Correspondence

Benefits and limitations of pathology databases to cancer registries

In a recent article, Brewster *et al*[1] presented data on estimated completeness of case ascertainment for 2335 cases of malignancy. Results are given for all subjects and for each of 36 anatomical sites. In each case, the observed percentage completeness is given, together with what purports to be a 95% confidence interval. A glance at these intervals is sufficient to suggest they are so narrow as to constitute wishful thinking.

On closer examination, it is evident that the authors intended to use the simplest method for calculating a confidence interval for a proportion, $p = r/n$, viz. $p \pm z \ (p \ (1 - p)/n)$ where $z = 1.96$ for a 95% interval. In fact they have failed to take the n appearing in the denominator inside the square root, with the result that the intervals are narrower than intended, by a factor n, which can be large.

Moreover, this method is a very poor one,[2 3] in several respects: it gives an inappropriate zero width interval when p is 0 or 1 (as in the article); when the numerator r is very small, the calculated lower limit can be an absurd value less than 0, with corresponding violation of the boundary at I when $n - r$ is very small; and the average coverage of what is claimed to be a 95% interval is only around 88%.

The accompanying table shows confidence intervals calculated by the currently accepted standard method.[4] These intervals may be obtained from Geigy scientific tables for small to moderate n; they are not as yet particularly accessible to researchers, but can be accessed by statistical software including Minitab and SAS indirectly via the inverse of the β integral. An alternative, accepted method that is of closed form and thus readily programmed is $(2np + z^2 \pm z \ (z^2 + 4np \ (1 - p))/(2(n + z^2))$.[5]

Furthermore, it is difficult to appraise whether the heterogeneity of observed completeness rates is any greater than would be expected on the basis of chance variation; it is noteworthy that the lowest rates tend to have small denominators. A hypothesis test would be helpful here—it would necessitate a judicious, anatomically or pathologically meaningful combination of groups into broad categories each containing 100 cases or more.

ROBERT G NEWCOMBE

Department of Medical Computing and Statistics, University of Wales College of Medicine, Heath Park, Cardiff CP4 4EE

1 Brewster DH, Crichton J, Harvey JC, Dawson G, Nairn ER. Benefits and limitations of pathology databases to cancer registries. *J Clin Pathol* 1996;**49**:947–9.
2 Vollset SE. Confidence intervals for a binomial proportion. *Stat Med* 1993;**12**:809–24.
3 Newcombe RG. Two-sided confidence intervals for the single proportion. A comparative evaluation of seven methods. *Stat Med* [in press].
4 Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* 1934;**26**:404–13.
5 Wilson EB. Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc* 1927;**22**:209–12.

Dr Brewster comments:

We are grateful to Robert Newcombe for drawing our attention to the error in our calculation of confidence intervals which was replicated by the use of a spreadsheet and which, in retrospect, we should have identified ourselves. We are also persuaded by his arguments for using other methods of calculating confidence intervals in this situation. Nevertheless, the recalculation of confidence intervals, by whatever method, does not alter the substance or conclusions of our paper.

Histological patchiness and sparing of the rectum in ulcerative colitis: refuting the dogma

We were delighted to see that Levine *et al*[1] have confirmed the findings of our prospective study reported in 1995.[2] Thus, refuted is the conventional wisdom that ulcerative colitis is always diffuse and most severe in the rectum, and that it tapers in severity proximally. Our findings were welcomed by many of our colleagues who knew this from their clinical experience but were often dismayed

Estimated completeness of case ascertainment by the Scottish Cancer Registration Scheme, based on an independent comparison with pathology data, 1992

ICD-9 code	Site	No of cases (x)	No of completely "missed" cases (y)	Estimated % completeness (x/x+y×100)	95% Confidence intervals Lower	95% Confidence intervals Upper
140–149	Lip, oral cavity, and pharynx	39	1	97.5	86.8	99.9
150	Oesophagus	58	0	100.0	93.8	100.0
151	Stomach	89	1	98.9	94.0	100.0
153, 154	Colon, rectum	192	20	90.6	85.8	94.1
155, 156	Liver, gall bladder	30	0	100.0	88.4	100.0
157	Pancreas	57	1	98.3	90.8	100.0
152, 158, 159	Small intestine, retroperitoneum, other digestive sites	14	0	100.0	76.8	100.0
160	Nasal cavities, middle ear, sinuses	1	1	50.0	1.3	98.7
161	Larynx	21	1	95.5	77.2	99.9
162	Trachea, bronchus, lung	360	2	99.4	98.0	99.9
163	Pleura	10	0	100.0	69.2	100.0
170, 171	Bone, connective tissue	9	0	100.0	66.4	100.0
172	Skin (melanoma)	41	0	100.0	91.4	100.0
173 (M-807)	Skin (squamous cell)	103	12	89.6	82.5	94.5
173 (M-809)	Skin (basal cell)	330	46	87.8	84.0	90.9
173 (M-8832)	Skin (dermatofibrosarcoma)	1	1	50.0	1.3	98.7
174, 175	Breast (female/male)	247	13	95.0	91.6	97.3
179	Uterus	2	1	66.7	9.4	99.2
180	Cervix uteri	33	6	84.6	69.5	94.1
182	Corpus uteri	24	2	92.3	74.9	99.1
183	Ovary	46	1	97.9	88.7	99.9
184	Other, unspecified genital organs	11	0	100.0	71.5	100.0
185	Prostate	86	6	93.5	86.3	97.6
186	Testis	11	0	100.0	71.5	100.0
187	Penis	2	0	100.0	15.8	100.0
188	Bladder	115	2	98.3	94.0	99.8
189	Kidney	29	0	100.0	88.1	100.0
190	Eye	2	0	100.0	15.8	100.0
191, 192	Brain, central nervous system	30	0	100.0	88.4	100.0
193	Thyroid gland	6	1	85.7	42.1	99.6
194	Endocrine glands	1	0	100.0	2.5	100.0
196–199	Secondary sites	83	0	100.0	95.7	100.0
200, 202	Non-Hodgkin's lymphoma	53	8	86.9	75.8	94.2
201	Hodgkin's disease	8	1	88.9	51.8	99.7
203	Multiple myeloma	16	2	88.9	65.3	98.6
204–208	Leukaemia	42	4	91.3	79.2	97.6
140–208	All malignant sites	2202	133	94.3	93.3	95.2

