

# SCIENTIFIC REPORTS



OPEN

## Sensitivity and specificity of Cobas TaqMan MTB real-time polymerase chain reaction for culture-proven *Mycobacterium tuberculosis*: meta-analysis of 26999 specimens from 17 Studies

Received: 06 October 2015  
Accepted: 12 November 2015  
Published: 09 December 2015

Nobuyuki Horita<sup>1</sup>, Masaki Yamamoto<sup>1</sup>, Takashi Sato<sup>1</sup>, Toshinori Tsukahara<sup>1</sup>, Hideyuki Nagakura<sup>1</sup>, Ken Tashiro<sup>1</sup>, Yuji Shibata<sup>1</sup>, Hiroki Watanabe<sup>1</sup>, Kenjiro Nagai<sup>1</sup>, Kentaro Nakashima<sup>1</sup>, Ryota Ushio<sup>1</sup>, Misako Ikeda<sup>1</sup>, Kentaro Sakamaki<sup>2</sup>, Takashi Yoshiyama<sup>3</sup> & Takeshi Kaneko<sup>1</sup>

Since 2010, studies on the diagnostic accuracy of COBAS TaqMan MTB (CTM) have been frequently reported with an unignorable discrepancy. The key inclusion criterion for this systematic review was original studies that could provide sufficient data for calculating the sensitivity and the specificity of CTM for *M. tuberculosis* (TB) or *M. tuberculosis* complex. The reference test was *Mycobacterium* culture. We used bivariate model for meta-analyses. Of the 201 candidate articles, we finally identified 17 eligible articles. Concerning the respiratory specimens, 1900 culture positive specimens and 20983 culture negative specimens from 15 studies were assessed. This provided the summary estimate sensitivity of 0.808 (95% CI 0.758–0.850) and the summary estimate specificity of 0.990 (95% CI 0.981–0.994). The area under curve was 0.956. The diagnostic odds ratio was 459 (95% CI 261–805, I<sup>2</sup> 26%). For the smear positive respiratory specimens, the sensitivity was 0.952 (95% CI 0.926–0.969) and the specificity was 0.916 (95% CI 0.797–0.968). For the smear negative respiratory specimens, the sensitivity and the specificity were 0.600 (95% CI 0.459–0.726) and 0.989 (95% CI 0.981–0.993), respectively. The diagnostic accuracy was poorer for the non-respiratory specimens, than for the respiratory specimens, but was acceptable. We believe that the information obtained from this study will aid physicians' decision making.

In 2013, approximately 1.5 million people died of *Mycobacterium tuberculosis* (*M. tuberculosis*). To provide appropriate antibiotic treatment, rapid and accurate detection of the *M. tuberculosis*, which has been a clinical challenge for more than a century, is crucial<sup>1</sup>. Classical acid-fast stain and mycobacterium culture still play a central role in the diagnosis of tuberculosis (TB), although they have considerable limitations: the acid-fast stain lacks specificity and the culture needs weeks of the incubation time<sup>2,3</sup>. Nucleic acid amplification tests, such as polymerase chain reaction (PCR) which was developed in 1983, are now an indispensable tool in the TB diagnosis as it provides quick and specific results to clinicians within hours<sup>2,3</sup>.

The COBAS Amplicor PCR system was the first commercially available automated nucleic acid amplification analyzer<sup>4</sup>. For the clinical evaluation, the COBAS Amplicor presented excellent sensitivity for smear-positive

