

GUIDANCE ON DATA CLEANING FOR THE NHS NATIONAL SURVEY OF LOCAL HEALTH SERVICES 2008

THE CO-ORDINATION CENTRE FOR THE
PRIMARY CARE TRUST SURVEY

Contacts

The Co-ordination Centre for the Primary Care Trust Survey
Picker Institute Europe
King's Mead House
Oxpens Road
Oxford
OX1 1RX

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

Author

Julia Martin

Updates

Before you use this document, check that you have the latest version, as there might be some small amendments from time to time (the date of the last update is on the front page). In the very unlikely event that there are any major changes, we will e-mail all trust contacts and contractors directly to inform them of the change.

This document is available from the Co-ordination Centre website at:

www.NHSSurveys.org

Questions and comments

If you have any questions or concerns regarding this document, or if you have any specific queries regarding the submission of data, please contact the Co-ordination Centre:

By e-mail: primarycare@PickerEurope.ac.uk

By phone: 01865 208127

Contents

1	Primary Care Trusts Survey 2008 – data cleaning.....	1
1.1	Introduction	1
1.2	The core and extended questionnaires	1
1.3	Definitions	1
2	Submitting raw ('uncleaned') data	3
3	Editing/cleaning data after submission	4
3.1	Approach and rationale.....	4
3.2	Filters	4
	Figure 1 - List of routing/filtering instructions	5
3.3	Additional filter	5
	Figure 2 – Additional filtering instruction	5
3.4	Dealing with multiple response questions	6
3.5	Dealing with demographics.....	7
3.6	Usability and eligibility.....	8
3.7	Missing responses	10
3.8	Non-specific responses	10
4	Appendix A: Example of cleaning	11
4.1	Incorrectly followed routing	11
	Figure 2: Example 'raw'/uncleaned' data	11
	Figure 3: Data from Figure 3 following cleaning	12
5	Appendix B: Non-specific responses.....	13
6	Submitting data	18

1 Primary Care Trusts Survey 2008 – data cleaning

1.1 Introduction

Once fieldwork for the 2008 primary care trusts survey has been completed, participating trusts and contractors will be required to submit data to the Co-ordination Centre in a **raw ('uncleaned')** format. Following this, data for all trusts in the survey will be collated and the full dataset will be cleaned together to ensure that cleaning taking place on data pertaining to different trusts is comparable.

This document provides a description and specification of the processes that will be used by the Co-ordination Centre to clean and standardise data submitted by contractors and trusts as part of the 2008 survey of local health services. By following the guidance contained in this document, it should be possible to recreate this cleaning process.

If you have any comments or queries regarding this document please contact the Co-ordination Centre on **01865 208 127**, or e-mail us at primarycare@pickereurope.ac.uk.

1.2 The core and extended questionnaires

For the 2008 primary care trusts survey, all trusts have the option to use either the 56 item 'core' questionnaire or to use an extended questionnaire with further questions available from the 'question bank' for the survey. The Co-ordination Centre only requires data to be submitted for the 56 core items, and so all cleaning undertaken by ourselves will involve only these 56 core items. As such, this document looks only at the cleaning required for the core survey. Nonetheless, the rules and principles of data cleaning described here can be applied to data from an extended questionnaire.

1.3 Definitions

Definitions of terms commonly used in this document, as they apply to the 2008 primary care trusts survey are as follows:

Raw/uncleaned data: 'Raw' or 'uncleaned' data is data that has been entered verbatim from completed questionnaires without any editing taking place to remove contradictory or inappropriate responses; thus, all responses ticked on the questionnaire should be included in the data entry spreadsheet¹ (see [Section 2: Submitting raw \('uncleaned'\) data](#), for detailed guidance on submitting raw data). The requirement for raw/uncleaned data does **not**, however, preclude the checking of data 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.

¹ Except where: a) multiple responses have been ticked - set these to missing (The **exceptions** to this are for the 'tick all that apply' questions **Q49** and **Q50** where respondents may tick more than one response option) b) year of birth has been entered in incorrect format - if the patient's *intended* response is unambiguous from the questionnaire, then enter this.

