International Journal of Infectious Diseases 113S (2021) S55–S62

Contents lists available at ScienceDirect

International Journal of Infectious Diseases

journal homepage: www.elsevier.com/locate/ijid

# Increasing smear positive tuberculosis detection using a clinical score – A stepped wedge multicenter trial from Africa

Frauke Rudolf[a,b,*], Ebba Abate[c,d], Binyam Moges[d], Antonio Mateus Mendes[a], Mezgebu Yitayal Mengistu[e], Armando Sifna[a], Hikma Fekadu[f], Segenet Bizuneh[f], Thomas Schön[g,h], Christian Wejse[a,b,i]

[a] Bandim Health Project, INDEPTH Network, Apartado 861, Bissau, Guinea-Bissau
[b] Department of Infectious Diseases, Aarhus University Hospital, Denmark
[c] Ethiopian Public Health Institute (EPHI), Addis Ababa, Ethiopia
[d] Tropical & Infectious Diseases Research Centre, University of Gondar, Ethiopia
[e] Department of Health Systems and Policy, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Ethiopia
[f] Department of Internal Medicine, College of Medicine and Health Science, University of Gondar, Ethiopia
[g] Department of Microbiology and Infectious Medicine, Kalmar County Hospital, Sweden
[h] Department of Biomedical and Clinical Sciences, Division of Infectious Diseases, Linköping University, Sweden
[i] GloHAU, Center for Global Health, School of Public Health, Aarhus University, Denmark

ARTICLE INFO

*Keywords:*
Tuberculosis
Case detection
Stepped wedge
TBscore

ABSTRACT

*Background:* The Bandim TBscore is a clinical score that predicts treatment outcome in Tuberculosis (TB) patients and proved useful as an indicator of which healthcare-seeking adults to refer for sputum smear microcopy. We aimed to test in a randomized trial if the TBscore could be used to enhance the detection of smear positive (SP) TB.
*Methods:* We carried out a stepped wedge cluster-randomized trial at six health centers in Bissau, Guinea-Bissau, and Gondar, Ethiopia. The primary outcome was diagnostic yield for SP TB. Secondary outcomes were successful treatment and effect on overall 12 months mortality. The study was registered at the Pan African Clinical Trials Registry (PACTR201611001838365).
*Results:* We included 3571 adults. Overall, there was no effect of the intervention on SP PTB detected (OR 1.39 (95%CI 0.75 – 2.56). Analysis stratified by country, showed that the TBscore increased case detection in Gondar (OR 4.05 (95%CI 1.67 – 9.85)) but no effect was found in Bissau (OR 0.47 (95%CI 0.22 – 1.05)) where take-up was much lower. Overall mortality decreased during the intervention (HR 0.31 (95%CI 0.13-0.72)).
*Conclusion:* Using the TBscore for triage before smear microscopy may improve case detection and decrease mortality if there is sufficient laboratory capacity to increase sputum smears.


## Background

Despite an effective treatment and promising new developments easing diagnosis, Tuberculosis (TB) is still the leading cause of death from a single infectious agent (WHO, 2018). Although one of the key measures to control TB is early detection (Harries and Kumar, 2018) around three million patients remained undiagnosed in 2018 (WHO, 2019). Studies in health care seeking adults show a large load of undiagnosed and unsuspected TB (Claassens et al., 2013; Harries and Kumar, 2018; Kweza et al., 2018). This is underlined by autopsy studies (Bates et al., 2015) finding undiagnosed TB post mortem, indicating the failure to recognise and diagnose TB in the cascade of healthcare. Robust algorithms to ensure proper referral for diagnostics and awareness towards TB are needed (Waitt et al., 2011).

We have developed a simple clinical tool, which has predictive potential for TB diagnosis, can assess disease severity and identify patients with a high mortality risk (Rudolf et al., 2014; Rudolf et al., 2017; Wejse et al., 2008). The Bandim TBscore was developed in Guinea Bissau with the aim to ease clinical decisions based on systematic assessment of symptoms and signs, which may be difficult to oversee (Getnet et al., 2017; Ho et al., 2016; Rudolf et al., 2017). It has been validated in Guinea Bissau (Rudolf et al., 2013)

* Corresponding author at: Palle Juul-Jensens Boulevard 99, Department of inefctious diseases, 8200 Århus N.
*E-mail address:* frauke.rudolf@clin.au.dk (F. Rudolf).

https://doi.org/10.1016/j.ijid.2021.03.041

and in Ethiopia (Janols et al., 2012) and we have recently assessed its diagnostic potential using it to triage presumed TB (preTB) patients (cough regardless of duration/ sputum production/weight loss) at local health care centres (Rudolf et al., 2014; Rudolf et al., 2017). The TBscore requires a systematic approach and standardizes the clinical evaluation in settings were this is often not done (Getnet et al., 2017; Ho et al., 2016; Rudolf et al., 2017); this is helpful in the diagnostic process of a complicated disease as TB.

In Bissau, a prevalence survey indicated a considerable burden of undiagnosed TB (Bjerregaard-Andersen et al., 2010). In a recent study one third of the smear positive patients did not meet the outdated (Harries and Kumar, 2018; Kweza et al., 2018) standard criteria for being preTB cases, i.e. cough for more than two weeks (Rudolf et al., 2014). In Ethiopia it was shown in a major prevalence survey that there is a considerable burden of undiagnosed TB among HIV infected and that two thirds of smear positive TB cases in the community were undetected (Tadesse et al., 2011). Even if GeneXpert MTB/Rif is rolled out even in hig endemic settings, it still only contributes to less than 30% of bacteriologically confirmed cases in most sub-Saharan Afrian countries leaving smear microscopy as the major diagnostic tool at hand in most rural areas (WHO, 2019).

