

Continuing controversies over “risks and rates” – more than a century after William Farr’s “On prognosis”

Dr. Vandembroucke is professor of clinical epidemiology at the Department of Clinical Epidemiology of the Leiden University Medical Centre

More than a century after William Farr’s 1838 publication “On prognosis”, the difference between “risk and rate” was rediscovered by epidemiologists in the 1970s. The concept of the incidence rate over person-time continues to be misunderstood and has led to recent controversies.

In the second part of his 1838 publication “On prognosis”, William Farr explained the “force of mortality”, which nowadays we call an “incidence rate”. He distinguished from “mortality”, which we call “risk” or “cumulative incidence”. The first is calculated over person-time and ranges from 0 to infinity; the second is a number between 0 and 1 without dimensions (Rothman & Greenland 1998). In this commentary, I will trace the recent history of the distinction between “risk and rate”, and recount how these concepts still lead to confusion and controversy.

A rediscovery

The distinction between “risk” and “rate” that was so well known to Farr was rediscovered in the 1970s in the USA (Vandembroucke 1985). In itself this is strange, since incidence rates had been used earlier in the 20th century in the UK and the USA. In the 1950s Richard Doll and Austin Bradford Hill used incidence per person-years in their studies of British doctors and smoking (Doll & Hill 1956). However, they did not publish about the theory of this calculation. Hill did explain risks and life tables in his influential textbook on medical statistics (Hill 1937), but only from the 7th edition onwards a few lines were devoted to incidence rates (Hill 1984). In the USA the 1960 first edition of MacMahon’s textbook on epidemiology used Doll and Hill’s calculations on smoking and lung cancer to explain the person-years denominator (MacMahon et al. 1960). The example was retained in all later editions of that textbook.

Nevertheless, the conceptual difference between “risk” and “rate” and their respective uses escaped attention.

Like Gerstman (2003), I think that Elandt-Johnson (1975) should be credited for bringing the topic very clearly to the attention of “modern” epidemiologists in 1975. The distinction became the basis for Miettinen’s (1976) proposal that odds ratio calculations in case-control studies could be exactly the same as a ratio of incidence rates, without need for the “rare disease assumption”. Thereafter, the distinction gained wide acceptance (Greenland 1987).

Continuing debates

Many persons remain confused about events per person-time. This was witnessed by a paper in the British Medical Journal in 1995, which called person-years “A dubious concept” (Windeler & Lange 1995). The authors stated that incidence rates are meaningless because they are not probabilities. They described that person-years allow to use multiple events over time, which they thought was wrong. Actually, this is an advantage, and incidence rates are used for that purpose in infectious diseases. Also, they thought that the calculation is used to hide loss to follow-up – which in fact gives the same problem of bias, whatever incidence measure is used. Finally, they repeated the often-heard argument that incidence rates can not distinguish between a few persons followed for a long time vs many persons followed for a short time. On the contrary, Farr demonstrated very beautifully in “On prognosis” how the time-dependency of incidence rates can be traced by dividing time into small periods (see the “Table of the deaths and recoveries out of 1000 patients – this number being constantly kept up – in fourteen stages of Small Pox”). Farr showed more than a century ago that it is perfectly possible to allow for the

duration of follow-up and varying incidences over time with the person-years method.

Another strange episode on “risks and rates” was a critique on William Farr and Florence Nightingale in the Annals of Internal Medicine in 1996, entitled “100 apples divided by 15 red herrings” (Iezzoni 1996). The author criticised publications on death rates in hospitals by Farr and Nightingale in the 1860s. The author rehashed the 130 year old accusation that Farr and Nightingale had used a “wrong” rate to overstate their political message about differences in death rates between hospitals. Farr and Nightingale had published death rates up to “90 per 100 per year” for particular hospitals for which they envisaged reforms. Interested parties at the time thought that these rates were impossible and thus wrong. Still, the explanation is straightforward: in hospitals several persons in succession occupy a single bed over a year, which yields 1 person-year of observation.

Several persons might have died in that bed, which might even lead to death rates larger than unity. A simple solution to restore “intuitive credibility” is to calculate death rates per person-month or person-week: a death rate of 90 per 100 patient-years becomes 1.73 per 100 patient-weeks. The latter number would not as quickly have drawn accusations of “political spin”, although it remains exactly the same incidence. Like the 19th century critics of Farr and Nightingale, the 1996 author concluded that these high mortality rates were used because of Nightingale’s political agenda. Vandenbroucke-Grauls and myself took the defence of Farr and Nightingale (Vandenbroucke & Vandenbroucke-Grauls 1996, 1997).