Data cleaning: The Co-ordination Centre uses the term 'data cleaning' to refer to all editing processes undertaken upon survey data once the survey has been completed and the data has been entered and collated.

Routing questions: These 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. For the 2008 local health services survey, the routing questions in the core questionnaire are **Q1, Q2, Q6, Q8, Q10, Q18, Q23, Q25, Q29, Q31, Q38, Q47** and **Q49**.

Filtered questions: Items on the questionnaire that are not intended to be answered by all respondents are referred to as 'filtered' questions. Whether individual respondents are expected to answer filtered questions depends on their responses to preceding routing questions. For the 2008 local health services survey, the filtered questions in the core questionnaire are **Q2—Q5¹, Q7—Q9², Q11—Q17, Q19—Q22, Q24, Q26—Q28, Q30—Q33³, Q39, Q48, and Q50-Q51**.

Non-filtered questions: these are items in the questionnaire which are not subject to any routing/filtering and which should therefore be answered by all respondents. For the 2008 local health services survey, the non-filtered questions are **Q1, Q6, Q10, Q18, Q23, Q25, Q29, Q34—Q38, Q40—Q47, Q49, and Q52—Q56**.

Out-of-range data: This refers to instances where data within a variable has a value that is not permissible. For categorical data – as in the case of the majority of variables in this survey – this would mean a value not allowed in the data, for example, a value of '3' being entered in a variable with only two response categories (1 or 2). For scalar data – eg year of birth – data is considered to be out-of-range if it specifies a value that is not possible (for instance, year of birth as 983 or 2983). Out-of-range responses entered into the dataset should not be automatically (eg, algorithmically) removed prior to submitting the data to the Co-ordination Centre (see [Section 2: Submitting raw \('uncleaned'\) data](#)).

Non-specific response: This 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 considered 'non-specific' – for example, responses such as "I did not need to discuss anything" or "No treatment or action was needed". A full listing of such responses for the 2008 local health services survey can be found in [Appendix B: Non-specific responses](#).

¹ The range Q2-Q5 includes three separate sets of filtered questions; Q2-Q5, Q3-Q4 and Q4.

² The range Q6-Q9 includes two separate sets of filtered questions; Q6-Q9, and Q9.

³ The range Q30-Q33 includes three separate sets of filtered questions; Q30-Q34, Q32-Q33 and Q33.

2 Submitting raw ('uncleaned') data

For the 2008 primary care trust survey, trusts and contractors are required to submit raw ('uncleaned') data to the Co-ordination Centre. For clarification, raw data is created as follows:

- i) All responses should be entered into the dataset, regardless of whether or not the respondent was meant to respond to the question (eg, where patients answer questions that they have been directed to skip past, these responses should still be entered)
- ii) Where a respondent has ticked more than one response category on a question, this should be set to 'missing' in the data. The **exceptions** to this are for the 'multiple response' questions, **Q49** and **Q50**, where respondents may tick more than one response option (ie. 'tick all that apply')
- iii) 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
- iv) 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 **Q53**, but written their *year* of birth in at the side of this, then the respondent's year of birth should be entered
- v) For the year of birth question, unrealistic responses should still be entered *except* following **iv)** above. For example, if a respondent enters '2008' 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
- vi) 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. This includes 'out-of-range' responses, which must **not** be *automatically* removed from the dataset. Responses in the dataset should only be changed before submission to the Co-ordination Centre where they are found to have been entered inconsistently with the respondent's intended response.

3 Editing/cleaning data after submission

3.1 Approach and rationale

The aim of the Co-ordination Centre in cleaning the data submitted to us is to ensure an optimal balance between data quality and completeness. Thus, we seek to remove responses that are known to be erroneous or inappropriate but do this in a relatively permissive way so as to enable as many responses as possible to contribute to the overall survey results.

3.2 Filters

Some of the questions included in the survey are only relevant to a subset of respondents, and in these cases filter instructions on the questionnaire are used to route respondents past questions that are not applicable to them. For example, people who tick “no” to **Q25** (“In the last 12 months, has anyone at your GP practice/ health centre referred you to a specialist (e.g. a hospital consultant)?”) are instructed to skip all further questions on referrals (e.g. **Q26**, **Q27** and **Q28**).

