website (under the ‘National Surveys’ section).

6.3 Dealing with multiple response questions

For most questions, each column corresponds to one survey question. However, there are some exceptions to this rule. For multiple response questions that give the instruction “Tick all that apply”, each response option is treated as a separate question, as shown in the example in section 5.2.

The last response to each of the two multiple response questions on long-standing conditions is an exclusive option. If a respondent ticks option 7 (“No, I do not have a long-standing condition”) to the question ‘Do you have any of the following long-standing conditions?’, options 1-6 should not have also been ticked; if any of these options have been ticked, they should be recoded from ‘1’ to ‘0’ when cleaning the data. The same applies for the follow-up question (‘Does this condition(s) cause you difficulty with any of the following?’); if response option 8 (“No difficulty with any of these”) is ticked, options 1-7 should not have also been ticked. If they have been ticked they should be recoded from ‘1’ to ‘0’. (See example below).

¹ **Routing questions** are items on the questionnaire which instruct respondents to either continue on to the next question or to skip past irrelevant questions depending on their response to the routing question.

Example: how to clean a multiple response question**K10. Do you have any of the following long-standing conditions? (Tick all that apply)**

- 1 ☒ I have deafness or severe hearing impairment
- 2 ☐ I have blindness or are partially sighted
- 3 ☐ I have a long-standing physical condition
- 4 ☐ I have a learning disability
- 5 ☒ I have a mental health condition
- 6 ☐ I have a long-standing illness
- 7 ☒ No, I do not have a long-standing condition

BEFORE CLEANING: K10 is coded as follows:

Column headings	K10_1	K10_2	K10_3	K10_4	K10_5	K10_6	K10_7
Codings for this example	1	0	0	0	1	0	1

AFTER CLEANING: K10 is coded as follows:

Column headings	K10_1	K10_2	K10_3	K10_4	K10_5	K10_6	K10_7
Codings for this example	0	0	0	0	0	0	1

When the data is cleaned, the responses to K10_1 and K10_5 are re-coded as '0' because option 7 ("No, I do not have a long-standing condition") has also been ticked.

6.4 Dealing with demographics

Basic demographic information, such as the age, sex, and ethnicity of patients are included in the sample section of the data. The 'About You' section at the end of the questionnaire also asks respondents to provide this information. In a minority of cases, the information contained in the sample file and that provided by the respondents themselves may not correspond – for example, the sample may identify an individual as male, but they report being female in the survey.

Because of this, and because questions about demographics tend to produce relatively high rates of non-response, it is not appropriate to rely on either source of data alone.

Where responses to demographic questions are present, it is assumed these are more likely to be accurate than sample information, since it is assumed that respondents are best placed to know their own sex, age, and ethnic group¹. Where responses to demographic questions are missing, however, sample data can be used in their place.

For demographic analysis on groups of cases, it is necessary to use a combination of the information supplied in the sample frame and by respondents. To do this, we first copy all valid responses to survey demographic questions and use this data to create a new variable. Where responses are missing we then add in the relevant sample information. For a very small number of patients, demographic information may be missing in both the sample and response sections; in such cases data must necessarily be left missing in the new variable.²

A common error when completing year of birth questions on forms is for respondents to accidentally write in the current year – thus responses of ‘2012’ should be set to missing during cleaning. Out-of-range responses should also be set to missing.³

6.5 Usability and eligibility

Sometimes questionnaires are returned with only a very small number of questions completed. For the national patient surveys, questionnaires containing fewer than five responses are considered ‘unusable’. When this occurs, the Co-ordination Centre deletes all responses pertaining to such cases and the outcome code of 1 (‘returned completed questionnaire’) relating to these cases is changed to 6 (‘questionnaire not returned’). Please note that the number of responses per questionnaire must be counted **after** all other cleaning, and that all responses remaining at this stage will be counted (this includes responses to the demographic questions)⁴. This should only affect a very limited number of cases, and so should not have a significant impact on response rates.

¹ Despite this assumption, please note that any respondent recorded as over the age of sixteen in the sample information, but who reports themselves as under the age of sixteen in their response to the survey, should *not* be considered as ineligible. This is because of the difficulty of inferring the source of errors when year of birth from sample and response sections are mismatched. We cannot be *certain* whether this mismatch occurs due to an error in the sample file, an error in the patient’s completion of the questionnaire form, or an error in data entry. See section 6.5.

