

Optimal tuberculosis case detection by direct sputum smear microscopy: how much better is more?

A. Van Deun,* A. Hamid Salim,† E. Cooreman,† Md. Anwar Hossain,† A. Rema,† N. Chambugonj,†
Md. A. Hye,† A. Kawria,† E. Declercq‡

* Mycobacteriology Unit, Institute of Tropical Medicine, Antwerpen, Belgium; † Damien Foundation Bangladesh, Dhaka, Bangladesh; ‡ Damien Foundation Belgium, Brussels, Belgium

SUMMARY

SETTING: A tuberculosis control project in Bangladesh.
OBJECTIVE: To define the efficiency of numbers of microscopic fields screened and the sputum collection scheme used for diagnostic smear examination.

DESIGN: Quality controllers noted cumulative numbers of acid-fast bacilli per 100 fields screened. The incremental diagnostic yield of different sputum sampling strategies was determined. Doubtful series were re-checked and/or further samples examined.

RESULTS: Acid-fast bacilli were found in 99.6% of 1412 positive and in 79.3% of 576 scanty slides in the first 100 fields. Examination of a third specimen yielded a maximum of 2.7% positives incrementally. The most efficient strategy, using three morning specimens, yielded 94.2% positives on the first and 1.0% on the third sputum; although 10% of suspects did not return, only

1.5% of the positives were among them and more cases were confirmed and treated. The positive predictive value of a single positive or scanty smear was very high (99.2%).
CONCLUSIONS: Reading more than 100 fields per smear or examining a third sputum has insufficient marginal returns to justify the workload. Examining morning samples only is more efficient, and their collection does not necessarily inconvenience patients. Treatment can be started on the basis of one positive smear. Provided that a well functioning system of smear-microscopy quality control is in place, we propose a strategy based on examination of two morning sputum samples for negative suspects, with the diagnosis based on a single positive result.

KEY WORDS: acid-fast; smear examination; tuberculosis; diagnosis

IN HIGH-PREVALENCE COUNTRIES, tuberculosis case detection is largely based on microscopic examination of sputum for acid-fast bacilli (AFB). As this is the most efficient way to detect the sources of transmission, it represents one of the five pillars of the DOTS strategy.¹ The technical guidelines of the World Health Organization and those of the International Union Against Tuberculosis and Lung Disease (IUATLD) specify that this should be done by examination of three samples, the first and the third being spot specimens taken at the centre, and the second one an early morning sputum.¹⁻³ It has been recommended that a minimum of 100 microscopic high-power fields should be examined for maximum yield.^{2,3} A minimum of 10 AFB per 100 fields is taken as the threshold for considering a result as positive, and a definite case should have at least one such result confirmed by a second smear examination or by a suggestive chest radiograph, or alternatively there should be one positive mycobacterial culture result.⁴

Some of these guidelines are based on older studies, while for others the arguments are unclear. In particu-

lar, the examination of three sputum samples per suspect has been severely criticised from a public health viewpoint,^{5,6} and several recent trials have all documented the relative inefficiency of the third smear.⁷⁻⁹ Most of these trials have been conducted in Africa, in populations with a high human immunodeficiency virus (HIV) prevalence. While the necessity for confirmation of a positive smear has also been contested,^{5,6} there has been little or no criticism concerning the number of fields to be examined.

Damien Foundation Bangladesh projects have a well functioning network of AFB microscopy, with external quality assurance (EQA) through re-checking of routine smears since 1996, as described previously.^{10,11} The population covered is rural, but with relatively easy access to the diagnostic centres (small distances and a service entirely free of cost), but the people are very poor and hidden costs may be prohibitive. HIV is virtually absent. According to National Tuberculosis Control Programme (NTP) reports, our case detection rates are the highest in Bangladesh. An early analysis showed an extremely low yield for the

third sputum examination in suspects,¹² while EQA revealed very few errors in both false negatives and false positives. For these reasons, we undertook some operational trials to find the optimal strategy for case detection by sputum microscopy in this setting.

The objectives of these studies were 1) to determine the efficiency of examining an increasing number of microscopic fields; 2) to determine the incremental yield of positives using different sputum sampling strategies in a routine field programme setting, as well as their relative efficiency for case detection and starting treatment; and 3) to check the necessity of confirming a positive suspect smear on a second specimen.

METHODS

The numbers of fields to be checked were studied during routine re-checking quality control in positive (at least 10 AFB per 100 fields) or scanty (1–9 AFB per 100 fields) smears from over 50 peripheral microscopy centres, as described previously.¹⁰ For all smears of one quarter, all the controllers were asked to note their results incrementally after every 100 high-power fields (1000 × magnification). Peripheral centres as well as controllers checked 300 fields, or at times more, in case of doubt. Only slides confirmed as positive by both first controllers and countercheckers were included in the analysis. The frequency at which each 100 field line yielded the first AFB was determined, and stratified by positivity grading of the original result (or as specified by the countercheckers in case of discrepancy). The scale recommended by WHO and the IUATLD^{2,3} was used, with a split-up of scanty results into two groups: those with <1 AFB per 100 fields, and those with 1–9 AFB per 100 fields. The latter constitute the 1+ group in the also widely used American Thoracic Society (ATS) scale.¹³