It is necessary to clean the data to remove inappropriate 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 question instructing them to skip these questions (eg a respondent ticking “No” to **Q25** but then answering the three questions about referrals as in the example above). Responses to ‘filtered’ questions are not removed where the response to the routing question is missing. For example, **Q11-Q17** are filtered by the response to **Q10** (e.g. if **Q10=2**), but if a respondent does not answer **Q10**, or if the **Q10** response is missing for any reason, then responses to **Q11-Q17** should not be removed.

[Figure 1](#) (overleaf) shows a summary of all routing questions, and the filtered questions they relate to, that are included on the 2008 local health services survey. Please note that these instructions should be followed sequentially in order to be consistent with the procedures applied by the Co-ordination Centre.

Figure 1 - List of routing/filtering instructions

ROUTING QUESTION	RESPONSE VALUES	FILTERED QUESTIONS
<i>if</i> Q1 = 2	<i>then delete responses to:</i>	Q2 – Q5
<i>if</i> Q2 = 1	<i>then delete responses to:</i>	Q3 – Q4
<i>if</i> Q2 = 2, 5 OR 6	<i>then delete responses to:</i>	Q3
<i>if</i> Q6 = 2	<i>then delete responses to:</i>	Q7 – Q9
<i>if</i> Q8 = 2	<i>then delete responses to:</i>	Q9
<i>if</i> Q10 = 2	<i>then delete responses to:</i>	Q11 – Q17
<i>if</i> Q18 = 2 OR 3	<i>then delete responses to:</i>	Q19 – Q22
<i>if</i> Q23 = 2	<i>then delete responses to:</i>	Q24
<i>if</i> Q25 = 2	<i>then delete responses to:</i>	Q26 – Q28
<i>if</i> Q29 = 2	<i>then delete responses to:</i>	Q30 – Q33
<i>if</i> Q31 = 1	<i>then delete responses to:</i>	Q32 – Q33
<i>if</i> Q31 = 2, 6 OR 7	<i>then delete responses to:</i>	Q32
<i>if</i> Q38 = 1	<i>then delete responses to:</i>	Q39
<i>if</i> Q47 = 5	<i>then delete responses to:</i>	Q48
<i>if</i> Q49_7 = 1	<i>then delete responses to:</i>	Q50_1, Q50_2, Q50_3, Q50_4, Q50_5, Q50_6, Q50_7, Q50_8, Q51

Please note that these instructions should be followed sequentially in the order shown above.

A worked example of the cleaning process for removing unexpected responses to filtered questions is included in [Appendix A: Example of cleaning](#).

3.3 Additional filter

In addition to the questions that having routing instructions given on the questionnaire, two further questions have also had the response options filtered. If a respondent says that they have not visited their GP practice/ health centre in the last 12 months (i.e. Q6=2) responses to Q34 and Q35 are deleted. If the respondent does not answer Q6 or if the the Q6 response is missing for any reason, then responses to Q34 and Q35 should not be removed.

Figure 2 – Additional filtering instruction

ROUTING QUESTION	RESPONSE VALUES	FILTERED QUESTIONS
<i>if</i> Q6 = 2	<i>then delete responses to:</i>	Q34 – Q35

3.4 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 (**Q49** and **Q50**) that give the instruction “tick all that apply”, each response option is treated as a separate question in the data entry spreadsheet.

Example

Q49. 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 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 ticked¹

Q49 takes up seven columns in the data file, labelled as follows:

Column headings	Q49_1	Q49_2	Q49_3	Q49_4	Q49_5	Q49_6	Q49_7
Codings for this example	1	0	0	0	1	0	0

However, the last response options to Q49 and Q50 are exclusive options. If a respondent ticks option 7 to **Q49** (i.e. “I do not have a long-standing condition”), then 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 **Q50**; 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, then they should be recoded from ‘1’ to ‘0’

¹ Please note: if a respondent does not answer any part of a multiple response question, (ie does not tick any of the response options) then it should be left blank or coded as a full stop (.)

