

Exercise: critical appraisal of a diagnostic test study

Read the attached short report 'Simple tests for septic bursitis: comparative study' (Stell and Gransden 1998). Apply the guidelines of Greenhalgh and the QUADAS instrument (given below) to this paper. What glaring statistical error is in this paper?

Just in case you don't know, the conditions being studied are defined as follows (shamelessly plagiarised from the WWW). Olecranon bursitis causes fluid to collect in a sac that lies behind the elbow, called the olecranon bursa. A bursa is a slippery, sac-like tissue that normally allows smooth movement around bony prominences, such as the point behind the elbow. When a bursa becomes inflamed for some reason, the sac fills with inflammatory fluid. This can cause pain and a noticeable swelling behind the elbow. Common names include 'student's elbow' when it occurs in people who study with their elbows leaning on a desk and 'miner's elbow' or 'plumber's elbow' when the job involves crawling a lot using elbows. Prepatellar bursitis is a common cause of swelling and pain above the kneecap. The name 'housemaid's knee' for this inflammation of the bursa overlying the patella comes from its association with individuals whose work necessitates kneeling for extended periods of time. Sometimes the bursa becomes infected, often due to an external injury, leading to septic bursitis. In this study the authors are trying to identify septic bursitis among patients with bursitis.

In this paper, there are two new tests: culture in liquid media and cell counts. Direct culture in solid media is an existing part of the diagnostic process.

Reference

Stell IM and Gransden WR. (1998) Simple tests for septic bursitis: comparative study. *BMJ* 316, 1877.

Greenhalgh guidelines

- 1: Is this test potentially relevant to my practice?
- 2: Has the test been compared with a true gold standard?
- 3: Did this validation study include an appropriate spectrum of subjects?
- 4: Has workup bias been avoided?
- 5: Has expectation bias been avoided?
- 6: Was the test shown to be reproducible?
- 7: What are the features of the test as derived from this validation study?
- 8: Were confidence intervals given?
- 9: Has a sensible 'normal range' been derived?
- 10: Has this test been placed in the context of other potential tests in the diagnostic sequence?

QUADAS criteria

1. Was the spectrum of patients representative of the patients who will receive the test in practice?
2. Were selection criteria clearly described?
3. Is the reference standard likely to correctly classify the target condition?
4. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?

5. Did the whole sample or a random selection of the sample, receive verification using a reference standard?
6. Did patients receive the same reference standard regardless of the index test result?
7. Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?
8. Was the execution of the index test described in sufficient detail to permit replication of the test?
9. Was the execution of the reference standard described in sufficient detail to permit its replication?
10. Were the index test results interpreted without knowledge of the results of the reference standard?
11. Were the reference standard results interpreted without knowledge of the results of the index test?
12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?
13. Were uninterpretable/ intermediate test results reported?
14. Were withdrawals from the study explained?

What glaring statistical error is in this paper?