Over a period of about 12 months, groups of at least 10 diagnostic centres located in different administrative areas tested different sampling strategies in their daily work. The preceding period of 7 months, when all centres had used the standard NTP spot-morning-spot strategy, was used for comparison. The following variants were studied:

- ‘SMS’: the classical spot-morning-spot sampling, with the patient attending on two successive days; this was used by 72 centres;
- ‘SMM’: one spot and two morning samples, the latter produced at home on the first and second morning after the first visit, and both delivered by the patient on the second day; this was used by 21 centres;
- ‘MMM’: no spot but three morning samples, produced at home on the first, second and third morning after the first visit, and delivered together by the patient on the third day; this method was used by 19 centres;

- ‘SM(2)’: one spot and one morning sample, the latter produced at home on the first morning after the visit, and brought to the centre by the patient on that day. This morning sample was examined twice, once fresh and again after sedimentation—homogenisation by standing overnight at ambient temperature; this method was used by 10 centres;
- ‘SMM(2)’: one spot and two morning samples, the latter produced at home on the first and second morning after the first visit, and both brought to the centre by the patient on the second day; the second morning sample was examined twice—once direct, and once after sedimentation, as described under SM(2); this method was used by 27 centres.

With all the variants, additional specimens could be requested in case of inconclusive results after completion of the prescribed procedure based on three (or four) smears.

During one quarter, all centres subsequently tried a variation with two morning samples. Sampling was as described for SMM, but unlike in SMM the spot sputum was examined only after the two morning specimens, or in case the patient had not returned, about one week later (MMS). Non-returning suspects found to be AFB-positive on the spot sputum had to be visited at home for further sputum checks.

Suspects eligible for sputum AFB smear were defined as usual for NTPs (cough of more than 3 weeks, or haemoptysis, or a combination of symptoms suggestive of progressive tuberculosis). Smears were made, stained and examined by the hot Ziehl-Neelsen method according to WHO/IUATLD guidelines,² except that a higher (1%) basic fuchsin and a lower (0.1%) methylene blue concentration were used. In this later phase of the study, only 100 fields had to be examined to declare a smear to be negative, but 300 had to be examined for quantification of low positive and scanty smears. Cases could be started on treatment either by paramedical personnel based on at least one result of a minimum of 4 AFB per 100 fields (interpretation formerly used by WHO¹⁴), plus another positive or scanty sputum, or on the decision of a medical officer based on AFB microscopy plus other evidence. Once-positive suspects who did not return to complete the diagnostic series or to start treatment had to be followed up by a home visit.

Re-checking of series with discrepant results (positive or scanty as well as negative) was done at the central laboratories. Central laboratory routine re-checking of a random sample from all diagnostic and follow-up smears was done continuously for all centres throughout the study period. Results were registered in the usual AFB microscopy registers, but sputum containers and register columns were marked in such a way that each result could be identified by type of specimen or smear. All suspect series with at least one non-negative result were entered into a computer Epi-Info

Table 1 Cumulative yield per 100 high-power fields examined

Number of fields read when AFB were first seen	Percentage of smears found to contain AFB, by quantification group*		
	Scanty <1 AFB/100 fields (n = 64)	Scanty 1–9 AFB/100 fields (n = 512)	IUATLD/WHO positives (at least 10/100 fields) (n = 1412 [†])
100 fields	48.4%	83.2%	99.6%
200 fields	68.7%	94.1%	99.9%
300 fields	89.1%	99.2%	100%
>300 fields	100%	100%	

* All non-negative smears confirmed by routine re-checking quality assurance during one quarter are shown.

[†] 'IUATLD/WHO positives' refers to the quantification scale of the International Union Against Tuberculosis and Lung Disease and the World Health Organization.²

AFB = acid-fast bacilli.

file, in a format allowing analysis of incremental yield by sampling strategy, as well as counts by type of sputum, positivity grading, start of treatment and diagnostic centre. Analysis was done using Pearson's χ^2 or Fisher's exact test, as applicable.

RESULTS

The proportions of confirmed positive or scanty smears found to contain AFB after every 100 fields screened by the controllers, are shown as cumulative percentages in Table 1. Examination of the first 100 fields was sufficient to identify over 99% of all positives according to the IUATLD/WHO scale, or 83% of the 1–9/100 fields group. Two hundred fields had to be screened to find 94% of scanty results 1–9/100, but by then only about two-thirds of the group containing one or two AFB per 300 fields had been found. Furthermore, the quality control results of this period showed that after screening 300 fields, the first controllers had completely missed 0.2%, 0.8% and 8% of the IUATLD/WHO scale positives, 1–9/100 and <1/100 scanty, respectively (data not shown).

