

Bristol histopathology review: response to FOI request dated 9 November 2009

Request details

1. Full details of the methodology used and the statistical workings to show how the 3,500 sample size was arrived at, including, but not restricted to:

a. how the random samples was chosen - is it only specialist cases, or are GP cases included? Full details of how account was taken of the need to investigate four specialist areas - respiratory, gynaecology, breast, skin.

b. Show by mathematical modelling exactly how 3,500 cases (which forms a tiny proportion of the workload over the period during which alleged errors have arisen 2000 - end 2008 - and possibly to the present) is proved to be sufficient to obtain a reliable assessment of any significant error rate in the four areas of concern - respiratory, breast, gynaecology, skin.

2. All expert statisticians' reports, produced both prior to and after 20th July that validate, comment on and advise on the adequacy of the 3,500 sample size and methodology used to arrive at it.

Response

The selection of 3,500 random selected cases was the result of discussions between the Trust Medical Director and the Royal College of Pathologists (Professional Standards Unit), expert pathologists previously involved in the management of reviews to detect errors, and independent statisticians at the University of Bristol.

The Royal College of Pathologists recently revised its Guidance on Histopathological reviews¹ to classify stages for review as:

- (1) Histopathological Error
- (2) Duty of care potential for patient harm

Error in histopathological diagnosis has a clear definition but studies to measure diagnostic accuracy generate widely divergent results.²

Following the allegation that there was a high error rate in the Bristol Royal Infirmary Histopathology Department, initial discussions were held with the Royal College of Pathologists (RCPATH) Professional Standards Unit. They recommended an audit of a random sample work of done in one of the last three years (2006, 2007, 2008) in order to give results relevant to current clinical practice. After consideration of the number of adult cases for these years, and in discussion with the Medical Director, it was decided to take the samples from year 2007, because this was the most recent year prior to concerns being formalised by the North Bristol NHS Trust Medical Director, but was before the process changes in respiratory pathology agreed between the Trusts in August 2008. Any 2007 cases from the original 26 cases (2002-2008) flagged by North Bristol Trust in the original allegations were excluded as each of them is subject to a separate individual case review.

18,482 pathological specimens referred during 2007 were available for review. The RCPATH Professional Standards Unit initially recommended a random sample of 376 specimens for 95% confidence limits and 643 for 99% confidence limits.

The Medical Director was advised by the RCPATH Professional Standards Unit that there is no national standard for an upper limit on acceptable error rates. Following a review of published literature, he therefore chose a hypothetical potential clinically significant patient harm rate of 1% on which to base the sample size determination.³ Advice was then sought from pathologists who have previously been involved in reviews of this nature, and independent medical statisticians at the University of Bristol.

Specifically, the Medical Director consulted with Dr Paul Ewings, Medical Statistician based in the Regional NHS Research and Development team in Taunton. On his advice, he then consulted with Dr Chris Rogers and Professor Jonathan Sterne who are independent statisticians employed by the University of Bristol. Their advice was that:

- the lower the error rate to be detected, the higher would be the required sample size.
- further, a larger overall sample size was required if it was intended to estimate error rates separately by pathologist or by system (eg, Skin, Gynaecological, Musculo-skeletal, etc).

It should be noted that no written reports pertaining to the Trust's review of histopathology services were commissioned or received from the Royal College of Pathologists or the external experts, other than the table and graph reproduced below.

The final choice of sample size was based on calculations by Professor Jonathan Sterne of the 95% and 99% confidence intervals that would be produced if the error rate were 1%. For a percentage p , the standard error of p (denoted $se(p)$) can be calculated as:

$$se(p) = \sqrt{\frac{p(100 - p)}{n}}$$

where n is the chosen sample size

The 95% confidence interval (CI) can then be calculated as:

$$95\% \text{ CI} = p - 1.96 \times se(p) \text{ to } p + 1.96 \times se(p)$$

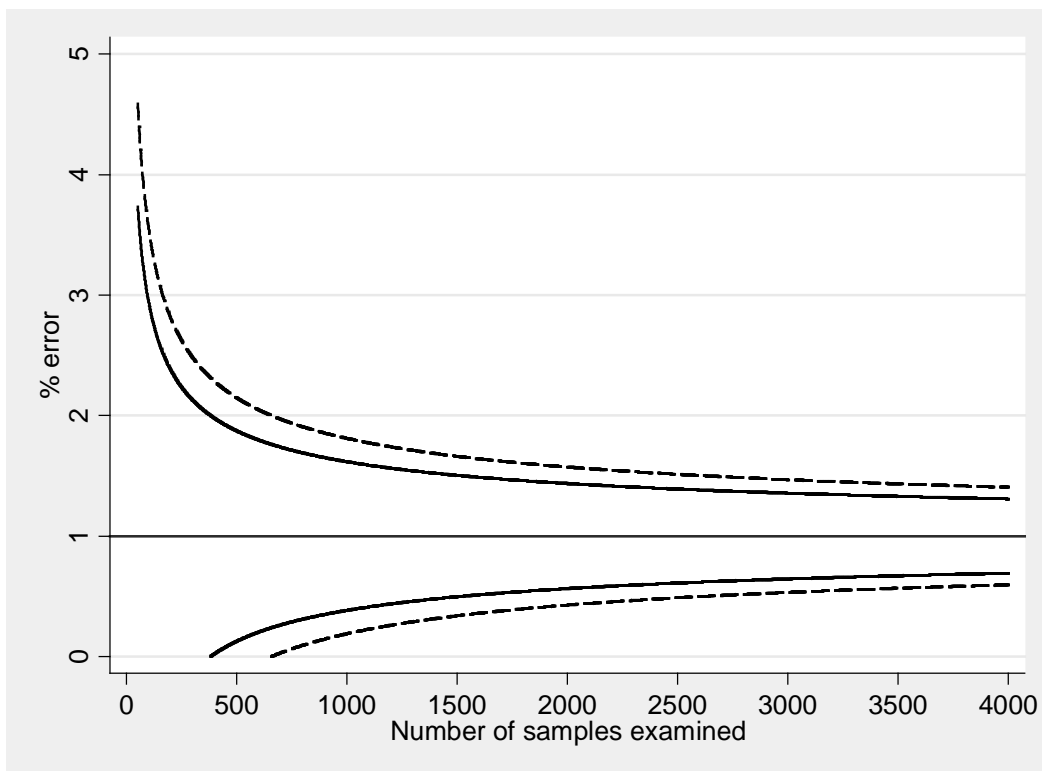
The 99% CI can be calculated in a similar way, as:

$$99\% \text{ CI} = p - 2.57583 \times se(p) \text{ to } p + 2.57583 \times se(p)$$

Based on these formulae, the table below shows the upper limits of the 95% CI and the 99% CI, for different chosen sample sizes.

N (number of samples)	Upper limit of 95% CI	Upper limit of 99% CI
200	2.378957	2.812257
400	1.97507	2.281459
600	1.796141	2.046307
800	1.689478	1.906128
1000	1.616688	1.810466
1200	1.562957	1.739851
1400	1.521197	1.684969
1600	1.487535	1.640729
1800	1.459652	1.604086
2000	1.436064	1.573086
2200	1.415771	1.546416
2400	1.398071	1.523153
2600	1.382454	1.50263
2800	1.368542	1.484346
3000	1.356045	1.467923
3200	1.344739	1.453064
3400	1.334446	1.439537
3600	1.325023	1.427153
3800	1.316354	1.41576
4000	1.308344	1.405233

The same information, together with the corresponding lower limits of the 95% and 99% CIs, is presented in the graph below. It can be seen that with sample sizes below about 500 the confidence intervals are wide, with the lower limits less than the minimum possible value of zero for the percentage error rate.



The final choice of sample size was based on the table and graph above. 3,500 was chosen as it gave 99% upper and lower limits within 0.5% of the estimated error rate, while still allowing error rates to be estimated separately by pathologist or by system. It can also be seen that for sample sizes above 3500 there would be little additional gain in precision (narrowing of the CIs) for the additional time and expense involved, because the CI curves are very flat.

For the chosen sample size, the 95% and 99% confidence intervals around an error rate of $p=1\%$ are as follows:

$$se(p) = \sqrt{\frac{1 \times 99}{3500}} = 0.1682$$

$$95\% \text{ CI} = 1 - 1.96 \times 0.1682 \text{ to } 1 + 1.96 \times 0.1682 = 0.6704 \text{ to } 1.3296$$

$$99\% \text{ CI} = 1 - 2.57583 \times 0.1682 \text{ to } 1 + 2.57583 \times 0.1682 = 0.5668 \text{ to } 1.4332$$

To select the 3,500 specimens for review, details of all 18,482 of the 2007 adult histopathology specimens (both GP and hospital referred) were exported to an excel spreadsheet. Cases with diagnoses referred from elsewhere, and synovial fluid biopsies were removed from the list. Sequentially, every fourth case was selected for the audit. During sampling an imbalance was discovered, with more selected specimens from one pathologist. Their specimens were therefore omitted from later stages of sampling, and sequential sampling continued until all pathologists had contributed approximately 550 cases to the audit. This sampling procedure ensures that the selected sample of specimens for review is representative of practice during 2007.