Example

Q49. 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 I do not have a long-standing condition

BEFORE CLEANING: Q49 is coded as follows:

Column headings	Q49_1	Q49_2	Q49_3	Q49_4	Q49_5	Q49_6	Q49_7
Codings for this example	1	0	0	0	1	0	1

AFTER CLEANING: Q49 is coded as follows:

Column headings	Q49_1	Q49_2	Q49_3	Q49_4	Q49_5	Q49_6	Q49_7
Codings for this example	0	0	0	0	0	0	1

When the data is cleaned, the responses to Q49_1 and Q49_5 are re-coded as '0' because option 7 (i.e. "I do not have a long-standing condition") has also been ticked.

3.5 Dealing with demographics

Basic demographic information, including age and sex of patients are included in the sample section of the data, but the 'About You' section at the end of the questionnaire also asks respondents to provide this information. In a minority of cases, the information provided from the sample frame and by the respondents does not correspond – for example, the sample may identify an individual as male only for them to report being female (eg **Q52=2**).

Because of this, and because questions about demographics tend to produce relatively high item non-response rates, 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 frame information (since it is assumed that respondents are best placed to

know their own sex or age)¹. Where responses to demographic questions are missing, however, sample data are used in their place.

For demographic analysis on groups of cases, then, it is necessary to use some combination of the information supplied in the sample frame and by the respondents. To do this, we first copy all valid responses to survey demographic questions into a new variable. Where responses are missing we then copy in the relevant sample information (note that 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

Certain demographic variables require special consideration during data cleaning:

Age (Q53)

A common error when completing year of birth questions on forms is for respondents to accidentally write in the current year – thus responses to Q53 of ‘2008’ will be set to missing during cleaning. Out-of-range responses will also be set to missing³. For the 2008 primary care trusts survey, out-of range responses are defined as **Q53≤1880 OR Q53≥2009**.

3.6 Usability and eligibility

Sometimes questionnaires are returned with only a very small number of questions completed. For the local health services survey 2008, questionnaires containing fewer than five responses are considered ‘unusable’ – we will delete all responses pertaining to such cases and outcome codes of 1 (‘returned useable questionnaire’) relating to these cases will be changed to 6 (‘questionnaire not returned’). Please note that the number of responses per questionnaire will be counted after all other cleaning, and that all responses remaining at this stage will be counted (including responses to the demographic questions and so on)⁴. This should only affect a very limited number of cases, and so should not have a significant impact on response rates.

¹ Please note though that respondents will *not* be considered ineligible for the survey solely on the basis of their response to the year of birth question indicating that they are under the age of sixteen. This is because of the difficulty of inferring the source of errors when age from the sample information and year of birth from the response section are mismatched – in other words, because 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 *overleaf*].

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 in that 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 be checked against the completed questionnaire forms, and data corrected if necessary, **prior** to submission of data to the Co-ordination Centre.

⁴ Please note that the multiple choice questions, Q49 and Q50 are only counted once. So for example, even if Q49_1 and Q49_4 are ticked, this would count as only one response for the purpose of determining if a questionnaire is usable.

Outcome codes for respondents will also be changed if the respondents are believed to be under the age of sixteen and therefore ineligible for participation. Since the sample files for the survey have been drawn directly from the NHAIS system, this is unlikely to affect more than a handful of cases throughout the survey. Sample members will not, however, be removed from the sample if data on their age is missing from the sampling frame. If sample information on a respondent's age is missing, though, and their response to **Q53** indicates that they are under 16 (specifically, if **Q53 > 1992**) 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 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. 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 in different instances either of these may be accurate or inaccurate) the benefit of the doubt is given in any assessment of eligibility.

3.7 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 missed 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 codes missing responses in the data with the value 999¹. For results to be consistent with those produced by the Co-ordination Centre, missing responses should be presented but should not be included in the base number of respondents for percentages.

3.8 Non-specific responses

As well as excluding missing responses from results, the Co-ordination Centre also removes non-specific responses 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. For a full listing of ‘non-specific’ responses in the 2008 local health services survey, please see [Appendix B: Non-specific responses](#).