Table 2 shows incremental yields of positive smears (IUATLD/WHO cut-off) for the various suspect sampling and examination strategies that were tested. The table takes into account only the first three results (or less, as incomplete series are included), as a fourth smear had been examined systematically only for the SMM(2) series. Considering only the first three results, the SMM and SMM(2) strategies were identical, and thus their results have been combined. The first smear yielded already 79.8%–94.2% of positives. For the second smear, the range was 4.8%–17.9%. The third smear invariably yielded very few additional positives, 1.0%–2.7%. All the differences of first smear yields were statistically significant, as was the yield of the third smear in the MMM strategy. Considerable variations were observed between individual centres using the same strategy (the ranges are shown in Table 2; centres with fewer than 15 cases have been excluded). In the SMM(2) strategy, a fourth smear made from the overnight sediment of the second sputum increased the number of cases detected by the first three smears by 1.5% (details not shown). Table 2 also shows the proportions of con-

Table 2 Incremental yield of positives using various suspect sputum examination strategies, IUATLD/WHO positivity cut-off

Sampling and examination strategy	Total positive suspects identified*	Percent positives identified on first smear [†] (range)	Percent positives identified on second smear [†] (range)	Percent positives identified on third smear [†] (range)	Percent cases confirmed by at least two positive results [‡]	Percent cases confirmed by at least one positive and one scanty result [‡]
SMS (spot-morning-spot)	4865	82.5 (60–98.3)	15.2 (1.7–36.9)	2.3 (0.0–6.2)	85.0	7.3
SMM+SMM(2) (spot-morning-morning)	4774	79.8 (64.3–94.4)	17.9 (1.7–35.3)	2.3 (0.0–7.7)	84.6	6.1
MMM (morning-morning-morning)	3077	94.2 (89.1–98.3)	4.8 (1.2–9.5)	1.0 (0.0–2.4)	92.3	3.7
SM(2) (spot-morning-sedimented morning)	1864	85.3 (77.4–94.7)	12.0 (5.3–22.6)	2.7 (0.0–6.3)	88.4	4.5

* The total number of suspect series with at least one smear containing AFB (positive or scanty) was 5147 for the SMS strategy, 5094 for SMM+SMM(2), 3159 for MMM and 1946 for SM(2).

[†] Percent positives identified refers to the first smear found to be positive above the cut-off, out of all series of suspects with at least one positive among the first 1–3 smear results.

[‡] Percent cases confirmed refers to suspects with at least two smears containing AFB (either at least two positive, or one positive and one scanty result). Series consisting of only two sputum samples have been included. The denominator is all suspects with at least one positive or scanty result in a series of two or three smears.

Table 3 Relative yield of positives according to type of sample and smear, shown differentially for the cut-offs of the ATS and the IUATLD/WHO

Type of sample/smear comparison*	Number of comparisons	Percent increase† (IUATLD/WHO cut-off)
Fresh morning over spot increase	16 965	+8.5
Homogenised morning over fresh morning increase	11 231	+2.7
Sedimented morning over homogenised morning increase	1 329	-1.9

* The results of series collected with different sampling strategies have been combined according to type of sputum and smear, irrespective of sample sequence.

† Percent increase refers to the number of positives identified by the second smear of the pair compared to the number identified by the first smear.

ATS = American Thoracic Society.

firmed cases among all those with at least one non-negative result. Based on at least two positive results, 84.6%–92.3% were confirmed as cases, depending on the sampling strategy used. Accepting cases identified by one positive plus one scanty result yielded another 3.7%–7.3%.

The strategy based on three morning sputum samples was significantly more efficient than the others in identifying the maximum proportion of cases on the first two specimens ($P < 0.001$). A variation of this strategy (MMS) applied in all centres, using two morning specimens with a spot sputum examined only after these two, yielded the same results as when only the two morning specimens were considered. Of the 3511 MMS series with at least one positive result, 96.7% were identified with the first and 3.3% with the second sputum examination, against respectively 97.2% and 2.8% with MMM (details not shown; non-significant).

An analysis of yields by type of specimen/smear is shown in Table 3. For these calculations, identical types of samples/smears (spot sputum, morning sputum, morning sputum at least one day old and homogenised by autolysis, or an overnight sediment of the former) were pooled over the different strategies. Only the summary relative increments in the specified order (spot-morning-homogenised morning-sediment) are shown. The morning specimen yielded 8.5% more

positives than the spot in the same series; a 1-day-old homogenised morning sample yielded 2.7% more than a fresh morning sample, but the sediment of a homogenised morning sample gave 1.9% less positives than the latter.

Table 4 shows the percentages of patients who started treatment by sampling strategy. Out of the whole study population with at least one non-negative result, 96.7% started treatment. The highest proportion, 98.6%, was found in the MMM group ($P < 0.001$). Comparing the non-starters with various sampling strategies, 0.3%–1.1%, or overall 0.6% of confirmed cases, did not start treatment ($P < 0.001$). Overall, 498 patients with at least one positive result in their series did not start treatment. Only 18.7% of all non-starters were cases confirmed by at least two positive results or one positive plus one scanty result (data not shown). The others were not confirmed as cases, as they had only one positive or only scanty results in the diagnostic series (either complete or incomplete). The reason was discovered for 134/222 (60.4%) with an incomplete diagnostic series. Of these, 37.3% had died very soon after giving the first positive specimen, while the remainder had refused to continue the investigations, preferring other services (private practitioners or village doctors), or refusing the provisional diagnosis of tuberculosis. These proportions

