From

Evolution of Epidemiologic Ideas: Annotated Readings on Concepts and Methods
Sander Greenland (Editor)
Epidemiology Resources; (July 1987)

II–4.

Barnet Woolf: On Estimating the Relation Between Blood Group and Disease. Annals of Human Genetics 1955; 19:251–253.

Woolf's "On Estimating the Relation Between Blood Group and Disease" is most often cited for its introduction of what was probably the first common odds-ratio estimator for a set of two-by-two tables (or studies), and a test for heterogeneity of the odds ratio across the tables; these are now known as Woolf's estimate and Woolf's test of homogeneity (see, e.g., Schlesselman [1982]). Though these statistics have been supplanted by Mantel-Haenszel and maximum-likelihood methods, the paper contains more than just bare mechanics, as it touches on a number of important points which were generally unappreciated in the early literature.

Woolf was motivated by concern about the then-prevalent practice of analyzing and comparing case-control results by computing the difference in exposure proportions between case and control groups. As he notes, this difference would vary with the background frequency of exposure, and artifactual differences between study results could be expected. Instead he chose to work with the ratio of incidence rates, and recognized (independently of Cornfield) that this ratio could be estimated from case-control data, even if the absolute rates could not. (He did not, however, delineate sufficient conditions for the sample odds ratio to estimate the incidence ratio.) He also noted some subtle points that many later users of his methods overlooked, one being that his variance formula for the common odds ratio applied only under the homogeneity assumption.

Reference:
Schlesselman JJ. *Case-control Studies: Design, Conduct, Analysis*. New York: Oxford University Press, 1982.

ON ESTIMATING THE RELATION BETWEEN BLOOD GROUP AND DISEASE

By BARNET WOOLF

Department of Animal Genetics, University of Edinburgh

Following the demonstration of a significant excess of blood group A in patients with cancer of the stomach (Aird, Bentall & Roberts, 1953) and of group O in sufferers from peptic ulcer (Aird, Bentall, Mehigan & Roberts, 1954) and from toxaemia of pregnancy (Pike & Dickins, 1954) it seems certain that many more studies will be made on the relation between blood groups and disease. It is therefore important that the best possible statistical methods should be used. The procedure recommended by Aird *et al.* (1954) is very efficient, but it is open to criticism on one rather important point. These workers take as criterion the difference in proportion of a given blood group in the disease and the control series. Denote the two blood types α and β. Suppose the disease series contains h patients of type α and k of type β, where $h+k=n$, and the control series has H of type α and K of type β, where $H+K=N$. Aird and associates calculate $d=h/n-H/N$. This is tested for significance against its sampling variance, combined with estimates from other bodies of data to give a weighted mean estimate, and compared with these other estimates in tests for heterogeneity.

Unfortunately, d will differ from one community to another even when the specific attack rate within any given blood group stays constant. This can be shown by a simple example. Consider a community of 10,000 people in which H and K are each 5000. Then if $h=100$ and $k=50$, $d=100/150-0{\cdot}5$, or $0{\cdot}1667$. Now consider another community in which H is 9000 and K is 1000. In this case $h=180$ and $k=10$, so $d=180/190-0{\cdot}9$, or $0{\cdot}0474$. Even when the essential biological conditions are identical, differences in blood-group frequencies in the population will introduce spurious heterogeneity. This kind of artefact is avoided if one works with incidence rates in the various blood groups. The data usually do not permit calculation of absolute rates, nor are they needed. What is wanted and readily obtained is an estimate of the ratio of one rate to another. The incidence in group α will be $h/H \times$ some constant, and that in group β will be $k/K \times$ the same constant. If the ratio is taken as x to 1, an estimate of x will be hK/Hk, and it may readily be shown that this is the maximum-likelihood estimate. The use of x is recommended instead of d as a criterion of differential incidence of disease in relation to blood group.

In all statistical computations it is best to transform x into its logarithm. This avoids difficulties due to asymmetry. If comparison of α with β gives $x=2$ say, comparison of β with α will give $x=\frac{1}{2}$; but $\log x$ will retain its numerical value, merely changing in sign. Moreover, the sampling variance of $\log x$ is a very simple expression free of 'nuisance parameters'. This is especially true if one transforms into $y=\log_e x$. If V is the sampling variance of y, then

$$V=1/h+1/k+1/H+1/K,$$

and w, the weight of y, is of course $1/V$. If the attack rate is the same for both blood types the expected value of y will be 0, so the null hypothesis is not to be rejected unless y differs significantly from zero. This is tested by χ^2 will be y^2/V or wy^2 for one degree of freedom. Combination of data from different communities proceeds as described by Aird *et al.* (1954). The weighted mean, Y,