¹ This is an arbitrary value chosen because it is ‘out-of-range’ for all other questions on the survey.

4 Appendix A: Example of cleaning

4.1 Incorrectly followed routing

Figure 2: Example 'raw'/'uncleaned' data

Record	Outcome	Q1	Q2	Q3	Q4	Q5
Sample Reference Number	Outcome of sending questionnaire (N)	Have you made an appointment with a doctor from your GP practice/ health centre in the last 12 months?	The last time you saw a doctor from your GP practice/ health centre, how long did you wait for an appointment?	What was the main reason you waited?	How do you feel about the length of time you waited for an appointment with a doctor?	If you want to make a doctor's appointment 3 or more working days in advance does your GP practice/ health centre allow you to do that?
A	6					
B	1	2
C	1	1	1	3	2	1
D	4					
E	1	2	2	.	.	.
F	6					
G	1	2	5	2	3	1
H	1	1	3	1	1	1

Figure 2 shows hypothetical raw/uncleaned data for eight sample members, five of whom have responded to the survey. It can be seen from this data that some of the respondents have followed filter instructions from routing questions incorrectly:

Respondents 'E' and 'G' have reported that they haven't made a doctors appointment in the last 12 months (**Q1=2**), but have both responded to filtered questions ('E' has answered the first question after the filter (**Q2**) before skipping the remaining questions, whilst 'G' has answered **Q2, Q3, Q4 and Q5**).

Following the cleaning instructions above will remove these inappropriate responses. Firstly, the filter instructions specify that:

<i>if</i>	Q1 = 2	<i>then delete responses to:</i>	Q2 – Q5
-----------	---------------	----------------------------------	----------------

In accordance with this, all responses for **Q2, Q3, Q4 and Q5** must be removed in cases where the respondent has ticked **Q1=2** ('waiting list or planned in advance'). Looking in column **Q1** we can see that three respondents, 'B', 'E' and 'G', have ticked **Q1=2**, so any responses they gave to questions two through to five should be removed. This will lead to one response being removed for patient 'E' (**Q2**) and four responses being removed for respondent 'G' (**Q2, Q3, Q4, and Q5**), who have for whatever reason followed the routing instructions incorrectly and continued to answer the section on making a doctors appointment.

[Figure 3](#) (below) shows how the data would look following cleaning by the co-ordination centre to remove responses to filtered questions that should have been skipped – cells where responses have been removed are shaded.

Figure 3: Data from Figure 3 following cleaning

Record	Outcome	Q1	Q2	Q3	Q4	Q5
Sample Reference Number	Outcome of sending questionnaire (N)	Have you made an appointment with a doctor from your GP practice/ health centre in the last 12 months?	The last time you saw a doctor from your GP practice/ health centre, how long did you wait for an appointment?	What was the main reason you waited?	How do you feel about the length of time you waited for an appointment with a doctor?	If you want to make a doctor's appointment 3 or more working days in advance does your GP practice/ health centre allow you to do that?
A	6					
B	1	2
C	1	1	1	3	2	1
D	4					
E	1	2
F	6					
G	1	2
H	1	1	3	1	1	1

5 Appendix B: Non-specific responses

The following table lists all 'non-specific responses' included in the 2008 primary care trusts survey. Please note that this table also includes items from the question bank which are not included in the minimal 'core' questionnaire. Numbers in the final column indicate the response options that should be considered non-specific. Where the 'non-specific responses' column contains only a dash, the relevant question has no such response options.