Table 4 Treatment start by sampling strategy

	SMS (spot-morning- spot)	SMM+SMM(2) (spot-morning- morning)	MMM (morning- morning- morning)	SM(2) (spot-morning- sedimented morning)	Total study population
Overall percentage started on treatment**†	96.9	95.2	98.6	97.0	96.7
Number who did not start treatment	159	236	45	58	498
Percent non-starters out of all confirmed cases	0.5	1.1	0.3	0.5	0.6
Percent non-starters of all those non-confirmed with incomplete series	72.0	83.2	69.6	87.5	78.6
Percent non-starters of all those with an isolated positive or scanty result	40.1	44.4	32.1	34.6	39.9
Percent non-starters of all those with repeatedly scanty results	35.9	20.0	30.8	0.0	27.2

Denominators used in the table vary and are specified in the first column.

* Cases could be started on treatment either by paramedical personnel based on at least one result of a minimum of 4 AFB per 100 fields plus another positive or scanty sputum, or on decision of a medical officer based on AFB microscopy plus other evidence.

† Diagnosis was considered confirmed when at least two results were positive or one positive and one scanty prevailed. When all results were scanty, or when only one positive or scanty result was found in a series of at least two sputum smears, diagnosis was not confirmed. Incomplete series means series with less than three results.

varied only little with the sampling strategy used, except for the MMM series, where all nine such cases refused to provide more samples.

Suspects who were requested to submit sputum samples but who did not do so were counted from the consultation registers for one quarter. Their percentage varied from 1.6% with the SMM+SMM(2) strategy and 3.6% with the SM(2) strategy to 13.1% with the MMM sampling strategy ($P < 0.001$; data not shown). The MMS phase of the trial made it possible to assess the proportion of cases among suspects who did not return when only morning specimens were requested. Forty-nine positive (plus eight scanty) spot specimen cases did not return with morning sputum, out of 3865 with at least one positive or scanty result (1.5%). After a successful home visit, 19 (0.5%) could be put on treatment after examination of more sputum samples, seven were registered as 'died' after submission of the first sputum sample, and two were found to be on treatment at other NTP centres. The remaining 29 were either not found, or they refused to be treated at the centre.

During the period of the SMS strategy (used in all centres), routine random quality assurance re-checking showed 2.8% false positives (136/4846 smears re-checked), and 2.9% false negatives (229/7999), scanty results of any grade included. Later on, when the other strategies were tried, the respective figures were 1% (41/4156) and 1.5% (148/9877), ranging from 0.75% (SM(2)) to 1.2% false positives (MMM), and from 1.2% (SMM(2)) to 1.8% (MMM) false negatives (detailed data not shown). Comparing the SMS period with the sum of the subsequent strategies, the differences of both false positive and false negative proportions were highly significant ($P < 0.001$), but there was no significant difference in error rates between the subsequent strategies.

Table 5 shows the results of the re-reading of the original smears, as well as the results of the examination of a fourth or even more sputum specimens. These are classified as series with one positive plus one scanty

result, series with a single positive or scanty result, and series with exclusively scanty results, the latter subdivided in groups with more than 3, 1–3 and <1 AFB per 100 fields (corresponding to cut-offs in the main quantification scales^{2,13,14}). Slides of 31%–82% of all such series were re-read. Only 2.5% of the series with one positive and at least one scanty result could not be confirmed, against 5.1% of the series with a single positive and 31.9% with a single scanty result. For the series with only scanty results, this varied from 6.1% to one in three. Proportions of the same groups for which additional sputum specimens had been examined ranged from 0% to 48% (Table 5). None of the series with two smears containing more than three AFB per 100 fields (one positive plus one scanty or two scanty results) proved consistently negative on further examination. However, 10.3% of series with repeatedly low scanty (1–3/100), 16.3% with an isolated positive and even 24.7% with an isolated scanty result, did not show any AFB in additional specimens and might thus have been false positives. Taken together with all cases confirmed by at least one positive and one scanty result in their series of sputum examinations, this suggests that a diagnosis based on a single positive or scanty sputum examination result would have been false in 0.8% of cases in our study.

DISCUSSION

While the number of serial smears that need to be examined per clinically identified suspect to identify a bacteriologically confirmed case of tuberculosis has been severely criticised, the optimal number of fields per smear has not been the subject of much reflection regarding costs and workload. However, it is clear that the choice between 100 and 300 fields may have an even greater effect in terms of workload and savings than the choice between two or three smears. Older mycobacteriology manuals often mention that 300 fields or even more ('the whole smear') should be examined,^{15,16} and such recommendations may also

Table 5 Results of re-checking and of examination of further sputum specimens for complete series with fewer than two positive results, by grading

Results of further checking	Non-negative results in the first three smears*					
	One positive plus one scanty	Isolated positive	Isolated scanty	Two scanty >3/100 fields	Two scanty 1–3/100 fields	Two scanty <1/100 fields
Re-checking of the first three smears [†]						
Number (%) re-checked	160 (31)	99 (82)	157 (82)	82 (39)	26 (52)	3 (100)
% negative	2.5	5.1	31.9	6.1	19.2	33.3
Examination of further specimens [†]						
Number (%) examined further	127 (14)	80 (48)	93 (31)	92 (28)	39 (45)	0
% negative	0	16.3	24.7	0	10.3	—

* A positive result is defined here as at least 10 AFB/100 fields (IUATLD/WHO cut-off). Scanty results have been stratified according to cut-offs in other scales: ATS scale positivity threshold (1/100 fields) and former WHO threshold for interpretation as positive (4/100 fields).

