INT J TUBERC LUNG DIS 21(11):1127–1133
© 2017 The Union
http://dx.doi.org/10.5588/ijtld.17.0139

Accuracy of giant African pouched rats for diagnosing tuberculosis: comparison with culture and Xpert® MTB/RIF

C. Mulder,*† G. F. Mgode,* H. Ellis,*‡ E. Valverde,*§ N. Beyene,* C. Cox,* S. E. Reid,¶# A. H. van't Hoog,† T. L. Edwards‡

***Anti-Persoonsmijnen Ontmijnende Product Ontwikkeling, Sokoine University of Agriculture, Morogoro, Tanzania; †Department of Global Health, Academic Medical Centre, Amsterdam Institute for Global Health and Development, University of Amsterdam, Amsterdam, The Netherlands; ‡Waikato University, Hamilton, New Zealand; §Vanderbilt University Medical School, Nashville, Tennessee; ¶Department of Medicine, University of Alabama at Birmingham, Birmingham, Alabama, USA; #Tuberculosis Department, Centre for Infectious Disease Research in Zambia, Lusaka, Zambia**

SUMMARY

SETTING: Enhanced tuberculosis (TB) case finding using detection rats in Tanzania.

OBJECTIVES: To assess the diagnostic accuracy of detection rats compared with culture and Xpert® MTB/RIF, and to compare enhanced case-finding algorithms using rats in smear-negative presumptive TB patients.

DESIGN: A fully paired diagnostic accuracy study in which sputum of new adult presumptive TB patients in Tanzania was tested using smear microscopy, 11 detection rats, culture and Xpert.

RESULTS: Of 771 eligible participants, 345 (45%) were culture-positive for *Mycobacterium tuberculosis*, and 264 (34%) were human immunodeficiency virus (HIV) positive. The sensitivity of the detection rats was up to 75.1% (95%CI 70.1–79.5) when compared with culture, and up to 81.8% (95%CI 76.0–86.5) when compared with Xpert, which was statistically significantly higher than the sensitivity of smear microscopy. Corresponding specificity was 40.6% (95%CI 35.9–45.5) compared with culture. The accuracy of rat detection was independent of HIV status. Using rats for triage, followed by Xpert, would result in a statistically higher yield than rats followed by light-emitting diode fluorescence microscopy, whereas the number of false-positives would be significantly lower than when using Xpert alone.

CONCLUSION: Although detection rats did not meet the accuracy criteria as standalone diagnostic or triage testing for presumptive TB, they have additive value as a triage test for enhanced case finding among smear-negative TB patients if more advanced diagnostics are not available.

KEY WORDS: diagnostic; olfactory detection; *Cricetomys ansorgei*; sensitivity; specificity

TUBERCULOSIS (TB) DETECTION using mammals, insects, and 'electronic noses' may be unique solutions to fill TB diagnostic needs in developing countries.[1–6] Trained African giant-pouched rats (*Cricetomys ansorgei*) are being used as a triage test for enhanced case finding in DOTS centres in Tanzania and Mozambique.[7] The rats work in a central laboratory and screen 100 sputum samples in 20 min, rapidly identifying false-negative smear samples at a cost of about US$1 per sample screened.[8]

Samples from smear-negative presumptive TB patients are screened using detection rats, and rat-positive samples are confirmed by concentrated light-emitting diode fluorescence microscopy (LED-FM). Employing this algorithm, over 10 000 additional patients have been confirmed since 2007,[8] translating into an annual 40% increase in smear-positive patients detected.[9,10]

The high-throughput and low cost of using detection rats position this method as an ideal triage test for enhanced or active case finding, rather than replacing existing diagnostic tests. Whether rats are valuable as a triage test for enhanced case finding in smear-negative presumptive patients depends on whether the accuracy of the diagnostic pathway is improved by adding them to the diagnostic cascade.[11,12]

The present study involved adults with the symptoms of pulmonary TB. The study objectives were 1) to determine the diagnostic accuracy of rats compared with culture and Xpert® MTB/RIF (Cepheid, Sunnyvale, CA, USA), and 2) to compare the accuracy of diagnostic algorithms that included rats, followed

Correspondence to: Christiaan Mulder, Department of Global Health, Academic Medical Centre, Amsterdam Institute for Global Health and Development, University of Amsterdam, Paasheuvelweg 25, 1105 BP Amsterdam, The Netherlands. e-mail: christiaan.mulder@kncvtbc.org

Article submitted 22 February 2017. Final version accepted 19 July 2017.

by either concentrated LED-FM or Xpert in smear-negative presumptive TB patients.

STUDY POPULATION AND METHODS

Study design and population

This was a fully paired accuracy study. Sputum was tested with rats, smear microscopy, Xpert, solid culture and liquid culture. Presumptive TB patients aged ⩾18 years were recruited consecutively between May 2014 and July 2015 at the DOTS centre, Mbagala Rangi Tatu in Temeke District, a high-throughput centre located in a low-income, densely populated area in Dar es Salaam, Tanzania.

Individuals were eligible if they were considered to be a presumptive TB patient as defined by the National Tuberculosis and Leprosy Programme guidelines. Individuals were excluded if their sputum samples were of insufficient quality (<5 ml or saliva only), they were on anti-tuberculosis treatment, they did not have documented HIV status or they were unwilling to be tested for HIV.

The study was powered to determine a 70% sensitivity of rats with 95% confidence intervals (CIs), allowing for 10% imprecision; 81 patients with culture-confirmed TB were needed, and assuming a 20% TB prevalence at least 405 were needed. A total of 1009 presumptive TB patients were needed to stratify the results by HIV status, assuming a 40% co-infection rate.