<sup>1</sup>Department of Pulmonology, Yokohama City University Graduate School of Medicine. 3-9, Fukuura, Kanazawa, Yokohama, Japan. <sup>2</sup>Department of Biostatistics, Yokohama City University Graduate School of Medicine. 3-9, Fukuura, Kanazawa, Yokohama, Japan. <sup>3</sup>Division of Respiratory Medicine, Japan-Anti-Tuberculosis Association Fukujiji Hospital. 3-1-24, Matsuyama, Kiyose, Tokyo, Japan. Correspondence and requests for materials should be addressed to N.H. (email: nobuyuki\_horita@yahoo.co.jp)

specimens, moderate sensitivity for smear-negative specimens, generally good specificity, and acceptable PCR inhibition rate<sup>5–7</sup>. Later, Roche Diagnostics also developed a novel real-time PCR system, COBAS TaqMan, which is more widely accepted now. For the diagnosis of TB, COBAS TaqMan MTB (CTM) focuses on a segment of the 16S rRNA gene<sup>8</sup>. Since 2010, studies on the diagnostic accuracy of CTM have been frequently reported with an unignorable discrepancy<sup>9–15</sup>. Some previous studies suggested that these heterogeneous results may be caused by inconsistent specimen types and smear status<sup>9–11</sup>. Even though manufacturer instruction limited CTM application only for the respiratory specimens<sup>16</sup>, researchers often evaluated the diagnostic accuracy of CTM for the non-respiratory specimens. It is because we have limited resources for extra-pulmonary TB. Therefore, in the current systematic review and meta-analysis, we assessed the diagnostic accuracy of CTM for the non-respiratory specimens in addition to the respiratory specimens.

## Methods

We conducted this meta-analysis following pre-specified protocol (UMIN000018272) following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement<sup>17</sup> and standard guidelines for systematic review of diagnostic test accuracy<sup>18</sup>. Quality of included studies was assessed by Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS2)<sup>19</sup>. Approval of Institutional Review Board was waived for reviewing nature of the current study.

## Study Search

Two investigators (HN, NK) independently and systematically searched PubMed, PMC, Cochrane Library, and Web of Science database. Our last search for these medical data base was done on June 13<sup>th</sup>, 2015. Hand search was conducted by checking references of already included studies and published guidelines. In addition, Google Scholar was searched. Articles in authors' reference list were also considered candidates. The inclusion criterion for this study was original studies that could provide sufficient data for calculating the sensitivity and/or the specificity of CTM for *M. tuberculosis* or *M. tuberculosis* complex. The reference test was culture proven *M. tuberculosis*. The composite reference standard combining the results of several reference tests was not allowed. The sensitivity was defined as true positives/(true positives + false negatives). The specificity was defined as true negatives/(true negatives + false positives). Cobas TaqMan probes other than CTM was excluded. For instance, TaqMan probe designed for IS6110 was excluded. No age limit was set. No publication-date limit was set. No language restriction was set, as long as an article provided sufficient data in English-written title, abstract, figure, or table. Besides full articles, short articles were also considered for inclusion. Report with sample size less than 20 were excluded. HIV status was not considered. We did not included a study that reported the respiratory and the non-respiratory specimens collectively as we have to evaluate them separately. A study that reported only the sensitivity or the specificity was excluded because we could not perform the bivariate meta-analysis for such a study. Duplicate use of the same data was examined carefully and excluded.

## Search formulas

We used following formulas. PubMed: (TaqMan OR (Cobas real PCR)) AND tuberculosis AND (sensitivity OR specificity OR positive OR negative OR (predictive value)). Cochrane library: (tuberculosis OR mycobacterium OR TB) AND (diagnosis OR sensitivity OR specificity OR PCR OR TaqMan). Web of Science: TaqMan AND tuberculosis AND (sensitivity OR specificity OR (predictive value)). PMC: TaqMan MTB.

## Cobas TaqMan MTB preparation

Regardless of specimen types, i.e. the respiratory or the non-respiratory specimens, specimens were generally liquefied and decontaminated with N-Acetyl-L-Cysteine-Sodium Hydroxide (NALC/NaOH), according to manufacturer instruction<sup>16</sup>. However, some investigators used other preparation methods (Table 1).

## Outcome

The numbers of positive and negative results were counted specimen based, and not person based.

