Health Professions Council Meeting – 26 March 2009

Proposal to reduce continuing professional development (CPD) audit sample size from 5% to 2.5% from June 2009 and CPD update

Executive Summary and Recommendations

Introduction
Since January 2007, the CPD project has been project managed by Claire Reed (HPC Project Manager), led by Richard Houghton (Head of Registration), with Greg Ross-Sampson (Director of Operations) as the project sponsor.

Decision

- The Council is requested to note the attached paper that was presented to the Education and Training Committee meeting on the 25 March 2009.

Background information
See attached Education and Training Committee paper.

Resource implications
Nil

Financial implications
Nil

Background papers
See attached Education and Training Committee paper.

Appendices
See attached Education and Training Committee paper.

Date of paper
12 March 2009

Appendix A

Health Professions Council (HPC)
Advice on Sample Size for CPD Audit Process V2.0

Polly Grove and Savitri Wilson
Statistical Services Centre
University of Reading
24th February 2009

Contents

Background

There are 13 Health Professions within the remit of the proposed Continuing Professional Development (CPD) Audit conducted by the Health Professions Council (HPC).
Auditing all registrants across all professions would be extremely burdensome and costly, therefore a sampling design and sample size needs to be determined in such a way as to be confident that the results of auditing process can be relied upon. In particular, the HPC wish to report findings on the register as a whole and be confident that any sampling scheme would be representative of all 13 professions. To date, a proportional sample size of 5% has been undertaken across two professions, ODP and Chiropodists.

There are a number of outcomes that are of interest and for which the HPC has information from the two samples that have already been conducted. These are proportions in each sample of the following:

- acceptances
- deferrals
- voluntary deregistrations
- non-renewals by cut-off (lapsed)
- removals from the register due to failure of the CPD Standards

Another concern of the HPC is the impact of a reduced sample size, in particular, taking a sample size of 2.5% as opposed to 5% from the remaining professions.

The HPC needs advice about the sampling design and the sample size requirements for continuing the above CPD audit of the remaining professions in order to report on all registrants.

Terms of reference

The Statistical Services Centre will

i. Propose a sampling scheme that allows estimates to be derived for the register as a whole.
ii. Present a number of options for sample sizes indicating the impact on the precision of the outcome estimates of the audit. Sample size results will explicitly describe assumptions made for the calculation of these estimates. Any limitations associated with the use of a 2.5% sample will also be highlighted.
iii. Present and discuss findings in a manner that allows options to be fully understood.

Sample Size Calculations

Recall that HPC are interested in outcome proportions of accepted profiles, deferrals, voluntary de-registrations, non renewals (lapsed) and removals. Their wish is to report on these for the entire register of all professions. To date, HPC has undertaken sampling audits across two professions, Chiropodists and Operating Department Practitioners. The initial findings from these suggest that the above mentioned outcome proportions range from about 1% to 70%.
In this report, we consider basing sample size calculations on a population of 183,912 registrants. We then look at the effects of taking a 2.5% sample and consider the risks of non-detection in this scenario.

Fixed Precision - Effect on Sample Size

Here we consider attaching a fixed *margin of error, d,* to our estimate, e.g. our estimate could be of the proportions of deferrals in our sample. In simple terms, *d* will be the distance between the estimated value based on sample results, and the true population proportion that we are trying to estimate. The smaller the value of *d,* the closer will be our sample estimate to the corresponding population value. We fix *d* to be a specified fraction of the proportion (or percentage) of interest, e.g. $d = 0.1 \times P$, where P is the percentage of interest in the population. Our aim is to be 95% confident that the estimate we get lies within $\pm$ **10%** of the true population value.

Table 1 presents the sample size calculation results based on varying proportions of interest. In this table, the first column represents the population percentage (call it P), that we wish to estimate using the sample data. The second column represents this percentage as a proportion. These columns are not shown beyond 50% (or 0.50) because results would be identical in reverse thereafter, i.e. results for 35% would be identical to results of 65%.