[†] 'Re-checking' and 'further specimens' do not have the same denominator, as re-checking of discordant series was not done for the SMS group. Some series have both been re-checked and extended by checking more specimens. AFB = acid-fast bacilli.

be found in some NTP technical guidelines.¹⁷ The first edition of the IUATLD Technical guide for sputum examination was ambiguous on this point, recommending 100 fields, but then stating that it was better to examine 300 further on in the text.¹⁸

Our study shows that clear positives (at least 10 AFB/100 fields) were virtually all recognised in the first 100 fields examined. Expanding the range to 1 AFB/100 fields, i.e., 10 times lower than the threshold for positivity set by the IUATLD and WHO, this was still true for 95% of positives detectable examining 300 fields. Moreover, examination of additional sputum specimens can logically be expected to further reduce the fraction that was missed in the first 100 fields of the first smear. The workload involved in checking 200–300 fields thus seems to be out of proportion to the gain that can be obtained. This is certainly true in our situation in Bangladesh, where only 10% of diagnostic smears fall into the scanty category using the IUATLD/WHO scale (own unpublished data). These proportions, and thus the potential utility of examining larger numbers of fields, might be different in highly HIV prevalent areas, where smears are known to be more often paucibacillary.¹⁹

Nevertheless, several recent publications from such areas on the yield of successive smears suggest that in HIV-ridden areas positive smears are also not very difficult to recognise. The third sputum yielded only 4–7% of total positives in Malawi and Zambia,^{7,8} and as few as 2.5% in Tanzania.²⁰ The classic studies on incremental yield, also discussed by Toman,²¹ were from India.^{22,23} These found an incremental yield of only around 3% after the second sputum examination. In fact, some of these authors concluded that the examination of more than two sputum samples was not rewarding.

The higher positivity rate of early morning compared to spot sputum smear examinations has been reported repeatedly.^{22,24} Urbanczik shows extreme results of two studies where this was over 60%–100% higher.²⁴ Furthermore, it has been known for a long time that the results of AFB microscopy for tuberculosis are not affected by delays, even when the sputum is kept at tropical temperatures.^{25,26}

Considering all these studies it is difficult to understand the rationale of a sampling strategy with three sputum samples, two of which are spot specimens. Requesting two spot specimens was a compromise proposed by Andrews and Radhakrishna when they were trying to improve accessibility by reducing the number of visits required, while sacrificing sensitivity only slightly.²² However, they counted on simultaneous cultures as well as on heavily positive specimens. This strategy may not work in other situations, with many low positive sputum samples (HIV), no cultures and overburdened services, and where the diagnostic delay is often much longer than one day. As the human factor is of prime importance, and overload and fatigue have

a disastrous effect on the quality of AFB microscopy, the examination of more sputum samples per patient might in fact lead to less positives being detected. This aspect has not yet been studied sufficiently, although Harries et al. have found that the positivity rate among suspects in one hospital was exactly the same when two sputum samples were examined as during the preceding period when three were examined.²⁷ Finally, while prolonging the diagnostic process to detect a maximum number of smear-positive cases, a considerable proportion of these cases may not even be put on treatment. For instance in Malawi, 14% of AFB-positive cases were not registered for treatment.²⁸

Our results, obtained by a highly proficient service in a virtually HIV-free population, indicate once again that the yield of the third sputum is very low. Despite the variations in sampling that were tried, none yielded more than 2.7% of total AFB-positives at the IUATLD/WHO cut-off point to designate positivity. The most efficient scheme, using only morning sputum samples delivered by the patient all at one time, left only 1.0% for the third sputum to detect, while the yield of the two first sputum samples ranged from 97.6% to 100% in the different centres. The efficiency of this scheme can probably be explained by the higher yield from morning specimens, especially when already homogenised (1 or 2 days old). On the other hand, an overnight sediment of homogenised sputum was less often positive. We do not fully understand why the older, homogenised sputum samples yielded a higher proportion of positives. Dispersion of AFB throughout the smear, as well as a better smear with an easy, amorphous background are certainly contributing factors. Thick, poor smears are a frequently observed deficiency (though not in DF Bangladesh centres), and carefully picking a thick, purulent particle, as is recommended in all textbooks and also taught during training, is the exception rather than the rule in daily routine. So, although we do not wish to dispute the recommended technique, in practice a strategy which uses old, homogenised sputum can give acceptable results even when such rules are neglected. Another consideration when using old sputum is their unpleasantness for the technicians, which should be weighed against the inconvenience for the patient if each morning specimen has to be delivered on the day it was produced.