CTM rarely output invalid results suggesting PCR inhibition, contamination, internal control signal being out of range, or low positive control. Therefore, samples with invalid results were excluded preceding the sensitivity and specificity evaluations because the manufacturer instruction recommended to repeat the entire run for the PCR procedure, and because most original researches reported only the numbers of positive and negative results. The pooled invalid rate was estimated separately from diagnostic accuracy.

We divided specimens in each study based on specimen types because the respiratory and the non-respiratory specimens should be evaluated independently. This was because CTM was originally designed only for the respiratory specimens<sup>19</sup>. We classified the specimens into the respiratory and the non-respiratory categories based on the description in each original report. The respiratory specimens included sputum, bronchial/tracheal aspirate, bronchial/tracheal lavage, broncho-alveolar lavage. The non-respiratory specimens included lymph node, articular fluid, ascites fluid, abscess/pus, urine, cerebrospinal fluid, other tissue, and other body fluids. Gastric fluid and pleural fluid were classified either into the respiratory or the non-respiratory specimen in each original report.

In addition to classification based on the specimen origin mentioned above, we performed subgroup analyses focusing on the smear status because we had acknowledged that the smear status affects the diagnostic accuracy of CTM.

## Statistics

We drew a paired forest plot and bivariate model summary receiver operating characteristics curves (SROC), which showed potential trade-offs between the sensitivity and the specificity. The summary estimate of the sensitivity and the specificity was also obtained from the bivariate model<sup>20</sup>.

