Commentary: In pursuit of the ‘best’ possible evidence in investigations of breast implants and connective tissue disease

Ganesa Wegienka

Department of Biostatistics and Research Epidemiology, Henry Ford Health Hospital, 1 Ford Place, 3E, Detroit, MI 48202, USA.
E-mail: gwegien1@hfhs.org

Accepted 16 September 2010

Clinical medicine and public health organizations rely on epidemiology to play an integral role in evidence-based medicine. As epidemiologists, we are constantly striving to provide the best evidence with the goal of influencing public health and clinical practice for the better. With this overall objective in mind, studies are planned in excruciating detail and executed with painstaking care. After the data are collected and all analyses have been conducted, results are prepared for public dissemination. Sometimes, even when the optimal study has been conducted given the study population and subject matter, the investigators are left to excoriate their work in the ‘strengths and limitations’ section of the paper. The investigators do this even if their evidence is the best possible evidence and an improvement over prior evidence. In the current edition of the *International Journal of Epidemiology*, Lee et al. provide an excellent example of how they present the best possible evidence when faced with complicated study design issues. They investigated the risk of connective tissue disease (CTD) associated with breast implants. CTDs have poorly understood aetiologies and vast, complex symptomatologies. Considering the exposure, not only are breast implants a sensitive topic in society, but also the methods to study them are complex. There are no registries to easily identify women with implants, thus any study participant has her motivation questioned—perhaps she is only participating because of adverse symptoms she is experiencing. Those with implants are a mix of those who are electively choosing them (assumed physically ‘healthy’) and women who have been affected by cancer. Additionally, availability of implant types (saline and silicone) has varied over time and an individual woman may have multiple types over time. Presenting the best possible evidence becomes a massive challenge when confronted with the burden of dealing with judgement in the public health arena, in which: (i) randomized control trials (RCTs) are held by many clinicians and other individuals as a ‘gold standard’ and the only ‘real’ evidence that...
they will consider; and (ii) the results will be fed to the public, who commonly hear about years of scientific work in a 30-s sound bite.

Facing an outcome that is difficult to assess and has a latency that is not short, as well as having no option for an RCT (would not be ethical), Lee et al. conducted an optimal study—and they were completely honest about the associated limitations. They did many ‘good’ things: use of validated questionnaires and disease classification criteria reviewed by blinded physicians; inclusion of incident cases to establish temporal relationships; use of multiple approaches for outcome definitions based on various levels of evidence (self-report, validated questionnaire and medical record); and consideration of implant type.

While Lee et al. studied a rare outcome, their study population is the second largest ever examined in a study of this kind—24,000 women. They described how response rates varied by implant status for those who completed a CTD screening questionnaire (response rates higher for those with implants) and those who gave permission to review medical records (higher for women without implants). They also acknowledged the low case confirmation rates based on medical records—a rate that was lower among those with implants (but similar by implant status when considered as a percentage of all reported cases). They also noted that the overall rate of rheumatoid arthritis was lower in their study population than in other population-based studies, indicating a potentially ‘healthier’ study population.

Their results were not congruent with a previous meta-analysis in which no increases in risk rates were found to be associated with breast implants. However, their results were similar to those of their earlier retrospective cohort study, as well as a more recent cohort study in which elevated, but not large, hazard increases were reported for women with breast implants.

After detailing all of these potential biases, the authors and the readers are challenged to judge the net simultaneous effect of all of these potential biases. How should this be done? There is no formula or model. There is no weighting scheme. The authors also suggest that future observational studies will not likely be able to detect small to moderate increases in risk given all the methodological issues involved in studying breast implants and CTDs. However, the authors make a thoughtful statement in this dilemma: ‘Perhaps the best advance that we can make is merely to exclude the likelihood of large increases in CTD risk associated with breast implants’.

Ideally, every epidemiology study would result in clearly actionable conclusions based on important observed effects; however, sometimes this is not feasible. Instead, the study provides the best possible evidence and it comes as no surprise to the readers of this journal that sometimes the best possible evidence does not provide a definitive conclusion.

The situation encountered by Lee et al. in which they have a study with many potential biases, but no alternatives, is not unique and will continue to remain a challenge for epidemiologists. Medical treatments and other technologies move and develop quickly and rarely have registries or prospective studies that result in quickly available results to evaluate them. Rare outcomes with long latencies always present a challenge for researchers. Mobile phone use and brain tumours are an example in which rapidly changing technology (cell phones) and the human interaction with this technology is difficult to study in relation to a rare outcome with a potentially long latency.

The effectiveness of RCTs themselves can also be challenged. For example, a study in which women are randomized to hysterectomy or continued medical treatment for heavy bleeding may not result in findings that are applicable to the general population of women.

Lee et al. provide a nice example of how transparency in approach and clear presentation of study limitations combined with thoughtful interpretation of the best available data provide the best possible evidence. Despite the lack of a definite quantification of effects of breast implants on the risks of CTD, the conclusion of Lee et al. is one that can indeed provide evidence to women and their providers who make decisions daily based on the best possible evidence available to them.

Conflict of interest: None declared.

References


5. Saracci R, Samet J. Commentary: Call me on my mobile phone...or better not?—a look at the INTERPHONE study results. *Int J Epidemiol* 2010;39:695–8.