We here report a multi-centre cluster-randomised clinical trial, testing the effect of implementing the use of the Bandim TBscore at the health center level in high endemic areas on the diagnostic yield of sputum smear microscopy, the only widely available diagnostic test in low resource settings (Harries and Kumar, 2018; WHO, 2015). The stepped wedge design was applied to ensure that all participating health centres implemented the intervention. We hypothesize that screening health care seeking adults with the TBscore, adding an active case finding approach in the patient initiated pathway, may improve the diagnostic yield for smear positive pulmonary TB.

## Methods

### *Setting*

From January 2017 until end of March 2018 we recruited health care seeking adults at three public health centers (clusters) in Bissau, Guinea-Bissau, with a TB-incidence rate of 279/100.000 person years of observation (Lemvik et al., 2014), and Gondar, Ethiopia, with a TB-incidence rate of approximately 113/100.000 person years of observation (Alene and Clements, 2019).

### *Laboratory structure*

All clusters carried out their own smear microscopy except cluster 3 in Bissau. Here, samples were sent to cluster 1 by motorbike courier. None of the clusters used sputum concentration methods. All clusters in Bissau and cluster 2 in Gondar used direct sputum smear microscopy (DM) to detect acid-fast bacilli while cluster 1 and 3 in Gondar used LED-fluorescence microscopy (FM). In Bissau spot-morning samples were collected, while Gondar collected same day spot-spot samples. Patients with at least one positive smear were defined smear positive pulmonary TB (SP PTB)-cases.

No assessment on or changes to laboratory structure was done prior to the study since we aimed to assess the implementation of the TBscore in screening for TB in existing structures.

### *Participating patients*

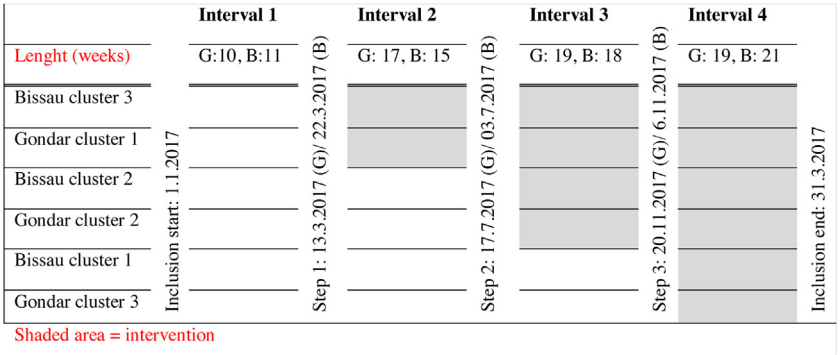
Patients were eligible when: ≥15 years old (≥18 years old for Gondar), seeking health care for current cough and/or weight loss and/or expectoration. All patients attending consultations at one of the six participating health centers were asked for the above symptoms and were offered enrollment upon confirmation of at least one.

### *Applied routines and follow-up*

All included patients were interviewed in the local language and clinically examined by the attending doctor or nurse at the health centers adult consultation upon first encounter. This included measurements for the TBscore, demographic questions on prior TB contacts, co-morbidities, health care seeking and previous TB treatment. In both phases of the trial, the patients were assessed in the same manner and if referred to diagnostic testing the sequence was the same, ie the trial in itself did not cause diagnostic delays. In Bissau all included patients were bled (venous blood sample (collected in a BD vacutainer, Spray-coated K2-EDTA)) for HIV-testing, while HIV-testing in Gondar followed the local guidelines. During the study period, there were only general guidelines to test adults seeking health care and it was up to the health professional at the Gondar clusters to decide whom to test.

One year after enrollment, all included patients were visited by local tracers to assess all-cause mortality. Patients with unknown address were called using the telephone number registered at inclusion.

**Table 1**
Study diagram.




Shaded area = intervention.

*Trial design and implementation of the TBscore*

We applied a cluster randomized stepped wedge design, where the three health centers in each country constituted the clusters. All clusters started in the control condition named enhanced usual care (EUC), which consisted of standard TB program diagnostics (i.e. sputum smear microscopy (SM), chest X ray and GeneXpert MTB/RIF following local guidelines) but with ensured availability of all reagents and access to X-ray at no cost to the patient during the first period of the study. The clusters were then switched to the intervention phase (I phase) following a random sequence (See Table 1). In total, there were four intervals, interval one with all clusters in EUC phase, one by one switching to the I phase through interval two to four and ending with all clusters in the I phase in interval four. The total length of each interval per setting is stated in Table 1.

During the EUC phase the health professionals at each health center decided whom to refer to further diagnostic work up, while this decision was dictated by a TBscore $\geq 3$ in the I phase. In both settings it is recommended to send patients for TB diagnostics when there are signs and symptoms or contact history suggesting possible TB disease.