Laboratory procedures

Study participants were shown a poster to instruct them in how to produce a good quality sputum sample. Two sputum samples, one spot and one early morning, were collected and checked immediately for quality and quantity by a laboratory technician. If disqualified, participants were requested to provide another sample. Qualified samples (⩾5 ml non-salivary sputum) were refrigerated.

Both smears were prepared the same day for routine microscopy using Ziehl-Neelsen (ZN) at the DOTS Centre laboratory and graded using World Health Organization (WHO) criteria. Both samples were then homogenised mechanically using sterile glass beads and vortex mixing. Three millilitres of each homogenised sample were aliquoted and refrigerated until same-day shipping in cool boxes to the Central Tuberculosis Reference Laboratory (CTRL), Muhimbili National Hospital, Dar es Salaam, for ZN smear microscopy and culture inoculation. The remainder of each sputum sample was frozen at –20°C and shipped within 1 week in cool boxes to the Anti-Persoonsmijnen Ontmijnende Product Ontwikkeling (APOPO) Laboratory, Sokoine University of Agriculture, Morogoro, Tanzania, for rat evaluation, concentrated LED-FM and Xpert testing.

Solid and liquid cultures were prepared according to standard procedures at CTRL. Negative and positive control samples were included in each batch. All cultures were performed after blinding to clinical and diagnostic information.

Samples transported to the APOPO laboratory were heat-inactivated at 90°C for 30 min,[13] and then stored at –20°C until the day of rat evaluation. Eleven male rats ($n = 8$) assessed all samples. Animals were between 2.5 and 8 (median 4.2) years of age at study commencement. Animals begin training at 6 weeks, and progress through training phases as the criteria for diagnostic accuracy are met. Training requires 8–9 months for all animals, after which they are considered to have passed APOPO's internal accreditation process. All animals had been assessing operational samples for ⩾1.3 years to a pre-defined standard (>70% sensitivity compared with smear microscopy as the reference standard) before study commencement. There was no difference in performance between individual rats: false-negatives and false-positives were idiosyncratic and at a comparable rate across animals.

The 11 rats evaluated the samples sequentially in a rectangular chamber (205 cm long × 55 cm wide × 55 cm high) with 10 holes in the floor. A positive rat response ('indication') was taught through operant conditioning during training, and was defined as the rat holding the nose in the scent hole for ⩾3 s.[5] Handlers observed the rat as it moved along the line of holes sniffing the samples, recording all positive indications. Correct positive indications had to be rewarded occasionally to maintain accurate responding, so positive and negative samples collected routinely from DOTS centres were included in the evaluation sessions, and indication of DOTS-positive samples was rewarded as they are in routine screening operations.[14] Of the 100 samples presented to the rats per daily evaluation session, 20 were study samples, whereas the remaining 80 were smear-positive and smear-negative control samples from other DOTS centres in Dar es Salaam.

After rat evaluation, samples were analysed at the APOPO laboratory using Xpert according to the manufacturer's instructions, and by concentrated LED-FM through chemical processing (sodium hydroxide and sodium chloride) and concentration (centrifugation at 3000 × g).[15] Sputum sediments were used to prepare smears stained using auramine-O. All laboratory personnel were blinded to clinical information and other test results.

Statistical analysis and case definitions

In the primary analysis, a TB patient was defined as having at least one positive culture for *Mycobacterium tuberculosis*, including co-infections with non-tuberculous mycobacteria (NTM). A TB-negative patient was defined as having negative culture on

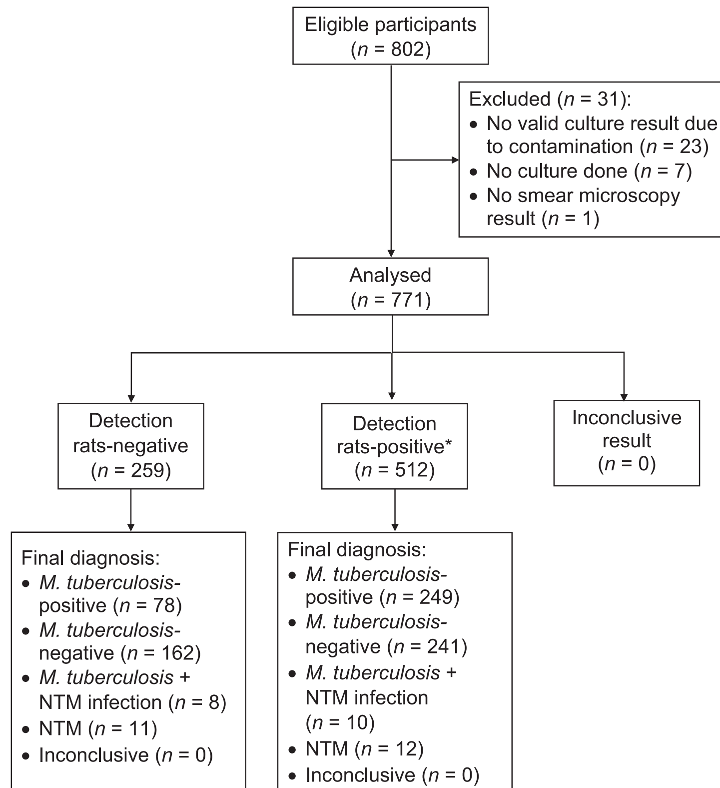

Figure 1 Flowchart of study participants. *Assuming a team of 11 rats with a threshold of 1 rat to consider someone as being positive. NTM = non-tuberculous mycobacteria.

both samples or being exclusively infected with NTM. The sensitivity, specificity and predictive values of teams consisting of either seven or 11 rats were assessed, as these team sizes are standard in APOPO's routine operation and research.[16]

The results of seven rats out of the total group of 11 were randomly selected 10 000 times to derive mean accuracy estimates with 95%CIs. Two rat thresholds were used to consider someone as 'rat-positive'—either $\geqslant 1$ or $\geqslant 2$ rats the team of rats had to indicate on a sample. The mean individual rat performance was also calculated. Individuals with missing culture, rat or smear results were excluded.

