

Community Mental Health Survey 2015: Sampling & Care Cluster Errors Report

1. Introduction

Survey contractors and in-house trusts were required to submit samples to the Co-ordination Centre for final quality control checks before the mailing out of survey packs could begin. These sample files did not include patient identifiable data such as names or addresses, but did include demographic and clinical information about each service user:

Demographic information:

- gender;
- year of birth;
- ethnic category.

Clinical information:

- date of last contact;
- CPA status;
- CCG code.

This year the Chief Executive Declaration was replaced by a Sample Declaration, to be signed both by the person drawing the sample and the Caldicott Guardian. Trusts were required to indicate the total eligible population of service users at the trust and the breakdown of the number of service users on new CPA, not on new CPA and where CPA status was not known, along with the CPA status of users in the 850 sample. Trusts were also required to confirm that the sample had been through DBS checks, that the sample had been checked as per the guidance manual, no service user identifiable information had been included, care cluster data had been removed, only data relating to the required sample variables was included and that service users who had dissented from the survey had been removed from the sample (*in addition, trusts were asked to declare the number who had been removed*). The Sample Declaration for trusts using a contractor can be found here: <http://www.nhssurveys.org/survey/1506>

This document outlines the types of errors made in Sample Declarations and in samples submitted to the Co-ordination Centre for checking. Sample errors are divided into major (*those requiring the sample to be redrawn*) and minor (*those that could be corrected using the same sample*). It is important to note that these are only the errors caught by the Co-ordination Centre; many samples would have had errors which were identified during contractors' checks. The types of errors found in care cluster data is also discussed, including Section 251 breaches. Note that as part of the requirement of the Section 251 approval for the survey, care cluster data was submitted separately to the main sample file and directly to the Co-ordination Centre. Finally, there were two trusts that were not able to participate in this year's survey: an explanation for the rationale behind this decision is given.

This document should be used by trusts and contractors to become familiar with past errors and to prevent these from happening in future surveys. If further assistance is required, please contact the Co-ordination Centre on 01865 208127.

2. Major Sampling Errors

There were six trusts that incurred major errors in drawing their samples. If not corrected then they can invalidate a trust's participation in the survey, meaning that the trust's survey data cannot be used by the Care Quality Commission in its assessment of mental health trusts in England. Of the 56 trusts who were eligible to take part and submitted samples, when the survey fieldwork period was near-completed it was found that one trust had incorrectly drawn its sample. This invalidated its inclusion and the 2015 survey therefore includes data from 55 trusts:

- One trust's sample showed a large drop in the total eligible population, and initially through discussions with the trust that it became apparent that in previous years there were certain ineligible service user groups that had been included. The person responsible for drawing the sample in 2014 had since left the trust. The trust provided reassurances that the current survey sample was correct with all inclusion and exclusion criteria correctly applied, with improved processes and data structures in place. However, further investigation into the scale of the difference in population size revealed that in re-writing the code, though the trust had been careful not to include any current inpatients, anyone who had ever been an inpatient had been excluded in error. It was found that this mistake in compiling the 2015 survey sample failed to account for around a third of the trust's eligible service users. Unfortunately, the true nature of the error was not realised until well into the fieldwork period, and therefore the 2015 data could not be used due to it not being representative of the trust's user population;
- One trust originally submitted a declaration form that declared twice as many service users in the eligible population than in the previous year (*to prevent delays to the sample submission process it was agreed that queries of this nature would be highlighted at the sample checking stage*). One error the trust had made was in including people who had been in telephone-only contact;
- The eligible population declared by one trust was a third lower than in the previous year, and the age distribution showed fewer older service users in comparison. The trust explained that this was as a result of the trust being in the process of migrating service users across to one system, with older persons' services being migrated across first. Because of this it was not possible to establish whether some service users had attended more than just an initial appointment, as the service users' personal details may have been moved across but not their attendances (*and therefore it was not possible to establish whether they were eligible as it is a requirement of the survey that service users must have had more than an initial assessment*). As it stood, the sample would have been skewed both in terms of the service user demographic and in terms of the services that were represented. The trust was asked to re-identify their eligible population and re-draw their sample once all service users had been migrated across to one system. This had to be done manually so took some considerable time though is not likely to have had an overly detrimental impact on response rates as the trust was still able to go into field within a reasonable timescale;
- A sample was submitted that looked odd when compared to the previous year's sample, chiefly in terms of the age profile which showed a large increase in numbers of older service users, but also with regards to gender and to a lesser degree, ethnic coding. The trust could not identify how an error had occurred but conceded that the sample did not appear to have been randomly selected from the eligible population (*which was comparable with the previous year's population profile*) and so it was re-drawn. This type of error suggests that either the original sample had not been randomly drawn, or an error had been made when