CORE	BANK	Question	Non-specific responses
Q1	A1	Have you made an appointment with a doctor from your GP practice/ health centre in the last 12 months?	-
	A2	Was the last contact with a doctor from your GP practice about a medical problem for yourself or for a child in your care?	-
Q2	A3	The last time you saw a doctor from your GP practice/ health centre, how long did you wait for an appointment?	5, 6
Q3	A4	What was the main reason you waited?	-
Q4	A5	How do you feel about the length of time you waited for an appointment with a doctor?	-
Q5	A6	If you want to make a doctor's appointment 3 or more working days in advance does your GP practice/ health centre allow you to do that?	3
Q6	B1	Have you visited your GP practice/ health centre in the last 12 months?	-
Q7	B2	When you arrived, how would you rate the courtesy of the receptionist?	-
	B3	In the reception area, could other patients overhear what you talked about with the receptionist?	4
	B4	Were you offered the chance to speak to the receptionist in a private area?	-
Q8	B5	How long after your appointment time did you have to wait to be seen?	1, 6
Q9	B6	Did someone tell you how long you would have to wait?	4
Q10	C1	Have you seen a doctor from your GP practice/ health centre in the last 12 months?	-
Q11	C2	Did the doctor listen carefully to what you had to say?	-
Q12	C3	Were you given enough time to discuss your health or medical problem with the doctor?	4
Q13	C4	Were you involved as much as you wanted to be in decisions about your care and treatment?	-
Q14	C5	If you had questions to ask the doctor, did you get answers that you could understand?	4
Q15	C6	Did the doctor explain reasons for any treatment or action in a way that you could understand?	4, 5
Q16	C7	Did you have confidence and trust in the doctor?	-
	C8	Did the doctor know enough about your condition or treatment?	4
Q17	C9	Did the doctor treat you with respect and dignity?	-
	C10	Were you able to discuss any emotional issues that might be affecting your health (e.g. anxiety, depression)?	4
	C11	Were you able to discuss how your family and/or living situation might be affecting your health (e.g. housing problems, family responsibilities, work-related problems)?	4
Q18	D1	In the last 12 months, have you had any new medicine(s) (including tablets, suppositories, injections) prescribed to you by a doctor or nurse practitioner from your GP practice/ health centre?	3

CORE	BANK	Question	Non-specific responses
Q19	D2	Were you involved as much as you wanted to be in decisions about the best medicine for you?	-
Q20	D3	Were you given enough information about the purpose of the medicine?	4, 5
Q21	D4	Were you given enough information about any side-effects the medicine might have?	4, 5
Q22	D5	Were you given enough information about how to use the medicine (e.g. when to take it, how long you should take it for, whether it should be taken with food)?	4, 5
Q23	D6	Have you been taking any prescribed medicine(s) for 12 months or longer?	-
Q24	D7	In the last 12 months, have you seen anyone at your GP practice to check how you are getting on with this medicine (i.e. have your medicines been reviewed)?	3
	D8	In the last 12 months, have you asked a pharmacist for any advice on medicines?	-
	D9	Was the pharmacist's advice helpful?	4
	E1	In the last 12 months, have you had any tests (e.g. blood tests, swabs, smear tests) carried out by anyone from your health centre?	3
	E2	Was the purpose of the test(s) explained in a way you could understand?	4
	E3	Did someone tell you how you would get the results of your test(s)?	3
	E4	Did someone tell you when you should expect to get the results of your test(s)?	3
	E5	Did you get your test results on time?	3
	E6	Did someone explain the results of the tests in a way you could understand?	4, 5
Q25	F1	In the last 12 months, has anyone at your practice referred you to a specialist (e.g. a hospital consultant)?	-
Q26	F2	When you were referred to see a specialist, were you offered a choice of hospital for your first hospital appointment?	3
	F3	Did your GP tell you why you were not offered a choice about where you were referred to?	4
	F4	How many different choices were you offered? (Your choices may have included different hospitals or community based services, such as GPs with special interests or community clinics)	5
	F5	Overall, were you given enough information by your GP or another health care professional at your GP practice/ health centre to help you make your choice?	4
	F6	Who or what were your source or sources of information about the different places?	7
	F7	Were you given a choice about who you were referred to (i.e. which specialist)?	4
Q27	F8	When you first saw the person you were referred to, did he/she seem to have all the necessary information about you and your condition or treatment?	4, 5
Q28	F9	Did you receive copies of letters sent between the specialist and your GP?	4, 5
Q29	G1	Have you seen anyone else from a GP practice/ health centre other than a doctor in the last 12 months?	-
Q30	G2	The last time you saw someone other than a doctor from a GP practice/ health centre, who did you see?	-