Analyses were stratified using smear microscopy (at the DOTS centre) and HIV status. Receiver operating characteristics (ROC) curves and the areas under the curve were calculated. Differences in accuracy between smear-positives and smear-negatives, and between HIV-positives and HIV-negatives, were compared using Pearson χ^2 tests. Statistical differences between rats and other diagnostics were compared using McNemar's tests. In a similar secondary analysis, Xpert was used as the reference standard. Finally, the number of true- and false-positives and negatives of diagnostic algorithms with and without rats were compared using the McNemar tests. For these analyses, only smear-negatives were considered, with culture as the reference standard. Statistical analyses were performed using SPSS v20 (IBM, Armonk, NY, USA) and R v3.3.1 (R Computing, Vienna, Austria). TDR Diagnostics Evaluation Expert Panel[17] and Standards for Reporting of Diagnostic Accuracy (STARD) guidelines were followed.[18]

Ethical considerations

The study protocol was approved by the Medical Research Coordinating Committee, Dar es Salaam (NIMR/HQ/R.8c/Vol.II/495). The Office of Laboratory Animal Welfare from the National Institutes of Health, Bethesda, MD, USA, approved an Animal Welfare Assurance (A5720-01). Written informed consent was obtained from literate study participants. Oral informed consent was attested by an impartial witness in cases of illiteracy. The study was conducted according to Good Clinical Practice guidelines, the principles of the Declaration of Helsinki, and Tanzanian laws and regulations.

RESULTS

A total of 802 eligible participants were enrolled; 30 individuals without culture or contaminated results and 1 without the DOTS centre smear result were excluded (Figure 1). Of the 771 participants, 52% were female; the mean age was 38.9 years (standard

Table 1 Sensitivity, specificity and predictive values compared with culture for different rat thresholds (not stratified for HIV)

Number of rats in team and indication thresholds	Sensitivity % (95%CI) All culture-positive (n = 345)	Smear-positive, culture-positive (n = 54)	Smear-negative, culture-positive (n = 291)	Specificity % (95%CI) Non-TB (n = 426)	PPV %	NPV %
Random 7 rats						
⩾1 rat	69.6 (66.1–72.8)	83.1 (79.6–85.2)	67.0 (62.9–70.8)	48.2 (46.2–50.7)	52.1	66.2
⩾2 rats	55.0 (51.3–58.6)	76.2 (70.4–81.5)	51.0 (46.7–54.6)	67.0 (63.4–70.7)	57.4	64.8
All 11 rats						
⩾1 rat	75.1 (70.1–79.5)	85.2 (72.3–93.0)	73.2 (67.7–78.1)	40.6 (35.9–45.5)	50.6	66.8
⩾2 rats	63.5 (58.1–68.5)	81.5 (68.1–90.3)	60.1 (54.2–65.8)	57.7 (52.9–62.5)	54.9	66.1
Mean 1 rat	38.8 (36.5–41.1)	63.6 (61.1–66.1)	34.2 (31.8–36.7)	79.5 (71.6–84.8)	60.5	61.6

HIV = human immunodeficiency virus; CI = confidence interval; TB = tuberculosis; NPV = negative predictive value; PPV = positive predictive value.

deviation 13.3), and HIV prevalence was 34%. In total, 345 participants (45%) were TB culture-positive, 54 (16%) of whom were DOTS centre smear-positive.

The highest sensitivity (85.2%, 95%CI 72.3–90.0) was found with a team of 11 rats, with a threshold of one among smear-positive patients (Table 1). Irrespective of the rat team size, the lower the threshold of rat indications used, the higher the sensitivity, whereas the higher the threshold, the higher the specificity (Figure 2). Rat sensitivity was statistically significantly higher in smear-positives than in smear-negatives, regardless of the threshold used (all $P \leqslant 0.05$), except if a threshold of one rat was used in a team of 11 rats.

There was no statistically significant difference in the sensitivity or specificity of rats comparing HIV-positives and HIV-negatives; this was independent of the size of the rat team (Figure 3, Appendix).*

With Xpert as the reference standard, the rat sensitivity was between 76% (95%CI 72.4–78.7) and 82% (95%CI 76.0–86.5), and was as high as 94% (95%CI 82.1–98.4) in smear-positive patients, whereas the specificity was similar to when culture was considered the reference standard (Table 2).

With a one-rat threshold, the sensitivity of the rats was statistically significantly higher than that of smear microscopy at the DOTS centre (16%, $P < 0.001$), ZN smear microscopy conducted at CTRL (24%, $P < 0.001$), concentrated LED-FM (33%, $P < 0.001$) and Xpert (52%, $P < 0.001$), irrespective of the team size of the rats (Table 3). The corresponding specificity of the rats was statistically significantly lower than those of the other diagnostics ($P < 0.001$).

The use of detection rats, followed by concentrated LED-FM, would result in a slightly lower yield than using LED-FM only (23% vs. 25%), but would reduce the number of false-positive LED-FM results from 3% to 2%, and would save 33% of LED-FM examinations needed for enhanced case finding in smear-negatives (Table 4). Rat detection followed by Xpert would lead to a statistically significantly lower yield than using Xpert only (39% vs. 45%, $P < 0.001$), but would reduce false-positives from 11% to 6%, which is statistically significant ($P < 0.001$), and would save 31% of the Xpert cartridges needed. Rat detection followed by Xpert would result in a statistically significantly higher yield than rats followed by concentrated LED-FM (39% vs. 23%, $P < 0.001$), but also in three times more false-positives (6% vs. 2%, $P = 0.003$).

