Review of BME coverage in the Healthcare Commission's patient survey programme

Patten Smith

August 2006

Summary

The primary purposes of this review are (i) to assess whether trust level patient surveys provide accounts of the experience of black and minority ethnic patients that are as accurate as the accounts they provide of the experiences of white patients and (ii) to make recommendations about how accuracy might be improved. A secondary purpose is to assess how successful, to date, has been the *communication* of information about group differences that are present in the data with a view to addressing equal opportunity issues.

The surveys covered by the review are:

- Acute trusts: inpatients;
- Acute trust: outpatients;
- Acute trust: emergency department;
- PCT patients;
- Acute trust: young patients;
- Mental Health trust Service users;
- Ambulance trust patients.

Representativeness review
In the representativeness review we attempt to assess how representative are the patient survey datasets.

In the first part of representativeness review we compare response rates across different ethnic groups. Conclusions drawn from these comparisons need to be treated with some caution because (i) ethnic group is not recorded on the sample frame for the PCT and Ambulance trust surveys, (ii) in the other surveys information on ethnic group is missing for substantial proportions of patients (between 20% and 47% depending on the survey) and (iii) there are inconsistencies between patient sample frame ethnic coding and patients' own self-categorisation on the questionnaires.

Response rates are consistently lower for minority ethnic patients than they are for white ones across all five surveys for which comparisons were made. Multivariate analysis demonstrates that ethnic differences in response propensity remain after controlling for respondent age and location of trust. It is noted that response rates were also lower for black and minority ethnic groups than they were for white ones in the "gold standard" Health Survey for England.

We have very little information that would allow us to estimate non-response bias arising specifically from these higher levels of non-response amongst minority ethnic sample members, although we conclude that whole-sample estimates are unlikely to be much affected. We are more concerned about levels of bias for individual minority ethnic groups, because lower response rates are associated with greater risk of bias. Furthermore, we are specifically concerned that relatively poor literacy levels amongst South Asians may increase non-response bias amongst this group.

In the second part of the representativeness review we discuss the case for comparing the profiles of our patient samples against benchmark data. Although the procedure is sound in principle, it is not possible to implement in practice because reliable and comparable benchmark data are not available.

Process review
This part of the review examines survey processes with a view to assessing whether any might be vulnerable to errors that differentially impact on different ethnic groups.

We conclude that errors in sampling, sample frames, survey specification, or processing are unlikely to affect different ethnic groups differentially.

We then discuss how survey procedures might affect response rates and conclude that there are a number of general methods for improving postal survey response rates that are recognised in the literature which could, in principle, be applied to the patient surveys. In improving response rates generally, these methods might also improve response rates specifically for minority ethnic groups. However, there may be practical reasons (including cost constraints), as to why some of these methods cannot be introduced.

We then discuss how non-response arising from poor literacy levels might be addressed. Ideally we would identify which languages that are appropriate for sample members in advance and then mail translated questionnaires. But this would require considerable improvements in the information held on the sample frames, and in any case would reduce rather than eliminate the problem because levels of poor literacy in *any* language are relatively high in South Asian groups.

We briefly discuss whether disengagement from "official" institutions may be one factor causing poor minority ethnic response rates. We have little evidence relating to this hypothesis. A non-response study would help us considerably in assessing this.

We have no information on whether measurement error may affect some ethnic groups more than others. However, we recommend that separate programmes of cognitive testing be undertaken amongst members of different ethnic groups during questionnaire development to try to mimimise the chances of this happening.

Reporting review

Reporting formats have developed which address the main survey objective, which is to enable trust performance to be monitored. In addition, multivariate analyses have been undertaken with a view to teasing out more complex relationships in the data.

To date the reporting has not sought to address equal opportunity issues. We recommend that discussions be held with relevant professionals with a view to assessing whether, and how, the data may be better used to do this. It needs, however, to be recognised that data collected for one purpose may not be adequate for addressing a different set of objectives without significant changes to survey procedures.

Main conclusions

1. Minority ethnic patients have lower survey response rates than other patients in the patient surveys.
2. There are currently no criterion data sets that can be accurately used to check the representativeness of the patient surveys.
3. We have no grounds for supposing that estimates for minority ethnic sub-samples are more, or less prone than others to errors in respect of sampling, sample frames, survey specification, or processing.
4. In the absence of information to the contrary, we have to suppose that, as a result of lower response rates, they *are* subject to a higher risk of non-response bias. However, if it were possible to conduct a study of non-responders, our knowledge of levels of non-response bias in different ethnic groups would be significantly improved.
5. We have grounds for thinking that the surveys will under-represent patients with poor literacy skills, that this is especially a problem for some South Asian groups, and that this may lead to bias in estimates for variables correlated with literacy.
6. We feel that there are measures that could be taken to boost response rates generally which would probably also have a positive impact on minority ethnic response rates.
7. If information about languages read or spoken could be collected or inferred from sample frame information, it would probably be worthwhile sending appropriately translated questionnaires to sample members.
8. Doing this would probably reduce non-response bias arising from poor literacy, but would not eliminate it.

9. We have no evidence on different ethnic groups' proneness to measurement error, but feel that in future measures should be taken during questionnaire development to help guard against this possibility.
10. Survey reports do not currently provide information in a manner well suited to equal opportunity monitoring, but this is because the surveys were not set up with this as a primary objective. However, new forms of survey reporting could be introduced; this would best be done after consultation with relevant professionals.

In addition we put forward a set of detailed methodological recommendations.

1 Introduction

1.1 The patient survey programme

The Healthcare Commission is responsible for overseeing a national programme of patient surveys. This programme covers a range of surveys which are developed and analysed centrally, but implemented locally. Individual healthcare trusts are responsible for funding the local surveys, and for ensuring that they are carried out following national guidance and timescales. There is scope for trusts to include on the questionnaire additional questions covering topics of local interest, and also to survey samples of patients that are larger than those prescribed by the Healthcare Commission.

Rather than conduct the surveys in-house, most trusts commission specialist contractors to carry out them out. In an effort to ensure that trusts employ contractors that are competent and suitably experienced, the Healthcare Commission vets candidate contractors, and encourages trusts to use only those that have been approved for the work.

Results from these surveys are used by individual trusts for identifying areas for improvement, by the Healthcare Commission for performance assessment, including the new annual health check, and by the Department of Health for monitoring wider government objectives on racial equality.

Table 1 shows the surveys that have been / are being undertaken as part of this programme to date (including those undertaken by Healthcare Commission's predecessor, the Commission for Health Improvement) .

Table 1 Patient surveys

Survey	Year(s)
Acute trust: Inpatients	2002* 2004 2005
Acute trust: Outpatients	2003 2004/5
Acute trust: Emergency Department	2003 2004/5
PCT	2003 2004 2005

Acute trust: Young patients	2004
Mental Health	2004 2005 2006
Ambulance	2004

* overseen =by the Department of Health

Information about the surveys within the existing programme can be found at **www.nhssurveys.org** (questionnaires, guidance and development reports) and **www.healthcarecommission.org.uk** (results).

The surveys are developed and overseen by central survey Coordination/Advice Centres under guidance from the Healthcare Commission. Until 2005 the Picker Institute Europe played this role for all surveys in the patient survey programme. In 2005 the Healthcare Commission started the process of re-tendering the contracts for coordinating the various patient survey programmes, and, as a result, the National Centre for Social Research (NatCen) and the Picker Institute Europe now both play Coordination Centre roles for different patient survey programmes.

1.2 Background to the review

The health needs of black and minority ethnic groups are diverse, and variations in health status, use of healthcare and outcomes between ethnic groups have been demonstrated repeatedly in recent years (Nazroo, 1999; Erens, Primatesta and Prior, 2001; Sproston and Mindell, 2006). A range of factors may contribute to variation in health status, including socio-economic factors, genetic factors, and access to health services.

NHS services must, of course, to be fully responsive to these varied needs, and the patient surveys potentially provide a useful tool for assessing the extent to which this is actually the case. The patient surveys will only be able adequately to fulfill this task if they are able to provide reasonably accurate accounts of the experiences of different black and minority ethnic groups. The primary purposes of this review are (i) to assess whether the patient surveys provide accounts of the experience of black and minority ethnic patients that are as accurate as the accounts they provide of the experiences of white patients and (ii) to make recommendations about how such accuracy might be improved.

Although reasonable accuracy is a necessary condition that must be fulfilled if patient surveys' data are to be effective for assessing responsiveness of health services to the needs of different ethnic groups, it is by no means a sufficient one. The mere *collection* of such data will be of little help. The data need to be analysed and presented in such a manner that (i) key differences between groups are highlighted, and (ii) where such differences are found, they are communicated in appropriate fashion to relevant health service audiences. A secondary purpose of this review is to assess how successful, to date, has been the communication of information about group differences that are present in the data.

By way of context we note a number of recent developments that are pertinent to this review.

First, the Commission has recently reported on variations in patient experience between ethnic groups, based on analysis of the five national patient surveys carried out in 2003/4 (Healthcare Commission 2005a), and the results from the analysis of the surveys carried out in 2004/5 are due shortly. Previously the Commission for Health Improvement (CHI) published a similar report based on surveys conducted in 2002/3 (Commission for Health Improvement, 2004). The Healthcare Commission's latest report into variations in patient experience (by ethnicity amongst other factors) is a deliverable output in the Commission's Race Equality Scheme

Second, the Race Relations (Amendment) Act (2001) mean that from May 2002, all public bodies had a statutory duty to demonstrate that they were promoting race equality. This has implications for the patient survey programme which, as a minimum, needs to ensure that there is an equal opportunity for all users to provide feedback on NHS services. Most recently, the Department of Health published a document, *National standards, Local Action* (DH, 2004), which set out the government's expectation that, as part of the annual healthcheck, the Commission will assess all healthcare organisations against four national targets, including a patient experience target, based on the Department of Health's Public Service Agreement (PSA). The patient experience target includes the statement that '*the experiences of black and minority ethnic groups will be specifically monitored as part of these surveys*'.

Third, within both the Healthcare Commission and the Department of Health concerns have recently been expressed about black and minority ethnic (BME) *response rates* to the patient surveys generally and to the mental health survey in particular. In response to these concerns the surveys and analysis teams have:

- written a joint note (for the Executive Team and Commissioners), which specifically addresses the complexities around mental health (Seccombe and Raleigh, 2005);
- updated the Race Equality Scheme action plan which describes, more generally, actions being taken to address these issues;
- commissioned this review.

Fourth, the Healthcare Commission Executive Team has recently announced that it will ensure that future patient survey programmes will undergo a race equality impact assessment.

Finally, in January 2006 the White Paper on care outside hospital (Our health, our care, our say) announced that DH would commission a review of the national patient experience surveys programme. The Healthcare Commission is working jointly with the Department of Health on the terms of reference for this review and have asked for it to include issues of black and minority ethnic group coverage and response.

1.3 Nature of this review

Ideally the review would systematically assess the overall accuracy of data collected for different ethnic groups, and, as such, would cover all survey processes. In practice, most concern that has been expressed has explicitly been centred on relative response rates, although it has implicitly also touched on the issue of error in answering questions, notably for those for whom English is not the first language. The reviewer concurs with this assessment of the likely weak points in the survey processes, and this is reflected in the focus of this report.