is $\Sigma wy/\Sigma w$, and its antilogarithm, X, is taken as the combined estimate of x. Significance of Y is tested by $\chi^2=(\Sigma wy)^2/\Sigma w$ or $Y^2\Sigma w$ for one degree of freedom. Heterogeneity is tested by $\chi^2=\Sigma wy^2-(\Sigma wy)^2/\Sigma w$ or $\Sigma wy^2-Y^2\Sigma w$ with degrees of freedom one less than the number of sets of data combined. The standard deviation of y is $V^{\frac{1}{2}}$, and the approximate fiducial limits at the 95% point are $y\pm 1{\cdot}96V^{\frac{1}{2}}$. Provided there is no significant heterogeneity the standard deviation of Y is $1/(\Sigma w)^{\frac{1}{2}}$ and the 95% fiducial limits are approximately $Y\pm 1{\cdot}96/(\Sigma w)^{\frac{1}{2}}$. By taking antilogarithms these can be transformed into fiducial limits for x or X. This is a 'large-sample' treatment, and the formulae cease to be applicable if any of the observed frequencies is small.

Table 1. *Calculation of combined estimate of incidence ratio of peptic ulcer in groups O and A*

City	Peptic ulcer		Control		$x=\frac{hK}{Hk}$	$y=\log_e x$	$w=\frac{1}{\frac{1}{h}+\frac{1}{k}+\frac{1}{H}+\frac{1}{K}}$	$wy^2=\chi^2$
	Group O (h)	Group A (k)	Group O (H)	Group A (K)				
London	911	579	4578	4219	1·4500	0·3716	304·9	42·11
Manchester	361	246	4532	3775	1·2224	0·2008	136·6	5·50
Newcastle	396	219	6598	5261	1·4418	0·3659	134·5	18·01
					$\Sigma wy=189{\cdot}94$		576·0	65·62

$Y=\Sigma wy/\Sigma w=0{\cdot}3289$.
$Y^2\Sigma w=62{\cdot}63$.
S.D. of $Y=(\Sigma w)^{-\frac{1}{2}}=0{\cdot}0417$.
95% fiducial limits of $Y=0{\cdot}2472$–$0{\cdot}4106$.
$X=$ antilog $Y=1{\cdot}39$.
95% fiducial limits of $X=1{\cdot}28$–$1{\cdot}51$.

χ^2 analysis

	D.F.	
Y	1	62·63
Heterogeneity	2	2·99
Total	3	65·62

Table 2. *Incidence ratios of some diseases in relation to blood group*

Disease	Comparison	X or x	95% fiducial limits	Reference
Cancer of stomach	Group A with group O	1·22	1·12–1·32	Aird *et al.* (1953)
Peptic ulcer	Group O with group A	1·39	1·28–1·51	Aird *et al.* (1954)
Toxaemia of pregnancy	Group O with all others	1·38	1·15–1·66	Pike & Dickins (1954)

Table 1 shows the calculations for comparing incidence of peptic ulcer in groups O and A using combined data from London, Manchester and Newcastle. This is the example worked out by their method by Aird *et al.* (1954, Table VII). The χ^2 values, 62·63 for significance and 2·99 for heterogeneity, agree closely with 66·21 and 3·01 found by the d method. Similarly, combined data from six centres on cancer of the stomach in groups O and A gave χ^2 values by the x method of 21·28 (1 D.F.) for significance and 2·63 (5 D.F.) for heterogeneity, against 21·49 and 2·69 by the d method. These close concordances are to be expected, since the observations all come from cities in England with very similar population blood-group frequencies. If data from different ethnic groups were combined, the d method would in general be expected to return unduly high χ^2 figures for heterogeneity. This might be of some biological or medical importance as

tending to be confounded with possible genuine heterogeneity arising either from environmental factors or from differential attack rates for the diverse genotypes that may go to make up a single blood group. Even when heterogeneity is not an issue, x is preferable to d because it has a direct medical meaning. Table 2 gives some estimated incidence ratios of diseases in relation to blood group, together with fiducial limits. A blood-group difference appears able to increase the risk of disease by as much as 39 %.

REFERENCES

AIRD, I., BENTALL, H. H., MEHIGAN, J. A. & ROBERTS, J. A. F. (1954). The blood groups in relation to peptic ulceration and carcinoma of colon, rectum, breast and bronchus. *Brit. Med. J.* **2**, 315.

AIRD, I., BENTALL, H. H. & ROBERTS, J. A. F. (1953). The relationship beteeen cancer of stomach and the *ABO* blood groups. *Brit. Med. J.* **1**, 799.

PIKE, L. A. & DICKINS, A. M. (1954). *ABO* blood groups and toxaemia of pregnancy. *Brit. Med. J.* **2**, 321.