The early morning strategy (MMM) allowed a higher proportion of cases to be confirmed on the basis of two positive (or at least one positive and one scanty) results and almost all the confirmed cases to be started on treatment. Overall, this scheme made the most efficient use of microscopy, as 98.6% of all patients with at least one non-negative smear started treatment. Some uncertainty remained because of the considerable proportion of suspects who did not return with sputum at all, until we found that those

who did not return comprised only 1.5% of all cases. Moreover, only one third of those could be put on treatment after a home visit. During this extension of the study, it proved possible to maintain the same efficiency of a two-morning sputum based strategy, applying it in all our routine centres.

A strategy using a collection of two successive morning samples delivered simultaneously might be most efficient. When a maximum yield rather than efficiency or cost-effectiveness is valued, more sputum samples can be examined provided that ample manpower is available. It might be more rewarding to do this by examining a second series of two morning sputum samples a few weeks later, for example after a course of antibiotics has failed to cure the patient. However, although this is an often recommended strategy, its relative efficiency in terms of yield of cases starting treatment has not been properly determined.

The need for systematic confirmation of an AFB-positive result by a second positive result to diagnose a case of tuberculosis has been challenged.²⁹ Critics have argued that this practice, besides causing an increase in workload, constitutes an unnecessary burden for the patient, and may actually result in fewer cases being put on treatment, considering the numbers that do not return after a first positive smear. Our results show that they may be partly right. On re-checking smears, the proportion of false positives was indeed negligible. However, this does not exclude mislabelling of specimens, slides or reports. Examining further sputum specimens, we found that no more than 10%–25% of series with an isolated positive/scanty or with repeatedly low scanty results could not be confirmed. Thus, out of all 15 098 positive series in this study with more than one sputum examined, less than 1% might have led to false diagnoses based on the first positive smear result. Furthermore, the majority of those with an isolated positive or scanty smear result, or with repeatedly scanty results, were nevertheless started on treatment. As this is the generally accepted policy for non-confirmed and even completely negative cases (provided the decision is made by a medical officer), routinely requesting confirmation by a second positive microscopy result seems rather futile.

It is not quite clear how many positives in our study would have been lost to treatment because of delays in waiting for confirmation on another specimen, as an unknown proportion of non-returners were confirmed and started on treatment after a home visit. Perhaps for this reason, only 222 suspects who were positive or scanty positive on the first sputum did not return, but these incomplete series constituted almost 40% of cases not starting treatment. However, it would be wrong to conclude that starting treatment on the basis of only one positive/scanty result would have avoided these losses, as we found that many of these patients had died after the first sputum, while

many others refused to be treated by our service. In Malawi, a similar proportion of cases who were not registered for treatment were found to have died, but refusal of treatment was very rare.²⁸ Another 37% of the 498 non-starters had a single positive or scanty result on the first three smears. As only about 20% of series with such results were found negative when additional smears were examined, it might be justified not to try to confirm such results.

While our results were obtained by an excellent microscopy network, it may not be correct to assume that in a less well functioning system more fields or more sputum samples should be examined. Prerequisites for strategies based on examination of fewer samples (for confirmation of negative suspects as well as cases) might rather be the existence of an efficient system of smear microscopy quality assurance, including regular supervision. Simply emphasising examination of more sputum samples or more fields may very well be counterproductive. In our experience, poor or overloaded technicians will often not comply and will stop after just a few fields, and they have been seen to copy the result of the first sputum to numbers two and three. For those who do comply, the examination of a large number of fields and/or too many smears risks being done too superficially to be rewarding.

CONCLUSIONS

Screening of only 100 microscopic fields per smear, rather than the sometimes recommended 300, reduces microscopy workload to one third while only a marginal proportion of positives is missed. Considering the pronounced negative effect of fatigue on the accuracy of AFB microscopy, 100 fields may be an optimal number.

The incremental yield of the third smear was small to negligible with all of the sampling schemes tested. Sampling only morning specimens, even older ones, can be highly efficient in terms of workload and proportions of cases detected and put on treatment. The NTP would then have to decide whom to inconvenience—the technicians having to process foul smelling sputum brought in by the patient all at one time, or the patients if they are requested to deliver one morning specimen on each of two days.

Provided that the laboratory network is reasonably proficient, which should be ascertained by re-checking the EQA, the proportion of false positive suspect smears is extremely low. If maximum accuracy is desired, low scanty smears might better be confirmed on a second specimen, but routinely requesting confirmation of a first positive or even scanty smear on the basis of a second one is not efficient, and may lead to an excessive loss of true cases.

We propose that one spot specimen should be sampled at presentation and two successive morning specimens should be requested to be delivered together 2

days later. However, except for suspects in a very poor clinical condition, the spot sputum should only be examined if no morning specimens are delivered at all. In that case, if found positive, the patient would have to be traced. As a standard, however, only the morning specimens would be examined and the spot specimen discarded, preferably with the suspect waiting for the results. In case the first morning sputum contains AFB, the second would only be examined for confirmation if the first result remains below the positivity threshold set by the NTP.