This report consists of three unequal sections followed by conclusions. In the first section, the *representativeness review*, we look for evidence which might help us to assess whether or not our survey samples under-represent black and minority ethnic group (BME) groups. Second, in the *process review*, we critically examine key survey processes in order to assess to what extent they may prove to be less effective in collecting data from some ethnic groups than from others. And third, in the *reporting review*, we briefly investigate whether or not there might be potential for further exploiting available data in order to compare the experiences of different ethnic groups, especially with a view to addressing equal opportunity issues.

The review was undertaken in the knowledge of the constraints under which the survey programme operates, notably:

1. the requirement that the methods used should be affordable to NHS organisations;
2. the requirement that the methods should be capable of being implemented by trusts (or their representatives) locally in a devolved but consistent manner;
3. the requirements that only methods capable of producing data of reasonably high quality should be considered.

The review involved the following activities:

1. examination of survey reports;
2. examination of key survey documentation (eg guidance notes, questionnaires);
3. investigation of other relevant literature;
4. interviews with two advice centres, stakeholders at the Healthcare Commission and at the Department of Health[1], and one external expert on minority ethnic health issues[2].

2 Representativeness review

Although concerns have been expressed that minority ethnic groups are under-represented in the patient surveys, relatively little evidence has been adduced to support this view. In our view two sorts of evidence that are pertinent to this issue. First, for the patient surveys in which ethnic group flags are present on the sampling frame, we can compare response rates for different ethnic groups. Second, it may be possible to compare the distribution of the patient survey samples across ethnic groups against corresponding distributions in benchmark data sets. Below we look at each approach in turn.

2.1 Response rates for different ethnic groups

2.1.1 Levels of non-response

Ethnic group is flagged on the sample frame for some, but not all, the patient surveys (table 2).

---

[1] Terry Day and Veena Raleigh at the Healthcare Commission, and Surinda Sharma at the Department of Health.

[2] Professor James Nazroo.

Table 2 Whether or not ethnic group was flagged on the sample frame

Survey / survey series	Whether or not ethnic group flagged
Acute Trusts: adult inpatients	Yes
Acute Trusts: young patients	Yes
Acute Trusts: outpatients	Yes
Acute Trusts: emergency departments	Yes
Ambulance Trusts	No
Primary Care Trusts	No
Mental Health Trusts	Yes

Even in those surveys where ethnic group is meant to be flagged, in practice the level of flagging is poor: no information on ethnic group is available for between about one in five and nearly half of the sample (see table 3). As a result, if response rates for patients whose ethnic group was not flagged differed substantially from response rates for patients in corresponding ethnic groups whose ethnic group *was* flagged, the response rates calculated for different ethnic groups could be misleading indicators of true propensities to respond.

Table 3 Proportion of issued sample for which no flag available for ethnic group

Survey	Issued sample (n)	No ethnic group flag (n)	No ethnic group flag (%)
Inpatients	143,322	41,602	29.0
Young patients	125,827	35,171	28.0
Outpatients	142,568	53,208	37.3
Emergency Department	129,085	60,530	46.9
Mental health trusts	66,544	13,457	20.2

Furthermore, there is considerable concern about the quality of the sample frame ethnic flag. When we compared ethnic group from the two sources for survey respondents for whom the sample frame ethnic flag was available, we found that correspondence between the two ethnic categorisations was high for the sample as a whole (ranging from 92% agreement for young patients to 98% for outpatients) and high for white respondents, but that it was considerably lower for minority ethnic group respondents (see Annex A, tables A.1 and A.2). Correspondence was especially low for those categorised as other on the sampling frame, but was also poor for those categorised as mixed on the sampling frame and for those categorised as mixed or other on the questionnaire. However, these inaccuracies do not nullify the worth of making response rate comparisons for two reasons. First, of those respondents who were categorised as Asian, Black and Chinese on the sample frame, fairly substantial majorities (around four fifths for Asian and Black groups, and somewhat lower for Chinese) turned out to be categorised in the same way on the questionnaire. Second, substantial majorities of those respondents categorised as Mixed, Asian, Black and Chinese on the sample frame coded themselves as belonging to *one* of the non-white groups on the questionnaire (see table A.3).

Despite the data quality problems highlighted in tables 3 and A.1 and A.2, we consider it highly probable that any large differences in response rate observed between Asian, Black, Chinese and White sample members are likely to reflect real underlying differences.

Tables A.4 to A.8 show response rates broken down by ethnic group. Response rates are calculated by dividing the number of eligible completed questionnaire returns by the total number of sample members mailed less those known (from their returned questionnaires or from other contacts) to be ineligible (see Annex D). We have calculated these afresh rather than rely upon published figures because (i) the latter are not available for the latest implementation of the inpatients and mental health surveys and (ii) in some cases the latter have been calculated inappropriately (see Appendix B for details).

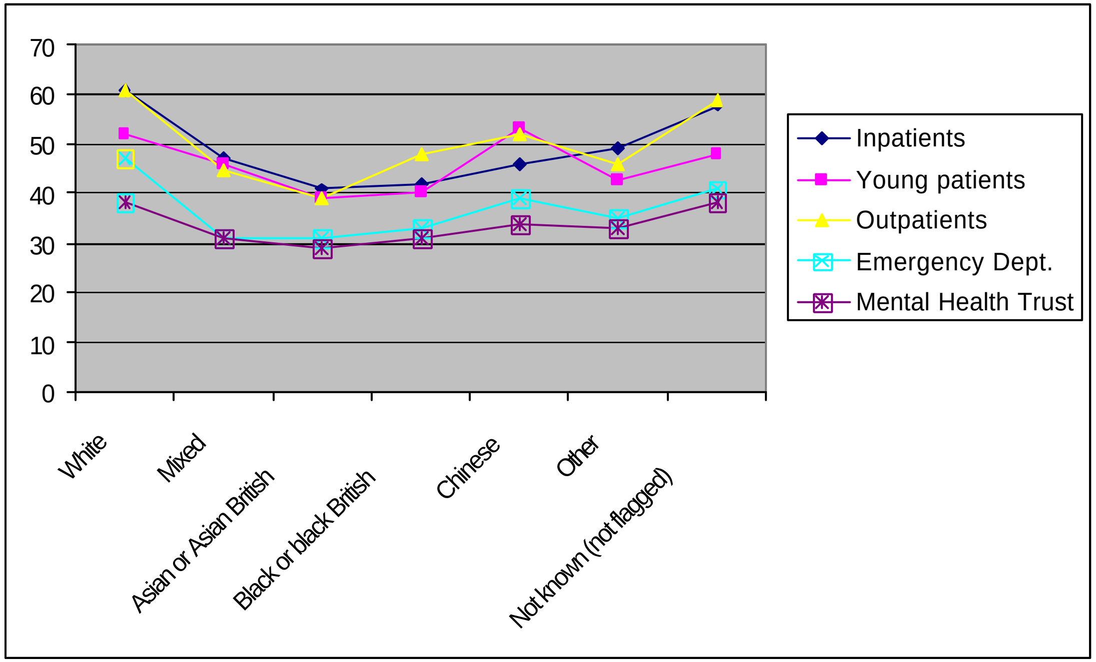
As can be seen in table 4 and Figure 1 (which summarise tables A.4 to A.8), with the single exception of Chinese patients in the young patients' survey, response rates were lower for patients coded as mixed, Asian, Black, Chinese or other than for patients coded as white. These response rate differences were generally substantial (over 10% in a majority of comparisons). On the face of it this indicates that black and minority ethnic group members were under-represented in the patient surveys.

Table 4 Response rates by ethnic group compared across all surveys

Ethnic group	**Inpatients, 2005**	**Young patients, 2004**	**Outpatients, 2004/5**	**Emergency Dept., 2004/5**	**Mental Health, 2006**
	%	**%**	**%**	**%**	**%**

White	61	52	61	47	38
Mixed	47	46	45	31	31
Asian or Asian British	41	39	39	31	29
Black or black British	42	40	48	33	31
Chinese	46	53	52	39	34
Other	49	43	46	35	33
Not known (not flagged)	58	48	59	41	38
Overall	59	50	59	43	37

Figure 1 Response rates by ethnic group compared across all surveys



In interpreting these differences it needs to be borne in mind that the minority ethnic patient populations are highly concentrated in London and other conurbations, and have younger age profiles than the white population (see table 5). This is of interest (i) because response rates are generally lower in London and inner city areas, and (ii) because in the patient surveys there was a strong positive relationship between age (as recorded on the sample frame) and response rate (see table 6), meaning that it is possible that the lower response rates observed for minority ethnic individuals arose, in whole or in part, because larger proportions of these groups than of whites were younger and lived in London and other conurbations.

Of course, even if response rate differences could be wholly or mainly attributed to age and area of residence, this would not in any way affect our fundamental concern that, because response rates are lower for minority ethnic groups than for white ones, there is a greater danger of non-response bias (and hence of samples being less representative) for the former groups than for the latter one. However, any such finding could help us explain *why* minority response rates are lower.

Table 5 Mean age by sample frame ethnic group

Ethnic group	Inpatient survey	Outpatient survey	Emergency Dept. survey	Mental Health Trust survey
White	59.9	59.1	48.5	51.5
Mixed	45.2	46.9	37.0	39.9
Asian	49.6	50.0	41.4	43.2
Black	46.6	51.6	39.9	43.7
Chinese	48.4	50.4	39.8	44.8
Other	52.2	49.6	40.2	42.7
Not known	57.5	57.0	44.9	52.5

Table 6 Response rate by age

Age	Inpatient survey	Outpatient survey	Emergency Dept. survey	Mental Health Trust survey
Up to 30	34.8%	34.5%	26.4%	27.9%
31-40	44.6%	43.4%	34.0%	34.0%
41-50	55.0%	52.5%	43.1%	40.7%
51-64	68.3%	65.2%	58.8%	45.8%
65 +	65.4%	68.2%	60.5%	35.4%

We used logistic regression techniques to examine the relationship between response rate and ethnic group after controlling for trust identity and age using the Inpatient, Outpatient, Emergency Department and Mental Health Trust surveys (we excluded the young patients' survey from the analysis because this sample is confined to a very limited age range). This analysis, which is summarised in tables A.9 to A.12, demonstrated that even after applying these controls, in comparison to response rates for white patients, response rates were significantly lower for respondents categorised as mixed, Asian and Black in all four surveys examined, were significantly lower for respondents categorised as other in three out of the four surveys and were significantly lower for respondents categorised as Chinese in two of the four surveys. We therefore have no grounds for concluding that lower minority ethnic response rates can be attributed entirely to the fact that minority ethnic and white patients differ in age and in where they live[3].

It is worth noting at this point that these ethnic differences in response rates are generally consistent with findings from high quality face-to-face interview surveys in which response rates have been compared across different ethnic groups. Probably the clearest comparisons are provided by the two rounds of the Health Survey for England which included minority ethnic boost samples because here (unlike other recent surveys with minority ethnic boost samples such as the British Crime Survey) special efforts were made to estimate individual response rates for different ethnic groups. These are shown in table 7 (For full details see Erens, Primatesta and Prior, 2001; Sproston and Mindell, 2006).
.