no explanatory text would have difficulty in understanding it. Equally, I found a completely new system of numbering diseases based on the chapter number unnecessarily confusing. Metabolic disease has its own jargon and abbreviations, more are not required.

It is important that metabolic disease is considered as part of differential diagnosis but I doubt whether its investigation could be effectively done using this book alone. Surely the laboratory has a role to play. The tests are usually performed in specialist centres that not only do the analysis but offer interpretation and can put the general clinician in touch with the appropriate expert.

This book is not for reading and it fails in my opinion to reach its target audience. Perhaps it would have been better to drop the first word of the title. Nevertheless, it contains a lot of useful information much of which is not easily found in other text books or sources. It could sit comfortably on a shelf next to *The Metabolic and Molecular Bases of Inherited Disease* in a library or in the specialist paediatricians office, and even the general laboratory.

G M ADDISON

Greenfield's Neuropathology. 6th edn. Graham D, Lantos P, eds. (2 volumes; £345.00.) Arnold. 1996. ISBN 0 3405 9809 3.

You don't need to get as far as page 1 to realise that the sixth edition of *Greenfield* is the result of a major overhaul, and is essentially a new book. Gone is the short fat, unwieldy volume that very quickly got separated from its covers; the book is now in two manageable volumes in larger format, each around 1000 pages long. The silver print on the outside announces completely new chapters: "schizophrenia", "prion diseases", "movement disorders", and, most notably, "tumours" (previously excluded because they were dealt with in the companion volume by

Russell and Rubinstein, from the same publishers). Inside, the list of contributors is impressive: well over half the 64 names are those of overseas authors, many of them leading names in their field, many of them clinicians. About a third of the chapters are completely new; all have been extensively rewritten.

So how can one do justice to such a book in 500 words? In brief: it is brilliant. No longer merely an anthology of macroscopic and microscopic morbid anatomy, the book aims to set classic descriptive neuropathology within the framework of molecular and cellular events. As a result, this edition has a much wider appeal than earlier ones. Neurosurgeons and neurologists will find it an invaluable reference work, while for neuroradiologists—whose imaging techniques have made them into neuropathologists, whether they like it or not—it should be required reading. I found the most successful chapters to be those that integrated normal anatomy, physiology, and cellular pathology with clinical information and neuropathology. Special mention must go to the masterly discussion of modern concepts of rising intracranial pressure, brain swelling and oedema in the chapter by Miller and Ironside; the timely update on the pathophysiology of traumatic brain injury by Graham and Gennarelli; the chapter on hypoxia, which includes an excellent account of events in hypoglycaemia, by Auer and Benveniste; and the clear and succinct discussion of malformations by Harding and Copp.

However, such selection is invidious. Almost all the contributions are of exceptionally high quality, and the information is as up to date as is possible in a large textbook such as this. What reservations do I have? A few, centring on the wisdom of including tumours. There are already a number of excellent texts on the subject, and the long chapter by Lantos and colleagues falls short of being useful as a reference work for surgical neuropathology.

However, where the authors are not dealing with macroscopic and microscopic pathology of the tumours, as in the sections on molecular biology and cytogenetics, and in their long discussions of oligodendrogliomas and meningiomas, they give us much valuable information unobtainable elsewhere.

In summary, I really cannot fault this new edition. It has regained its place as the leading textbook of neuropathology, and should be in the libraries of the histopathology and neurosciences departments of every teaching hospital. It is unfortunate that the huge pricetag will prevent it from being part of many personal collections.

J F GEDDES

Correction

Square root symbols were inadvertently omitted from three places in this letter. The second paragraph should have read:

On closer examination, it is evident that the authors intended to use the simplest method for calculating a confidence interval for a proportion, $p = r/n$, viz.

$p \pm z \sqrt{(p (1 - p)/n)}$ where $z = 1.96$ for a 95% interval. In fact they have failed to take the n appearing in the denominator inside the square root, with the result that the intervals are narrower than intended, by a factor $\sqrt{n}$, which can be large.

The equation in the fourth paragraph should have read:

$$(2np + z^2 \pm z \sqrt{(z^2 + 4np (1 - p))/(2(n + z^2))}$$

These errors are regretted.