Yield of serial sputum specimen examinations in the diagnosis of pulmonary tuberculosis: a systematic review


CRD summary
The authors concluded that the increase in average incremental yield and/or sensitivity gained by examining a third sputum specimen for the diagnosis of tuberculosis appeared to be low across diverse studies. The authors’ conclusion appears to be supported by the evidence, but the limited evidence from generally poor-quality and diverse studies suggests that a more cautious conclusion may be more appropriate.

Authors' objectives
To evaluate the diagnostic yield of the microscopic examination of three sputum specimens for acid-fast bacilli recommended for the diagnosis of pulmonary tuberculosis (TB).

Searching
PubMed (to May 2005) and BIOSIS Previews, EMBASE and Web of Science (all to December 2004) were searched for studies reported in English; the search terms were reported. In addition, conference abstracts, indices of two specified journals, and the reference lists of reviews and primary studies were screened, and authors were also contacted.

Study selection
Studies that evaluated patients suspected of having TB using microscopic examination of at least three separate sputum specimens were eligible for inclusion. Studies had to report sufficient data to enable the calculation of the incremental yield and/or the increase in sensitivity using the second and third specimens. Case reports and reviews were excluded.

In the review, a smear-positive case was defined as a positive result in at least one of the three specimens, and any positive smear was assumed to be a true positive even when no reference standard culture was performed. In the included studies, where reported, the reference standard was a mycobacterial culture on the same specimen submitted for microscopic examination. The studies differed with respect to study design (prospective and retrospective), stain used (fluorescence and Ziehl-Neelsen) and the methods used to process sputum (direct and processed). A minority of studies included patients with the human immunodeficiency virus (HIV).

Two reviewers independently screened citations and a third reviewer independently assessed all full-text reports.

Assessment of study quality
Study quality was assessed using the following criteria: blinded interpretation of smear and culture results; assessment of subsequent smears blinded to the results of previous smears; and the enrolment of prospective consecutive patients.

One reviewer assessed quality and a second reviewer checked 27% of the included studies. Any discrepancies were resolved by consensus.

Data extraction
For each study, the incremental yield in smear positivity and/or increase in sensitivity of serial sputum specimens were calculated.

One reviewer extracted the data and a second reviewer checked extracted data from 27% of the included studies. Any discrepancies were resolved by consensus.

Methods of synthesis
Weighted averages for incremental yield in smear positivity and/or increase in sensitivity (where cultures were reported) of the second and third sputum specimens were calculated; the studies were weighted by sample size. Separate analyses were conducted with studies grouped by design, stain used, methods used to process sputum, use of culture as a reference standard and HIV status.

Results of the review
Thirty-seven studies reported in 31 papers were included (the total number of patients was not reported). The sample size ranged from 5 to 11,650.

Twenty studies used culture as the reference standard. Eighteen studies were prospective and 19 were retrospective. Thirty-five studies enrolled consecutive patients. In only one study was interpretation of subsequent smears blinded to the results of previous smears. None of the studies reported smear readings blinded to culture results. Studies often did not provide information about the characteristics of the patients or details of sputum collection or quality assurance.

The overall weighted average incremental yield was 2.3% (95% confidence interval, CI: 1.8, 2.9; 29 studies) for the third sputum specimen. The overall weighted average percentage of all cases found with the first specimen was 85.8%, while the overall weighted average incremental yield of the second specimen was 11.9%.

For studies using culture as a reference standard, the overall weighted average increase in sensitivity was 3.1% (95% CI: 2.1, 4.2; 20 studies) for the third sputum specimen. The weighted average sensitivity of the first specimen was 53.8%, while the weighted average increase in sensitivity of the second specimen was 11.1%.

Subgroup analysis showed mean incremental yield in positive smear and/or mean increases in sensitivity ranging from 2 to 5% (results of subgroup analyses were reported).

Authors' conclusions
The increase in average incremental yield and/or increase in sensitivity gained by examining a third sputum specimen for the diagnosis of TB appeared to range from 2 to 5% across diverse studies. Examining only two sputum specimens could benefit TB programmes but should be used in conjunction with an operational research programme.

CRD commentary
The review question was stated clearly. Several relevant sources were searched and attempts were made to minimise publication bias but not language bias. Some attempts were made to reduce reviewer error and bias. Validity was assessed using specified criteria and the results of this assessment reported. Weighted averages were used to pool the data and the influence of various relevant factors was examined. However, there was no comment on any differences in results across studies, so it was not certain whether the average values reflect findings of all studies. In addition, the quality of the included studies appeared poor and this weakened the evidence base. The authors did discuss the limitations of the review. The authors' conclusion appears to be supported by the review, but the limited evidence from generally poor-quality and heterogeneous studies suggests that a more cautious conclusion may be more appropriate.

Implications of the review for practice and research
Practice: The authors stated that there are factors to consider before national TB control programmes change to examining only two sputum specimens. The following factors need to be taken into account: smear microscopy workload and human resources; reduction in the number of patients dropping out of the diagnostic pathway; potential savings that could be used to improve methods and equipment used for diagnosis; potential reduction in the number of smears required for quality assurance; potential decreases or increases in case detection; methods of obtaining a third specimen to meet the World Health Organization smear-positive case definition; and methods for following-up patients with two negative smears.

Research: The authors stated that studies should evaluate any new diagnostic strategy in different settings (including settings with low and high HIV prevalence) and standardise studies with respect to definitions, key variables and outcomes. Aspects to examine include the need for follow-up of patients with two negative sputum specimens; the need for follow-up of patients with one positive specimen out of two; identification of the best strategy to detect cases and best methods of specimen collection. In addition, the cost-effectiveness and the logistics of implementing different
strategies for finding cases of TB need to be assessed. Patient costs also need consideration.

**Funding**

**Bibliographic details**

**PubMedID**
17439669

**Other publications of related interest**


**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Humans; Sensitivity and Specificity; Specimen Handling; Sputum /microbiology; Tuberculosis, Pulmonary /diagnosis

**AccessionNumber**
12007001891

**Date bibliographic record published**
09/08/2008

**Date abstract record published**
23/12/2008

**Record Status**
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.