Table 7 HSE response rates for different ethnic groups

Group	**HSE 1999**	**HSE, 2004**
Black Caribbean	55%	57%
Black African	Not included	62%
Indian	59%	60%
Pakistani	60%	57%
Bangladeshi	66%	66%
Chinese	62%	55%
General population	70%	66%

As can be seen, in the 1999 survey response rates were lower for all the minority ethnic group boost samples than they were for the general population sample, and in the 2004 survey they were lower for all minority ethnic group boost samples except the Bangladeshi one. If even these "gold standard" face-to-face interview surveys exhibit lower response rates for minority group samples than they do for general population ones, we should not be surprised that the patient surveys do likewise.

---

[3] However, in analyses not shown in this paper, we found that the coefficients for minority ethnic groups were *reduced*, often fairly substantially, by the inclusion of trust and age variables.

2.1.2 Non-response bias

If we accept that the surveys under-represent minority ethnic group patients, we need to assess what sort of impact this might have on the accuracy of our estimates. In doing this we need to make a distinction between impact on the sample as a whole (in individual trusts, nationally or in aggregations of trusts) and the impact on individual minority group samples. In either case our main concern with under-representation will be with the extent to which this introduces bias into our estimates: in the former case we are concerned with bias for whole sample estimates and in the latter with estimates for individual minority groups.

Non-response bias for a particular variable is determined by both response rate *and* the extent to which respondents and non-respondents differ in their mean values for the variable[4] (Groves 1989; Groves and Couper 1998). However, because we rarely know non-respondent mean values, we are usually uncertain as to the level of non-response bias, and we therefore focus our attention on what *can* be measured, namely response rate. It is worth reiterating, however, that response rate does *not* measure non-response bias, but rather measures the *risk* of its occurrence. The survey literature demonstrates that the link between non-response and non-response bias is by no means straightforward, and at times high response rates have been found to be associated with substantial bias (eg Teitler, Reichman and Sprachman, 2003; Stoop, 2006), while at other time low response rates have been found to be associated with low levels of bias (eg Keeter, Miller, Groves and Presser, 2000; Merkle and Edelman, 2002). it should be clear from this discussion, assessing levels of non-response bias in the patients' surveys is not a straightforward task.

Bias in whole sample estimates

Looking at whole sample estimates first, we can attempt to quantify the impact of under-representation of minority ethnic groups by comparing the issued and achieved sample ethnic group distributions in the patient surveys that would hold if (i) the proportion of cases with valid ethnic group flags was constant across all ethnic groups, and (ii) *relative*[5] response rates observed for different ethnic groups in flagged cases could be generalised to non-flagged cases. Although we do not know how close these assumptions come to the truth, we consider it implausible that the

---

[4] Non-response bias = Respondent mean – true mean
= non-response rate × (respondent mean – non-respondent mean)

[5] Relative differences would be preserved if, say, flagged case response rates were 60% for group1 and 40 % for group 2, and unflagged case response rates were 48% for group 1 and 32% for group 2. In each case the ratio of group1 to group 2 response rates is 1.5 to 1.

general pattern of observed differences in response rates across ethnic groups for flagged cases would not also hold for unflagged ones[6]. Table 8 shows the issued and achieved distributions that would result in each survey were the above assumptions to hold.

As can be seen there, if we were using the patient survey respondents to estimate proportions in each ethnic group in the population of patients, we should be very concerned about bias. Across the surveys the proportion of patients who were white would be over estimated by some two to three per cent, the proportion who were Asian would be under-estimated by one to one and a half per cent, and the proportion who were black would be under-estimated by half to one per cent. Similarly, possibly substantial overall non-response biases resulting from under-representation of minority ethnic groups would be expected for any measures that are very highly correlated with ethnic group membership (eg religious affiliation). It is, however, very unlikely that there are many such variables in the patient data sets.

For most variables including those associated moderately with ethnic group, we expect the biasing impact of under-representation of minority ethnic groups to be far smaller simply because minority ethnic patients represent a relatively small proportion of the overall patient population. All an illustration we have looked at the proportion of the 2004/5 Emergency Department patients in different ethnic groups (according to sample frame flag) who answered “yes, completely” to the question:

> While you were in the Emergency Department, did a doctor or nurse explain your condition and treatment in a way you could understand?

These proportions were:

White	63.9%
Asian	57.9%
Black	59.4%
Mixed/Chinese/Other	62.4%

---

[6] By which we mean that if group A has a substantially higher repose rate than group B amongst flagged cases, we think it plausible that unflagged members of group A will be more likely to respond than will unflagged members of group B.

Table 8 Profiles of issued and achieved samples that would be found if observed response rates were generalised

	Inpatients		Young patients		Outpatients		Emergency Department		Mental health trusts	
	Issued (%)	Achieved (%)	Issued (%)	Achieved (%)	Issued (%)	Achieved (%)	Issued (%)	Achieved (%)	Issued (%)	Achieved (%)
White	89.1	91.2	82.6	85.6	90.4	92.9	89.5	92.5	90.2	92.1
Asian or Asian British	3.9	2.9	6.4	5.0	4.3	2.8	5.1	3.5	4.0	3.1
Black or black British	2.4	1.9	3.3	2.7	3.4	2.7	3.8	2.8	3.9	3.1
Mixed race, Chinese or other ethnic group	4.5	4.0	7.7	6.8	2.0	1.5	1.5	1.1	1.9	1.7
All	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

If the distribution of the *whole* issued sample across ethnic groups were as shown in the Emergency Department "issued" column in table 8, if the whole of the Emergency Department issued sample had participated, and, if the percentages answering "yes, completely" were as shown above for the different ethnic groups, the proportion of the total sample answering "yes, completely" would have been 63.4%, and we would consider this to be a good estimate of the true proportion. If, on the other hand, our estimate came from *respondents* only, our estimate would be based on the Emergency Department achieved sample (again, shown in table 8) which rather under-represents minority ethnic groups, and this would have the impact of increasing the proportion answering "yes completely" to 63.5% - that is by only 0.1%. Even in trusts with considerable minority ethnic populations, on any reasonable assumptions, the impact on whole-sample estimates are likely to be small.

The moral here is that, even for variables which vary across ethnic groups, the impact on whole-sample estimates are likely to be small. This is *not* to say that whole-sample non-response bias may not be substantial (it may well be), but rather that what bias there is unlikely to be attributable to a major degree specifically to the lower observed response rates amongst minority ethnic patients

Bias in estimates for particular ethnic groups

Having discussed the impact on whole sample estimates of lower minority ethnic response rates, we need now to ask what will be their impact on estimates for individual ethnic groups. We are concerned here with the question not of whether estimates for individual minority ethnic groups are subject to non-response bias (they almost certainly are), but with the question of whether they are likely to be subject to *more* bias than are corresponding estimates for white respondents. (This is a more precise way of asking whether the samples of minority ethnic respondents are less representative than the samples of white ones as a result of their lower response rates.)

As stated above, response rates do not measure this bias directly but rather give an indication of the *risk* of non-response bias. In an ideal world, the patient surveys would be accompanied by one or more non-response assessment studies (almost certainly using a face-to-face interviewing methodology) which would allow us to assess likely levels and directions non-response bias for different ethnic groups, and we recommend that such studies should be carried out. In the absence of such studies, we are left in the unsatisfactory position of having to make judgements on the basis of observed response rates and available contextual information.

On this basis, it is fair to say that there are grounds for worrying about non-response bias for all surveyed groups, and that, other things being equal, this worry should be greater for both surveys and population groups with lower response rates. And, because minority ethnic groups tend to have lower response rates, our worries about non-response bias will be greater for these groups than for others. However, we should also ask ourselves whether there is any contextual information which gives us reason to be concerned that non-response bias amongst one or more minority ethnic groups will be especially high. A "special" reason in this sense is one that gives us grounds for thinking that on some survey variables *respondents and non-respondents in one ethnic group are more likely to be systematically different from one another than they are in other ethnic groups.* Below, and again in section 3.2.2 we discuss one such reason – poor literacy - which is likely to have an more biasing impact on estimates for South Asian sample members than on estimates for others. The fact that we discuss it does not mean that we necessarily consider it to be a more important source of bias than others we do not know about, or that we necessarily consider survey estimates for South Asians to be more biased than those for other groups. We discuss it because (i) we have better evidence (albeit circumstantial) relating to it than we have for other potential bias sources and (ii) we feel that it is incumbent upon us to try to address the likely biases that we *do* know about even if there remain many that we do not know about.

There is considerable evidence that literacy amongst older South Asian people – mainly women - is particularly poor. We discuss this further in section 3.3.2, but for the moment the key point is that any survey using a self-completion questionnaire methodology will almost certainly substantially under-represent those with poor literacy skills. We know that individuals with poor literacy skills are more likely than others to be older and to be women, and we also know that reported patient experience varies by age and gender (see, for example, Commission for Health Improvement, 2004; Healthcare Commission, 2005a). We also strongly suspect that reported patient experience will vary by literacy *within* age and gender categories. To the extent that this is true, low survey response by individuals with poor literacy skills will lead to bias in survey estimates, and this bias will affect South Asian sub-samples more than it will affect other sub-samples. And it will not be possible to correct for such bias using weights based on age-sex categorisation.

Non-response bias summary
We can summarise this analysis of non-response bias thus:

1. the under-representation of minority ethnic groups is unlikely to have a significantly biasing effect on estimates for populations of patients as a whole;
2. *other things being equal*, lower response rates give us greater concern about potential for non-response bias, and because minority ethnic groups show relatively low response rates, we should, probably, worry more about non-response bias from these groups than for others;
3. we are concerned that South Asian samples may be particularly susceptible to non-response bias arising from the fact that their parent populations exhibit relatively high non-literacy levels (although we do *not* infer from this that overall non-response bias for South Asian groups will necessarily be greater than that for other groups).

2.2 Comparison against criterion data sets

The logic of this approach is easily stated. A search is made for sources of highly reliable[7] benchmark data on ethnic group relating to the same populations as are sampled by the patient surveys. A comparison is then made between the distribution across ethnic groups taken from the patient survey and that taken from the benchmark data. Because of the known reliability of the benchmark data, differences are

---

[7] ie generally accepted as being less error prone than the data taken from the patient surveys.

interpreted as arising primarily from errors in the patient survey data. On the face of it, this method of evaluation is attractive, and we are aware of two attempts to use it for patients surveys: (i) for the Mental Health Trust survey (Seccombe and Raleigh, 2005) and (ii) for a recent Healthcare Commission of Stroke patients (Healthcare Commission, 2005b)

However, although the logic of the approach is sound, in practice it is of limited use because of difficulties in finding highly reliable benchmark data.