Incidence calculations per person-week or per person-month are used nowadays for nosocomial infection rates in hospi-

tals. This also took some time. Up to the early 1980s the Centers for Disease Control (CDC) in the USA propagated a so-called “rate” for the calculation of hospital infections: the number of infected divided by number of admissions (Haley et al. 1981). That calculation was already frowned upon in 1947 by Major Greenwood, at the London School of Epidemiology and Tropical Medicine (Vandenbroucke & Vandenbroucke-Grauls 1988). Dividing the number of infected by number of admissions amounts to dividing the number of deaths in a population by the number of newborns over the same time period – which is neither an incidence rate, nor a risk. Even leading institutions seemed unaware that Farr and Nightingale had shown the right way of calculating incidence rates in hospitals in the middle of the 19th century (Vandenbroucke & Vandenbroucke-Grauls 1988). This has changed (Freeman & McGowan 1978), and in the early 1990s the CDC used nosocomial infection rates per person-time (Gaynes et al. 1991).

The difference between risk and rate gave raise to confusion and controversy, more than a century ago, and it still does. To those who have difficulty with the use of rates, the beautiful examples and explanations in “On prognosis” are still worthwhile reading.

Acknowledgement

I am indebted to Sir Iain Chalmers, Oxford, for the use of his collection of the several editions of Sir Austin Bradford Hill’s “Textbook on medical statistics”, and to Professor Christina Vandenbroucke-Grauls, Amsterdam, for guiding me through the nosocomial infection literature.

Jan P. Vandenbroucke

References

- Doll R, Hill AB* (1956). Lung cancer and other causes of death in relation to smoking: a second report on the mortality of British doctors. *BMJ* 2: 1071–81.
- Elandt-Johnson RC* (1975). Definition of rates: some remarks on their use and misuse. *Am J Epidemiol* 102: 267–71.
- Freeman J, McGowan JE Jr* (1978). Risk factors for nosocomial infection. *J Infect Dis* 138: 811–19.
- Gaynes RP, Martone WJ, Culver DH, et al.* (1991). Comparison of rates of nosocomial infections in neonatal intensive care units in the United States. National Nosocomial Infections Surveillance System. *Am J Med* 91: 192S–6S.
- Gerstman BB* (2003). Comments regarding On Prognosis by William Farr (1838), with reconstruction of his longitudinal analysis of smallpox recovery and death rates. *Soz Präventiv Med* 48 (5): in press.
- Greenland S, ed.* (1987). Evolution of epidemiologic ideas: annotated readings on concepts and methods. Chestnut Hill, MA: Epidemiology Resources Inc.: 180.
- Haley RW, Hooton TM, Culver DH et al.* (1981). Nosocomial infections in U.S. hospitals, 1975–1976: estimated frequency by selected characteristics of patients. *Am J Med* 70: 947–59.
- Hill AB* (1937). Principles of medical statistics. London: The Lancet. (Postgraduate series; vol. 3).
- Hill AB* (1984). A short textbook on medical statistics. 11th ed. London: Hodder & Staughton.
- Iezzoni LI* (1996). 100 apples divided by 15 red herrings: a cautionary tale from the mid-19th century on comparing hospital mortality rates. *Ann Intern Med* 124: 1079–85.
- MacMahon B, Pugh TF, Ipsen J* (1960). Epidemiologic methods. Boston, MA: Little, Brown & Co.: table 33.
- Miettinen OS* (1976). Estimability and estimation in case-control studies. *Am J Epidemiol* 103: 226–35.
- Rothman KJ, Greenland S* (1998). Modern Epidemiology. 2nd ed. Philadelphia, PA: Lippincott-Raven: 29–46.
- Vandenbroucke JP* (1985). On the rediscovery of a distinction. *Am J Epidemiol* 121: 627–8.
- Vandenbroucke JP, Vandenbroucke-Grauls CMJE* (1988). A note on the history of the calculation of hospital statistics. *Am J Epidemiol* 127: 699–702.
- Vandenbroucke JP, Vandenbroucke-Grauls CMJE* (1996). In defense of Farr and Nightingale. *Ann Intern Med* 125: 1014.
- Vandenbroucke JP, Vandenbroucke-Grauls CMJE* (1997). A return to Farr and Nightingale. *Ann Intern Med* 127: 170–1.
- Windeler J, Lange S* (1995). Events per person year – a dubious concept. *BMJ* 310: 454–6. (<http://bmj.com/cgi/content/full/310/6977/454>). (accessed May 2003).

Address for correspondence

Jan P. Vandenbroucke, MD PhD
Department of Clinical Epidemiology
Leiden University Medical Centre, C9-P
NL-2300 RC Leiden
Tel.: +31-71-5265230
Fax: +31-71-5266994
e-mail: J.P.Vandenbroucke@lumc.nl



To access this journal online:
<http://www.birkhauser.ch>