Epidemiologic research with administrative databases: red herrings, false alarms and pseudo-epidemics

David A. Grimes*

Department of Obstetrics and Gynecology, University of North Carolina School of Medicine, Chapel Hill, NC, USA

*Correspondence address. Tel: +1-919-656-7227; E-mail: david_grimes@med.unc.edu

Submitted on May 13, 2015; resubmitted on May 13, 2015; accepted on June 1, 2015

In 1994, Altman (1994) startled the medical world by declaring that poor research is as unethical as fraudulent research. His reasoning was simple: both kinds of research mislead which, in turn, harms patients. A decade later, others (von Elm and Egger, 2004) found that the problem had persisted. With the growing indiscriminate use of administrative databases for epidemiologic research (Grimes, 2010) the problem has only gotten worse since then. The record-linkage study of assisted reproductive technologies (ART) and cancer in this issue (Reigstad et al., 2015a) illustrates the problem.

Fuzzy exposure

In any cohort study, the exposure must be clear, specific and measurable (Grimes and Schulz, 2002). In the report by Reigstad and colleagues (Reigstad et al., 2015a), the exposure was a disparate group of ill-described techniques to promote fertility: the administrative database had scant information about them. The regimens used, drugs administered, number of treatment cycles, and combinations of treatments were unknown or unclear. Aggregating these heterogeneous exposures as ‘ART’ is analogous to combining all methods of fertility control as ‘contraception’ then studying cancer as an outcome. Any result is incapable of interpretation.

Inadequate control for potential confounding factors

In these Norwegian databases (Reigstad et al., 2015a) information about potential confounding factors was inadequate. The databases were not designed for epidemiologic research. The only factors for which adjustments could be made were ‘age at start of follow-up, parity, region of residence and calendar year at follow-up’ (Reigstad et al., 2015a). Attempting to study the epidemiology of specific cancers without considering known risk factors is untenable.

Many risk factors for cancer are known. Cigarette smoking (Winkelstein, 1990) and sexual behavior (Wakeham and Kavanagh, 2014) are risk factors for cervical cancer. The report from Reigstad et al. (2015a) had no information about smoking, number of sexual partners or use of barrier contraceptives (Lam et al., 2014). Obesity is strongly related to endometrial cancer (Nevadunsky et al., 2014); Reigstad et al. (2015a) reported no information on weight, body mass index (BMI) or other measures of adiposity. Oral contraceptive use is associated with a lower risk of ovarian cancer (Ness et al., 2000) but Reigstad et al. (2015a) had no information on this common exposure. Trying to study the epidemiology of cancers with these Norwegian databases is analogous to trying to study the epidemiology of automobile accidents with a Department of Motor Vehicles database. Age, height, weight, and eye and hair color cannot explain why accidents occur (Grimes, 2010).

Quick-and-dirty research

Research using existing administrative databases has obvious appeal. Instead of doing a proper (and thus expensive) study of a specific a priori hypothesis with a written protocol, epidemiologists pore over huge data sets in search of something statistically significant, a process derisively termed ‘risk-factorology.’ (Smith, 2001). The more comparisons are made, the greater is the likelihood that something will pop up as statistically significant (even when no association exists). Indeed, after adjustment for the scores and scores of comparisons...
made (Reigstad et al., 2015a) statistical significance vanished. No safety signals remained (Silver, 2012).

Journals are enamored of large administrative databases as well. Huge numbers of patients produce narrow confidence intervals and tiny $P$ values. However, because these databases were not designed for epidemiologic research, the results are often wrong: garbage in, garbage out (Grimes, 2010). Large numbers increase the precision of results but do not address the more important issue of validity. A big database study can give an invalid answer with a narrow confidence interval. Stated alternatively, the result can be precisely wrong (Lidegaard et al., 2012a,b). That is dangerous.

**Weak associations**

The need for circumspection is critical with weak associations. In Hill’s classic paper on judging causal associations (Hill, 1965), the first consideration was strength of the association. In a cohort study, strength is measured by the relative risk. Hill gave two examples: cigarette smoking and death from lung cancer and polluted water and cholera. The smallest relative risks in these two examples were 8 and 14, respectively (Hill, 1965). Relative risks of that size are likely due to causation, not bias.

The lower limit of discrimination for cohort studies is widely considered to be a relative risk of about 2, or its reciprocal, 0.5 (Fig. 1) (Grimes and Schulz, 2012). Relative risks between 0.5 and 2 are likely due to bias, which is present in all observational studies. Values above 2 or below 0.5 (Fig. 1) may merit attention (Grimes and Schulz, 2012). In the report from Reigstad et al. (2015a) the adjusted hazard ratio was 1.16 overall and 1.50 for tumors of the central nervous system. Hazard ratios (interpreted like relative risks) of this size are more likely due to bias than to causation (Shapiro, 2000). Only a properly done randomized controlled trial free of bias could discern effect sizes this small (Schulz and Grimes, 2006).

**Pseudo-epidemics in administrative databases**

Other Scandinavian studies linking administrative databases have recently caused false alarms. Several reports on medical abortion have used the Finnish Abortion Registry (Ninimaki et al., 2009, 2011). The reported adverse event rates were nonsensical: 20 to 100 times higher than rates reported by others (Grimes and Raymond, 2011). When challenged in a letter to the editor (Fjerstad et al., 2010), the Finnish authors acknowledged that, ‘many of the “complications” are not really such, but rather concerns…that bring women back to the health care system.’ The reports had counted office visits by the ‘worried well’ as adverse events, which is misleading.

In Denmark, studies of venous thromboembolism using the National Registry of Patients have produced red herrings as well (Grimes, 2010). For example, reports in high-profile general medical journals claimed that the levonorgestrel intrauterine system was associated with statistically significant protection against venous thrombosis (Lidegaard et al., 2012b) and thrombotic stroke (Lidegaard et al., 2012a). While this intrauterine device (IUD) has several established non-contraceptive health benefits (Fraser, 2013), prevention of clots and prevention of stroke are not among them. Without validation of outcomes (Severinsen et al., 2010; U.S. Food and Drug Administration) and adequate control for potential confounding (Dinger and Shapiro, 2012), these database studies (Lidegaard et al., 2012a,b) are not credible.

**Epidemiologic hubris**

Epidemiologists need to approach their work with greater humility and circumspection (Boffetta et al., 2008). Few seem to understand that most published research findings are false (Ioannidis, 2005). Today, any result coming from an observational study is most likely wrong (Young and Karr, 2011). Among the reported associations that are real, most are exaggerated (Ioannidis, 2008). Few reports, including Reigstad et al. (2015a) adequately explain their weaknesses (Shapiro, 2008). Sadly, these flawed manuscripts usually get published somewhere (Smith, 2006; Ioannidis et al., 2010). As noted by Sir Iain Chalmers, a great deal of rubbish is published in medical journals these days (Chalmers, 2007).

**Damage from false alarms**

Modern epidemiology has a checkered history (Taubes, 1995). Among its many false alarms have been purported links between cigarettes and suicide, reserpine and breast cancer, menopausal estrogen and reduced cardiovascular disease, coffee and pancreatic cancer, beta-carotene and reduced lung cancer, IUD use and infertility, and oral contraceptives and pituitary adenoma (Grimes and Schulz, 2012). The claim that ART is related to an increased risk of brain tumors (Reigstad et al., 2015a) and breast cancer (Reigstad et al., 2015b) can be added to the list.

Infertile couples face many challenges: social, emotional and financial (Cousineau and Domar, 2007; Chambers et al., 2013). Compounding their suffering by frightening them about brain tumors (Reigstad et al., 2015a) and breast cancer (Reigstad et al., 2015b) is cruel. Newspapers love sensational medical claims, no matter how implausible (Russell, 1999). These media scares hurt women (Jones et al., 1980). For example, the 1995 media blitz about the alleged (and subsequently refuted) hazards of certain oral contraceptives (Dinger et al., 2007, 2014) was temporally associated with a preventable epidemic of unplanned pregnancies and abortions (Skjeldestad, 1997).

**Research priorities**

A finite amount of money and resources can be spent on women’s health research. Hence, no research project can be considered in a vacuum. One must always ask what benefits would have accrued had the same
amount of money and resources been devoted to other health problems. If the other problems have greater importance, then the research in question has a net negative effect on public health. Linking and dredging administrative databases in search of a problem (rather than addressing a real one) wastes resources and is thus unethical (American College of Obstetricians and Gynecologists, 2004).

The way ahead

Medicine today needs less research, better research and research done for the right reasons (Altman, 1994). The incentives, however, are still wrong; researchers are rewarded for the quantity and not the quality of their publications. Medicine also needs ‘better training and more circumspection on the part of investigators, tougher editorial standards on the part of journals, and hefty skepticism on the part of referees and readers’ (Grimes and Schulz, 2012). As a Harvard epidemiologist ruefully noted, ‘We [epidemiologists] are fast becoming a nuisance to society. People don’t take us seriously anymore, and when they do take us seriously, we may unintentionally do more harm than good.’ (Taubes, 1995).

Author’s roles

D.A.G. wrote, reviewed, critically revised and approved the final manuscript.

Funding

No external funding was either sought or obtained for this study.

Conflict of interest

D.A.G. serves on several Scientific Monitoring and Advisory Committees for studies sponsored by Bayer Pharma AG.

References


Fraser IS. Added health benefits of the levonorgestrel contraceptive intrauterine system and other hormonal contraceptive delivery systems. Contraception 2013;87:273–279.
Skjeldstad FE. Increased number of induced abortions in Norway after media coverage of adverse vascular events from the use of third-generation oral contraceptives. Contraception 1997;55:11–14.