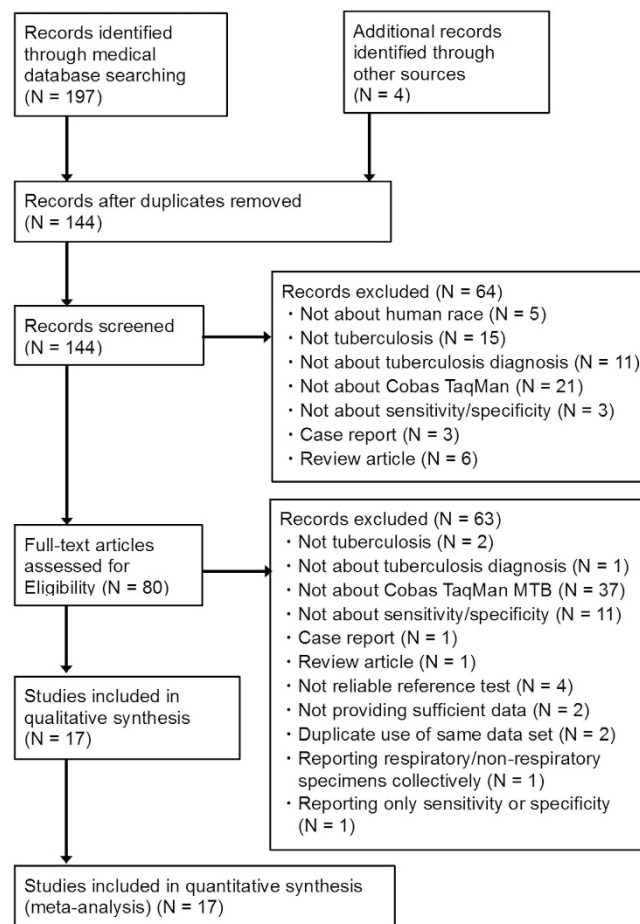
Author Year	Country (Income class)	Recruitment	Study design	Acid-fast stain	Specimen preparation for Cobas TaqMan	Culture	TB confirmation	Respiratory specimen	Non-respiratory specimen	Specimen number
Antonenka 2013	Germany (A)	Specimen bank	R, F, I	AR	NALC-NaOH	MGIT, LJ	Genotype CM, Genotype MTBC	sp, ba		R 116 NR 0
Bloemberg 2013	Switzerland (A)	Tertiary care diagnostic center	P, I	AR, ZN	NALC-NaOH	MGIT, 7H11	16SrRNA	sp, bal, ba, other	ti, ur, csf, ln, as, ab	R 829 NR 280
Causse 2011	Spain (A)	?		AR	NALC-NaOH	7H9, LJ, liquid medium	?		csf, pf, ar, as, ln, ti, gf	R 0 NR 340
Chandran 2010	India (C)	Tb Control Program	Short article	?	NALC-NaOH	LJ	Genotype MTBC	sp		R 72 NR 0
Cho 2015	Korea (A)	University hospital	I	?	NALC-NaOH	2%Ogawa, MGIT	multiplex PCR	sp, ba, bl, bal, ts	pus, pf, csf, ur, ti, ab	R 2384 NR 626
Choi 2013	Korea (A)	?	F, Korean, #1	AR, ZN	NALC-NaOH	MGIT, 3%Ogawa	Seeplex	sp, bal		R 619 NR 0
Huh 2015	Korea (A)	Tertiary care hospital	R #2	AR, ZN	NALC-NaOH (2%)	Liquid, solid medium	cordF, MPT64, MTB-ID V3	?		R 6772 NR 0
Ikegame 2012	Japan (A)	?	R	AR	NaOH	?	?	sp, bal, bl		R 1745 NR 0
Jönsson 2015	Sweden (A)	University hospital laboratory	R, I	?	NALC-NaOH	LJ, MGIT	AccuProbe, 16SrRNA GenoType MTBC,	sp, bal, bl	bf, ti, ur, gf, csf	R 2388 NR 1005
Kim JH 2011	Korea (A)	Tertiary care hospital	P, I, #2	ZN	NALC-NaOH	3%Ogawa	Amplicor, history	sp, bal, pf	ar, as, bf, ti	R 92 NR 0
Kim JW 2013	Korea (A)	?	F, #1	?	NALC-NaOH	MGIT	PCR assay Smear	bal, ba, sp	ab, csf, ar, as, pf, ti, ur	R 360 NR 65
Lee 2013	Taiwan (B)	University hospital & General hospital	P, I	AR, Ki	NALC-NaOH	MGIT, LJ	Middlebrook7H11	sp		R 586 NR 0
Lim 2014	Korea (A)	University hospital	P, B	ZN	NALC-NaOH (4%)	3%Ogawa	?	sp, bal		R 1167 NR 0
Linasmita 2012	Thailand (B)	University hospital	P, I	?	QIAGEN	MGIT	?		cervical ln	R 0 NR 73
Park 2013	Korea (A)	Tertiary care hospital	P, B, #2	AR, ZN	NALC-NaOH (2%)	MGIT, 3%Ogawa	cord F, conventional PCR MPT64, MTB-ID V3,	sp, bal		R 320 NR 0
Tortoli 2012	Italy (A)	Reference center laboratory	R	AR	NALC-NaOH (1%)	MGIT, LJ	GenoType, 16SrRNA	bl, gf, sp	ti, bf, csf, pf, ab, ur	R 4340 NR 1727
Yang 2011	Taiwan (B)	University hospital laboratory	P	AR, Ki	NALC-NaOH	LJ, MGIT, 7H11	MPT64, GenoType, Mycobacterium CM assay	su, ba, bal		R 1093 NR 0