DISCUSSION

The present study showed that teams of rats detected TB in up to 75% of culture-positive patients, including 73% of smear-negatives. Up to 82% of the Xpert positives were detected by the rats, including 78% of smear-negatives. Our study suggests that the sensitivity of rats is independent of HIV and is statistically significantly higher in smear-positives than in smear-negatives.[16] The ability to rapidly detect most Xpert-positives, including among smear-negative TB cases and people living with HIV, could provide a major improvement in TB control in high TB-HIV prevalence areas.

The present study confirms the notion that rat teams are more sensitive than smear microscopy for TB detection.[14] The low sensitivity of microscopy (24%) underscores the urgent need to replace microscopy at the point-of-care level.[19] The Xpert sensitivity was low, as most patients in our study were smear-negative, which affected its sensitivity.[20] Because this is the reality in sub-Saharan Africa, more TB diagnostic research is needed to ensure more smear-negative patients are detected.

An enhanced case-finding algorithm in smear-negative presumptive TB patients using rats as a triage test and Xpert as the diagnostic test would detect 66% more patients than when rat detection is

* The appendix is available in the online version of this article, at http://www.ingentaconnect.com/content/iuatld/ijtld/2017/00000021/00000011/art00007

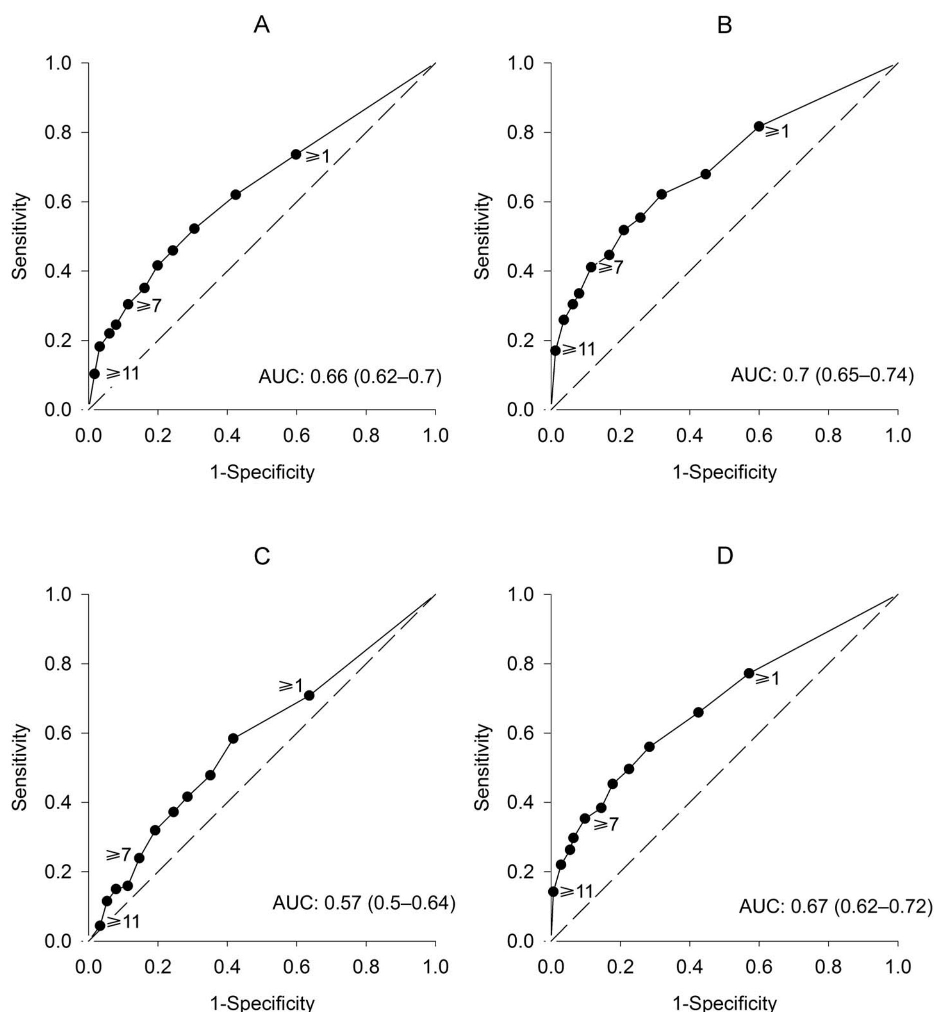

Figure 2 ROC curves for 11 rats and associated AUC with 95% confidence intervals: **A)** compared with culture; **B)** compared with Xpert® MTB/RIF; **C)** among HIV-positive patients, with culture as the reference standard; and **D)** among HIV-negative patients, with culture as the reference standard; ⩾1, ⩾7 and ⩾11 refer to the number of rats out of a team of 11 rats that at least need to give a positive indication to consider someone to be positive. AUC = area under the curve; ROC = receiver operating characteristic; HIV = human immunodeficiency virus.