The third column in Table 1 represents the degree of accuracy ($\pm d$) we wish for our estimate, i.e. we wish our estimate to lie within $\pm$ d percentage points of the true value P with a high level of confidence (say 95%). In Table 1 we have taken d to be 10% of P. Thus values in the third column are a tenth of the values in column 1. Likewise, column 4 represents a tenth of the values in column 2.

The final column shows the sample sizes required to ensure 95% confidence that the percentages shown in column 1 can be estimated to lie within $\pm d$ percentage points (column 3) of values shown in column 1.

To illustrate, say we were interested in being able to estimate a population proportion which corresponds to 15%. This could represent the expected rate of deregistration, for example. In this scenario, we fix the level of accuracy to be 10% of 15%, i.e. $\pm$ 1.5 percentage points, and the level of precision required to be 95%. In other words, if the true population value of interest is 15%, we want our estimate to lie within 13.5% and 16.5% with 95% confidence. To achieve this degree of accuracy and precision, we would need a sample of size n=2242.

As a second example, consider the last row of Table 1. With a sample size of n=402, we can be 95% confident that the estimate we get, if the true population proportion is P=50%, will lie within 5 percentage points of 50%, i.e. the estimate will lie between 45% and 55%.

Table 2, on page 6, presents sample size results based on a degree of accuracy of **20%** of the population value. Here *d* is 20% of P. We now see that values of *d* in column 3 are one fifth of those in the first column. As an example, say we are interested in being able to estimate a population proportion of 15% as in the example above based on Table 1. Now we are fixing the level of accuracy to be 20% of 15%, i.e. $\pm$ 3 percentage points, and the level of precision required to be 95%. In other words, if the true population value of interest is 15%, we want our estimate to lie within 12% and 18% with 95% confidence. To achieve this degree of accuracy and precision, we would need a sample of size n=576. With a sample size of n=102 (the final row of Table 2), we can be 95% confident that the estimate we get, if the true population proportion is P=50%, will lie within 10 percentage points of 50%, i.e. the estimate will lie between 40% and 60%.

The sample size results between these two precision levels diverge greatly as the population proportion we are interested in being able to detect gets smaller (See Figures 1 and 2).

Table 1: Sample Size Results for various Proportions with fixed Precision (d=0.1xP)

Population %	Population Proportion	d as %	d as proportion	Resulting Sample Size
0.5	0.005	0.05	0.0005	78493
1	0.01	0.1	0.001	39053
2	0.02	0.2	0.002	19333
3	0.03	0.3	0.003	12760
4	0.04	0.4	0.004	9473
5	0.05	0.5	0.005	7501
10	0.1	1	0.01	3557
15	0.15	1.5	0.015	2242
20	0.2	2	0.02	1585
25	0.25	2.5	0.025	1191
30	0.3	3	0.03	928
35	0.35	3.5	0.035	740
40	0.4	4	0.04	599
45	0.45	4.5	0.045	490
50	0.5	5	0.05	402

Table 2: Sample Size Results for various Proportions with fixed Precision (d=0.2xP)

Population %	Population Proportion	d as %	d as proportion	Resulting Sample Size
0.5	0.005	0.1	0.001	20157
1	0.01	0.2	0.002	10029
2	0.02	0.4	0.004	4965
3	0.03	0.6	0.006	3277
4	0.04	0.8	0.008	2433
5	0.05	1	0.01	1927
10	0.1	2	0.02	914
15	0.15	3	0.03	576
20	0.2	4	0.04	407
25	0.25	5	0.05	306
30	0.3	6	0.06	239
35	0.35	7	0.07	191
40	0.4	8	0.08	154
45	0.45	9	0.09	126
50	0.5	10	0.1	104

Figure 1 shows the results of Table 1 and Table 2 graphically. As the proportion of what we may be interested in increases, so our required sample size decreases. An interest in detecting a rate less that 5% causes a very steep increase in the required sample size. The effect of choosing to tolerate a greater margin of error is extremely apparent when the population proportion is small. Figure 2 shows a magnification of a section from Figure 1 to allow us to see more clearly the impact on sample size of varying the proportion of interest both for an error margin of 10% of the estimate and 20% of the estimate.