**Table 1. Characteristics of the included studies.** Income class: The World Bank income classification. A, high-income economy; B, upper-middle-income economy; C, lower-middle-income economy; D, low-income economy. Taiwan was classified as China for this table. Study design: P, prospective study; R, retrospective study; F, using frozen specimen; B, TaqMan MTB was conducted blindly; I, invalid rate was reported; if these information was not provided, left blank. Short article and non-English article were also mentioned. #1, Choi and Kim JW used data from same hospital in different time. #2, Huh, Kim JH, and Park used data from same hospital in different time. Acid-fast stain: ZN, Ziehl-Neelsen; AR, auramine-rhodamine; Ki, Kinyoun. Specimen preparation for Cobas TaqMan: NALC-NaOH, N-acetyl-L-cysteine with sodium hydroxide; NaOH, sodium hydroxide. Culture: MGIT, Mycobacteria Growth Indicator Tube; LJ, Löwenstein-Jensen; 7H9, Middlebrook 7H9 Broth; 7H10/11, Middlebrook 7H10/11 Agar; BacT/A, BacT/Alert. Tb confirmation: 16SrRNA, 16S rRNA gene sequence; cordF, cord formation. Respiratory specimen: sp, sputum; ba, broncheal/tracheal aspirate; bl, broncheal/tracheal lavage; bal, bronchialalveolar lavage; gf, gastric fluid; ts, throat swab. Non-respiratory specimen: ln, lymph node; pf, pleural fluid; ar, articular fluid; as, ascite fluid; ab, abscess/pus; ur, urine; csf, cerebrospinal fluid; bf, body fluid; bl, blood; gf, gastric fluid; ti, other tissue sample. Specimen number: R, number of the respiratory specimens; NR, number of the non-respiratory specimens.

To assess the overall accuracy, we calculated diagnostic odds ratio (DOR) using DerSimonian-Laird random-model and the area under SROC curve (AUC) using the bivariate model. We estimated the pooled invalid rate using the simple random model with generic inverse variance.

The heterogeneities for the DOR among studies and between subgroups were evaluated using  $I^2$  statistics with the following interpretation:  $I^2 = 0$ , no heterogeneity;  $0 < I^2 < 25$ , mild heterogeneity;  $25 \leq I^2 < 50$ , moderate heterogeneity;  $50 \leq I^2 < 75$ , strong heterogeneity;  $75 \leq I^2 < 90$ , considerable heterogeneity;  $90 \leq I^2$ , extreme heterogeneity<sup>21</sup>. Publication bias was not evaluated as this is not usually recommended in the meta-analysis for the diagnostic test accuracy<sup>18</sup>.

The paired forest plot and the random-effect model meta-analysis were performed using Reviewing Manager ver. 5.3 (Cochrane Collaboration, Oxford, UK). Following commands of “mada” package on free software R were used: “madauni” for DOR and “reitsma” for SROC, AUS and summary estimate for the sensitivity and the specificity<sup>22,23</sup>.



**Figure 1. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).**

Specimen type	Study n	Specimen n	Sensitivity (95% CI)	Specificity (95% CI)	DOR (95% CI), I <sup>2</sup>	AUC
Respiratory	15	22883	0.808 (0.758–0.850)	0.990 (0.981–0.994)	459 (261–805), 26%	0.956
Respiratory, smear+	7	821	0.952 (0.926–0.969)	0.916 (0.797–0.968)	269 (66–1104), 0%	0.961
Respiratory, smear–	7	11621	0.600 (0.459–0.726)	0.989 (0.981–0.993)	132 (71–243), 14%	0.971
Non–respiratory	7	4116	0.586 (0.437–0.721)	0.984 (0.955–0.994)	86 (34–217), 7%	0.898

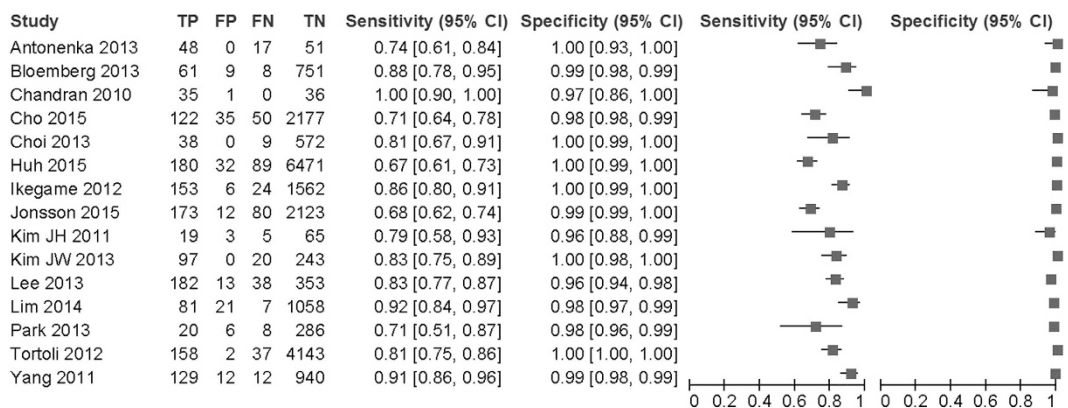
**Table 2. Summary of the results.** DOR: diagnostic odds ratio. I<sup>2</sup>: Higgins' heterogeneity. AUC: area under receiver operating characteristics curve.