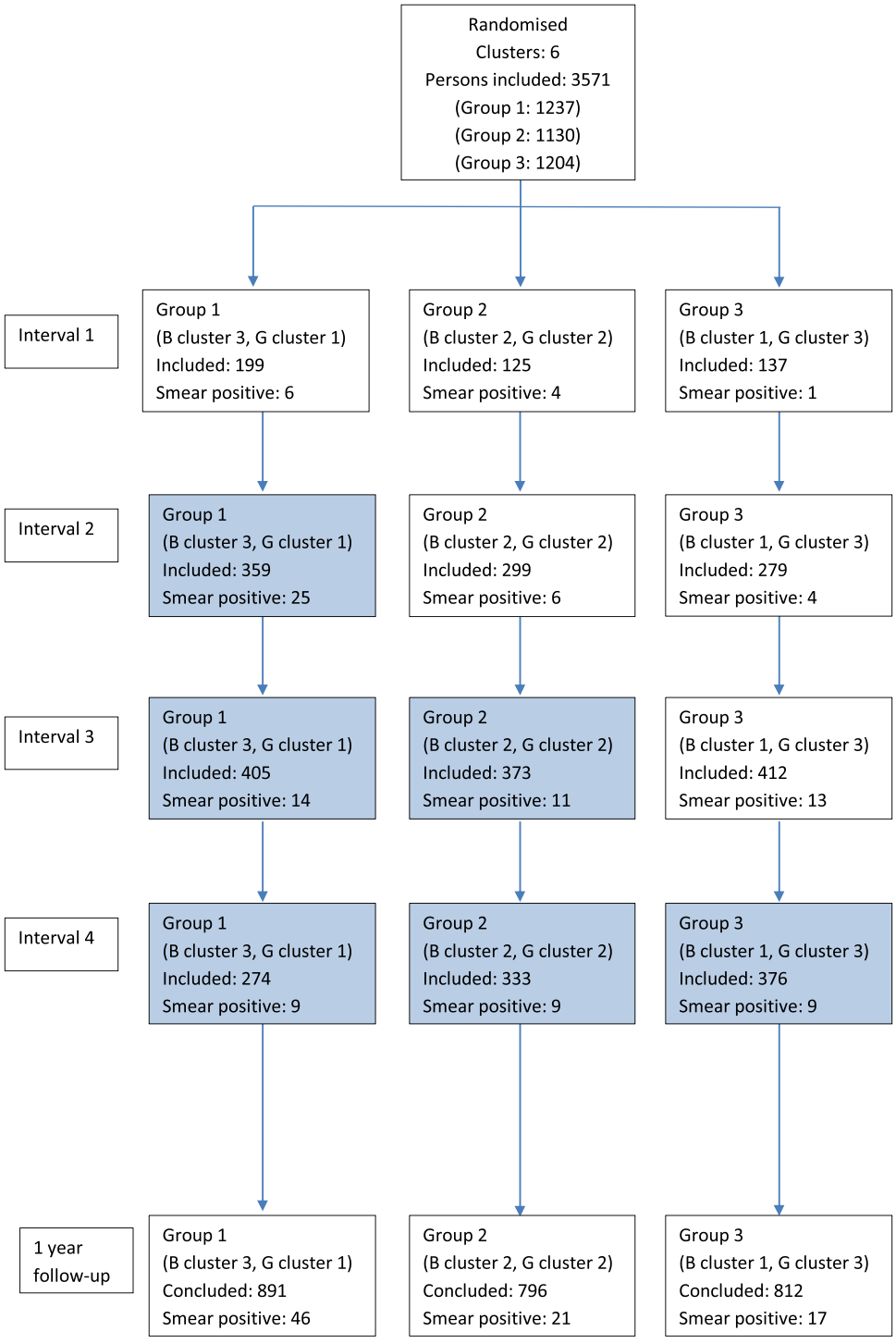



**Figure 1.** Diagnostic cascade by phase.

At each switch, workshops were held in both settings. Here, the TBscore was explained and the need for referral to SM was stressed.

All patients included were followed up after two weeks, to asses for persisting symptoms. If symptoms persisted and the initial sputum smear was negative or not done further diagnostic testing was carried out. The results of this diagnostic cascade will be published in a separate paper.

*The Bandim TBscore and TBscoreII*

The Bandim TBscore is based on five self-reported symptoms (cough, hemoptysis, dyspnoea, chest pain, night sweats) and six clinical signs (anemia, tachycardia, positive finding at lung auscultation, temperature > 37 °C, Body Mass Index (BMI) <18 and <16 and Mid-Upper-Arm-Circumference (MUAC) <220 mm and <200 mm) (Wejse et al., 2008). For the recently revised version, TBscoreII, we kept cough, dyspnoea and chest pain, anemia, BMI < 18 and <16, MUAC<220 mm and <200 mm (Rudolf et al., 2013). Each clinical variable contributes with one point while BMI and MUAC contribute with an additional point if <16 and <200 mm, respectively, hence the maximum score of 13 for the TBscore and eight for TBscoreII.

*Measurements and definitions*

In Bissau HIV antibody status was obtained with Determine HIV-1/2 (Alere Inc., Massachusetts, USA) and confirmed with First response (Premier Medical Corporation, Daman, India). In Gondar DetermineR HIV-1/2 (Ag/Ab Combo, Orlando, Florida, USA) and Colloidal Gold Device (Beijing Wantai Biological Pharmacy Enterprise Co., Ltd.) were used and positive results confirmed with Vikia HIV-1/2 (bioMérieux SA, Lyon, France) and Unigold (Trinity Biotech USA INC, NY, USA).

Pale inferior conjunctivae and palm were considered signs of anemia. The radial artery pulse was counted for 30 s and multiplied by two. A pulse >100 beats/minute was defined as tachycardia. Crepitation, rhonchi and subdued or complete absence of respiratory sounds were defined as positive findings at lung auscultation. A roll-up tape measure was used to assess height. MUAC was measured on the non-dominant arm at the mid-point between acromion and olecranon with a non-stretchable tape (TALC, Herts, UK). Temperature was measured axillary.

Following WHO's (WHO, 2019) definition, failure was defined as death during treatment, loss to follow up and/or treatment failure while treatment completed and cured were defined as non-failure. Patients transferred out were counted as missing outcomes.