An important reason for this is that it is hard to find criterion data which relate to the same populations as are used in the patient surveys. Table 9 defines the survey population for each of the patient surveys. It can be immediately seen that the way in which most of these survey populations have been defined leads to special difficulties for the search for good benchmark data sources. Of the seven surveys mentioned, three draw flow samples from *populations of patients with a defined type of contact* over a reference period (which itself is sometimes variable), and three draw samples from populations of *events* (attendances/incidents) (ie sample each patient with probability proportional to his/her attendance frequency within the reference period). Some important considerations arise from this. *Flow samples* will represent patients in proportion to their frequency of contact to the extent that their visits are separated by periods longer than the reference period: for example, the patient visiting an outpatient clinic once a month will have a six times greater chance of being included in the outpatient sample (with its one-month reference period) than will the patient who visits it every six months. Samples of *events* will represent each patient in proportion to the number of events (eg attendances) associated with him/her *however* short are the gaps between visits. One challenge, then, is to find benchmark data that allow us to estimate the proportions of these flow / event populations that fall into different ethnic groups. (It is worth noting that to do this benchmark datasets do not themselves necessarily have to comprise flow or event samples/populations, but they *do* need to include data on event frequencies or timings so that weights can be applied to allow us to estimate the characteristics of flows and events.)

Table 9 Survey populations

Survey	Definition of population
Adult inpatients in acute trusts	• Aged 16+ • Flow sample of patients who stayed for 1+ nights over variable reference period (varying in length depending upon number of inpatients in trust) • Excluding maternity patients, psychiatry patients, patients admitted for termination of pregnancy, current inpatients, private patients, patients treated as NHS patients at private hospitals, patients without UK address.
Young patients in acute trusts	• Aged 0-17 • Flow sample of day patients or patients who stayed for 1+ nights • Excluding maternity patients (babies from delivery suite), newborn babies where mother is primary patient, psychiatry patient, children “primarily treated by mental health services”, patients admitted for termination of pregnancy, private patients, current inpatients, patients without a UK postal address or patients who were only admitted to a Neonatal Intensive Care Unit (NICU) or a SCBU (Special Care Baby Unit).
Outpatients in acute trusts	• Aged 16+ • Sample of outpatient department attendances during a single reference month (meaning that patients’ probabilities of inclusion in the sample were proportional to the number of their attendances during the reference month) • Excludes current inpatients, patients without a UK postal address, private patients, patients attending GUM or STD clinics, patients whose attendance concerns termination of pregnancy, attendances at psychiatric outpatients clinics, attendances at maternity outpatient clinics,
Emergency Department patients in acute trusts	• Aged 16+ • Sample of emergency department attendances during a single reference month (meaning that patients’ probabilities of inclusion in the sample were proportional to the number of their attendances during the reference month) • Excludes minor injuries clinic attendances, current inpatients, planned attendances at outpatient clinics run within the emergency department (eg fracture clinics), patients without a UK postal address
Mental health trust service users	• Aged 16-64 • Flow sample of mental health service users registered as subject to care programme approach (CPA) and seen during a three-month reference period (or, if contact information unavailable, who had care review in last 12 months) • Excludes people seen once only for an assessment, current inpatients, patients without a UK postal address
Primary care trust patients	• Aged 16+ • Patients registered with a GP in PCT
Ambulance trust patients	• Aged 16+ • Sample of incidents for which patient report forms (PRFs) were completed during a reference week (meaning that patients’ probabilities of inclusion in the sample were proportional to the number of PRFs completed for them during the reference week)

So, what are the candidates for sources of benchmark data? On the basis of expert consultation we have found only two possible candidates for inclusion, the Health Episode Statistics and the Health Survey for England.

Hospital Episode Statistics (HES)
In theory, these should be useful: they are based on the full population of admissions, include information on ethnic group, and although the relate to episodes they can be aggregated to the patient level. Furthermore as they have been used by the Healthcare Commission as benchmark data to assess the ethnic group profile of the survey of stroke patients (Healthcare Commission, 2005b).

However, our understanding is that the recording of ethnic group in the HES data is currently very poor[8], and, in our view, this disqualifies HES from providing suitable criterion data.

Health Survey for England
The 1999 implementations of the Health Survey for England (HSE) included boost samples of Black and minority ethnic respondents and included questions on health service usage[9]. It may therefore, in principle, be possible to compare the relative usage of different health services by different minority ethnic groups with their representation in corresponding patient surveys. In practice, there are a number of difficulties with this approach.

First, the HSE only covered use of certain services: inpatients, outpatients and Emergency Department combined, and GPs.

Second, our ability to compare across surveys is further reduced by the fact that outpatient and ED attendances are combined in HSE but not in the patient surveys: this renders reliable comparisons impossible.

Third, HSE and patient surveys define *contact* with particular services differently as can be seen in table 10 below. As a result any direct comparison for GP or inpatient attendances would first require data from one or the other source to be modified. If we are to compare HSE GP attendance data with PCT patient survey data, we need information on when last visited a GP, or more generally on frequency of such visits for PCT patient survey sample members. Unfortunately this information was not collected in the survey. Similarly, we need further, and unavailable, information on attendance frequency in order to make safe comparisons between inpatient data from the two sources.

Despite this, we would argue that, if it proves possible, it is still worth making comparisons for GPs and inpatients, even though any conclusions drawn from them would have to be hedged with caveats – mainly relating to the key assumption that the attendance frequency distribution was unrelated to ethnic group membership.

---

[8] Indeed our understanding is that HES ethnic breakdowns and the survey sample frame ethnic flags will be based on the same data source.

[9] The 2004 HSE also included minority ethnic boost samples, but did not include questions on health service usage.

Table 10 HSE 1999 and Patient Survey definitions of contact

Patient type	HSE 1999	Patient survey
GPs	Seen by GP in past 2 weeks	GP registered patients
Outpatients	Patients attending in past 12 months (combined with ED)	Attendances in past month
Emergency Department (ED)	Patients attending in past 12 months (combined with outpatients)	Attendances in past month
Inpatients	Patients attending in past 12months	Patients attending over recent reference period (shorter than 12 months)

The forth difficulty with using HSE data as a criterion data set arises from the fact that HSE relates to a single national sample of adults, whereas patient survey data relate to multiple trust level samples. If we were to compare data from the two sources we would need to weight the patient survey data to make it representative of the population of adults in England. In order to construct such weights we would need to know the size of the survey population in each trust (for example, how many outpatient attendances there were in the survey reference month). To our knowledge such size data were not recorded at the time of sampling, are not readily available, and would probably have to be estimated from other sources such as HES. At present, we are unsure as to the feasibility of doing this.

In conclusion, it seems that currently no criterion data sets are available which can be used to provide accurate benchmark data for checking patient survey representativeness, but we would recommend that periodically checks should be made to see if the situation has changed.

3 Process review

In this section we discuss whether the patient survey *processes* might be vulnerable to errors that differentially impact on different ethnic groups.

Over the past decade or two a number of writers have documented the various forms of survey error, and, in so doing, have attempted to integrate diverse perspectives, and broaden attention from those types of error which are most easily measured (notably sampling error). (Groves, 1989; Groves, Fowler, Couper, Lepkowski, Singer and Tourangeau 2004; Biemer and Lyberg, 2003). Biemer and Lyberg (2003) presented a useful framework for the error types that can affect survey estimates which we shall use here:

1. sampling error;
2. specification error;
3. frame error;
4. non-response error;
5. measurement error;

6. processing error.

For the patient surveys we can see no plausible reason to suppose that errors of types 1, 2, or 6 will impact differently on members of different ethnic groups, and we do not discuss these further here. On the other hand, the patient survey processes clearly are susceptible to error types 3, 4 and 5, and it is possible that these might impact differently on different ethnic groups. We discuss these below.

3.1 Sample frame error

Sample frame error relates to how successfully a frame can be used to draw an unbiased sample of the population of interest. For a survey error point of view, the most important characteristics of a sampling frame are, its *coverage*, and its inclusion of *duplicate entries*.

Coverage
Frame coverage is measured as the proportion of the survey population that is included on the sample frame. Clearly any member of the survey population who is not included on the frame will have a zero chance of being included in any survey sample. And, to the extent that sample members included and not included on the sample frame differ in respect of key survey variables, non-coverage will lead to biased estimates.

Administrative records are used as sample frames in all the patient surveys. Centralised electronic lists are available for each trust for the various acute trust surveys and for the PCT survey. However, the Mental Health Trust survey and Ambulance Trust survey may require some bringing together of locally held lists and / or compiling electronic lists from original paper based ones.

To our knowledge, no serious effort has been made to measure levels of frame coverage for the various surveys, and we cannot, therefore, be confident about levels of frame coverage for the surveys. We do not, on the other hand, have any reason to doubt that high coverage is achieved for the Acute Trust surveys and for the PCT survey. Nor do we have any grounds for suspecting that some ethnic groups will be more susceptible to non-coverage than others. We are, however, less sanguine about the Mental Health Trust Survey and the Ambulance Trust survey coverage. Both these surveys require frames to be constructed from lists sometimes held in different parts of the trust, and, sometimes for data entry to be undertaken. These procedures may have to be undertaken by trust/contractor representatives who lack sampling knowledge, and without day to day accountability to the survey advice centre. They are, we suspect, susceptible to error. That said, we can see no reason why any such error would differentially affect different ethnic groups.

We note that the National Centre for Social Research is currently reviewing the sampling processes used for the Mental Health Trust surveys. Hopefully, after this review we will be in a significantly better position to assess the magnitude of non-coverage problems in this survey.

Duplicates
By duplicate entries on a sample frame we mean that the frame includes more than one entry for whatever unit is to be sampled[10]. The presence duplicate entries can compromise sample quality because it introduces variations in selection probabilities which, unless eliminated or corrected, are likely to introduce bias. In order to avoid this problem it is necessary to remove duplicates *before* (not after) sample selection.

The questions arise (i) as to whether or not the patient survey sample frames are likely to include duplicates and (ii) as to whether the presence of duplicates is likely to affect different ethnic groups differentially. In fact, the potential for duplication differs across the surveys. For the two surveys (inpatients and young patients) which use consecutive recent cases as samples the issue of duplication does not arise as no sampling of cases is undertaken. For the three surveys of *attendances/incidents* (outpatients, emergency department and ambulance), it is acceptable for the same patient to appear more than once on the frame for reasons already discussed. However, it would not be acceptable for the same attendance/incident to appear more than once on the frame, and to our knowledge no check is made to ensure that this does not arise; but, we suspect that this is unlikely to occur. The PCT sampling is undertaken by a third party, and we have no direct information on how likely it is that the sample frame will include duplicates or on what measures are taken to remove them before sampling. That said, our understanding is that the database used for PCT sampling is very well maintained, and we therefore suspect that there will be no serious problem with duplicates. The mental health survey sampling process appears to be less likely than the others to rely upon the presence of standard format lists across trusts, and probably requires the trust to engage in more work compiling the sampling frame. In doing this trusts are instructed to remove duplicates, but in the absence of central control over this work, we suspect that there is greater risk of duplication here than for the other patient surveys.

However, although we acknowledge that there is some risk of duplication, (notably in the mental health survey), we can see no reason why this would affect different ethnic groups in different ways.