## Results

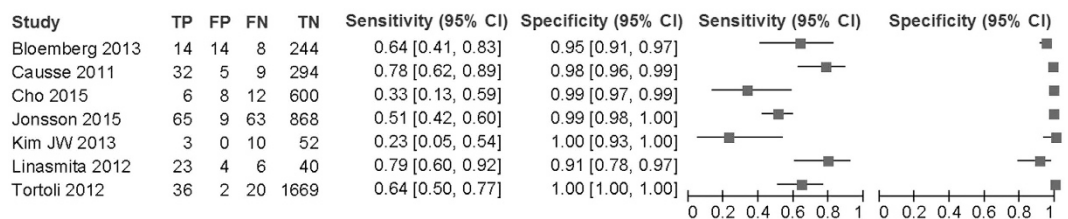
**Study search.** Of the 201 candidate articles, we finally identified 17 eligible articles (Fig. 1). Notably, two candidate articles were excluded due to overlapping data<sup>24,25</sup>. The finally eligible 17 articles included 16 English-language articles and one Korean-language article (Table 1). There were 16 full-length articles, and one non-full-length article. Most of the studies were reported from high-income or upper-middle income Asian countries, namely, Korea. From EU region, there was one article each from Germany, Italy, Spain, Sweden, and Switzerland. One article focused only on adult cases, three articles declared inclusion of both adult and child cases, and the others did not mention the age limitation. Three articles used frozen specimens and seven clearly wrote that the studies were conducted prospectively (Table 1).

The number of specimens in an article ranged from 72 to 6772 with a median of 619 specimens. The total number of specimens that provided valid results amounted to 26999, which consisted of 22883 respiratory specimens

## Respiratory specimens



## Non-respiratory specimens



**Figure 2.** The paired forest plot for the diagnostic accuracy. TP: true positive. FP: false positive. FN: false negative. TN: true negative.

and 4116 non-respiratory specimens (Table 1, Supplementary Table 1). In addition, our data included 68 specimens that provided invalid results.