*Statistics*

Data were analyzed with Stata Statistical Software version 12 (Stata Corporation, College Station, Texas, USA). The effect of the intervention was assessed by fitting a fixed effect spline model via logistic regression with robust variance estimates to account for clustering. The model was fitted overall and stratified by country. Chi square test was applied for categorical variables while differences in mean were assessed by a two independent samples t-test. Risk assessment was done fitting a binomial regression model.

Mortality risk was compared applying a Cox proportional hazards model. Proportionality of the categorical predictors was assessed with Kaplan–Meier curves and tested for equality with a log-rank test of equality across strata. Continuous predictors were assessed with univariate Cox proportional hazard regression and included if a p-value of 0.2 or less was found. Possible interactions were identified with the log rank test, none where found significant. The assumption of proportional hazards was met.

*Outcomes*

The primary outcome was to assess the diagnostic yield as of sputum smear positive TB in intervention compared with control EUC clinics.

Secondary outcomes were to assess the effect on treatment outcomes (for TB diagnosed patients) and all-cause mortality 12 month after first encounter (for all included).

*Sample size*

Sample size was determined using the "steppedwedge" function in Stata (Hemming and Girling, 2014). We estimated

**Table 2**
Baseline characteristics, by country.

| | | | Bissau | Gondar | Total | p-Value |
|---|---|---|---|---|---|---|
| Females, n (%) | | | 678 (55) | 1137 (49) | 1815 (51) | 0.001 |
| Mean age (IQR) | | | 33 (21–43) | 37 (23–50) | 36 (22–48) | <0.001 |
| Cough as inclusion-eligibility, n (%) | | | 1172 (95) | 2330 (99) | 3502 (98) | <0.001 |
| Consultation during the last 12 months, n (%) | | | 594 (48) | 604 (26) | 1198 (34) | <0.001 |
| Contact with TB patient | Household, n (%) | | 139 (11) | 164 (7) | 303 (8) | <0.001 |
| | Other, n (%) | | 101 (8) | 100 (4) | 201 (5) | <0.001 |
| Mean TBscore (95% CI) | | | 3·7 (3.6–3.9) | 4·0 (3.9–4.1) | 3.9 (3.8–4.0) | <0.001 |
| Mean TBscoreII (95% CI) | | | 2·4 (2.3–2.5) | 2·7 (2.7–2.8) | 2.6 (2.5–2.7) | <0.001 |
| Mean TBscore (95% CI) for patients sent for AFB, EUC | | | 5.1 (4.9–5.4) | 4.3 (4.1–4.5) | 4.5 (4.3–4.7) | <0.001 |
| Mean TBscoreII (95% CI) for patients sent for AFB, EUC | | | 3.0 (2.9–3.2) | 2.9 (2.7–3.0) | 2.9 (2.8–3.0) | 0.168 |
| Mean TBscore (95% CI) for patients sent for AFB, I | | | 3.8 (3.7–3.9) | 4.4 (4.3–4.5) | 4.2 (4.1–4.3) | <0.001 |
| Mean TBscoreII (95% CI) for patients sent for AFB, I | | | 2.6 (2.5–2.7) | 2.9 (2.8–2.9) | 2.8 (2.7–2.9) | <0.001 |
| HIV | Tested (%) | | 1151 (93) | 1300 (56) | 2451 (69) | <0.001 |
| | Infected | | 152 (12) | 35 (2) | 187 (5) | <0.001 |
| Sputum smear results | Overall | Sent (% of all incl) | 787 (64) | 1727 (74) | 2514 (70) | <0.001 |
| | | Done (% of all send) | 543 (69) | 1647 (95) | 2190 (87) | <0.001 |
| | | Positive (% of all done) | 42 (8) | 69 (4) | 111 (5) | 0.001 |
| | Enhanced usual care | Sent (% of all incl) | 178 (38) | 482 (49) | 660 (45) | <0.001 |
| | | Done (% of all send) | 148 (83) | 450 (93) | 589 (91) | <0.001 |
| | | Positive (% of all done) | 21 (14) | 13 (3) | 34 (6) | <0.001 |
| | Intervention | Sent (% of all incl) | 609 (80) | 1245 (92) | 1854 (88) | <0.001 |
| | | Done (% of all send) | 395 (65) | 1197 (96) | 1592 (86) | <0.001 |
| | | Positive (% of all done) | 21 (5) | 56 (5) | 77 (5) | 0.608 |

**Table 3**
Characteristics of smear positive TB patients, by country.

| | Bissau | | | | Gondar | | | |
|---|---|---|---|---|---|---|---|---|
| All included | | | | | | | | |
| | SP PTB patients | | Others | | SP PTB patients | | Others | |
| Mean Cough length (IQR) | 55 days (14–60) | | 20 days (3–10) | | 44 days (14–60) | | 36 days (7–30) | |
| Mean Age (IQR) | 33 years (26–40) | | 33 years (21–43) | | 29 years (26–33) | | 38 years (37–38) | |
| Male gender | 29 (69%) | | 531 (44%) | | 42 (61%) | | 1154 (51%) | |
| 12 months all cause mortality | EUC | I | EUC | I | EUC | I | EUC | I |
| | 1 (5%) | 0 | 15 (3%) | 8 (1%) | 0 | 2 (4%) | 3 (0·3%) | 0 |
| Tuberculosis patients | | | | | | | | |
| Cough less than 2 weeks | | 9 (21%) | | | | 9 (13%) | | |
| Referred to treatment | | 38 (90%) | | | | 60 (87%) | | |
| | EUC | | I | | EUC | | I | |
| Days with symptoms (IQR) | 73 (27–120) | | 116 (34–199) | | 100 (41–159) | | 38 (28–48) | |
| Mean TBscore (95% CI) | 6·4 (5·5–7·3) | | 5·8 (4·8–6·7) | | 6·0 (4·6–7.3) | | 6·1 (5·4–6·7) | |
| Mean TBscoreII (95% CI) | 3·6 (2·9–4·3) | | 3·6 (2·9–4·2) | | 4·1 (2·8–5.3) | | 3·7 (3·2–4·3) | |
| Successful outcome | 13 (62%) | | 10 (48%) | | 11 (85%) | | 44 (79%) | |
| Missing outcome | 3 (14%) | | 3 (14%) | | 2 (15%) | | 10 (17%) | |

that an increase in detection rate from 6% to 10% could be detected with a power of 80% and a significance level of 0.05. With three clusters in each country, we needed to include 152 patients per time interval (of 19 weeks) and cluster resulting in a target of 3648 total inclusions.