Acknowledgements

We wish to extend our sincere thanks to the personnel of the Damien Foundation Bangladesh projects. Only their enthusiasm has made these trials possible, which have been conducted without any additional support. We have also very much appreciated the encouragement and material support extended by the National Tuberculosis Programme of Bangladesh.

References

- 1 World Health Organization. Guidelines for national programmes. WHO/TB/97.220. Geneva: WHO, 1997.
- 2 Enarson D A, Rieder H L, Arnadottir T, Trébuq A. Management of tuberculosis: a guide for low income countries. 5th ed. Paris: IUATLD, 2000.
- 3 World Health Organization. Laboratory services in tuberculosis control. WHO/TB/98.258. Geneva: WHO, 1998.
- 4 WHO, IUATLD, KNCV. Revised international definitions in tuberculosis control. *Int J Tuberc Lung Dis* 2001; 5: 213–215.
- 5 De Cock K M. Tuberculosis control in resource-poor settings with high rates of HIV infection. Editorial. *Am J Public Health* 1996; 86: 1071–1073.
- 6 Dujardin B, Haelterman E, Van Damme W, Kegels G. The adequacy of one smear for diagnosis of pulmonary tuberculosis. *Am J Public Health* 1997; 87: 1234–1235.
- 7 Harries A D, Kamenya A, Subramanyam V R, Salaniponi F M, Nyangulu D S. Sputum smears for diagnosis of smear-positive pulmonary tuberculosis. *Lancet* 1996; 347: 834–835.
- 8 Walker D, McNeerney R, Kimankinda Mwembo M, Foster S, Tihon V, Godfrey-Faussett P. An incremental cost-effectiveness analysis of the first, second and third sputum examination in the diagnosis of pulmonary tuberculosis. *Int J Tuberc Lung Dis* 2000; 4: 246–251.
- 9 Wu Z L, Wang A Q. Diagnostic yield of repeated smear microscopy examinations among patients suspected of pulmonary TB in Shandong province of China. *Int J Tuberc Lung Dis* 2000; 4: 1086–1087.
- 10 Van Deun A, Portaels F. Limitations and requirements for quality control of sputum smear microscopy for acid-fast bacilli. *Int J Tuberc Lung Dis* 1998; 2: 756–765.
- 11 Van Deun A, Roorda F A, Chambugonj N, Hye Md A, Hossain Md A. Reproducibility of sputum smear examination for acid-fast bacilli: practical problems met during cross-checking. *Int J Tuberc Lung Dis* 1999; 3: 823–829.
- 12 Salim A H. The efficiency of sputum examination strategies: a case study from Bangladesh. MPH thesis. Antwerp, Belgium: Prince Leopold Institute of Tropical Medicine, 1998.
- 13 American Thoracic Society. Diagnostic standards and classification of tuberculosis and other mycobacterial diseases. 14th ed. New York: American Thoracic Society, 1981.
- 14 Tuberculosis Programme, World Health Organization. Managing tuberculosis at district level, a training course. Supporting laboratory services. WHO/TB/96.203. Geneva: WHO, 1996.
- 15 Tison F, Carbone B. Recherche, isolement et étude du bacille tuberculeux et des autres mycobactéries en pratique courante. Lille: Crouan & Roques, 1972.
- 16 Kent P T, Kubica G P. Public Health Mycobacteriology. A guide for the level III laboratory. Atlanta: Centers for Disease Control, 1985.
- 17 Programme National Anti-tuberculeux Intégré aux soins de santé de base 'P.A.T.I. 3'. Kinshasa: Bureau National de la Tuberculose, 1996.
- 18 International Union Against Tuberculosis. Technical guide for sputum examination for tuberculosis by direct microscopy. *Bull Int Union Tuberc* 1978; 53 (Suppl 2): 4–16.
- 19 Johnson J L, Vjecha M J, Okwera A, et al. Impact of the human immunodeficiency virus type 1 infection on the initial bacteriologic and radiographic manifestations of pulmonary tuberculosis in Uganda. *Int J Tuberc Lung Dis* 1998; 2: 397–404.
- 20 Ipuge Y A I, Rieder H L, Enarson D A. The yield of acid-fast bacilli from serial smears in routine microscopy laboratories in rural Tanzania. *Trans Royal Soc Trop Med Hyg* 1996; 90: 258–261.
- 21 Toman K. Tuberculosis case-finding and chemotherapy. Questions and answers. Geneva: WHO, 1979.
- 22 Andrews R H, Radhakrishna S. A comparison of two methods of sputum collection in the diagnosis of pulmonary tuberculosis. *Tubercle* 1959; 40: 155–162.
- 23 Nair S S, Gothi G D, Naganathan N, Rao K P, Banerjee G C, Rajalakshmi R. Precision of estimates of prevalence of bacteriologically confirmed pulmonary tuberculosis in general population. *Ind J Tub* 1976; 23: 152–159.
- 24 Urbanczik R. Present position of microscopy and of culture in diagnostic mycobacteriology. *Zbl Bakt Hyg* 1985; A 260: 81–87.
- 25 Paramasivan C N, Narayana A S L, Prabhakar R, Rajagopal M S, Somasundaram P R, Tripathy S P. Effect of storage of sputum specimens at room temperature on smear and culture results. *Tubercle* 1983; 64: 119–124.
- 26 Harries A D, Nyirenda T E, Banerjee A, Mundy C, Salaniponi F M. District sputum smear microscopy services in Malawi. *Int J Tuberc Lung Dis* 1998; 2: 914–918.
- 27 Harries A D, Mphasa N B, Mundy C, Banerjee A, Kwanjana J H, Salaniponi F M L. Screening tuberculosis suspects using two smears. *Int J Tuberc Lung Dis* 2000; 4: 36–40.
- 28 Nyirenda T, Harries A D, Banerjee A, Salaniponi F M. Registration and treatment of patients with smear-positive tuberculosis. *Int J Tuberc Lung Dis* 1998; 2: 944–945.
- 29 De Cock K M, Wilkinson D. Tuberculosis control in resource-poor countries: alternative approaches in the era of HIV. *Lancet* 1995; 346: 675–677.