sorting or filtering the sample data. It was likely that, if uncorrected, whole groups of service users or clinical areas would not have been accounted for;

- Another trust originally arrived at an eligible population that failed to account for two-fifths of its eligible service users. The trust then wrote an entirely new extract code and it became apparent that not all eligible service users were being picked up (*it is unclear which groups had been excluded*). The trust re-sampled and the resubmitted sample was approved;
- A fifth trust had excluded people with dementia. They were identified by a clinician who thought they were cognitively impaired to the extent that they would be unable to answer the questionnaire. The trust was asked to re-draw their sample to include service users with dementia.

3. Minor Sampling Errors

There were three minor errors found during sample checking. These do not require the sample to be re-drawn but must be corrected.

Trusts were asked to provide the date of last contact as part of the sample information. Two trusts made minor errors:

- Several service users at one trust had a last contact date that extended beyond the date the sample was drawn. It transpired that these were scheduled appointments - and not all of these were attended – the trust therefore provided the latest known attended appointment date up to the date the sample was drawn;
- Another trust had originally submitted a sample that did not account for all its eligible service users as records were held across two systems. These were later merged to enable the sample to be re-drawn once all eligible service users had been accounted for. We would expect to see an exponential curve of latest contact dates from September through to when the sample was drawn, but it was clear that this had been updated for a large cohort of service users but not others. The latest contact date was subsequently updated for the service users in the original sample, bringing the pattern of most recent dates of attendance in line with expectations;
- Finally, one trust's sample had to be resubmitted due to contractor error. In copying across ethnicity data to the sample file, a large chunk of data was not copied correctly and so the submitted file had missing ethnicity codes for 808 of its 850 service users.

4. Errors in Sample Declarations

Some trusts made straightforward errors on their Sample Declarations, such as declaring incorrect figures or neglecting to apply their initials to sections of the form. In instances where only a small change has been made, for example if a record has had to be replaced after the declaration had been signed off, then it is acceptable to explain these differences in a covering note or email that accompanies the declaration:

The following details the errors that were made:

- The numbers of service users falling into each of the CPA status categories at Section B of the form did not match the sample that was submitted by two trusts;

- Another trust made the same error so a second declaration was submitted that reflected the fact that some deceased service users had been removed after the declaration had been signed off. The new form however did not reflect this new figure at Section A which should have shown fewer service users in the eligible population than was declared initially. Rather than submit another form, the trust was asked to confirm what the numbers should be, however in discussions in the interim period it was evident that they had drawn their sample incorrectly and therefore a third declaration form was submitted;
- One trust declared a sample size at Section B of the form that totalled 851, rather than 850. Furthermore, what the trust declared as the number of dissenters removed was with reference to an earlier sample that was drawn and not the fresh sample submission;
- The person drawing the sample is asked to initial the box that confirms that only the required sample variables are included in the file and any patient identifiable information has been removed – one trust had left this blank;
- Another trust had made reference to an appended document at the section that asks for the trust's definitions of CPA status, but had neglected to attach the document;
- A further trust did not initial the section that states that the sample had been checked by trust staff as outlined in the guidance, and the person responsible for drawing the sample had not signed or dated the declaration.