*Ethical considerations and consent*

Approval was obtained from the Guinean Ministry of Health and the Ethics review board of the University of Gondar. Consultative approval was obtained from the Central Ethical Committee in Denmark. All patients provided written informed consent and were offered HIV-testing with pre- and post-test counseling. HIV-infected individuals were accompanied to nearby Anti-Retroviral Treatment centers.

## Results

*Patient flow and characteristics*

From January 2017 until end of March 2018 we included 1238 patients in Bissau and 2333 in Gondar; out of them 765 and 1347 respectively entered in the intervention phase (Figure 1). Half of the included patients were female (55% in Bissau and 49% in Gondar) and were older in Gondar (37 years, (IQR 23 - 50)) than in Bissau (33 years (IQR 21 – 43)) (Table 2). Overall, 76% of the included patients were HIV tested, with the highest test-rates in Bissau. Of those tested 12% were HIV infected in Bissau, significantly more females than males (15% vs 9%, not shown). In Gondar the HIV-infection rate was 2%, with no difference by gender (not shown).

More patients in Bissau had sought help during the previous 12 months compared with patients from Gondar and more stated to share a household with a TB patient. Of the symptoms screened for concerning inclusion-eligibility (current cough and/or weight loss and/or expectoration), most patients stated to have all three, 43% in Gondar (n = 1012) and 39% (n = 484) in Bissau. Almost all included patients in Gondar had cough as one of the inclusion-eligibility symptoms (n = 2330, 99%) while it was seen in 95% (n = 1172) in Bissau (Table 2).

The patients diagnosed with SP TB stated a mean cough length of 55 days (IQR 14 – 60) in Bissau and 21% stated to have coughed less than two weeks. In Gondar, the TB patients had coughed for an average of 44 days (IQR 14 – 60), with 13% having coughed less than two weeks (Table 3).

*Sputum smear referral*

In Bissau 64% of all included patients were referred for SM during the trial while 74% in Gondar were referred. The referral rate increased from 38% during the EUC phase to 80% during the I phase in Bissau and from 49% to 92% in Gondar (Table 2). In both settings, there was no significant difference in percentage of males and females referred.

Overall, TBscores were slightly higher in Gondar and among patients referred to SM, both during both phases of the trial. Compared between the setting, the patients referred to SM during the EUC phase in Bissau had a higher TBscore than patients in Gondar (5.1 (95%CI 4.9 – 5.4)) vs 4.3 (95%CI 4.1 – 4.5)) while the opposite was seen during the I phase (Table 2).

Of all patients referred to SM, only 69% of patients had their smear done in Bissau as compared with 95% in Gondar. (Table 2) The relative risk declined clearly in the I phase of the trial in Bissau (RR 0.78 (95%CI 0.71 – 0.85) while it was stable in Gondar (RR1.03 (95%CI 1.00 – 1.06)). (Table 4) The decrease in Bissau was especially pronounced in cluster 3 where there was no on-site access to SM (Table 4).

In Gondar, the rate of positive smears among all SM done for clusters using FM was 5%, which was not significantly higher than the 3% in the cluster using DM. Comparing the countries revealed a significantly higher smear positivity rate in Bissau (7%) compared with Gondar (4%). This was especially pronounced in the EUC phase of the trial (14% vs 3%) and levelled out during the intervention phase (5% in both settings).

*Primary study outcome*

There was no overall effect of the TBscore intervention on sputum smear positivity rate (1.39 (95%CI 0.75 – 2.56)). In a stratified analysis by country, the effect of using the TBscore on increasing detection of smear positive TB was clear in Gondar (OR 4.05 (95%CI 1.67 – 9.85)), while it was not in Bissau (OR 0.47 (95%CI 0.22 – 1.05)). Overall, the rate of smear positive TB was 3% (n = 42 in Bissau and n = 69 in Gondar) in both settings among all patients included; with a clear increase during the I phase in Gondar from 1% to 4% (n = 13 to n = 56). For Bissau the rate slightly decreased from 4% to 3% (n = 21 in both phases) switching from EUC to I phase.