Figure 1: Sample Size vs. Population %

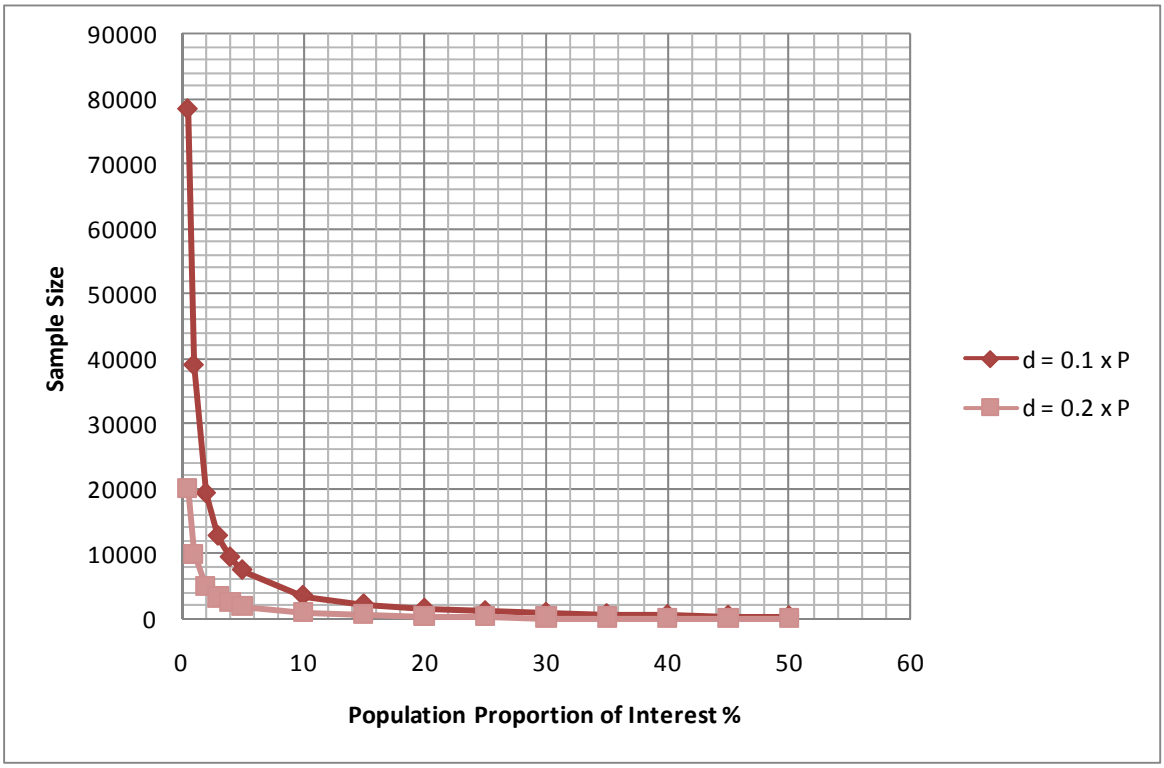
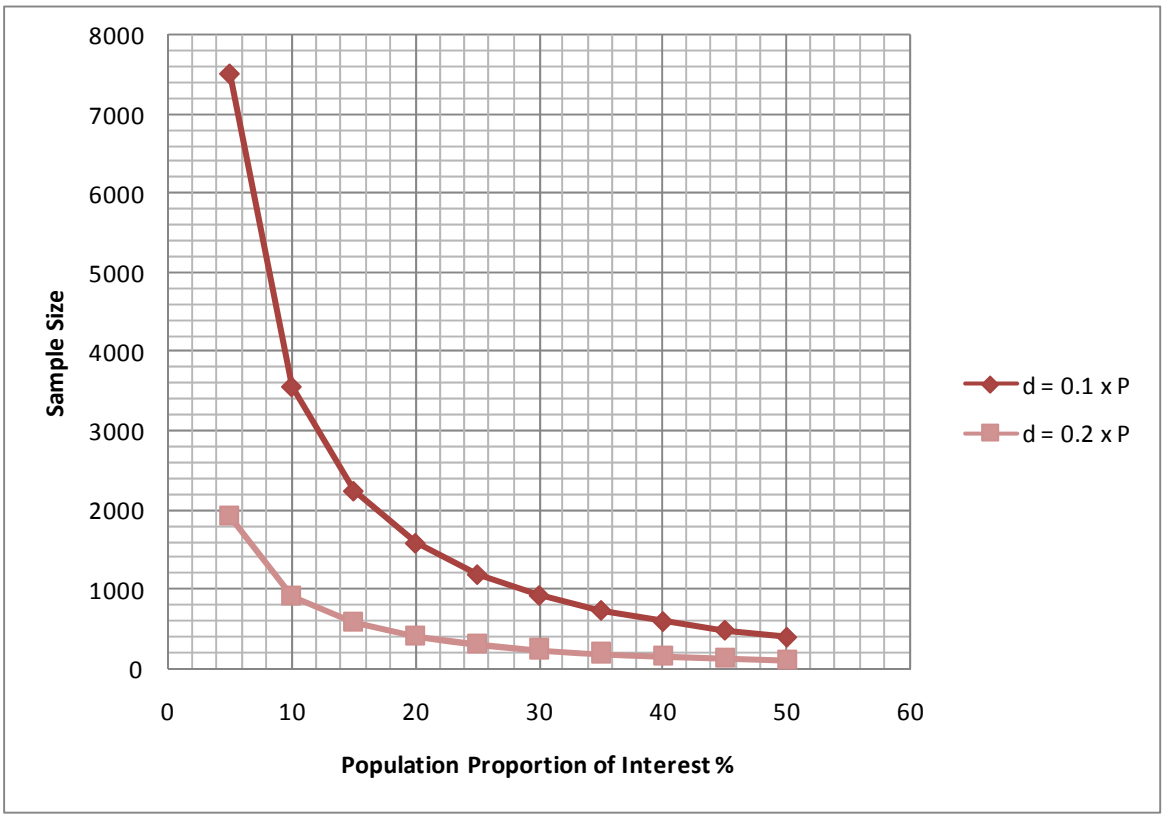