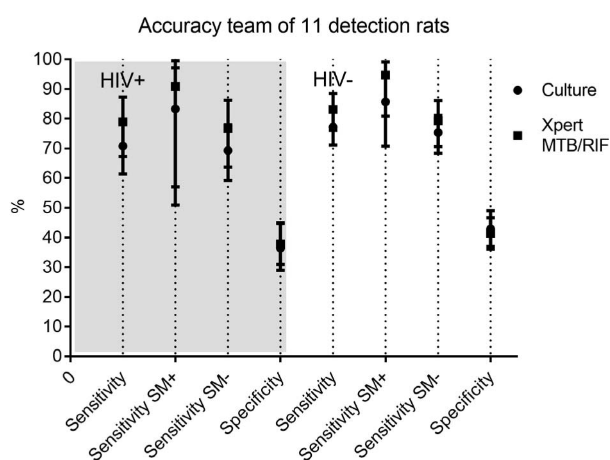

Figure 3 Sensitivity and specificity of a team of 11 detection rats with a rat threshold of 1 against culture and Xpert® MTB/RIF as the reference standard, stratified by HIV status (shaded: HIV-positive) and smear status. HIV = human immunodeficiency virus; + = positive; − = negative; SM = smear microscopy.

followed by concentrated LED-FM. In our study, using the rats before Xpert would have almost halved the number of false-positive Xpert tests, and would have saved 31% of the Xpert cartridges needed. However, the high sensitivity associated with large rat teams and liberal indication criteria comes at the expense of low specificity. Our study also showed that triaging with rats regardless of smear status is not justified, as the rats missed 15% of the smear-positive patients.

The accuracy of rat detection presented here does not meet the WHO-recommended standards to replace smear microscopy or serve as a point-of-care triage test in all presumptive TB patients.[19] Based on the present findings and those of other studies, the rats are best positioned in high-throughput situations in which evaluation of all samples by other technology, such as Xpert, would be too expensive.

Table 2 Sensitivity, specificity and predictive values compared with Xpert® MTB/RIF* for different rat thresholds

	Sensitivity % (95%CI)			Specificity % (95%CI)		
Number of rats in team and indication thresholds	All Xpert-positive (*n* = 225)	Smear-positive, Xpert-positive (*n* = 49)	Smear-negative, Xpert-positive (*n* = 176)	Non-TB (*n* = 540)	PPV %	NPV %
Random 7 rats						
⩾1 rat	75.9 (72.4–78.7)	91.9 (87.8–93.9)	71.5 (67.6–75.0)	47.1 (44.8–49.8)	37.4	82.4
⩾2 rats	62.0 (57.8–65.8)	85.7 (81.6–89.8)	55.5 (51.1–59.7)	65.2 (62.0–68.7)	42.6	80.5
All 11 rats						
⩾1 rat	81.8 (76.0–86.5)	93.9 (82.1–98.4)	78.4 (71.5–84.1)	40.0 (35.9–44.3)	36.2	84.1
⩾2 rats	68.0 (61.4–74.0)	89.8 (77.0–96.2)	61.9 (54.3–69.0)	55.2 (50.9–59.4)	38.7	80.5
Mean 1 rat	46.4 (44.0–48.8)	71.6 (69.4–74.1)	39.4 (36.9–42.0)	78.8 (70.6–84.0)	47.7	77.9

* Total = 765; 6 participants with missing Xpert results excluded.
CI = confidence interval; TB = tuberculosis; NPV = negative predictive value; PPV = positive predictive value.

Maintenance requirements for pouched rats are minimal.[21] Given that a group of rats can screen several hundred samples per day, per-sample evaluation costs can be extremely low. We have estimated the cost at approximately US$1 per sample, but the specific cost will depend upon sample throughput, as well as the cost of supplies and labour at each location. In the absence of an 'ideal' triage test, a very inexpensive test could be valuable as a triage test even at modest specificity (40–50%); a cost-effectiveness study is in preparation.[22]

Triaging with rats for active case finding in a low TB prevalence population is currently being studied. Regardless of the outcomes of this research, more studies are needed to optimise the accuracy of rat detection, especially the sensitivity of individual rats, which was only 39% in our study. Higher individual sensitivity would allow for improved specificity as stricter group indication criteria could be applied. This would make it more feasible and affordable to confirm with Xpert for both enhanced and active case finding.[23] Studies have shown that the specificity of rat detection is minimally affected by NTM.[16,24]

One study limitation was that previously treated patients were not excluded. This may have resulted in lower specificity for Xpert.[25] Over-decontamination of culture seemed not to have affected Xpert specificity, as only 3% of culture-negatives were smear-positive, and in total 45% of the participants were culture-positive. Another limitation is that the large volume of sputum required for study inclusion may have inflated the TB prevalence and may have excluded individuals with less severe symptoms. Finally, as the enrolled number of study participants was lower than anticipated, the study was not powered to stratify the results by HIV status.

CONCLUSION

Detection rats did not meet the accuracy criteria as a stand-alone TB diagnostic or as a triage test for presumptive TB. However, to accelerate TB elimination in low-resource settings, using rats to enhance case finding among smear-negatives should be promoted.[26] Until fast, accurate and sustainable TB diagnostic tests are developed, validated, endorsed and scaled up, detection rats can play a complemen-

Table 3 Sensitivity, specificity and predictive values of smear microscopy (ZN), concentrated LED-FM, and Xpert® MTB/RIF

	Sensitivity							
	Culture-positive *n/N* (%) (95%CI)	Smear-positive, culture-positive *n/N* (%) (95%CI)	Smear-negative, culture-positive *n/N* (%) (95%CI)	HIV+	HIV–	Specificity % (95%CI)	PPV %	NPV %
ZN at DOTS centre (*n* = 771)	54/345 (15.7) (12.1–20.0)	NA	NA	12/113 (10.6) (5.9–18.2)	42/232 (18.1) (13.5–23.8)	412/426 (96.7) (94.4–98.1)	79.4	58.6
ZN at CTRL (*n* = 771)	84/345 (24.3) (20.0–29.3)	NA	NA	21/113 (18.6) (12.1–27.2)	63/232 (27.2) (21.7–33.5)	413/426 (96.9) (94.7–98.3)	86.6	61.3
LED-FM at APOPO (*n* = 771)	115/345 (33.3) (28.4–38.6)	NA	NA	28/113 (24.8) (17.4–34.0)	87/232 (37.5) (31.3–44.1)	412/426 (96.7) (94.4–98.1)	89.1	64.2
Xpert at APOPO (*n* = 765)	177/340 (52.1) (46.6–57.5)	45/53 (84.9) (71.9–92.8)	132/287 (46.0) (40.1–51.9)	55/112 (49.1) (39.6–58.7)	122/228 (53.5) (46.8–60.1)	377/425 (88.7) (85.2–91.5)	78.7	69.8

