VERIFICATION BIAS

Dear Editor, I read with interest the article by Shoshtari and co-workers[1] on the diagnostic value of C- reactive protein (CRP) measurements in patients with acute appendicitis. Unfortunately, I feel various methodological flaws could have led the authors to invalid conclusions. I will focus on one main limitation i.e. partial verification bias.

Diagnostic accuracy studies determine the performance of a test in diagnosing a *target disorder*. The performance of a new test (*index test*) is validated against an established *reference (gold) standard* test. Improperly conducted and incompletely reported studies are prone to bias that, in turn, may lead to overly optimistic appraisal of evaluated tests.[2] The performance of a diagnostic test can be estimated in several ways, including sensitivity, specificity, receiver operating characteristic (ROC) curves, positive and negative predictive values, likelihood ratios, and diagnostic odd ratios.[3,4]

Ideally all participants or a random selection of participants receiving index test should proceed to receive the same reference standard test. *Differential verification bias* occurs when participants with a positive index test receive a robust (but expensive and/ or invasive) reference standard, while participants with a negative index test receive a less accurate (but cheap and/or less invasive) reference standard. *Partial verification bias* occurs when a non-random group of participants selectively receive reference standard test. An extreme example is when reference standard test is performed only on people who have tested positive with the index test. Shoshtari et al recruited patients who received the reference standard test i.e. histology examination of appendix following appendicectomy. As surgery is likely to proceed preferentially in patients with elevated biological markers including CRP, the group of "cases" will be enriched by a clinical factor which itself may be associated with the CRP levels. In the study 81 patients had a CRP level above threshold and only 17 with a CRP level below threshold. Hence, it is likely that majority of participants with a normal CRP level did not proceed to appendicectomy. It is important to report how many participants receiving index or reference standard test failed to receive the other test and the reason of failing to do so. A flow diagram is highly recommended to clearly explain this issue.[3,5]

Differential verification bias leads to overestimation of both sensitivity and specificity. However, in partial verification bias the number of false negative and true negative participants is reduced, leading to over-estimation of sensitivity, and under-estimation of specificity and negative predictive value (NPV).[3] In the study the NPV is only 47%; implying only half the participants with a normal CRP level were truly non-diseased.

We know that some reference standards are expensive or invasive, and it may be impossible, or even unethical, to subject participant with a negative index test to these reference standards. A slightly modified approach can be adopted to overcome this problem. Participants with a negative index test can be followed up for a certain time period to confirm they remain disease free. In case symptoms persist and a later diagnosis of the target disorder is made, or if a known complication of target disorder develops, these participants could be reclassified as false negatives. This approach is likely to improve the number of true negatives (and false negatives) within a study and consequently, reveal accurate specificity and NPV.

In summary, patients' recruitment in the study should not have been based on the fact that they received the reference standard test. To avoid work up bias in prospective cohort studies such as this one, all or random selection of participants should be assigned to receive both the index test and the reference

standard verification either by direct procedure or by clinical follow up.

REFRENCES

1. Shoshtari MH, Askarpour S, Alamshah M. Diagnostic value of quantitative CRP measurement in patients with acute appendicitis. Pak J Med Sci 2006; 22: 300-3.
2. Lijmer JG, Mol BW, Heisterkamp S. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA 1999; 282: 1061–6.
3. Bossuyt PM, Reitsma JB, Bruns DE. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Clin Chem 2003; 49: 7–18.
4. Deeks J. Systematic reviews in health care: systematic reviews of diagnostic tests. BMJ 2001; 323: 157–62.
5. Bossuyt PM, Reitsma JB, Bruns DE. Standards for Reporting of Diagnostic Accuracy steering group towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. BMJ 2003; 326: 41–4.

M. A. Rehman Siddiqui. MRCS(Ed), MRCOphth, MSc
Specialist Registrar & Clinical Lecturer,
University of Aberdeen,
United Kingdom, AB25 2ZN.

Note: There was no response from the authors

* * * * * *

ALERT: A NEW PARADIGM TO PLAGIARISM

Some researchers are willingly citing their primary emails on publications. They believe that email address is considered as a source of communication for all researchers. Indeed, this is a nice gesture and commendable. Every one of us is bombarded with a flood of emails in our inboxes every day. Therefore, we should be vigilant as our emails could be deemed as a source of plagiarism for others. Spectrum of penetration computer programs are available on the web pages, these programs such as" pro agent", "password recovery" and others are easy to use, free to download and sophisticated enough to be caught by the anti-virus scanning programs.

I am not an expert in computer, however, I would like to mention some modest recommendations to all researchers, especially juvenile. I am sure that they will cascade my suggestions to all of their colleagues to be watchful.

1. Use of sophisticated anti-virus scanning programs is a good deterrent for such piercing programs, for both authors and editors.
2. Some tricky emails which are designed using famous names, tempt nascent authors. Such wicked emails are avenues for oozing the piercing programs, which are attached to your email and they ask you to download the attached files. I do recommend authors to ignore and delete such types of emails.
3. Another type of email misconduct is to send a rosy message which offer a list of services to all junior researchers, including promises helps in both study design and statistics. They ask you to open a link and to follow up instructions. The tricky programs are concealed behind such links. Do not care about such emails and delete them.
4. Finally, such wicked programs can be leakage to your computer through scientific clubs. Be careful.

Younis A. Skaik
Y_skaik@hotmail.com

* * * * * *