Figure 2: Sample Size vs. Population % (section from Figure 1)



Fixed Sample of 4598 registrants (2.5% of 183912) – Effect on Precision

Here we consider the effect on the precision of our estimates by going with a fixed sample size from our population of registrants. Table 3 shows the precision results based on the proposed sample size of 2.5% of the registrants which corresponds to a sample size of 4598.

For illustration of these results, say the proportion of voluntary deregistered we find in our sample of 4598 registrants is 4% (Row 5, column 2 in Table 3). With this sample size and wishing a 95% confidence, the margin of error works out to be ± 0.58 percentage points i.e. we have 95% confidence that our sample estimate is within ± 0.58 percentage points of the true population proportion of voluntary deregistered (CI[3.42, 4.58]). We can then calculate this margin of error as a percentage of the estimate itself. A margin of error of 0.58 percentage points on an estimate of 4% corresponds to *d* being ± 14.4% of the estimate ((0.58/4)*100%=14.4%) which is shown in column 4 in Table 3.

If we find, for example, a proportion of 1% that are removed from the register at the end of the CPD audit process (Row 2, column 2 in Table 3) in our sample of 4598, we can report that this sample estimate is within ± 0.3 percentage points (i.e. ± 30% of the sample estimate of 1%) of the true value in the population (of all registrants). These margins of error become more reasonable as the sample proportion becomes greater. For example, say you find 35% deferrals from the sample (Row 12, column 2 in Table 3). Based on the sample size of 4598 and from Table 3 we may say that this sample estimate of 35% is within ± 1.4 percentage points (i.e. ± 4% of the sample estimate of 35%) of the true value in the population (of all registrants). Figures 3 and 4 represent Table 3 graphically.