3.2 Maximising response rates

We have already seen that rates of non-response differ between ethnic groups, and discussed the associated risk of non-response bias (see 2.1 above). In the next sections we discuss how survey procedures might impact on response rates. In 3.2.1 we ask whether all is being done that can reasonably be done to maximise response rates in general, and in section 3.2.2 we examine whether there are further measures that might be taken specifically to address low response rates for particular minority ethnic groups.

---

[10] Note that sometimes the same frame will include duplicates for one kind of sample but not for another; for example, the sample frame for the outpatient survey is one of outpatient *attendances* during a reference month - meaning that the patient who attends three times in a month will have three entries on the frame, whereas the patient who attends only once will have only one entry. As such, the frame will include duplicates when regarded as a frame for a survey of patients attending over the reference month, but may well be free of duplicates when regarded as a frame of *attendances* over the month.

3.2.1 Maximising response rates in general

In this section we briefly ask to what extent do the patient surveys are following accepted canons of good practice for mail surveys, and examine whether there are any procedural modifications from which they might benefit. This is important because our efforts to address low response rates for minority ethnic groups should draw on the full range of techniques likely to impact on response and not just on those specifically focused on minority ethnic groups.

In making this assessment we shall rely heavily on the work of Donald Dillman (Jenkins and Dillman 1997; Dillman 2000) who is widely acknowledged as the leading expert in the world on postal survey methodology.

Dillman's prescriptions cover both questionnaire construction and survey procedures. Here we do not discuss questionnaire construction further because the patient survey questionnaires are professionally designed, well laid out and follow the spirit if not the letter of most of Dillman's prescriptions, and, as Dillman points out, in such circumstances implementation procedures have a far greater influence upon response rates than questionnaire design.

In table 11, we summarise the main procedures advocated by Dillman, and indicate, for each, whether it has been applied to the patient surveys.

Table 11 Dillman's prescribed methods

Procedure	**In patient surveys?**
1. Pre-notice letter	No
2. Questionnaire mailing plus two reminders, second of which has extra copy of questionnaire	Yes
3. Above return envelopes with first class stamps	No
4. Third reminder using phone or special type of mail (eg recorded delivery)	No
5. Personalised correspondence	No
6. Token financial incentive with first questionnaire mailing	No
7. Reminders with escalating urgency	Yes

As can be seen, a number of Dillman's suggested procedures are not currently part of the survey process, and would perhaps be worth considering in future implementations of the survey.

In this context, it is worth noting that in the Autumn of 2003 the Picker Institute conducted an experiment as part of the mental health trust survey development process in two mental health trusts (Osborn, Reeves, Howell and Magee, 2004). This tested the efficacy of two of Dillman's suggested approaches in full and one in part. Seven hundred and eighty one service users in the two trusts were divided into three groups:

group 1: initial questionnaire mailing plus two reminder mailings (2nd questionnaire copy with the second);
group 2: as for group 1 but with addition of an advance letter;
group 3: as for group 2 but with addition of a Co-op gift voucher for £2 in the first questionnaire mailing.

All three groups were then sent a third reminder slip (although not using special mail as advocated by Dillman[11]). On the assumption that all replies received after the 3rd reminder can be attributed to its being sent, this design allows six conditions to be compared, in effect testing the efficacy of sending an advance letter, of sending an advance letter plus incentive, and of sending a third reminder. The response rate associated with each condition is shown in table 12.

Table 12 Response rates associated with interventions in mental health survey pilot work

	Condition	Response rate (%)
1	Initial questionnaire mailing plus two reminder mailings	50.8
2	Advance letter, initial questionnaire mailing plus two reminder mailings	52.5
3	Advance letter, initial questionnaire mailing plus two reminder mailings; incentive in first questionnaire mailing	58.5
4	Initial questionnaire mailing plus two reminder mailings; 3rd reminder	52.0
5	Advance letter, initial questionnaire mailing plus two reminder mailings; 3rd reminder	54.9
6	Advance letter, initial questionnaire mailing plus two reminder mailings; incentive in first questionnaire mailing; 3rd reminder	59.3

It would appear from this table that all three procedures, but especially sending incentives, had a positive impact on response rate.

We feel that it should be possible to boost response rates overall in the patient surveys, by adopting procedures with proven track records. Response rates were increased by 8.5% overall by adopting three enhancements in the mental health survey pilot, and we have no reason to think that these interventions would not be effective in other patient surveys, nor to think that other enhancements could not raise response rates further.

Although we can be confident that response rates overall could be increased through the adoption of various methodological enhancements, we do not know how (or indeed whether) these increases would translate into response rates changes for members of individual ethnic groups. It seems probable that individual group response rates would increase, but we cannot be confident of this without further methodological research designed to investigate this directly.

---

[11] This was felt to be threatening by members of the survey advisory group.

It is worth noting that the Healthcare Commission has recently awarded a new contract for the acute survey coordination centre, and in so doing, specifically required the new centre to review what steps could reasonably be taken to maximise response rates in general, and response rates for BME and other hard to reach groups in particular. The centre was specifically asked to review both the applicability of Dillman's prescriptions and the likely impact of introducing other possible methodological enhancements.

3.2.2 Interventions to increase response rates for specific ethnic groups

Obviously, if we are to diagnose effective methods to increase response rates of minority ethnic groups specifically, it will be of considerable help to know what are the main reasons they do not respond. Unfortunately, no hard and fast information is available on this issue, and, as already discussed, we recommend that efforts be made to collect such information.

That said, in discussions with survey contractors, two reasons for lower minority ethnic response rates were suggested, lower literacy rates and disengagement. We discuss literacy at length, and then turn briefly to disengagement. We conclude the section with a brief discussion of tailoring.

Literacy

Lower literacy rates amongst some minority ethnic groups appears to be a real problem. Table 13 shows figures from the 1991 and 1994 HSE Health and Lifestyle Surveys for proportions of South Asians able to read English, and for proportions unable to read *any* language (Rudat, 1994; Johnson, Owen , Blackburn , Rehman and Nazroo, 2000).

Table 13 Literacy in English and in any language

	Women 30-49	**Men 30-49**	**Women 50-74**	**Men 50-74**
Able read English(1991)				
Indian	67	83	34	71
Pakistani	31	77	7	54
Bangladeshi	15	60	4	38
Unable to read any language(1991)				
Indian	4	2	25	6
Pakistani	31	7	68	16
Bangladeshi	24	3	52	19
Able read English(1994)				
Indian	78	98	42	79
Pakistani	55	82	31	55
Bangladeshi	37	87	13	62

Of course these data need to be interpreted with caution because (i) they are fairly old and the situation has probably changed since their collection, and (ii) they do not relate directly to the survey populations sampled by the patient surveys[12]. Despite these reasons for caution, it seems very likely that considerable proportions of South Asian people, especially those who are Bangladeshi or Pakistani, women, or older, will not be able to read English. Furthermore, the table also strongly suggests that substantial (although considerably smaller) proportions will be unable to read *any* language.

Three conclusions emerge from this. First, any survey using English language self-completion questionnaires is likely to under-represent significantly the substantial proportions of South Asian patients who cannot read English, thereby lowering response rates and increasing non-response bias for South Asian groups. Second, South Asian representation would probably be improved if questionnaires translated into relevant languages were sent to non-English readers. Third, even if translated questionnaires were sent, a substantial proportion of women, older South Asians and Bangladeshis/Pakistanis would be unable to complete questionnaires without assistance.

It would appear, then, that there is a problem here with *any* self-completion questionnaire methodology. This conclusion is supported by figures presented in the 2004 HSE technical report (reworked and shown in table 14 below) showing that Bangladeshi and Pakistani interview respondents were less likely than others to complete a short self-completion supplement handed out during the interview (Sproston and Mindell, 2006). In interpreting these figures it should be borne in mind: (i) that questionnaires were translated into Urdu, Punjabi, Gujarati, Hindi, Bengali, Mandarin and Cantonese; (ii) because interviewers handed over the questionnaire during the face-to-face interview, they were able to check which language (if any) was appropriate and (iii) that the Pakistani and Bangladeshi groups, who show the lowest completion rates, were not characterised by especially low *interview* response rates (indeed Bangladeshis had the *highest*). It seems reasonable to interpret the lower Pakistani and Bangladeshi completion rates in terms of these groups' lower literacy levels.

---

[12] A third reason for caution might be that some of the figures appear to have changed by such a large factor over a three year period as to cast some doubt on the robustness of the data collected in the Health and Lifestyle surveys.

Table 14 Proportion of HSE 2004 interview respondents who completed self-completion supplement by ethnic group and sex

	Men (%)	**Women (%)**
General population sample	94	95
Black Caribbean	84	85
Black African	83	84
Indian	88	88
Pakistani	78	73
Bangladeshi	69	68
Chinese	91	90

To summarise, it would appear that poor literacy in English is an identifiable factor lowering response rates amongst some South Asian groups. This is likely to bias results for these groups more than for others because literacy is almost certainly correlated with survey relevant characteristics. It also seems likely that non-response arising from poor English literacy could be mitigated if it were possible to identify those who cannot read English, but who can read another language, and send them an appropriately translated questionnaire. But it is also clear that, even if this were possible, the presence of poor literacy in any language amongst some South Asian groups would remain a substantial problem for any self-completion questionnaire survey.

The conclusion is unsatisfactory, but inevitable when using self-completion questionnaire data collection methods. The only way to eliminate poor literacy as a direct cause of non-response would to move to an interviewer administered survey. But changing data collection mode in this way would increase survey costs by several times, and our understanding is that this would be unacceptable.

That said, even if we have to accept that those with poor literacy will inevitably be under-represented, it is important to make efforts to minimise these problems, and one way to do this might be to translate questionnaires. However, this could only be effective if translated questionnaires could be targeted to the appropriate individuals. Unfortunately, our patient sampling frames do not contain data on whether or not sample members are literate in English or, if not, in what other language they are literate. This means we are unable to know in advance which translated version of the questionnaire to include in the questionnaire mailing. In such circumstances in the past, a commonly used approach has been to include in the first questionnaire mailing a note (translated into many languages) encouraging sample members to contact the research team so that they can be sent a translated questionnaire. (A variant of this is to supply a freephone number by means of which the questionnaire can be administered over the phone by a relevant language speaking interviewer.) Unfortunately, in practice, only small numbers of sample members ever request such translated questionnaire copies. In other words, offering translated questionnaires leads to no discernable improvement in response rate or reduction in non-response error.

It seems likely that some improvements in response rate amongst relevant South Asian groups would arise if researchers were able to identify preferred written language in advance of the first mailing. In theory this might be done by asking sample members for this information at the time patient details are entered on the sample frame, although in practice this is almost certainly infeasible. Short of this, it may still be possible to use local knowledge of languages spoken by minority ethnic groups in combination with sample-frame coding of ethnic group to determine which translated versions of the questionnaire to include (presumably always accompanied by an English one). However, this would require trusts to improve considerably the coding of ethnic group in the records forming the basis of the survey sample frames. Furthermore, even if we had information on preferred language readily available, we do not *know*, that our sending translated questionnaires would substantially improve response rates and/or reduce response error. It seems likely that there would be improvements, but we cannot be sure of this without empirical testing.