**Table 4**
Patients send for sputum smear microscopy, number of smears done and results. By cluster.

| **Bissau** | | | | | | |
|---|---|---|---|---|---|---|
| Cluster | | Sent/total (%) | SP (% of all sent) | SN (% of all sent) | Not done: Dry cough/No sample/Not registered (% of all sent) | Relative risk (95%CI) |
| 1 | Control | 104/274 (38) | 11 (11) | 68 (65) | 3 (3)/21 (20)/1(1) | 0·73 (0·61–0·88) |
| | Intervention | 135/172 (78) | 4 (3) | 71 (53) | 6 (4)/52 (39)/2 (1) | |
| 2 | Control | 59/127 (46) | 6 (10) | 49 (83) | 1 (2)/3 (5)/0 | 0·81 (0·73–0·89) |
| | Intervention | 240/272 (88) | 11 (5) | 170 (71) | 14 (6)/ 42 (17)/3 (1) | |
| 3 | Control | 15/72 (21) | 4 (27) | 10 (67) | 1 (6)/0/0 | 0·64 (0·54–0·76) |
| | Intervention | 234/321 (73) | 6 (3) | 133 (57) | 17 (7)/72 (31)/6 (2) | |
| **Gondar** | | | | | | |
| 1 | Control | 73/134 (54) | 2 (3) | 66 (90) | 5 (7)/0/0 | 1·05 (0·99–1·12) |
| | Intervention | 696/710 (98) | 42 (6) | 639 (92) | 15 (2)/0/0 | |
| 2 | Control | 119/301 (40) | 4 (4) | 98 (82) | 17 (14)/0/0 | 1·06 (0·98–1·15) |
| | Intervention | 361/430 (84) | 9 (2) | 319 (88) | 33 (10)/0/0 | |
| 3 | Control | 290/551 (53) | 7 (2) | 273 (94) | 10 (4)/0/0 | 1·05 (1·05–1·05) |
| | Intervention | 188/207 (91) | 5 (3) | 183 (97) | 0/0/0 | |

**Table 5**
Quantity of smears done and positivity rate per cluster and phase.

| Average quantity/month | Bissau 1 | Bissau 2 | Bissau 3 | Gondar 1 | Gondar 2 | Gondar 3 |
|---|---|---|---|---|---|---|
| Prior to the study | | | | | | |
| Sputum smear | 25·3 | – | 6·1 | 43·5 | 20·4 | 20·3 |
| Positive sputum smears | 4·0 | – | 1·2 | 3·3 | 0·8 | 1·3 |
| During enhanced usual care | | | | | | |
| Sputum smear | 30·0 | 20·3 | 18·0 | 59·5 | 25·4 | 31·9 |
| Positive sputum smears | 5·0 | 2·7 | 1·0 | 2·0 | 0·2 | 0·9 |
| During intervention | | | | | | |
| Sputum smear | 40·0 | 39·7 | 18·6 | 74·2 | 44·9 | 45·8 |
| Positive sputum smears | 5·0 | 3·0 | 1·7 | 3·5 | 1·4 | 0·4 |
| After the study | | | | | | |
| Sputum smear | 23·5 | 15·3 | 6·2 | 20·4 | 12·6 | 21·8 |
| Positive sputum smears | 2·3 | 1·8 | 0·6 | 1·8 | 0·2 | 1·0 |

*Secondary study outcomes - TB diagnosis, treatment outcomes and one-year mortality*

Overall, 111 patients were diagnosed with SP PTB, 71 males and 40 females. In both countries, the share of males was higher (69% in Bissau vs 64% in Gondar) (Table 3). The percentage of patients referred to treatment was 90% in Bissau, while it was 87% in Gondar. While the percentage of pre-treatment loss to follow-up was equal among males and females in Gondar, more females (23%) than males (3%) were lost in Bissau. There was no difference concerning age between the patients referred and the patients lost pre-treatment in neither setting. The treatment success (i.e. cured or treatment completed) was higher in Gondar with 85% (11/13) during the EUC phase and 79% (44/56) during the I phase as compared to Bissau, where it was 62% (13/21) and 48% (10/21) respectively. Fourteen to seventeen percent of the patients had an unknown outcome due to transferal outside of the covered area (Table 3).

We were able to obtain survival status at 12 months for 79% (n = 981) of the included patients in Bissau and 65·5% (n = 1518) in Gondar. In Gondar five patients (0.2%) had died at 12-month follow-up, while 24 (1.9%) had died in Bissau. There was no difference in the mortality rate between EUC and I phase in Gondar (0.7%, n = 3 vs 0.4%, n = 2) while the mortality in Bissau was significantly lower during the I phase compared with the EUC phase (3%, n = 16 vs 1%, n = 8) (Table 3). A subgroup analyses showed that this difference was especially pronounced among the HIV infected, where the mortality was 19% (10/52) during the EUC compared with 2% (2/100) during the I phase. However, neither the rate of smear positivity among HIV infected patients nor the rate of HIV infection in general was different between EUC and I phase in Bissau. (not shown) The mortality risk decreased 69% during the intervention (HR 0.31 (95%CI 0.13-0.72)) compared with the EUC phase, controlled for HIV-infection, smear positivity, sex and age.

*General quantity and smear positivity-rate at the health center laboratories before, during EUC, during intervention and after end of inclusion*

We assessed the quantity of smears done and positive sputum smears detected one year prior to the study, during the study and five months after the study at every participating health center. At all sites except cluster 2 in Bissau, where we were unable to obtain data, the number of sputum smears increased during the start of the study and further when switching to intervention. For 2/3 sites this was accompanied by an increase in positive sputum smears during the intervention period of the study. After the study was concluded the quantity of smears done and the average number of positive smears dropped except for cluster 3 in Gondar (Table 5).

## Discussion

Overall, there was no diagnostic yield in terms of sputum smear microscopy of the TBscore. However, the Bissau site did not meet the capacity needed to expand microscopy, which makes it difficult to interpret the results of the intervention here. In contrast, HCs in Gondar could cope with the increase in smear which resulted in an overall 4-fold yield of TBscore in terms of positive sputum smears. This provides useful lessons for programmatic management of TB case detection.

The demographics for the two sites differ (WHO), which is reflected in our data. In Bissau, the patients were younger, more

frequently female and more often HIV-coinfected. More of them had sought help for their symptoms and could recall having had contact with a TB patient.