ZN = Ziehl-Neelsen; LED-FM = light-emitting diode fluorescence microscopy; CI = confidence interval; HIV = human immunodeficiency virus; + = positive; – = negative; NPV = negative predictive value; PPV = positive predictive value; NA = not available; CTRL = Central Tuberculosis Reference Laboratory; APOPO = Anti-Persoonsmijnen Ontmijnende Product Ontwikkeling.

Table 4 Diagnostic outcomes of four enhanced case finding diagnostic algorithms in smear-negative culture-positive/negative presumptive TB patients

	True-positive (yield) *n/N* (%)	False-negative *n/N* (%)	False-positive *n/N* (%)	True-negative *n/N* (%)	Tests saved *n*	Positive likelihood ratio	Negative likelihood ratio
LED-FM	74/291 (25)	217/291 (75)	12/412 (3)	400/412 (97)	NA	8.7	0.8
Use of detection rats, followed by LED-FM	68/291 (23)	223/291 (77)	9/412 (2)	403/412 (98)	234/703 (33)	10.7	0.8
Xpert	132/291 (45)	159/291 (55)	44/412 (11)	367/412 (89)	NA	4.2	0.6
Use of detection rats, followed by Xpert	113/291 (39)	175/291 (60)	25/412 (6)	387/412 (94)	218/703 (31)	6.4	0.7

TB = tuberculosis; LED-FM = light-emitting diode fluorescence microscopy.

tary part in increasing case detection in low-resource, high TB-HIV burden populations for the foreseeable future.

Acknowledgements

The authors thank all study participants, clinicians, laboratory technicians and rat handlers. The study was supported financially by the UBS Optimus Foundation, Zurich, Switzerland, and the Carraresi Foundation, Vancouver, BC, Canada.

Conflicts of interest: none declared.

References

1 Suckling D M, Sagar R L. Honeybees *Apis mellifera* can detect the scent of *Mycobacterium tuberculosis*. Tuberculosis (Edinb) 2011; 91: 327–328.
2 Phillips M, Basa-Dalay V, Blais J, et al. Point-of-care breath test for biomarkers of active pulmonary tuberculosis. Tuberculosis (Edinb) 2012; 92: 314–320.
3 Bruins M, Rahim Z, Bos A, van de Sande W W, Endtz H P, van Belkum A. Diagnosis of active tuberculosis by e-nose analysis of exhaled air. Tuberculosis (Edinb) 2012.
4 Nakhleh M K, Jeries R, Gharra A, et al. Detecting active pulmonary tuberculosis with a breath test using nanomaterial-based sensors. Eur Respir J 2014; 43: 1522–1525.
5 Weetjens B, Mgode G F, Machang'u R S, et al. African pouched rats for the detection of pulmonary tuberculosis in sputum samples. Int J Tuberc Lung Dis 2009; 13: 737–743.
6 Kolk A H, van Berkel J J, Claassens M M, et al. Breath analysis as a potential diagnostic tool for tuberculosis. Int J Tuberc Lung Dis 2012; 16: 777–782.
7 Beyene N, Mahoney A, Cox C, et al. APOPO's tuberculosis research agenda: achievements, challenges, and prospects. Tanzan J Health Res 2012; 14: 121–130.
8 Mulder C, Mgode G, Reid S E. Tuberculosis diagnostic technology: an African solution…think rats. Afr J Lab Med 2017; 6: a420.
9 Poling A, Mahoney A, Beyene N, et al. Using giant African pouched rats to detect human tuberculosis: a review. Pan Afr Med J 2015; 21: 333.
10 Poling A, Valverde E, Beyene N, et al. Active tuberculosis detection by pouched rats in 2014: more than 2,000 new patients found in two countries. J Appl Behav Anal 2017; 50: 165–169.
11 Bossuyt P M, Irwig L, Craig J, Glasziou P. Comparative accuracy: assessing new tests against existing diagnostic pathways. BMJ 2006; 332: 1089–1092.
12 Hayen A, Macaskill P, Irwig L, Bossuyt P. Appropriate statistical methods are required to assess diagnostic tests for replacement, add-on, and triage. J Clin Epidemiol 2010; 63: 883–891.
13 Doig C, Seagar A L, Watt B, Forbes K J. The efficacy of the heat killing of *Mycobacterium tuberculosis*. J Clin Pathol 2002; 55: 778–779.
14 Mahoney A, Weetjens B, Cox C, et al. Pouched rats' detection of tuberculosis in human sputum: comparison to culturing and polymerase chain reaction. Tuberc Res Treatment 2012; 1–6. https://www.hindawi.com/journals/trt/2012/716989/ Accessed August 2017.
15 Steingart K R, Ng V, Henry M, et al. Sputum processing methods to improve the sensitivity of smear microscopy for tuberculosis: a systematic review. Lancet Infect Dis 2006; 6: 664–674.
16 Reither K, Jugheli L, Glass T R, et al. Evaluation of giant african pouched rats for detection of pulmonary tuberculosis in patients from a high-endemic setting. PLOS ONE 2015; 10: e0135877.
17 TDR Diagnostics Evaluation Expert Panel, Banoo S, Bell D, Bossuyt P, et al. Evaluation of diagnostic tests for infectious diseases: general principles. Nat Rev Microbiol 2010; 8 (Suppl): S17–S29.
18 Bossuyt P M, Cohen J F, Gatsonis C A, Korevaar D A, STARD Group. STARD 2015: updated reporting guidelines for all diagnostic accuracy studies. Ann Transl Med 2016; 4: 85.
19 World Health Organization. High-priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting. WHO/HTM/TB/2014.18. Geneva, Switzerland: WHO, 2014.
20 Steingart K R, Schiller I, Horne D J, Pai M, Boehme C C, Dendukuri N. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. Cochrane Database Syst Rev 2014; 1: CD009593.
21 Cooper R G. Care, husbandry and diseases of the African giant rat (*Cricetomys gambianus*). J S Afr Vet Assoc 2008; 79: 62–66.
22 Van't Hoog A H, Cobelens F, Vassall A, et al. Optimal triage test characteristics to improve the cost-effectiveness of the Xpert MTB/RIF assay for TB diagnosis: a decision analysis. PLOS ONE 2013; 8: e82786.
23 World Health Organization. Systematic screening for active tuberculosis: an operational guide. Geneva, Switzerland: WHO, 2015.
24 Mgode G F, Weetjens B J, Nawrath T, et al. Diagnosis of tuberculosis by trained African giant pouched rats and confounding impact of pathogens and microflora of the respiratory tract. J Clin Microbiol 2011; 50: 274–280.
25 Steingart K R, Schiller I, Dendukuri N. In reply to 'False-positive Xpert® MTB/RIF assays in previously treated patients'. Int J Tuberc Lung Dis 2015; 19: 366–367.
26 World Health Organization. End TB Strategy. Geneva, Switzerland: WHO, 2014.