Disengagement

In both advice centre interviews conducted for this work it was suggested that disengagement from government, and "official" institutions (including the NHS) might contribute to lower response rates amongst younger minority ethnic respondents. Such disengagement might be associated with membership of particular socio-economic groups to which minority ethnic people belong disproportionately, or might be especially prevalent in certain ethnic groups after controlling for socio-economic circumstances.

Although, the disengagement hypothesis is plausible, we are aware of no hard evidence that bears on it. Without this, it is impossible to assess the extent to which this is a problem, or what might be done about it if it is. A well conducted non-response study would, of course, allow us to make these assessments.

Tailoring

It is perfectly possible that different factors affect different groups' propensities to respond in different ways. For example, the main impediment to response amongst young south Asians may be disengagement, whereas that for older south Asians may relate to literacy. An approach to dealing with this kind of situation was suggested during one of the advice centre interviews: this was that efforts should be made to *tailor* questionnaires, accompanying letters, style of approach, etc to the expected characteristics of respondents. The idea underlying this approach (which has been much discussed in the interview survey literature (eg Groves and Couper, 1998) is that if we can match way the survey is presented to the specific concerns of different types of sample members, response is likely to be encouraged.

This approach sounds sensible in principle, but is probably hard to implement in practice. It would require that (i) we know what are likely to be the concerns of different sub-groups and (ii) that we can identify these sub-groups accurately at the sampling stage. Achieving the first of these would require a non-response study, and the second would require considerable improvements in the quality of information held on the sample frame – especially information relating to ethnic group. However, if these obstacles could be overcome, tailoring would seem to be a promising approach worthy of serious consideration.

We note that because the patient surveys are devolved, there is some scope for trusts to tailor questionnaire content and sample size to local conditions. This is, however, rather different to the tailoring just discussed, which would require adapting (potentially radically) how respondents are approached, questionnaire format and content, etc to respondent characteristics. Most of this tailoring would be expected to take place *within* rather than *between* trusts.

3.3 Measurement error

Measurement error refers to the fact that answers to survey questions will often not deliver the information sought accurately for reasons such as poor question comprehension, memory distortion, social desirability pressures, “lazy” answering (termed “satisficing” in the literature), etc (Tourangeau, Rips and Rasinsky, 2000; Krosnick, 1991, 1999). It is possible, perhaps even likely, that measurement errors will affect different ethnic groups in different ways and to different degrees. For example, it seems likely that those for whom English is a second language will be more prone than others to comprehension difficulties. However, in common with other social surveys, the patient surveys do not collect systematic information on measurement error, and we therefore have no information on the nature of any ethnic differences. (Given the multiplicity of response effects, it is far from obvious as to how one would set about systematically measuring response effects in a particular survey even if one had the resources to do so).

That said, concerns that systematic differences in how questions are answered by different ethnic groups might compromise survey results could at least partly be addressed by taking measures during questionnaire development. In particular, we would suggest undertaking separate programmes of cognitive testing with respondents from different ethnic groups in order to establish whether questions are answered in the intended manner by members of different ethnic groups (see Willis 2005).

4 Reporting review

Currently results from the patient surveys are published by the Healthcare Commission in three standard formats:

1. a short key findings report which presents national level results for each question;
2. a benchmark report for each trust showing, for each relevant question, the mean *performance score* for its respondents after age-sex standardisation, and for comparison purposes, the distribution of corresponding performance scores[13] for all trusts;
3. spreadsheet tables showing percentage distributions of answers to all questions for each trust, for SHAs and for England as a whole.

None of these Healthcare Commission standard reports present data separately for sub-groups. However, survey contractors also report to commissioning trusts to a specification agreed locally, and these reports may include analyses by sub-group.

In addition to the standard reports, three further reports have been produced by the Commission for Heath Improvement / Healthcare Commission which explore

---

[13] These are derived scores. Actual answer distributions are not shown.

variations in patient experience “domain” scores[14] by a range of demographic and other factors (Commission for Health Improvement, 2004; Healthcare Commission, 2005a; Raleigh, Irons, Scobie, Cook, Reeves 2005) using, in the main, multivariate analysis techniques. Between them these reports indicate that patient experience does vary by ethnic group with minority ethnic respondents – particularly south Asians – showing lower scores. However, ethnic group membership was less important as a predictor of patient experience domain scores than a number of other predictor variables (notably age and self-reported heath status).

Concern has been expressed that, between them, these reports do not fully exploit the data’s potential for exploring issues of equal opportunities for different ethnic groups. This is argued on the grounds that (i) the basic reports and tables present no findings showing scores for different ethnic groups and (ii) the three reports in which ethnic differences have been explored are academic in style thereby rendering opaque their implications for practice.

On the face of it this argument is reasonable: the data are not reported in a way which facilitates their ready use for monitoring equal opportunities. But this is not surprising because the survey programme was set up primarily to allow individual trust performance to be monitored over time and compared with performance of other trusts, and not to explore equal opportunity issues. Thus the analytic focus of the programme has been on trusts as a whole, and this focus has been reflected in the standard reporting formats. Furthermore, the sub-group analysis that has been done was undertaken with a view to teasing out complex relationships in data, and was not designed to provide information specifically to inform equal opportunity policy. That it is academic in style is therefore unsurprising: the use of complex analytic techniques was integral to its purpose.

It is therefore unfair to criticise the ways in which the surveys have been reported to date for not explicitly addressing issues the that neither surveys nor reports were designed to address. However, this is not to say that the survey reporting should not be adapted in the future so that it can address equal opportunities issues directly. Given the importance of monitoring equal opportunities, it clearly makes sense to do this. But it must also be acknowledged that there will be limits to what can be done with data taken from a survey which was not designed primarily to monitor equal opportunities.

In theory an obvious enhancement would be to present trust level data broken down by main ethnic groups in trusts with large minority ethnic populations (with sub-sample size caveats attached). However, in practice trust sample sizes are almost always too small to allow reliable estimates to be made for individual ethnic groups at trust level[15]. Trusts interested in obtaining separate estimates for individual minority groups may, of course, be able to overcome this problem by boosting their overall sample size. Further, they might also consider boosting sample numbers of particular

---

[14] constructed by aggregating scores derived from answers to defined constituent questions.

[15] In a handful of trusts per survey there are enough respondents in one (or very occasionally two) of the minority groups to allow some comparisons with white respondents to be made; but even in these trusts inevitably sub-group confidence intervals will be large and the chances of finding statistically significant differences will be small.

minority ethnic groups, although most trusts would have to improve substantially their levels of routine ethnic flagging to make this a viable option.

Another approach might be to present simple *national* or *other sub-national* breakdowns by ethnic group for individual questions. Of course, as implicitly acknowledged in the earlier analyses by ethnic group (Commission for Health Improvement, 2004; Healthcare Commission, 2005a), this could lead to real dangers of misinterpretation because of members of minority ethnic groups tend to be concentrated in certain areas and because different ethnic groups also differ from one another demographically and socio-economically; but this, presumably, could be dealt with by the inclusion of appropriate caveats.

Ultimately, if data are to be presented in new ways designed to help monitor equal opportunities, there will need to be a prior consultation process in which relevant professionals are asked about their data needs. Hopefully it will be possible to accommodate many of these by broadening the range of data reports that are produced. However, there is also a strong possibility that it will not be possible to meet some these needs with the surveys in their current form. If this is the case decisions will need to be made about exactly what the focus of the survey programme should be. We should reiterate though, that if any of these decisions involve boosting numbers of minority ethnic respondents in individual trust samples, the possibility of their implementation will depend upon trusts' improving very substantially their recording of patient ethnic group[16].

5 Conclusions

Our main conclusions from this work are summarised below.

1. Minority ethnic patients have lower survey response rates than other patients in the patient surveys.
2. There are currently no criterion data sets that can be accurately used to check the representativeness of the patient surveys.
3. We have no grounds for supposing that estimates for minority ethnic sub-samples are more, or less prone than others to errors in respect of sampling, sample frames, survey specification, or processing.
4. In the absence of information to the contrary, we have to suppose that, as a result of lower response rates, they *are* subject to a higher risk of non-response bias.
5. We have grounds for thinking that the surveys will under-represent patients with poor literacy skills, that this is especially a problem for some South Asian groups, and that this may lead to increased bias for these groups in estimates for variables correlated with literacy.
6. We feel that there are measures that could be taken to boost response rates generally which would probably also have a positive impact on minority ethnic response rates.
7. If information about languages read or spoken could be collected or inferred from sample frame information, it would be worthwhile sending appropriately translated questionnaires to sample members.

---

[16] Without this, any boost sample would be potentially subject to very large non-coverage biases.

8. Doing this would probably reduce non-response bias arising from poor literacy, but would not eliminate it.
9. We have no evidence on different ethnic groups' proneness to measurement error, but feel that in future measures should be taken during questionnaire development to help guard against this possibility.
10. Survey reports do not currently provide information in a manner well suited to equal opportunity monitoring, but this is because the surveys were not set up with this as a primary objective. However, new forms of survey reporting could be introduced; this would best be done after consultation with relevant professionals.

On the whole it is fair to say that the patient surveys are well conducted postal surveys. They do under-represent minority ethnic groups relative to the white patient population, but in the main we feel that this can be attributed to (i) the fact that survey response rates generally tend to be lower for minority ethnic groups and (ii) use of a self-completion methodology which will inevitably achieve lower response rates with less literate populations.

It could, of course, be argued that an interview based methodology should be used instead of a self-completion one. A well conducted interview survey would certainly achieve higher response rates all round and would probably remove most of the non-response bias attributable to poor literacy. It would, however, be prohibitively expensive. In our view, for the allowable budget there is at present no alternative to the current methodology (although this may change over time).

We have a number of suggestions for methodological improvements, designed either to help improve the quality of data collected, or to help measure levels of error.

1. We recommend that considerable efforts should be made to persuade trusts to improve the recording of ethnic identity in the databases which serve as the basis of the survey sample frames. This (i) would allow better estimation of response rates for different ethnic groups (ii) would enable trusts to boost numbers sampled in particular ethnic groups (iii) would open up the possibility of using ethnic group as a non-response weighting factor in an effort to reduce non-response bias (iv) could serve as the basis for sending translated questionnaires (v) could serve as the basis for tailoring questionnaires to respondent characteristics. Of course, if it were possible to collect other related information (such as that relating to language spoken) this would be better still.
2. We recommend that trust population size data should be collected from the sample frame at the time of sampling so that, in principle, it becomes possible to weight survey data to represent national patient populations; producing estimates for these populations may be of interest in their own right and would also facilitate comparison with criterion datasets.
3. We recommend that response rates be calculated in a standard way, preferably one which adheres to currently accepted canons of good practice (see annex D)
4. We recommend that continued efforts should be made to conduct non-response studies to try to assess (i) why different groups (including ethnic groups) do not respond, and (ii) to what extent responders differ from non-responders on key survey variables. Such studies would provide invaluable information on how response rates might be improved (and would serve as the basis for a tailoring approach of that were

felt to be worthwhile) and would enable us to estimate likely levels of non-response bias[17].