CORE	BANK	Question	Non-specific responses
Q31	G3	The last time you saw this person, how long did you wait for an appointment?	6,7
Q32	G4	What was the main reason you waited?	-
Q33	G5	How do you feel about the length of time you waited for an appointment with this person?	-
	G6	Were you involved as much as you wanted to be in decisions about your care and treatment?	4
	G7	Did that person explain the reasons for any treatment or action in a way that you could understand?	4, 5
	G8	Did that person treat you with respect and dignity?	-
	G9	Did you have confidence and trust in that person?	-
	H1	In the past 12 months, have you tried to contact your local health centre/ GP practice about a medical problem when the practice was closed, either on your own behalf or for someone else?	-
	H2	The last time you called the practice out of hours, did you get through to someone?	-
	H3	If a doctor or nurse came to visit you at home, how long did you wait from the time you finished the telephone call?	6
	H4	Overall, was the main reason you contacted the practice out of hours dealt with to your satisfaction?	-
	H5	Have you heard of NHS Direct (a 24 hour helpline staffed by nurses)?	3
	H6	In the last 12 months, have you contacted NHS Direct, either on your own behalf or on behalf of someone else?	-
	H7	Was your call to NHS Direct dealt with satisfactorily?	-
Q34	J1	Was the main reason you went to your GP practice/ health centre dealt with to your satisfaction?	-
Q35	J2	In your opinion, how clean is the GP practice/ health centre?	5
	J3	How easy do you find it to move around inside the practice/ health centre?	4
Q36	J4	Have you had a problem getting through to your GP practice/ health centre on the phone?	4
	J5	In the last 12 months, have you talked over the phone to a doctor from your GP practice/ health centre about a medical problem?	-
Q37	J6	In the last 12 months, have you ever been put off going to your GP practice/ health centre because the opening times are inconvenient for you?	-
	J7	If it were possible for your GP practice/ health centre to open at additional times, which of these times would you most like it to be open?	-
	J8	If your GP practice/ health centre were to be open either earlier in the morning or later in the evening, how many days a week would you want this to happen?	4
	J9	If your GP practice/ health centre were to be open extra hours but had to close for some of its normal hours to allow this, would this be acceptable to you?	-
	J10	Do you need help understanding English?	-
	J11	The last time you saw someone from your local health centre/ GP practice, was there someone who could interpret for you?	-
Q38	K1	Do you visit a dentist regularly (that is at least once every 2 years)?	4
Q39	K2	Would you like to receive dental care as an NHS patient?	-
	K3	In the last 24 months, have you visited a dentist as an NHS patient?	3
	K4	Why did you go to a dentist?	-

CORE	BANK	Question	Non-specific responses
	K5	How long did it take to get an appointment?	-
	K6	Were you involved as much as you wanted to be in decisions about your dental care and treatment?	-
	K7	Did the dentist explain the reasons for any treatment or action in a way that you could understand?	4, 5
	K8	Did you have confidence and trust in the dentist?	-
	K9	Did dental staff do everything they could to help control any pain you experienced?	4, 5
	K10	Overall, was the main reason for this visit dealt with satisfactorily?	-
	K11	Have You tried to get out-of-hours dental treatment as an NHS patient during the past 12 months?	-
Q40	L1	In the last 12 months have you had your blood pressure taken by anyone from your GP practice/ health centre?	3
Q41	L2	In the last 12 months, have you been given advice from your GP practice/ health centre on your weight?	5
Q42	L3	In the last 12 months, have you been given advice or help from your GP practice/ health centre on eating a healthy diet?	4
Q43	L4	How regularly do you eat the recommended 5 portions of fruit or vegetables a day?	5
Q44	L5	In the last 12 months, have you been given advice or help from your GP practice/ health centre on getting enough exercise?	4
	L6	How many days a week on average, do you do at least 30 minutes of physical activity (e.g. brisk walking, cycling, sport, heavy housework/ gardening, or as part of a physically active job)?	5
	L7	Thinking about smoking, which of the following best applies to you?	-
Q45	L8	In the last 12 months, have you been given advice or help from your GP practice/health centre on giving up smoking?	1, 5
Q46	L9	In the last 12 months, have you been asked by someone at your GP practice/ health centre about how much alcohol you drink?	3
Q47	L10	In the last 12 months, have you been given advice or help from your GP practice/health centre on sensible alcohol intake?	4, 5
Q48	L11	How many units of alcohol do you normally drink in a week? (1 unit is roughly equivalent to one small glass of wine, half a pint of beer or one pub measure of spirits)	6
Q49_1	L12_1	I have a long-standing condition involving deafness or hearing impairment	-
Q49_2	L12_2	I have a long-standing condition involving blindness or are partially sighted	-
Q49_3	L12_3	I have a long-standing physical condition	-
Q49_4	L12_4	I have a long-standing condition involving a learning disability	-
Q49_5	L12_5	I have a long-standing mental health condition	-
Q49_6	L12_6	I have a long-standing condition involving an illness such as cancer, HIV, diabetes, CHD, or epilepsy	-
Q49_7	L12_7	I do not have a long-standing condition	-
Q50_1	L13_1	This condition causes me difficulty with everyday activities that people of my age can usually do	-
Q50_2	L13_2	This condition causes me difficulty at work, in education, or training	-
Q50_3	L13_3	This condition causes me difficulty with access to buildings, streets, or transport vehicles	-
Q50_4	L13_4	This condition causes me difficulty with reading or writing	-
Q50_5	L13_5	This condition causes me difficulty with people's attitudes to me because of my condition	-