Table 3: Fixed Sample Size (n=4598) with varying proportions-Resulting precision

	The Estimate				
Sample Size	Sample %	Sample Proportion	d as a % of the estimate	d as percentage points	d as proportion
4598	0.5	0.005	43.2	0.22	0.002
4598	1	0.01	29.9	0.30	0.003
4598	2	0.02	20.8	0.42	0.004
4598	3	0.03	16.8	0.50	0.005
4598	4	0.04	14.4	0.58	0.006
4598	5	0.05	12.8	0.64	0.006
4598	10	0.10	8.8	0.88	0.009
4598	15	0.15	7.0	1.04	0.010
4598	20	0.20	5.8	1.17	0.012
4598	25	0.25	5.1	1.26	0.013
4598	30	0.30	4.5	1.34	0.013
4598	35	0.35	4.0	1.39	0.014
4598	40	0.40	3.6	1.43	0.014
4598	45	0.45	3.2	1.45	0.014
4598	50	0.50	2.9	1.46	0.015

Figure 3: Margin of Error, *d*, vs. Sample Proportion

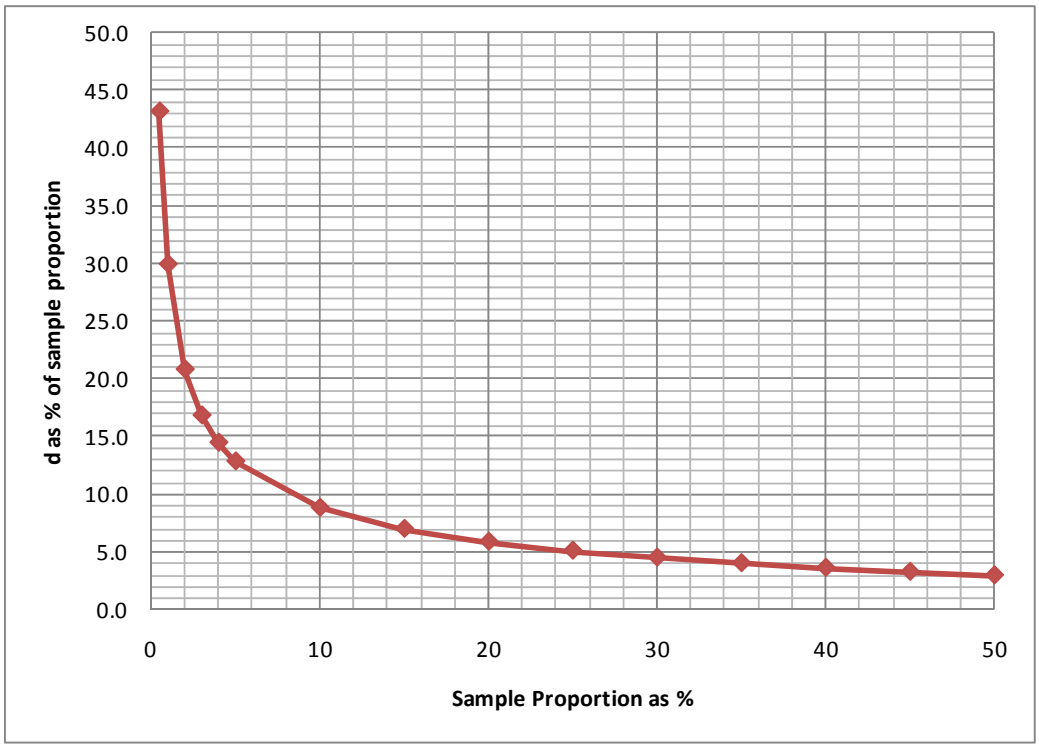
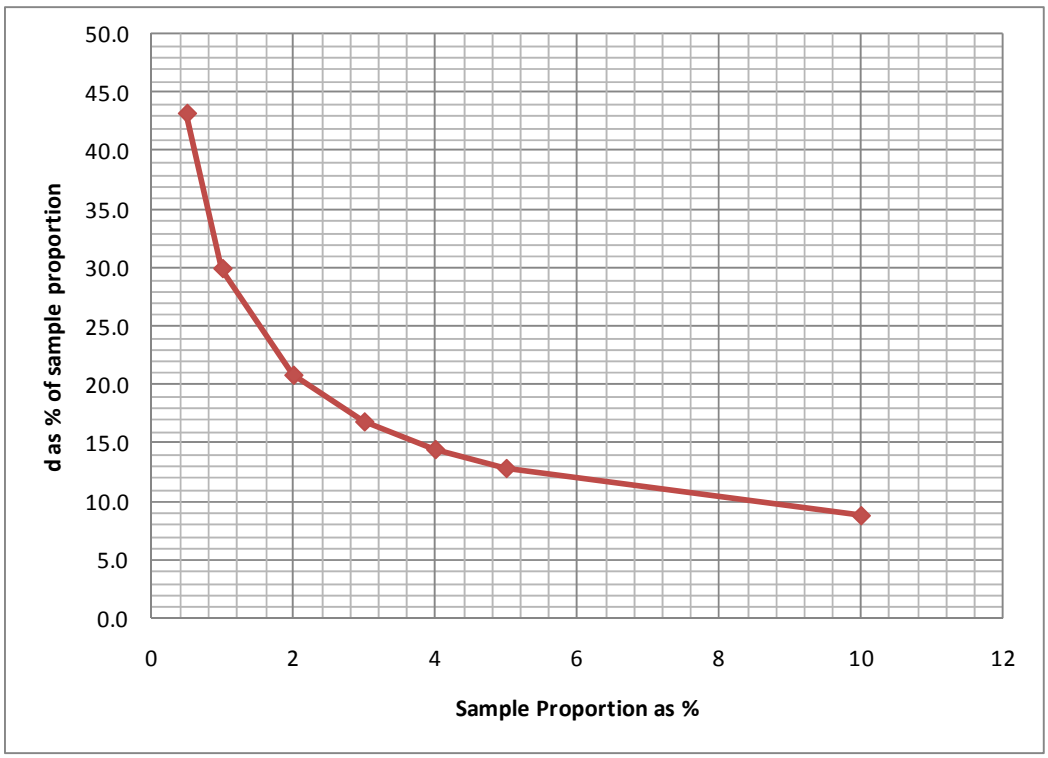


Figure 4: Margin of Error, *d*, vs. Sample Proportion (section from Figure 3)



Effects of small sample sizes on risk of non-detection

In the previous sections, we have restricted the discussion to the overall sample size without regard to the particular types of professions covered by the CPD study. Our understanding is that results will only be reported for the overall sample, and that there will be no breakdown of the results according to the profession concerned.

It must be recognised however, that estimating deferrals, deregistrations, removals etc., within a particular profession, e.g. Dieticians, will have a serious effect on the precision of the estimates concerned. This is an inevitable consequence of the reduced sample size that results for individual professions.

In this section, we look in particular at the effects that small sample sizes would have on the **risk of non-detection with small proportions**. Here there are two parameters of interest; first is the value of the true population proportion to be estimated, second is the sample size. In the discussion below, we restrict the sample sizes to those that would result for each profession if a 2.5% sample was drawn from each.

For each combination of sample size and population proportion (the latter varying from 0.01 to 0.50), we determine the risk that the given sample size will *not detect even a single person who falls into the category of interest*, e.g. the risk that the sample will not yield a single voluntary deregistered person, even if the true population proportion of voluntary deregistered people in the selected profession is (say) 10%. In other words, we are interested in estimating the chance (risk) of non-detection of a particular group (such as voluntary deregistered) in the sample.