5. If future non-response studies prove to be infeasible, we recommend that consideration be given to conducting one or more *national* high response rate face-to-face interview surveys of relevant patient populations. These would provide invaluable information on levels non-response bias generally, and (with a suitable sample design) should also be able to provide information on non-response bias for specific minority ethnic groups.

6. We recommend that further checks should be made of the properties of the sample frames for the ambulance and PCT surveys (we understand that such a review id already underway for the mental health survey).

7. We recommend that serious consideration be given to introducing methodological enhancements (eg providing incentives, sending advance letters, etc) which would be expected to boost response rates generally. Of course, any such consideration should consider carefully how realistic it will be for trusts to met the additional costs associated with some of these enhancements (notably with the provision of incentives).

8. We recommend that greater focus should be given to likely differences in measurement error between different ethnic groups during questionnaire development, perhaps through separate programmes of cognitive testing with different ethnic groups.

9. We recommend that discussions be continued with professionals involved in equal opportunities monitoring with a view to meeting their reporting needs (accepting that some of these may not be able to be met by the survey programme in its current form).

---

[17] Although it should be noted that a previous proposal for a non-response study was refused ethical approval. We hope that it will be possible in the future to convince an ethics committee that a well conducted non-response study does not contravene any reasonable ethical code, and that this has been demonstrated in the conduct of numerous non-response studies conducted in this country and abroad.

Annex A: Additional tables

Table A.1 Percentages with same ethnic group on questionnaire as on sampling frame (survey respondents)

Ethnic group on sample frame	% with same ethnic group on questionnaire				
	Inpatients	Young patients	Outpatients	Emergency Department	Mental Health Trusts
White	99.4	98.2	99.6	98.9	98.8
Mixed	36.2	68.5	45.3	48.8	44.5
Asian / Asian British	81.0	79.3	85.7	79.6	81.5
Black / Black British	86.0	80.0	76.5	77.7	69.1
Chinese	76.8	59.5	73.6	71.4	69.4
Other	4.4	1.9	10.5	7.2	9.1

Table A.2 Percentages with same ethnic group on sampling frame as on questionnaire (survey respondents)

Ethnic group on questionnaire	% with same ethnic group on sample frame				
	Inpatients	Young patients	Outpatients	Emergency Department	Mental Health Trusts
White	98.3	94.9	98.7	98.7	98.1
Mixed	22.0	34.4	26.3	25.4	29.2
Asian / Asian British	84.6	77.8	85.3	82.2	78.4
Black / Black British	83.8	74.6	87.9	82.3	88.0
Chinese	71.0	61.7	83.2	65.2	71.4
Other	40.2	31.8	52.9	20.8	14.5

Table A.3 Percentages coding themselves as belonging to any group apart from white on questionnaire by ethnic group on sampling frame (survey respondents)

Ethnic group on sample frame	% coding themselves as belonging to a non-white group on the questionnaire				
	Inpatients	Young patients	Outpatients	Emergency Department	Mental Health Trusts
White	0.6	1.8	0.4	1.1	1.2
Mixed	67.7	80.6	76.0	81.3	70.1
Asian / Asian British	86.5	85.7	90.8	85.3	91.8
Black / Black British	93.9	92.5	84.6	87.3	80.9
Chinese	91.9	86.5	86.8	90.5	86.1
Other	29.2	30.2	49.1	49.7	39.9

Table A.4 Response rate by ethnic group: inpatients 2005 (Acute Trusts)

Ethnic group	Base	Adjusted response rate (%)
White	99,946	60.9
Mixed	526	46.8
Asian or Asian British	4,540	40.9
Black or black British	2,963	41.6
Chinese	235	46.0
Other	1,960	48.8
Not known (not flagged)	30,419	58.3
Overall	140,589	59.0

Table A.5 Response rate by ethnic group: young patients 2004 (Acute Trusts)

Ethnic group	Base	Adjusted response rate (%)
White	73,999	52.1
Mixed	1,413	46.4
Asian or Asian British	5,700	39.3
Black or black British	2,994	39.8
Chinese	234	52.6
Other	5,264	43.0
Not known (not flagged)	35,602	48.4
Overall	51,207	49.7

Table A.6 Response rate by ethnic group: outpatients 2004/05 (Acute Trusts)

Ethnic group	Base	Adjusted response rate (%)
White	80,788	61.1
Mixed	444	45.0
Asian or Asian British	3,829	39.3
Black or black British	2,996	47.8
Chinese	248	52.0
Other	1,055	46.1
Not known (not flagged)	53,208	58.6
Overall	142,568	59.1

Table A.7 Response rate by ethnic group: emergency department 2004/5 (Acute Trusts)

Ethnic group	Base	Adjusted response rate (%)
White	61,368	46.5
Mixed	425	31.1
Asian or Asian British	3,517	31.0
Black or black British	2,633	32.7
Chinese	115	39.1
Other	497	35.4
Not known (not flagged)	60,530	40.5
Overall	129,085	42.9

Table A.8 Response rate by ethnic group: service users 2006 (Mental Health Trusts)

Ethnic group	Base	Adjusted response rate (%)
White	47,546	38.2
Mixed	495	30.9
Asian or Asian British	1,878	28.9
Black or black British	2,008	30.8
Chinese	107	33.6
Other	831	32.9
Not known (not flagged)	13,380	37.6
Overall	66,245	37.4

Table A.9 Logistic regression of whether or not responded on ethnic group, trust* and age: coefficients for ethnic group variables (inpatients 2005 (Acute Trusts))**

	Coef.	Std. Err.	P	Odds Ratio
Mixed	-0.308	0.091	0.001	0.735
Asian	-0.532	0.033	0.000	0.587
Black	-0.406	0.041	0.000	0.667
Chinese	-0.433	0.135	0.001	0.649
Other	-0.251	0.049	0.000	0.778
Not known	-0.158	0.016	0.000	0.854

* 167 dummy variables

** using white as reference category

Table A.10 Logistic regression of whether or not responded on ethnic group, trust* and age: coefficients for ethnic group variables (outpatients 2004/5 (Acute Trusts))**

	Coef.	Std. Err.	P	Odds Ratio
Mixed	-0.301	0.099	0.002	0.740
Asian	-0.585	0.036	0.000	0.557
Black	-0.229	0.041	0.000	0.795
Chinese	-0.095	0.131	0.467	0.909
Other	-0.299	0.065	0.000	0.742
Not known	-0.141	0.014	0.000	0.868

* 168 dummy variables

** using white as reference category

Table A.11 Logistic regression of whether or not responded on ethnic group, trust* and age: coefficients for ethnic group variables (Emergency Department patients 2004/5 (Acute Trusts))**

	Coef.	Std. Err.	P	Odds Ratio
Mixed	-0.285	0.109	0.009	0.752
Asian	-0.307	0.040	0.000	0.735
Black	-0.205	0.047	0.000	0.815
Chinese	0.001	0.198	0.995	1.001
Other	-0.242	0.097	0.013	0.785
Not known	-0.261	0.016	0.000	0.770

* 152 dummy variables
** using white as reference category

Table A.12 Logistic regression of whether or not responded on ethnic group, trust* and age: coefficients for ethnic group variables (service users 2006 (Mental Health Trusts))**

	Coef.	Std. Err.	P	Odds Ratio
Mixed	-0.252	0.098	0.010	0.777
Asian	-0.362	0.053	0.000	0.696
Black	-0.248	0.052	0.000	0.781
Chinese	-0.153	0.206	0.458	0.858
Other	-0.100	0.077	0.194	0.904
Not known	-0.009	0.022	0.699	0.991

* 78 dummy variables
** using white as reference category

Annex B: Published response rates

Tables B.1 to B.5 show response rates as reported in the survey key findings reports for different ethnic groups and for the overall sample in the most recent implementations of each of the four Acute Trust surveys for which these figures have been reported.

Note that the response rates in these tables have been adjusted by the advice centre by taking out of the base cases where questionnaires were returned undelivered or where respondents had died. If response rate is to be used as a measure of survey quality it should be calculated on the base of all *eligible* sample members (Lynn, Beerton, Nicolaas, Laiho and Martin, 2004). And as there is no reason for supposing that patients whose questionnaires were returned undelivered were not eligible for the survey (it would appear, that they had merely supplied poor address details), we dispute the decision to exclude these from the response rate base. In the response tables shown in the body of the text, only those sample members that have been found to be definitely ineligible have been removed from the response rate base.

Furthermore, there are inconsistencies in how published response rates have been calculated and presented across the different patient surveys. And indeed response rates published for outpatients, and Emergency Department Surveys are incorrect, for some groups at least, almost certainly by a substantial margin. These inconsistencies and errors relate to the treatment of sampled cases for which questionnaire ethnic group was different from sample frame ethnic group (the latter will include those not flagged on the sampling frame). The different treatment of these cases is summarised below.

	Survey	**Treatment of cases for which ethnic group not flagged on sample frame**
(a)	Adult inpatient survey	Excluded entirely from ethnic group response rate calculations
(b)	Young patients survey	Separate response rate presented for each ethnic group and for unknown cases
(c)	Outpatients survey Emergency department survey	If the ethnic group question on questionnaire was answered, case was included in both numerator and denominator totals for that group (irrespective of sample frame information).

Response rates are presented most satisfactorily for survey (b) (young patients). Separate figures are given for each ethnic group and, again separately, for cases where ethnic group is not known at the time of sampling. This enables clear comparisons to be made. The report for survey (a) presents separate response rates for cases for which ethnic group is known during sampling, but not for cases for which it is not known. However the latter can be (and is here) estimated by subtraction from figures presented in the report. The reports for surveys listed as being of type (c) present response rates that are biased (mainly upwards) to an extent that cannot be known or measured for different ethnic groups (see Annex C).

Table B.1 Published response rate by ethnic group: adult inpatients survey 2004 (Acute Trusts)

Ethnic group	Base	Adjusted response rate (%)
White	90,679	65
Asian or Asian British	3,961	47
Black or black British	2,454	50
Mixed race, Chinese or other ethnic group	4,626	56
All above groups	101,720	64
Not known (not flagged)	41,602*	57*
Overall	143,322	63

* Not shown in original table, but estimated from it by author.

Table B.2 Published response rate by ethnic group: young patients 2004 (Acute Trusts)

Ethnic group	Base	Adjusted response rate (%)
White	73,258	53
Asian or Asian British	5,645	40
Black or black British	2,937	41
Mixed race, Chinese or other ethnic group	6,816	45
Not known (not flagged)	35,171	49
Overall	125,827	50

Table B.3 Published response rate by ethnic group: outpatients 2004/05 (Acute Trusts)

Ethnic group	Base	Adjusted response rate (%)
White	110,037	71
Asian or Asian British	4,527	49
Black or black British	3,247	52
Mixed race, Chinese or other ethnic group	1,825	50
Not known (not flagged)	22,758	Not shown
All above groups combined	142,394	59
Overall	143,596	59

Table B.4 Published response rate by ethnic group: emergency department 2004/05 (Acute Trusts)

Ethnic group	Base	Adjusted response rate (%)
White	82,734	61
Asian or Asian British	4,199	44
Black or black British	3,044	44
Mixed race, Chinese or other ethnic group	1,389	53
Not known (not flagged)	35,747	Not shown
All above groups combined	127,113	44 (43)*
Overall	129,948	44

*The figure of 44% was presented in the report. It is hard to reconcile this figure with the response rates for the individual ethnic groups, and 43% seems more likely.