5. Errors in 2014 sampling

In checking the samples for this year's survey, it became apparent that the samples of four trusts submitted for the 2014 survey had been incorrectly drawn. This makes any year on year comparisons questionable:

- As discussed above, one trust submitted an incorrectly drawn sample in 2015, and although it only became clear much later into the fieldwork period that this was down to a major error in drawing this year's sample, it was evident that historically the trust had been including certain ineligible service user groups, for example, forensic services, learning disabilities and Improving Access to Psychological Therapies (IAPT) services;
- For one trust, a two-thirds drop in the size of the eligible population raised concerns initially with regards to this year's submission. It transpired that data requests had previously been outsourced and the trust admitted that there had been data quality issues that they were working to overcome. The new data contact at the trust had used the same extract code on the previous year's data but could not arrive at the same population figure as declared previously. The 2015 data extract was checked once again and it was concluded that the 2015 submission was correct but the previous year's sample submission had been incorrectly drawn. Since it was highly likely a sampling error had occurred in 2014, no trust-level historical comparisons were provided for this trust;
- After lengthy discussions with the trust it was discovered that the 2014 sample had been drawn incorrectly. This was in part evident in that the 2015 and 2013 samples were very similar, and the 2014 very different in terms of CPA breakdown and gender profiles. Coupled with this, the total eligible population in 2014 was significantly lower than in 2013 or 2015. The trust spent considerable time revisiting the 2014 data and concluded that while the 2015 was correct, an error had been made the previous year (*in 2014 the differences seen versus the 2013 sample were at the time attributed to investments in certain services*). Due

to the apparent error in drawing the 2014 sample, the trust did not receive trust-level historical comparisons in their report;

- Through discussions with the trust it became evident that some service users had been removed from their sample as it was thought they would be too cognitively impaired to take part. Unfortunately, this included service users who used memory clinics who are eligible. It was felt that this cohort of service users was likely excluded from the previous year's survey, though as the previous data contact had since left the trust, it was unclear. As it was estimated that this comprised only a few service users, their 2014 submission was accepted for inclusion in the trust-level historical reporting.

6. Errors in care cluster data

As in 2014, trusts were once again asked to submit data relating to their sampled service users' care clusters. The process for compiling and submitting care cluster data was outlined in the main guidance manual, with a more detailed guidance specific to sampling the care cluster data published separately.

Only three columns of data were required: the three-digit trust code, service user record number and care cluster. In the care cluster column, we would only expect to see numeric codes 0-21 (with 0 representing variance, and 1-21 representing the care clusters) and 99 (to replace any blanks, N/A etc.). Further to this, care cluster 9 is no longer used so we would not expect to see a service user assigned this code. Though these errors may seem small, fixing these issues takes up a large portion of Co-ordination Centre time in raising queries and re-formatting the data:

- Several trusts included blanks or 'Null' for care cluster, when these should have been replaced with code 99;
- Another trust erroneously coded all missing data as 0, which were later amended to code 99;
- Two trusts submitted double-digit codes which were also preceded by a 'P';
- One trust submitted codes plus long text descriptions relating to care cluster;
- Three trusts submitted data with the columns in the wrong order;
- One trust only submitted two columns of data, omitting the trust code column;
- Two trusts had service users with a code 9 for care cluster and as this code is no longer used, the trusts were asked if there was a more recently applied care cluster for these service users. All cases had a more recently applied different code or 99;
- One trust had neglected to make a service user replacement that had been done to the main file and so the service user record numbers did not match;
- Finally, two trusts submitted Service User Record Numbers which were not in the prescribed format, e.g. inserting hyphens or missing leading zeros for the record number.

Care cluster data requiring resubmission

- Several trusts had high proportions of service users without any care cluster assigned and so were asked to revisit their records for these service users to see if a care cluster had more recently been allocated (*as care cluster data was drawn at the same time as the main sample file and care cluster data was checked for all trusts once all trusts' sample files had been approved which, for some trusts, was some weeks after their sample approval*);
- Trusts were asked to submit data for all 850 service users, as per their sample file: one trust submitted data for 851 service users which included an original ineligible and replacement record and another submitted only 847 rows of data as they had removed three records relating to service users who had subsequently been identified as deceased. These two trusts were asked to resubmit their data with the full complement of 850 service users for completeness;
- Finally, one trust submitted care cluster data which included six pairs of duplicate Service User Record Numbers.