The diagnostic process for TB in high endemic areas often starts with sputum smear microscopy (Harries and Kumar, 2018; WHO, 2015). In our study, the number sent on to smear microscopy increased markedly during the intervention, as expected and intended. One of the main differences between this study's two settings was the laboratory structure. While Bissau in the recent years had centralized SM diagnostics, each health center in Gondar had its own laboratory carrying out SM. Splitting the data by setting and analyzing it for possible explanations revealed that not all ordered sputum smears were effectuated. This number of "missed" SM was especially high at the health center in Bissau which depended on transport of sputum samples (cluster 3) to another health centers (cluster 1) for analysis due to previous guidance towards centralizing diagnostic capacity (WHO, 2015). Centralization is known to be problematic and decreases the number of smears actually done considerably as shown in studies from Tanzania (Kilale et al., 2013) and Malawi (Harries et al., 2004), where only 10% respective 44% of the samples were analyzed.

Another reason for the difference in smears carried out between the two settings could be the difference in sample collection (i.e. morning-spot in Bissau and spot-spot in Gondar). Relying on the patients returning to hand in sputum is a known reason for loss before diagnostic testing (WHO, 2015).

We did not find a difference between clusters applying FM and clusters using DM in Gondar, although FM is known to be more sensitive than DM (WHO, 2011b). However, we found a higher smear positivity rate during the EUC phase in Bissau. This cannot be explained by the difference in sampling (morning-spot vs spot-spot), since it was not seen during the intervention. It could indicate, however, that the selection of patients sent on to sputum smear was more restrictive in Bissau than in Gondar, which is supported by a higher TBscore/TBscoreII among patients referred to AFB in Bissau compared with Gondar during the EUC phase.

The effect of the intervention was pronounced in Gondar where the use of the TBscore as diagnostic trigger increased the number of SP PTB patients 4-fold. This is in line with findings from South Africa (Kweza et al., 2018) and Ethiopia (Arega et al., 2019) where 71% and 67% bacteriological confirmed TB patients, respectively, went undetected when seeking health care for TB related symptoms. Both Guinea-Bissau and Ethiopia have high case detection gaps which have not improved as expected in the recent years (WHO, 2019). Among the various reasons for this, the high threshold for suspecting TB may be the essential one. Although the well-known criterion "cough for more than two weeks" is outdated and it is recommended to be as inclusive as possible (WHO, 2011a), it still is this symptom the healthcare workers are brought up with and remember when they screen for TB. It is the most frequently used cut off for when to suspect TB (Arega et al., 2019), even though it results in missed cases (Harries and Kumar, 2018, Kweza et al., 2018, Rudolf et al., 2014Harries and Kumar, 2018Harries and Kumar, 2018, Kweza et al., 2018, Rudolf et al., 2014).

Of all diagnosed patients 10% were lost before reaching a treatment facility, known as pre-treatment loss to follow-up. This is in line with findings with a recent review; where between 4 – 38% of patients in African settings were lost before initiating treatment and lower than in most other settings (MacPherson et al., 2014). Treatment outcome was overall better in Gondar as compared to Bissau, albeit in line with the latest WHO data it is based on limited patient sample (WHO, 2019).

The overall mortality rate in both settings was low and the intervention seemed to decrease mortality. This was especially pronounced among HIV-infected patients in Bissau. HIV testing was part of the inclusion process in Bissau and the early detection and start of anti-retroviral treatment has probably prevented deaths – both during EUC, and I phase of the trial. We have previously found slightly higher rates of mortality in Bissau among preTB patients (Rabna et al., 2009; Rudolf et al., 2017).

The low sensitivity of direct smear microscopy without concentration may have left TB cases undetected (WHO, 2015), but it is the most widely used method. This makes the study applicable in settings all over the world as it identifies the patients contributing the most to the spread of TB. Patients lost to follow up, in particular at the Gondar site may influence the interpretation of treatment outcome and mortality.

We used a clinical score with the aim to increase TB case detection by targeting symptomatic patients in a systematic manner. In Bissau, the overwhelmed laboratories and structural impairment may have obscured an effect of the intervention. We know from earlier studies that using the TBscore as smear microscopy trigger will increase case detection (Rudolf et al., 2014; Rudolf et al., 2017) and are thus certain, that the reason for the missing effect is the ineffective laboratory structure. At first glance, relaying in existing structures hampers the effect of the study but it highlights the shortcomings in routine practice and pinpoints where to adjust in order to improve case detection.

## Conclusion

Systematically assessing adults seeking help for TB related symptoms can increase diagnostic yield considerably in sites with sufficient capacity.

## Funding

The study was supported by grants from the Novo Nordisk Foundation, Scandinavian Society for Antimicrobial Chemotherapy Foundation, Gondar University and the Swedish research council.

## Conflict of interest

The authors declare that they have no competing interests.

## Acknowledgements

We thank the participating patients in Gondar and Bissau and the project assistants, nurses and physicians in Gondar and Bissau for their good and consistent work. We thank DrMedSci, PhD, MSc Henrik Støvring at the department for public health, Aarhus University, for statistical advice.

## Transparency declaration

This article is part of a supplement entitled Commemorating World Tuberculosis Day March 24th, 2021: "The Clock is Ticking" published with support from an unrestricted educational grant from QIAGEN Sciences Inc.

## References

Alene KA, Clements ACA. Spatial clustering of notified tuberculosis in Ethiopia: a nationwide study. PLoS One 2019;14(8)e0221027.

Arega B, Tilahun K, Minda A, Agunie A, Mengistu G. Prevalence rate of undiagnosed tuberculosis in the community in Ethiopia from 2001 to 2014: systematic review and meta-analysis. Arch Public Health 2019;77:33.