Table 4 on page 11, shows the risk (expressed as a percentage) for the range of values that a population proportion can take, when taking a 2.5% sample from each professional body. It will be observed that with samples of size larger than 500, the risk is negligible whatever the value of the proportion being estimated (the first four professions in column 1). For the Paramedic group, the risk is about 2.5% when the proportion being estimated is 1% or lower, but negligible otherwise.

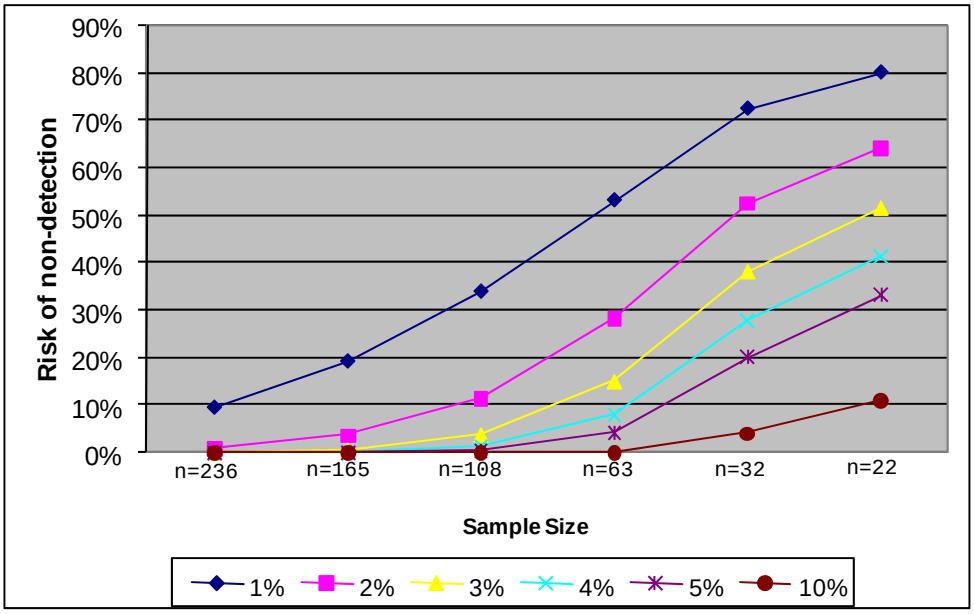
Moving down the table however, it can be seen that the risks increase dramatically and are extremely high for those professions from which small numbers have been sampled, unless the proportion to be estimated is above 20%. Figure 5 on page 12 shows a subset of the results for the last six professions of Table 4 for population proportions less than 10%.

Table 4: Risk of non-detection with a 2.5% sample for varying values of the (unknown) population proportion

Profession	Sample size(n)	Population Proportion													
		0.01	0.02	0.03	0.04	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
Physiotherapist	1057	0.003%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Occupational therapist	747	0.057%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Radiographer	630	0.18%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Biomedical scientist	556	0.38%	0.001%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Paramedic	368	2.5%	0.064%	0.002%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Chiropodist/Podiatrist	313	4.4%	0.191%	0.008%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Speech & Language therapist	301	4.9%	0.243%	0.012%	0.001%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Operating Dept. Practitioner	236	9.4%	0.892%	0.084%	0.008%	0.001%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Dietitian	165	19.2%	3.7%	0.708%	0.136%	0.026%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Clinical scientist	108	34.0%	11.5%	3.9%	1.3%	0.452%	0.002%	0.0%	0.0%	0%	0%	0%	0%	0%	0%
Arts therapist	63	53.3%	28.4%	15.1%	8.0%	4.3%	0.184%	0.008%	0%	0%	0%	0%	0%	0%	0%
Orthoptist	32	72.6%	52.7%	38.3%	27.8%	20.2%	4.1%	0.823%	0.166%	0.034%	0.007%	0.001%	0%	0%	0%