Table B.5 Published response rate by ethnic group: service users 2004/05 (Mental HealthTrusts)

Ethnic group	Base	Adjusted response rate (%)
White	45,922	53
Asian or Asian British	1,884	40
Black or black British	1,625	36
Mixed race, Chinese or other ethnic group	999	51
Not known (not flagged)	16,518	Not shown
All above groups combined	66,948	39*
Overall	66,948	41

* Estimated by author.

Annex C: Reason why reported response rates for Outpatients, Emergency Department and Metal Health Trust surveys are generally inflated

With the simplifying assumptions that (i) the sample frame accurately records ethnic group and that (ii) all respondents (whether initially flagged or not) will give an accurate account of their ethnic group in the questionnaire[18], the true response rate for ethnic group i can be given by:

$$RR_i = r_i/s_i$$

$$= r_i/(r_i + n_i)$$

where:

r_i = no. eligible patients responding in group i
s_i = no. eligible patients sampled in group i
n_i = no. eligible patients sampled in group i not responding

This can be expressed as:

$$RR_i = (r_{if} + r_{iu})/(r_{if} + r_{iu} + n_{if} + n_{iu})$$

where:

r_{if} = no. respondents in group i flagged as being in group i at time of sampling
r_{iu} = no. respondents in group i not flagged as such at time of sampling
n_{if} = no. non-respondents in group i flagged as being in group i at time of sampling
n_{iu} = no. non-respondents in group i not flagged as such at time of sampling

Because we do not know the value of n_{iu} we can only estimate r_i/s_i.

Surveys (a) and (b) do this with the formula:

$$RR_{iab} = r_{if}/(r_{if} + n_{if})$$

This will give a good estimate of RR_i if response rates for flagged and unflagged cases belonging to group i are the same (ie if = $r_{if}/(r_{if} + n_{if}) = r_{iu}/(r_{iu} + n_{iu})$). If this assumption is untrue, these response rate estimates will be either biased upwards or downwards depending on whether the flagged case response rate is greater or less than the unflagged case one.

Surveys (c) use the formula:

$$RR_{ic} = (r_{if} + r_{iu})/(r_{if} + r_{iu} + n_{if})$$

As can be seen, this formula solely differs from that for an unbiased measure of response rate in not including an n_{iu} term in the denominator. For this reason, to the

---

[18] These assumptions enable our exposition to be kept reasonably simple. Although they not strictly true, they are almost certainly close enough to the truth to enable the logic we present to stand up.

extent that responses are received from unflagged patients in group *i*, its estimate of response rate will be biased upwards[19]. The extent of this bias will depend upon (i) the extent to which unflagged cases belong to ethnic group *i* and (ii) the level of response from group *i* unflagged cases. Obviously, factors (i) and (ii) are likely to vary from group to group meaning that, in turn, the extent of upward bias is likely to vary from group to group.

Therefore, to the extent that the two assumptions stated at the beginning of this annex are true, response rates for different ethnic groups presented in survey reports for surveys of type (c) (and reproduced in tables B.1 to B.4) will be inflated to a greater or lesser extent.

Of course, we know that these assumptions are not strictly true, and we might therefore revise them to: (i) the assumption that those recorded as belonging to an ethnic group on the sample frame may or may not actually belong to that ethnic group and (ii) the assumption that respondents will give an accurate account of their ethnic group in the questionnaire[20]. On this basis we can say that the true response rate will be:

$$RR_i = (r_{if} + r_{in} - r_{ip} + r_{iu})/(r_{if} + r_{in} - r_{ip} + r_{iu} + n_{if} + n_{in} - n_{ip} + n_{iu})$$

where:

r_{if} = no. respondents in group *i* flagged as being in group *i* at time of sampling
r_{in} = no. respondents in group *i* that were incorrectly flagged as being in an ethnic group other than group *i* at time of sampling
r_{ip} = no. respondents *not* in group *i* that were incorrectly flagged as being in group *i* at time of sampling
r_{iu} = no. respondents in group *i* not flagged at time of sampling
n_{if} = no. non-respondents in group *i* flagged as being in group *i* at time of sampling
n_{in} = no. non-respondents in group *i* that were incorrectly flagged as being in an ethnic group other than group *i* at time of sampling
n_{ip} = no. non-respondents *not* in group *i* that were incorrectly flagged as being in group *i* at time of sampling
n_{iu} = no. non-respondents in group *i* not flagged at time of sampling

On these less stringent assumptions response rates reported for surveys of type (c) will have been calculated as:

$$RR_{ic} = (r_{if} + r_{in} - r_{ip} + r_{iu})/(r_{if} + r_{in} - r_{ip} + r_{iu} + n_{if})$$

This adjusts the formula for RR_{ic} given above ($RR_{ic} = (r_{if} + r_{iu})/(r_{if} + r_{iu} + n_{if})$) by adding ($r_{in} - r_{ip}$) to both numerator and denominator. The impact of adding this quantity on the response rate for an individual ethnic group will obviously depend upon the size and sign of the quantity, and on the magnitudes of ($r_{if} + r_{iu}$) and ($r_{if} + r_{iu} +$

---

[19] Except in the very unlikely circumstance that $n_{iu} = 0$.

[20] Using the standard approach to ethnic categorisation in Britain this is true by definition. See http://www.statistics.gov.uk/about/data/harmonisation/default.asp

n_{if}), and, as such, it can have an overall inflationary or deflationary impact on response rates. On the other hand, as discussed above, the addition of the r_{iu} to both numerator and denominator will have a consistently inflationary impact.

In other words, if response rates are calculated by the method used for surveys (c) they will almost certainly be biased upwards simply because of the method of calculation used. On the other hand the calculation method used for surveys (b) will deliver a response rate for an individual ethnic group which may or may not be biased depending on (i) the level and type of inaccuracy of recording of ethnic group on the sample frame (ii) the differences between the response rate for those members of an ethnic group for whom ethnic group was accurately recorded on the sample frame and the response rate for those for whom it was not accurately recorded.

Annex D: Recommended response rate calculation

Currently, the patient surveys records one of the following outcomes for each issued sample member:

1 Returned useable questionnaire
2 Returned undelivered or patient moved house
3 Patient died
4 Patient too ill, opted out or returned blank questionnaire
5 Patient not eligible to fill in questionnaire
6 Questionnaire not returned - reason not known

The most appropriate measure of response rate is the one that gives us the best estimate of how well we have covered eligible sample members. In our view this is best obtained by dividing the number of respondents for whom a returned (and eligible) questionnaire was received by the number of sample members sent questionnaires less those known to have died or to be definitely ineligible. In terms of the above codes, this means calculating it as:

Response rate = 1/(1 + 2 + 4 + 6)

This is the method we have used in this report, and the one we recommend for future calculations.

References

Biemer, P. and Lyberg, L. (2003), *Introduction to Survey Quality*, New York: Wiley

Commission for Healthcare Improvement (2004), *Unpacking the patients' perspective: variations in NHS patient experience in England*, London: CHI

Department of Health (2004), *National Standards Local Action*, London DH

Dillman, D. (2000), *Mail and Internet Surveys: the Tailored Design Method*, New York: Wiley

Erens, B,. Primatesta, P. and Prior, G. (2001), *The Health Survey for England: the Health of Minority Ethnic Groups*, London: HMSO

Groves, R. (1989) *Survey Errors and Survey Costs*, New York: Wiley

Groves, R. and Couper, M. (1998), *Nonresponse in Household Interview Surveys*, New York: Academic Press

Groves, R., Fowler, F., Couper, M., Lepkowski, J., Singer E. and Tourangeau, R. (2004) *Survey Methodology*, New York: Wiley

Healthcare Commission (2005a) *Variations in the experiences of patients in England: Analysis of the Healthcare Commission's 2003/2004 national surveys of patients*, London: Healthcare Commission

Healthcare Commission (2005b), "Comparison of ethnic composition of stroke survey sample respondents with ethnic composition derived from Hospital Episode Statistics", *Unpublished paper*

Jenkins, C. and Dillman, D. (1997), "Towards a Theory of Self-Administered Questionnaire Design, " in Lyberg, L., Biemer, P., Collins, M., de Leeuw, E., Dippo, C., Schwartz, N. and Trewin, D (eds), *Survey Measurement and Process Quality*, pp 165-196, New York: Wiley

Johnson M, Owen D. and Blackburn C. (2000), *Black & ethnic minority groups in England: the second health & lifestyles survey*, London: Health education Authority

Keeter, S., Miller, C., Kohut, A., Groves, R., & Presser, S. (2000). "Consequences of reducing nonresponse in a national telephone survey," *Public Opinion Quarterly*, 62(2), 125-148.

Krosnick, J., (1991), "Response strategies for coping with the Cognitive Demands of Attitude measures in Surveys," *Applied Cognitive Psychology*, 5, pp 213-236

Krosnick, J., (1999), "Survey Research," *Annual Review of Psychology*, 50, pp 537-567

Lynn, P., Beerten, R., Laiho, J. and Martin J. (2001) "Recommended Standard Final Outcome Categories and Standard Definitions of Response Rate for Social Surveys". Working Paper of Institute for Social and Economic Research, paper 2001-23 (PDF). Colchester: University of Essex.

Merkle, D. and Edelman, M. (2002), "nonresponse in Exit Polls: a Comprehensive Analysis," in Groves, R., Dillman, D., Eltinge, J. and Little, R. *Survey Nonresponse*, pp243-258, New York: Wiley

Nazroo, J., (1997), *The Health of Britain's Ethnic Minorities*, London: PSI

Osborn, C., Reeves, R., Howell, E. and Magee, H. (2004), *Development and pilot testing of the questionnaire for use in NHS Trust-based Mental Health service user survey*, Oxford: Picker Institute Europe

Raleigh, V., Irons, R., Scobie, S., Cook, A. and Reeves, R. (2005) *Ethnic variations in the experiences of mental health service users in England: results of a national survey programme*, London: Healthcare Commission

Rudat, K. (1994) *Health and Lifestyles: Black and Minority Ethnic Groups in England,* London, HEA

Seccombe, I. and Raleigh, V. (2005) "Black and minority ethnic response rates in the Healthcare Commission's Surveys of community mental health service users" *Unpublished paper*

Sproston, K. and Mindell, J. (2006) *Health Survey for England 2004:the health of minority ethnic groups,* London: HMSO

Stoop, I., (2006), "How to achieve high response rates", Paper presented at RSS Social Statistics Section meeting on non-response, 21 March, 2006

Teitler, J., Reichman, N., Sprachman, S. (2003) "Costs and Benefits of Improving Response Rates for a Hard-to-Reach Population." *Public Opinion Quarterly* 67(1): 126-138.

Tourangeau, R., Rips. L. and Rasinsky, K. (2000) *The Psychology of Survey Response,* Cambridge: Cambridge University Press

Willis, G. (2005), *Cognitive Interviewing*, Thousand Oaks: Sage