Bates M, Mudenda V, Shibemba A, Kaluwaji J, Tembo J, Kabwe M, et al. Burden of tuberculosis at post mortem in inpatients at a tertiary referral centre in sub-Saharan Africa: a prospective descriptive autopsy study. Lancet Infect Dis 2015;15(5):544–51.

Bjerregaard-Andersen M, da Silva ZJ, Ravn P, Ruhwald M, Andersen PL, Sodemann M, et al. Tuberculosis burden in an urban population: a cross sectional tuberculosis survey from Guinea Bissau. BMC Infect Dis 2010;10:96.

Claassens MM, Jacobs E, Cyster E, Jennings K, James A, Dunbar R, et al. Tuberculosis cases missed in primary health care facilities: should we redefine case finding?. Int J Tuberc Lung Dis 2013;17(5):608–14.
Getnet F, Demissie M, Assefa N, Mengistie B, Worku A. Delay in diagnosis of pulmonary tuberculosis in low-and middle-income settings: systematic review and meta-analysis. BMC Pulm Med 2017;17(1):202.
Harries AD, Kumar AMV. Challenges and progress with diagnosing pulmonary tuberculosis in low- and middle-income countries. Diagnostics (Basel) 2018;8 (4).
Harries AD, Michongwe J, Nyirenda TE, Kemp JR, Squire SB, Ramsay AR, et al. Using a bus service for transporting sputum specimens to the Central Reference Laboratory: effect on the routine TB culture service in Malawi. Int J Tuberc Lung Dis 2004;8(2):204–10.
Hemming K, Girling A. A menu-driven facility for power and detectable-difference calculations in stepped-wedge cluster-randomized trials. Stata J 2014;14 (2):363–80.
Ho J, Fox GJ, Marais BJ. Passive case finding for tuberculosis is not enough. Int J Mycobacteriol 2016;5(4):374–8.
Janols H, Abate E, Idh J, Senbeto M, Britton S, Alemu S, et al. Early treatment response evaluated by a clinical scoring system correlates with the prognosis of pulmonary tuberculosis patients in Ethiopia: a prospective follow-up study. Scand J Infect Dis 2012;44(11):828–34.
Kilale AM, Ngowi BJ, Mfinanga GS, Egwaga S, Doulla B, Kumar AM, et al. Are sputum samples of retreatment tuberculosis reaching the reference laboratories? A 9-year audit in Tanzania. Public Health Action 2013;3(2):156–9.
Kweza PF, Van Schalkwyk C, Abraham N, Uys M, Claassens MM, Medina-Marino A. Estimating the magnitude of pulmonary tuberculosis patients missed by primary health care clinics in South Africa. Int J Tuberc Lung Dis 2018;22 (3):264–72.
Lemvik G, Rudolf F, Vieira F, Sodemann M, Ostergaard L, Rodrigues A, et al. Decline in overall, smear-negative and HIV-positive TB incidence while smear-positive incidence stays stable in Guinea-Bissau 2004-2011. Trop Med Int Health 2014;19(11):1367–76.
MacPherson P, Houben RM, Glynn JR, Corbett EL, Kranzer K. Pre-treatment loss to follow-up in tuberculosis patients in low- and lower-middle-income countries and high-burden countries: a systematic review and meta-analysis. Bull World Health Organ 2014;92(2):126–38.
Rabna P, Andersen A, Wejse C, Oliveira I, Gomes VF, Haaland MB, et al. High mortality risk among individuals assumed to be TB-negative can be predicted using a simple test. Trop Med Int Health 2009;14(9):986–94.
Rudolf F, Haraldsdottir TL, Mendes MS, Wagner AJ, Gomes VF, Aaby P, et al. Can tuberculosis case finding among health-care seeking adults be improved? Observations from Bissau. Int J Tuberc Lung Dis 2014;18(3):277–85.
Rudolf F, Lemvik G, Abate E, Verkuilen J, Schon T, Gomes VF, et al. TBscore II: refining and validating a simple clinical score for treatment monitoring of patients with pulmonary tuberculosis. Scand J Infect Dis 2013;45(11):825–36.
Rudolf F, Wagner AJ, Back FM, Gomes VF, Aaby P, Ostergaard L, et al. Tuberculosis case finding and mortality prediction: added value of the clinical TBscore and biomarker suPAR. Int J Tuberc Lung Dis 2017;21(1):67–72.
Tadesse T, Demissie M, Berhane Y, Kebede Y, Abebe M. Two-thirds of smear-positive tuberculosis cases in the community were undiagnosed in Northwest Ethiopia: population based cross-sectional study. PLoS One 2011;6(12)e28258.
Waitt CJ, Peter KB, White SA, Kampmann B, Kumwenda J, Heyderman RS, et al. Early deaths during tuberculosis treatment are associated with depressed innate responses, bacterial infection, and tuberculosis progression. J Infect Dis 2011;204(3):358–62.
Wejse C, Gustafson P, Nielsen J, Gomes VF, Aaby P, Andersen PL, et al. TBscore: Signs and symptoms from tuberculosis patients in a low-resource setting have predictive value and may be used to assess clinical course. Scand J Infect Dis 2008;40(2):111–20.
WHO. Country profile.
WHO. Early detection of tuberculosis – an overview of approaches, guidelines and tools. 2011.
WHO. Fluorescent light-emitting diode (LED) microscopy for diagnosis of tuberculosis policy – policy statement. 2011.
WHO. Implementing tuberculosis diagnostics – a policy framework. 2015.
WHO. The top 10 causes of death. 2018.
WHO. Global tuberculosis report 2019. 2019 Available from: http://www.who.int/tb/publications/global_report/en/. [Accessed 15 February 2020]..