APPENDIX

Table A.1 Sensitivity and specificity values compared with culture for different rat thresholds, stratified by the HIV status of study participants

Number of rats in team and indication thresholds	HIV-positive patients				HIV-negative patients					
	Sensitivity in culture-positive % (95%CI)			Specificity % (95%CI)	Sensitivity in culture-positive % (95%CI)			Specificity % (95%CI)	Sensitivity *P* value*	Specificity *P* value*
	All culture-positives ($n = 113$)	Smear-positive ($n = 12$)	Smear-negative ($n = 101$)	($n = 151$)	All ($n = 232$)	Smear-positive ($n = 42$)	Smear-negative ($n = 190$)	($n = 275$)		
Random 7 rats										
⩾1 rat	64.9 (59.3–69.0)	79.4 (66.7–83.3)	63.3 (56.4–68.3)	45.3 (41.1–50.3)	71.8 (68.5–75.0)	84.2 (81.0–85.7)	69.1 (65.3–72.6)	49.9 (47.3–54.2)	0.187	0.362
⩾2 rats	49.7 (45.1–54.0)	69.1 (58.3–75.0)	47.3 (42.6–52.5)	64.6 (61.6–68.9)	57.5 (53.4–61.2)	78.1 (71.4–83.3)	53.0 (47.9–56.8)	68.3 (64.4–72.0)	0.173	0.431
All 11 rats										
⩾1 rat	70.8 (61.4–78.8)	83.3 (50.9–97.1)	69.3 (59.2–77.9)	36.4 (28.9–44.7)	77.2 (71.1–82.3)	85.7 (70.8–94.1)	75.3 (68.4–81.1)	42.9 (37.0–49.0)	0.194	0.192
⩾2 rats	58.4 (48.8–67.5)	75.0 (42.8–93.3)	56.4 (46.2–66.2)	58.3 (50.0–66.2)	66.0 (59.4–72.0)	83.3 (68.0–92.5)	62.1 (54.8–69.0)	57.5 (51.4–63.3)	0.168	0.863
Mean 1 rat	32.6 (29.5–35.4)	58.3 (54.8–61.9)	29.5 (26.0–32.7)	76.8 (68.4–81.8)	41.8 (39.7–43.8)	65.2 (62.6–67.7)	36.7 (34.2–38.9)	81.1 (72.4–87.0)	0.097	0.297

* Determined using the Pearson χ^2 test comparing individuals who were HIV-positive with those who were HIV-negative.
HIV = human immunodeficiency virus; CI = confidence interval.

Table A.2 Sensitivity and specificity of LED-FM and Xpert® MTB/RIF in detection rat-positive study participants

	Sensitivity in detection rat-positives*					Specificity % (95%CI)
	Culture-positive ($n = 259$) *n* (%) (95%CI)	Smear-positive, culture-positive ($n = 46$) *n* (%) (95%CI)	Smear-negative, culture-positive ($n = 213$) *n* (%) (95%CI)	Smear-negative, culture-positive, HIV+ ($n = 70$) *n* (%) (95%CI)	Smear-negative, culture-positive, HIV– ($n = 143$) *n* (%) (95%CI)	Rat+, culture-negative ($n = 253$) *n* (%) (95%CI)
LED-FM	108 (41.7) (35.7–48.0)	40 (87.0) (73.1–94.6)	68 (31.9) (25.8–38.7)	16 (22.9) (14.0–34.7)	52 (36.4) (28.6–44.9)	242 (95.7) (92.1–97.7)
Xpert	155 (59.9) (53.6–65.9)	42 (91.3) (78.3–97.2)	113 (53.1) (46.1–59.9)	36 (51.4) (39.3–63.4)	77 (53.9) (45.3–62.2)	224 (88.5) (83.8–92.1)

* A threshold of one rat used in a team of 11 rats.
LED-FM = light-emitting diode fluorescence microscopy; CI = confidence interval; HIV = human immunodeficiency virus; + = positive; – = negative.