CORE	BANK	Question	Non-specific responses
Q50_6	L13_6	This condition causes me difficulty with communicating, mixing with others, or socializing	-
Q50_7	L13_7	This condition causes me difficulty with other activities	-
Q50_8	L13_8	This condition does not cause me difficulty with any of these	-
Q51	L14	In the last 12 months, have you had enough support from local services or organisations to help you to manage your long-term health condition(s)? (Please think about all services and organisations, not just health services)	4,5
	L15	In the last 12 months, have you been offered a flu jab (influenza vaccination)?	4, 5
	L16	In the last 12 months, have you been given advice or help from your GP practice/ health centre on contraception/family planning?	4, 5
	L17	In the last 12 months, have you been given advice or help from your GP practice/ health centre on safer sex?	4, 5
	M1	When did you last have an eye test?	3
	M2	Do you have any problems with your hearing which affect your everyday life?	-
	M3	Have you ever been provided with advice/assistance to enable you to cope with your hearing problem?	-
	M4	Do you have any difficulty carrying out your daily activities (dressing, washing, going to the toilet, moving about your home, cooking a meal etc)?	-
	M5	Have you been provided with advice to help you cope with carrying out your daily activities?	-
	N1	Would you know how to get involved in making decisions about the NHS in your area? (e.g. attending meetings, becoming a member of a local patients group)	-
	N2	Have you changed your GP (family doctor) within the last 12 months?	-
	N3	What was the reason for this change?	-
	N4	How easy was it to register with another GP (family doctor)?	-
	N5	Have you received a copy of Your Guide to Local Health Services? (This is a leaflet on local NHS services and how to access them).	3
	N6	Did you find the Guide to Local Health Services useful?	4
Q52	P1	Are you male or female?	-
Q53	P2	What was your year of birth?	-
Q54	P3	How old were you when you left full-time education?	-
Q55	P4	Overall, how would you rate your health during the past 4 weeks?	-
	P5	Are you the parent or guardian of anyone aged under 18 who lives with you?	-
	P6	Do you look after, or give special help to anyone who is sick, has a disability, or is an older person, other than in a professional capacity?	4
Q56	P7	To which of these ethnic groups would you say you belong? (Tick ONE only)	-

6 Submitting data

Data may be submitted to the Co-ordination Centre either by e-mail or by post (with the data on a CD or floppy disk). There is no requirement for the anonymised dataset to be password protected. Data should be submitted to the following address:

By e-mail: primarycare@PickerEurope.ac.uk

or

By post: Co-ordination Centre for the Primary Care Trust Survey
Picker Institute Europe
King's Mead House
Oxpens Road
Oxford
OX1 1RX