RÉSUMÉ

CADRE : Un projet de lutte antituberculeuse au Bangladesh. **OBJECTIF :** Définir l'efficacité du nombre de champs microscopiques examinés et celle du schéma du recueil des expectorations utilisé pour l'examen des frottis pour diagnostic. **SCHÉMA :** Des contrôleurs de qualité ont pris note des

nombre cumulatifs de BAAR par 100 champs examinés. Le rendement supplémentaire de diagnostic de différentes stratégies d'échantillonnage des expectorations a été déterminé. Les séries douteuses ont été recontrôlées et/ou d'autres échantillons ont été examinés. **RÉSULTATS :** On a décelé des BAAR dans les 100 pre-

miers champs chez 99,6% de 1.412 frottis positifs et dans 79,3% de 576 lames contenant de rares bacilles. L'examen d'un troisième échantillon a fourni au maximum un rendement supplémentaire de 2,7% de positifs. La stratégie la plus efficace utilisant trois échantillons du matin a obtenu 94,2% de positifs sur le premier et 1,0% sur le troisième échantillon. Bien que 10% des suspects ne soient pas revenus, 1,5% seulement des positifs se retrouvaient parmi eux et un nombre plus grand de cas ont été confirmés et traités. La valeur prédictive positive d'un frottis unique positif ou contenant de rares bacilles est très élevée (99,2%).

CONCLUSIONS : La lecture de plus de 100 champs par frottis ou l'examen d'une troisième expectoration ont un rendement insuffisant pour justifier la charge de travail. L'examen des seules expectorations du matin est plus efficace et leur recueil n'entraîne pas nécessairement d'inconvénients pour les patients. Le traitement peut démarrer sur la base d'un frottis positif. Pour autant qu'un système performant de contrôle de qualité de la bacilloscopie soit en place, nous proposons une stratégie basée sur l'examen de deux expectorations matinales pour les suspects négatifs et sur un diagnostic reposant sur un seul résultat positif.

RESUMEN

MARCO DE REFERENCIA : Un proyecto de control de la tuberculosis en Bangladesh.

OBJETIVO : Definir la eficiencia del número de campos microscópicos examinados y de la modalidad de toma de muestras de esputo, para el diagnóstico por examen microscópico directo de la expectoración.

MÉTODO : Las personas que realizan el control de calidad anotaron los números acumulativos de BAAR por 100 campos examinados. Se determinó el rendimiento suplementario de diagnóstico producido por diferentes estrategias de toma de muestras de esputo. Las series dudosas fueron controladas de nuevo y/o se examinaron nuevas muestras.

RESULTADOS : Se encontraron BAAR en los 100 primeros campos en el 99,6% de un total de 1.412 frotis positivos y en el 79,3% de los frotis que contenían un número escaso de bacilos. El examen de una tercera muestra produjo un máximo de 2,7% de incremento de la positividad. La estrategia más eficaz, que utiliza tres muestras matinales, produjo un 94,2% de baciloscopias

positivas en la primera muestra y 1,0% en la tercera. Aunque el 10% de los sospechosos no retornaron, sólo el 1,5% de los positivos se encontraban entre ellos y un número más importante de casos fueron confirmados y tratados. El valor predictivo positivo de un frotis positivo o que contiene un escaso número de bacilos es muy alto (99,2%).

CONCLUSIÓN : El rendimiento de la lectura de más de 100 campos por lámina o el examen de una tercera muestra es insuficiente como para justificar la sobrecarga de trabajo. El examen sólo de las muestras matinales de esputo es más eficiente y la toma de muestras no trae necesariamente inconvenientes para el paciente. El tratamiento puede comenzarse en base a una sola baciloscopia positiva. A condición que exista un eficiente sistema de control de calidad de la baciloscopia, se propone una estrategia basada en el examen de dos muestras matinales de esputo para los sospechosos negativos y en el diagnóstico basado en un solo resultado positivo.