Prosthetist/Orthotist	22	80.3%	64.4%	51.7%	41.5%	33.3%	11.1%	3.7%	1.2%	0.409%	0.136%	0.045%	0.015%	0.005%	0.002%

Figure 5: Risk of non-detection for population proportions varying from 1% to 10% with a 2.5% sample of records



Recommendations and Limitations

It is of course recognised that the sample sizes may need to be driven by the costs involved. When it comes to estimating/detecting small expected proportions at the individual profession level, the above risks could be improved by cutting down the sample sizes in the larger professions and increasing the sample sizes in the smaller professions. We understand from our discussions with HPC that their presenting a sampling scheme in which there is a higher chance of selection in some professions than others would appear to be unfair and as such may not be a viable solution. Hence, the risks of non detection in some professions must be appreciated and accepted as a consequence of the smaller sample sizes i.e. sticking with the proportional sample sizes described above in Table 4. However, as we understand it, HPC wish to report on the register as a whole rather than draw inference from individual professions at this stage.

We have shown that with the current proposed sample size of 4598, proportions of interest that we expect to be less than 5% will be less reliably estimated than larger proportions (e.g. deferrals). If our aim were to be 95% confident that the estimate we get lies within ± **10%** of the true population value, with a sample size of 4598, we could hope to detect those proportions above roughly 8% (between rows 6 and 7 in Table 1 and see Figure 2) at this level of accuracy. Having a more tolerable margin of error for our resulting estimates of ± **20%** (Table 2), we find that a sample size of 4598 would allow detection of expected proportions in our population of a little over 2% (Table 2 row 3) whist maintaining this level of accuracy.

We are able to recommend taking a 2.5%, when HPC are willing to tolerate error margins highlighted in the Section 'Fixed Sample of 4598' when estimating very small proportions i.e. those less than approximately 8%. The audit of the first two professions has produced results with the proportions of interest lying in the region of 10% and above (e.g. accepted, deferred, deregistered) ignoring removals (see below). When this is the case, with a 2.5% sample size, we will obtain reliable population proportion estimates from the sample when reporting on the register as a whole. Such a sample size would result in a maximum margin of error of ± 10% of the true population proportion.

We make the assumption that the key proportions of interest in all professions would follow roughly the same pattern, i.e. we make an assumption that professions do not differ in any way in terms of what we expect from the CPD audit results.

Note that the risk of failing to detect removals, i.e. those that fail the CPD Standards and hence removed from the register, will inevitably be large with our recommendation above since we expect this rate to be very small indeed. This limitation will have to be accepted since the proportion of removals will be so small that it will never be cost effective to increase the sample size to estimate this quantity reliably.

This recommendation outlines a sample size justification with the associated limitations for the initial CPD audit to be carried out by HPC. The sampling scheme may be adapted and modified for future audits based on findings from the initial audit to, for example, gain insight into specific risk areas. Where concern may arise from the initial sample audit regarding those registrants who are removed from the register due to failure of the CPD Standards, a redesigned sampling scheme could be proposed that would aim to obtain reliable estimates on the removal rate for example or a scheme that would allow for more reliability for profession level reporting.

Sample Size Software:
PASS 2008

PASS cites the following References:
Fleiss, J. L., Levin, B., Paik, M.C. 2003. Statistical Methods for Rates and Proportions. Third Edition. John Wiley & Sons. New York.
Newcombe, R. G. 1998. 'Two-Sided Confidence Intervals for the Single Proportion: Comparison of Seven Methods.' Statistics in Medicine, 17, pp. 857-872.

Appendix B

Table 1 CPD audit results to date

CPD status	Number	%of total number selected for CPD
Accepted	806	71.9
Deferred	103	9.19
Deregistered	132	11.8
Removed	6	0.5

Note:

Total chiropodists and operating department practitioners selected for CPD audit was 1120