² Please note that whilst a combination of sample and response information is used for demographic analysis on groups of cases, **only the sample information is used to calculate response rates by demographic groups**. Unlike subgroup analysis of variables which is only carried out for respondents, response rates are calculated using all patients in the sample. Because response rates vary between groups with differing demographic variables, using response and sample data to calculate response rates would create a systematic source of bias because we are only able to amend information for the *respondents*. Therefore, only the sample information should be used to calculate response rates by demographic groups.

³ The majority of out-of-range responses present in data relating to year of birth questions result from errors in data entry (for example, not keying one of the digits – so ‘1983’ may become 983, 183, 193, or 198). In such cases it is important that the responses are checked against the completed questionnaire forms, and the data corrected if necessary.

⁴ Please note that the multiple choice questions are only counted once. So for example, if a respondent ticks three options at a multiple choice question, this would count as only one response for the purpose of determining if a questionnaire is usable.

Outcome codes for respondents should also be changed if respondents are believed to be under the age of sixteen and therefore ineligible for participation. This is unlikely to affect more than a handful of cases throughout the survey, as patients recorded as being aged under 16 will be identified and removed from the sample **before** the start of the survey. Sample members will not, however, be removed from the sample if data on their year of birth is missing from the sampling frame. If sample information on a respondent's year of birth is missing, though, and their survey response indicates that they are under 16, then the outcome code for that patient should be recoded from 1 ('returned completed questionnaire') to 5 ('ineligible for participation in the survey'). ***This should only be done where sample information is missing.*** If sample information indicates that a patient is aged 16 or over, but this is contradicted by the patient's response, then the patient's survey outcome should remain as 1 ('returned completed questionnaire'). This is to avoid removing legitimate responses because of an overly conservative approach to assessing eligibility. In other words, where the patient's age is uncertain (because sample and response information contradict each other, and it is uncertain in which data the error lies) the benefit of the doubt is given in any assessment of eligibility.

6.6 Missing responses

It is useful to be able to see the numbers of respondents who have missed each question for whatever reason. Responses are considered to be missing when a respondent is expected to answer a question but no response is present. For non-filtered questions, responses are expected from all respondents – thus any instance of missing data constitutes a missing response. For filtered questions, only respondents who have answered a previous routing question instructing them to go on to that filtered question or set of filtered questions are expected to give answers. Where respondents to the survey have not answered a routing question, they are not expected to answer subsequent filtered questions; thus only where respondents were explicitly instructed to answer filtered questions should such blank cells be coded as missing responses.

The Co-ordination Centre for the national acute patient survey programme codes missing responses in the data with the value 999¹. For results to be consistent with those produced in the national patient survey programme, missing responses should be presented but should not be included in the base number of respondents when calculating percentages.

6.7 Non-specific responses

'Non-specific responses' is a loose term for response options that can be considered as not being applicable to the respondent in terms of directly answering the specific question to which they are linked. Most commonly, these are responses such as "Don't know/Can't remember", which indicate a failure to recall the issue in question. Likewise, responses that indicate the question is not applicable to the respondent are also considered 'non-specific' – for example, responses such as "I did not need to find a place to park" or "I did not discuss my condition with a receptionist".

As well as excluding missing responses from results, non-specific responses should also be removed from base numbers for percentages. The rationale for this is to facilitate easy comparison between institutions by presenting only results from those patients who felt able to give an evaluative response to questions.

¹ This is an arbitrary value chosen because it is out-of-range for all questions in the survey.

Appendix 1: Template Freephone Query Sheet

Date: _____ Time: _____

Patient's survey number:

(On front of the questionnaire and on the patients' address label)

Name of NHS Trust or hospital/clinic etc. _____

Who called? Patient ☐ Relative/Friend ☐ Carer ☐

Problem/comment

Patient request

1. Received reminder but not mailing 1

Please send mailing 3 ☐

Don't send mailing 3 ☐

2. Returned questionnaire some time ago but just received reminder

Send mailing 3 – will fill it in again ☐

Don't send mailing 3 ☐

3. Patient wants to opt out*

No reason/not interested ☐

Too ill/frail to complete ☐

4. Should not have been included*

Deceased patient ☐

Patient not at this address ☐

Child ☐

** Advise callers that every effort will be made to remove them from the database, but to please ignore any further reminders they might accidentally receive*

Other reason to exclude:

5. How do I complete this questionnaire?

Details of query:

6. Questionnaire completed over the phone

☐

Caller required interpretation services ☐

Name of person dealing with call: _____

Date logged in database: _____

Code logged in database: