Journal of Plastic, Reconstructive & Aesthetic Surgery (2016) 69, 864—868

JPRAS
An International Journal of Surgical Reconstruction
www.JPRASurg.com

LETTER FROM AMERICA

The future of plastic surgery data collection, analysis and presentation

Silicone sequels

It should not surprise readers across the pond to learn that just as Hollywood regularly produces sequels to horror movies, e.g., *Jaws IV* or *Friday the 13th part 13*, so too do horror stories recur in the press for products that we use such as silicone gel filled implants.[1] You also should not be surprised that *Annals of Internal Medicine* is not on the list "Journals Regularly Read by American Plastic Surgeons."

The tag line for *Jaws* was "Just when you thought it was safe to go back in the water". One might have thought that after the silicone crisis of the early 1990s ended it was safe to use silicone gel filled breast implants for cosmetic purposes. However, the systematic review, funded by the Plastic Surgery Foundation, which appeared in the November 2015 issue of *Annals of Internal Medicine*, concluded that:

> "The body of comparative study evidence on breast implants still does not provide conclusive evidence regarding whether silicone gel implants affect the risk for cancer, CTDs [connective tissue diseases], or other health outcomes. Owing to the flaws and inconsistencies among the studies reviewed, further investigation is required to determine whether any true associations exist between silicone gel implants and long-term health outcomes. For all outcomes, better evidence is needed, even where there is apparent consistency among flawed studies. *If* sufficient data on confounders *can be made available in existing* large studies, *it may be possible* to reanalyze these studies to *clarify the strength of associations* between silicone gel implants and health outcomes."[2] [Emphasis added]

Notice the conditions that the authors included — "if", "can be" and "may be possible". There are more qualifying conditions in this conclusion than in the promises of any US presidential candidate. I am not very sanguine that sufficient data can be made available in existing large studies.

Mea culpa

Last year I wrote about the importance of standardized reporting guidelines in the preparation of articles.[3] I am guilty of having put the cart before the horse, because no matter how diligently authors follow guidelines; they must have data to analyze before they can write articles. It is like knowing the rules of the road and how to drive a car, but you will never drive your car if it has no fuel. Simply put, we need more data in large quantities than are currently available and we must be able to analyze the data accurately.

All registries are not created equal

The American College of Surgeons began a data registry called the National Surgical Quality Improvement Program (ACS NSQIP®) that originated in the 1980s when the US Congress became concerned with the high mortality rates at our Veterans Hospitals.[4] It then morphed into its current form as a voluntary program in which subscribing hospitals have trained nurses extract relevant data from a sample of patient charts and follow up for a 30 day post-operative period.[5] It is not an inclusive registry because only a sample of patient charts is extracted. This is probably a relic from the decades when data storage was dear. Now data storage costs are almost nil. The most costly portion of the data collection process is the time expended by practitioners sitting in front of screens entering data, rather than providing care.

While NSQIP® is a noble effort, it has limited applicability for plastic surgery. Most obviously, many of the procedures that we do are not done in hospitals. Also, many of the complications that are important to us and to our patients occur after a 30-day post-operative period. Last year, Luce and Pierce suggested that the NSQIP® registry data lacked validity for alloplastic immediate breast reconstruction since, in their own series, 60% of prostheses

http://dx.doi.org/10.1016/j.bjps.2016.03.020

doomed to removal were still *in situ* at 30 days.[6] Adding further concern to the accuracy of the NSQIP® database was a report published online only that suggested that there are serious issues in the data entry for breast reconstruction. The authors found that 2013 was the perfect year for breast reconstruction with not a single complication of free flap loss, pedicle flap failure of prosthetic loss having been reported.[7] Another recent article questioned the accuracy and validity of hospital discharge coding data for breast reconstruction finding that only 77% of breast reconstruction flaps were correctly coded over a 5-year period at a major university hospital.[8] As hospital discharge data is fed into the Medicare database that too is possibly polluted.

Accurately analyzing data

The purpose of collecting these data is to create models that aid us in either diagnosis or prognosis. Putting aside these disturbing facts about the input to databases, we should examine the output that researchers have been diligently dredging. For example, as of today, 1738 articles are indexed in MEDLINE® with either NSQIP® or National Surgical Quality Improvement Program in either their title or abstract, clearly a virtual goldmine for producing papers.[9] Fifty-three of these articles were published in plastic surgery journals.[10] Most of the plastic articles use sophisticated statistical modeling techniques including logistic regression modeling or propensity score matching. Frequently, the articles statistical methodology does not match the authors' enthusiasm. Too many of these articles use "stepwise regression" techniques based upon low p values to select the predictor variables that, while easy to do with available software, have been known by statisticians for 20 years to produce spurious results.[11] This problem is not limited to articles using the NSQIP® database or even plastic surgery. A recent systematic review found that poor predictor variable selection was present in two-thirds of the prediction studies published in six high impact general medical journals in 2008.[12]

Validation

It is insufficient to create a model. Models must be validated, either internally by using the same data with a sophisticated statistical technique called bootstrapping or externally by applying the model to an independent data set.[13] Not validating a model is akin to accepting solely on blind faith that algorithm that an article advocates "works", although that algorithm was based upon the retrospective review of a case series; in other words, like the political candidate says, "Trust me."[14]

Rarely have authors who created models taken the next step of validating their models. A recent systematic review of externally validated models concluded:

> "The vast majority of studies describing some form of external validation of a multivariable prediction model were poorly reported with key details frequently not presented. The validation studies were characterised by poor design, inappropriate handling and acknowledgement of missing data and one of the most key performance measures of prediction models i.e. calibration often omitted from the publication. It may therefore not be surprising that an overwhelming majority of developed prediction models are not used in practice, when there is a dearth of well-conducted and clearly reported (external validation) studies describing their performance on independent participant data."[15]

It appears that the NSQIP® goldmine is mining fools' gold. There is a recommended international standard for reporting models and it is called TRIPOD.[16] It describes how models should be created and validated. However, until authors use TRIPOD and peer-reviewers and editors enforce its use, a substantial amount of space will be wasted in journals that could have been occupied by more meritorious articles. Perhaps, more importantly, readers will be wrongly impressed by the plethora of p values in the process called decerebrate genuflection.[17]

The benefit of data disclosure

Perhaps you have noticed that recent papers in JPRAS have online supplements with the data appended? There are a number of reasons for this, including:

- Consistency with BAPRAS' support of transparency
- Aiding future systematic reviewers who can use the data for pooled analyses
- Complying with the International Committee of Medical Journals (ICMJE) recommendations for submissions that require authors:

> "Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to judge its appropriateness for the study *and to verify the reported results.*"[18,19] [Emphasis added]

Here is an example of why this new policy is beneficial. Recently, Leckenby et al. published their findings on PIP Implants. They concluded that: "The size of an implant has no correlation to the prevalence of rupture ($p = 0.065$)".[20] Because their prospective registry publicly was appended online, I was able to create a LOESS plot with 95% confidence bands.[21] (Figure 1) This paints a different picture from Leckenby et al.'s conclusion because one can see that the 95% confidence bands for the intact and leaking implants diverge approximately 1680 days on the x-axis. This means that there was a statistically significant difference between implants that had been implanted for only about 1680 days with there being more leaking implants amongst the larger volume implants. If, as suggested by the plot, one splits the implants into two groups with one group being <300 ml and the other being ≥ 300 ml and then draws Kaplan-Meier survival plots, the results are significant. Larger implants did not survive as long as the smaller implants (Figure 2). This suggests:

- A manufacturing defect arose in the final 1680 days or 4.6 years that the larger PIP implants were used by these plastic surgeons.