Table A.3 Sensitivity and specificity of detection rats using Xpert® MTB/RIF as the reference standard, stratified by HIV and smear status

	HIV-positive patients				HIV-negative patients					
	Sensitivity in Xpert-positive patients % (95%CI)			Specificity % (95%CI)	Sensitivity in Xpert-positive patients % (95%CI)			Specificity % (95%CI)		
Number of rats in team and indication thresholds	All Xpert-positives ($n = 71$)	Smear-positive ($n = 11$)	Smear-negative ($n = 60$)	Non-TB ($n = 191$)	All ($n = 154$)	Smear-positive ($n = 38$)	Smear-negative ($n = 116$)	Non-TB ($n = 349$)	Sensitivity *P* value*	Specificity *P* value*
All 11 rats										
⩾1 rat	78.9 (67.3–87.3)	90.9 (57.1–99.5)	76.7 (63.7–86.2)	37.7 (30.9–45.0)	83.1 (76.1–88.5)	94.7 (80.9–99.1)	79.3 (70.6–86.1)	41.3 (36.1–46.6)	0.443	0.419

* Determined using the Pearson χ^2 test comparing individuals who were HIV-positive with those who were HIV-negative.
HIV = human immunodeficiency virus; CI = confidence interval; TB = tuberculosis.

RÉSUMÉ

CONTEXTE : Amélioration de la détection des cas de tuberculose (TB) grâce à des rats détecteurs en Tanzanie.

OBJECTIF : Evaluer l'exactitude diagnostique des rats détecteurs comparés à la culture et à l'Xpert® MTB/RIF et comparer les algorithmes améliorés de recherche des cas, notamment les rats détecteurs chez des patients tuberculeux présumés à frottis négatif.

SCHÉMA : Une étude parfaitement appariée de fiabilité diagnostique dans laquelle les crachats de nouveaux patients adultes présumés tuberculeux en Tanzanie ont été testés par microscopie de frottis, par 11 rats détecteurs, par culture et par Xpert.

RÉSULTATS : Sur 771 participants éligibles, 345 (45%) ont eu une culture positive à *Mycobacterium tuberculosis* et 264 (34%) ont été positifs au virus de l'immunodéficience humaine (VIH). La sensibilité des rats détecteurs a atteint 75,1% (IC95% 70,1–79,5) comparée à la culture et jusqu'à 81,8% (IC95% 76,0–86,5) comparée à l'Xpert, ce qui a été statistiquement significativement plus élevé que la sensibilité de la microscopie de frottis. La spécificité correspondante a été de 40,6% (IC95% 35,9–45,5) comparée à la culture. La fiabilité des rats a été indépendante du statut VIH. Le recours aux rats comme système de triage suivi de l'Xpert aboutirait à un rendement statistiquement plus élevé que les rats suivis de LED-FM, tandis que le nombre de faux positifs serait significativement plus faible qu'en utilisant l'Xpert seul.

CONCLUSION : Bien que les rats détecteurs n'aient pas répondu aux critères de fiabilité comme méthode de diagnostic autonome ou test de triage de présomption de TB, ils ont une valeur supplémentaire comme test de triage pour améliorer la détection des cas parmi les patients tuberculeux à frottis négatif dans des contextes où des diagnostics plus avancés ne sont pas disponibles.

RESUMEN

MARCO DE REFERENCIA: La búsqueda intensificada de casos de tuberculosis (TB) mediante la utilización de ratas de detección en Tanzania.

OBJETIVOS: Evaluar la exactitud diagnóstica del método de detección con ratas en relación con el cultivo y la prueba Xpert® MTB/RIF y comparar los algoritmos de detección intensificada de casos que comportan este método, en los pacientes con presunción clínica de TB y baciloscopia negativa.

MÉTODO: Se llevó a cabo un estudio de exactitud diagnóstica plenamente apareado, en el cual se examinó la muestra de esputo de pacientes nuevos con presunción clínica de TB, mediante la baciloscopia, la exposición a 11 ratas de detección, el cultivo y la prueba Xpert.

RESULTADOS: De los 771 participantes idóneos, 345 obtuvieron un cultivo positivo (45%) para *Mycobacterium tuberculosis* y 264 fueron positivos frente al virus de la inmunodeficiencia humana (VIH; 34%). La sensibilidad de la detección con ratas alcanzó 75,1% al compararla con el cultivo (IC95% 70,1–79,5) y hasta 81,8% cuando se comparó con la prueba Xpert (IC95% 76,0–86,5); estos resultados fueron superiores a la sensibilidad de la baciloscopia, con una diferencia estadísticamente significativa. La especificidad fue 40,6% (IC95% 35,9–45,5) comparada con el cultivo. La exactitud de la detección con ratas fue independiente de la situación frente al VIH. La práctica de la detección con ratas seguida de la prueba Xpert aportaría un rendimiento estadísticamente superior al rendimiento del examen con las ratas seguido por una baciloscopia en microscopia de fluorescencia con LED y el número de positivos falsos sería significativamente menor que cuando solo se utiliza la prueba.

CONCLUSIÓN: Si bien la detección con ratas no alcanza los criterios de exactitud como medio diagnóstico independiente ni como prueba de selección en el lugar de la consulta de los casos con presunción de TB, este método puede ofrecer un valor agregado como prueba de selección en la estrategia de búsqueda intensificada de casos en los pacientes con TB y baciloscopia negativa, en los entornos donde no se cuenta con medios diagnósticos más avanzados.