Please note that the percentage of workload represented by the sample is not relevant to the choice of sample size, because it does not affect the precision of the analysis. Similarly, an opinion poll sample size is chosen based on a target precision, regardless of whether the underlying population size is 100,000 or 100,000,000.

Methodology for the ongoing review was discussed at a meeting on 12 August 2009 between Professor Andy Ness (University of Bristol Department of Oral and Dental Sciences), Professor Jonathan Sterne (University of Bristol Department of Social Medicine) and Dr Jonathan Sheffield (Trust Medical Director). Notes from that meeting, which were circulated and agreed immediately after the meeting, are appended to this document. Two issues discussed at that meeting are relevant to the questions raised in the FOI request.

First, it was agreed that site-specific error rates would be estimated, and that a statistical test for heterogeneity (between-site differences in error rates) would be done. It was agreed that if the p value from this test is <0.1 this will be accepted as evidence of heterogeneity requiring further investigation of error rates separately by site.

Secondly, the Trust's intention to ask an external statistical expert to review and comment on all analyses and the final report was formally agreed. Following the meeting, Professor Sterne approached Mr Ian White, who is a Senior Statistician at the MRC Biostatistics Unit in Cambridge and undertakes research with the London School of Hygiene and Tropical Medicine (see <http://www.mrc-bsu.cam.ac.uk/People/IWhite.html>). Mr White has agreed to act as the external statistical reviewer.

Finally, the purpose of the audit is estimate the rate of serious errors in the department. The results of the review will be published. The Royal College of Pathologists will be invited to comment on the acceptability of the estimated error rate based on their professional judgement.

References

1. Review of the categorisation of discrepancies in histopathology. The Royal College of Pathologists. November 2008.
2. Effectiveness of random and focussed review in detecting surgical pathology error. Rabb SS, Grzybicki DM, Mahood LK, et al. Am J Clin Pathol January 2009.
3. Many cases need to be reviewed to compare performance in surgical pathology. Renshaw A, Young ML, Jiroutek MR. Am J Clin Pathol 2003:119

Jonathan Sterne, Andy Ness, Jonathan Sheffield

7 November 2009

Andy Ness, Jonathan Sheffield, Jonathan Sterne

Notes

- Original concern over 16 errors and further cases added so total is 26 from 2000 to 2008
- Originally just lung but now widened to include all areas
- Diagnostic error is a discrepancy agreed by 2 pathologists after independent review
- No agreed error rate – college says down to clinical opinion
- Liam Donaldson says 1% of people suffer harm from medics so standard is 1%
- College recommended 350 cases for 95% confidence and ~650 cases for 99% confidence
- Statistical advice was for larger survey – 3500 being done across all areas for 2007
- No frozen sections or cytology reports
- To date about 1/3 completed.
- Due to complete review mid to late October
- These will then be reviewed to see if they had a detrimental impact on patient care
- UH Bristol handles 18,000 to 19,000 specimens per year
- College says no point going back more than 3 years
- Also processes for NBT lung pathology changed in 2008
- Existing National audit patchy
- This should be written up for publication – may want to do cost estimates for the paper

Analyses

- Need to calculate confidence interval of the proportion
- As small numbers we should use the SE of the log odds
- We need to compare different tumour groups by testing for heterogeneity of log odds
- If $p < 0.1$ accept evidence of heterogeneity and investigate further and present by site
- Sites are: respiratory, skin, gynaecological, gastrointestinal, head and neck, others
- Fields required are: sample ID, pathologist (number not name), site (see point above), error category (0-5), whether error resulted in problem with patient care, biopsy or resection
- Jonathan Sterne has new statistician who could help with analysis by pathologist

Action points

- Jonathan Sterne will identify external statistician who can confirm the analyses are correct
- Andy to ask Grace to sort meeting of Jonathan Sterne (secretary Claire Snadden), Jonathan Sheffield (PA Claudette Young), Sam Leary and Andy Ness for end of October
- Jonathan Sheffield to ask Phil Hall to liaise with Jonathan Sterne and Andy over preparation of the dataset.

Andy Ness, 12th August 2009