7. Section 251 breaches

As part of the Section 251 approval, it was mandatory for trusts to submit their care cluster data directly to the Co-ordination Centre's secure ftp server. In submitting their care cluster data, two trusts incurred a breach of Section 251 as they had uploaded their data to their contractors' ftp servers, rather than to that of the Co-ordination Centre. In neither case was the file opened by the contractor, however Section 251 approval by the Confidentiality Advisory Group (CAG) for the collection of care cluster data was granted on the condition that this data would be submitted directly to the Co-ordination Centre. Under these circumstances it was necessary to notify CAG of the breach and the CQC inspector for the two trusts in question.

8. Exclusions

As described in section '2. Major Sampling Errors', one trust's survey data was excluded as it was subsequently found that the 2015 sample had been drawn incorrectly, excluding around a third of its eligible population.

Two additional trusts were excluded from the Community Mental Health Survey 2015 prior to the sample submission stage as they did not have an appropriate consent mechanism in place and therefore had large proportions who could not be approached to take part in the survey. The first had adopted an opt-in consent system which is in contrast with the opt-out consent model employed by the Community Mental Health Survey, in which service users are assumed to consent to being contacted unless they explicitly opt out¹. This procedure has received full support from the Confidentiality Advisory Group under section 251 of the NHS Act 2006.

¹ Participating trusts must display posters, during the sample period, drawing service users attention to the national survey and how they can indicate dissent should they not wish for their details to be used for anything other than clinical care

Opt-in consent models lead to lower overall response rates² and introduce biases in respondent demographics that are not reflected in response/non-response patterns for opt-out surveys³. The trust was first excluded from the 2013 survey but subsequently took steps to change their consent model in readiness for the following year's survey. The new consent form did enable distinctions to be made between service users who had declined to complete the form, who had completed it and who had not been offered it - however, an opt-in question was also included. That meant that the trust was only able to invite those who had given consent for their data to be used and those who had not previously been asked to consent, which still left a large proportion of service users who could not be approached to take part in 2014.

This same consent mechanism was in place at the start of this year's survey and the trust had not addressed what should be done with the large group of service users for whom explicit or implicit dissent was already recorded. The trust was therefore once again deemed ineligible for the 2015 survey. In preparation for the 2016 survey and when a new opt-out consent process is in place, it is anticipated that all service users will need to be contacted once again, to inform them of the new system and to accurately record their wishes.

A second trust also had a high proportion of service users who had either indicated dissent or whose dissent status was unknown. To determine whether the data that would be obtained by running the survey would be representative of all service users, the demographic details of service users who would otherwise be eligible for the survey was requested from the trust. The aim was to identify whether the dissenters as a group differed from those giving consent or whose status was unknown, on particular characteristics that are known to have an impact on survey responses.

It was evident that there were significant differences between the two groups in terms of age, date of attendance, length of contact with the service, CPA status and, to a lesser extent, ethnicity, which would invariably affect survey responses. The potential for bias in the sample that would be obtained was too great to allow for the data to be accepted for Intelligent Monitoring and inspection purposes. It was decided therefore that the trust would not be able to take part in this year's survey.

The participation of a further two trusts was under scrutiny due to their high proportions of dissenters and / or those who had not been asked the dissent question. As for the trust above, demographic and clinical information was sought in order to compare the profiles of their eligible service users versus those who would be excluded. In both cases, the profiles were considered to be sufficiently comparable that the exclusion of these groups would introduce minimal sampling bias, and thus their results would be largely representative of eligible service users at their trust. It was agreed that they would be eligible for this year's survey, however it was recognised that, along with the excluded trusts, they would need to work with the Care Quality Commission and the Information Commissioner's Office in developing and implementing a more systematic consent system that accurately recorded service user dissent status.

² See, for instance, Angus VC et al (2003) The requirement for prior consent to participate on survey response rates: a population-based survey in Grampian. *BMC Health Services Research*; 3: 21.

³ See, for instance, Al-Shahi R et al (2005) Bias from requiring explicit consent from all participants in observational research: prospective, population based study. *BMJ*, 331, 942-5.

9. Technical problems

One trust had ongoing delays in compiling their sample due to their data being housed with another trust and their request not being prioritised. This put back the trust's sample submission by several weeks. The trust also encountered issues in gaining access to DBS, and as such a manual check had to be undertaken to prevent any further delay in going into field.