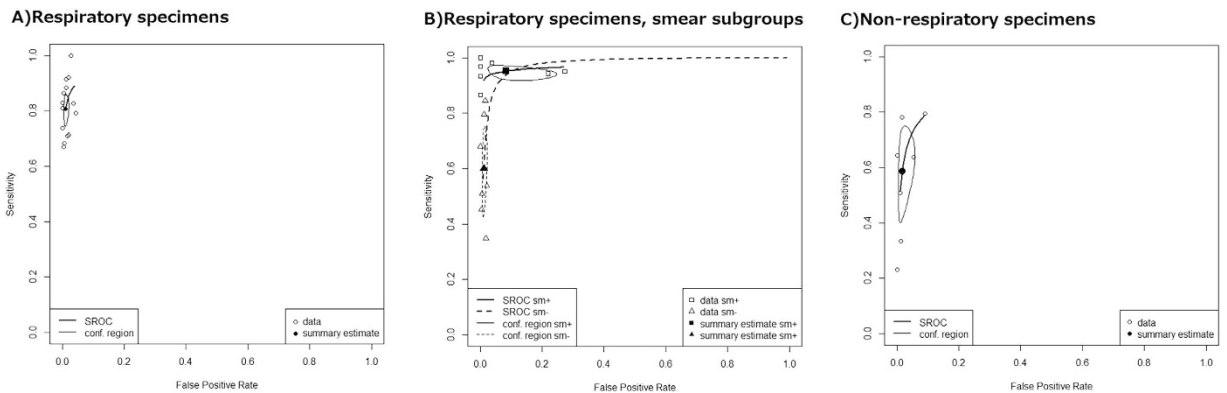
**Quality evaluation.** We assessed the quality of the studies using the QUADAS2. In two studies, samples were used from already diagnosed TB and non-TB individuals. These studies were considered to have a case-control design; thus, these two studies have a high risk of bias concerning patient selection. One study did not provide sufficient description regarding the methodology concerning the reference tests, mycobacterial culture, and subsequent TB confirmation. This study with insufficient description was considered to have a high risk of bias concerning reference test. No other domain had a high risk of bias or a high applicability concern (Supplementary Figure 1).

**Respiratory specimen.** From 15 studies, 1900 culture positive specimens and 20983 culture negative specimens were assessed (Table 2, Figs 2 and 3A). This provided the summary estimate sensitivity of 0.808 (95% CI 0.758–0.850) and the summary estimate specificity of 0.990 (95% CI 0.981–0.994). The AUC was 0.956. The DOR was 459 (95% CI 261–805) with a moderate heterogeneity of 26%.

We performed subgroup analyses based on the smear status (Figs 2 and 3B). When focused on the smear positive specimens, seven studies with 566 culture positive specimens and 255 culture negative specimens were assessed. This yielded AUC of 0.961 and DOR was 269 (95% CI 66–1104) without heterogeneity. Concerning the smear negative specimens, the same seven studies provided data for 498 culture positive specimens and 11123 culture negative specimens. These yielded the AUC of 0.971 and the DOR of 132 (95% CI 71–243) with least heterogeneity of 14%. These data suggested that CTM provided the similar overall diagnostic accuracies for both the smear positive and the negative specimens. However, the summary estimate sensitivity and specificity differed largely based on the smear status. For the smear positive specimens, the sensitivity was 0.952 (95% CI 0.926–0.969) and the specificity was 0.916 (95% CI 0.797–0.968). For the smear negative specimens, the sensitivity and the specificity were 0.600 (95% CI 0.459–0.726) and 0.989 (95% CI 0.981–0.993), respectively.

As a sensitivity analysis, we conducted a subgroup analysis based on whether the *M. tuberculosis* was confirmed after culture or not. For 13 studies with the TB confirmation, the sensitivity, the specificity, the AUC and the DOR were 0.790, 0.990, 0.954, and 401, respectively. For two studies without the TB confirmation, the sensitivity, the specificity, the AUC and the DOR were 0.895, 0.991, 0.966 and 978, respectively. The differences between subgroups were not so large. Overestimation of the specificity due to the lack of *M. tuberculosis* confirmation was not strongly suggested.

**Non-respiratory specimens.** According to seven studies with 307 culture positive non-respiratory specimens and 3809 culture negative specimens, the diagnostic accuracy was poorer than that for respiratory specimen but still acceptable. The AUC was 0.898 and the DOR was 86 (95% CI 34–217,  $I^2$  7%). The summary estimate sensitivity and specificity were 0.586 (95% CI 0.437–0.721) and 0.984 (95% CI 0.955–0.994), respectively (Table 2, Fig. 3C).



**Figure 3.** The summary receiver operating characteristics curves. SROC: summary receiver operation characteristics curve. conf. region: confidence region. sm+, smear positive. sm−, smear negative.

A subgroup analysis for the non-respiratory specimens based on the smear status was not conducted because only one study reported the diagnostic accuracy based on the smear-status subgroup.

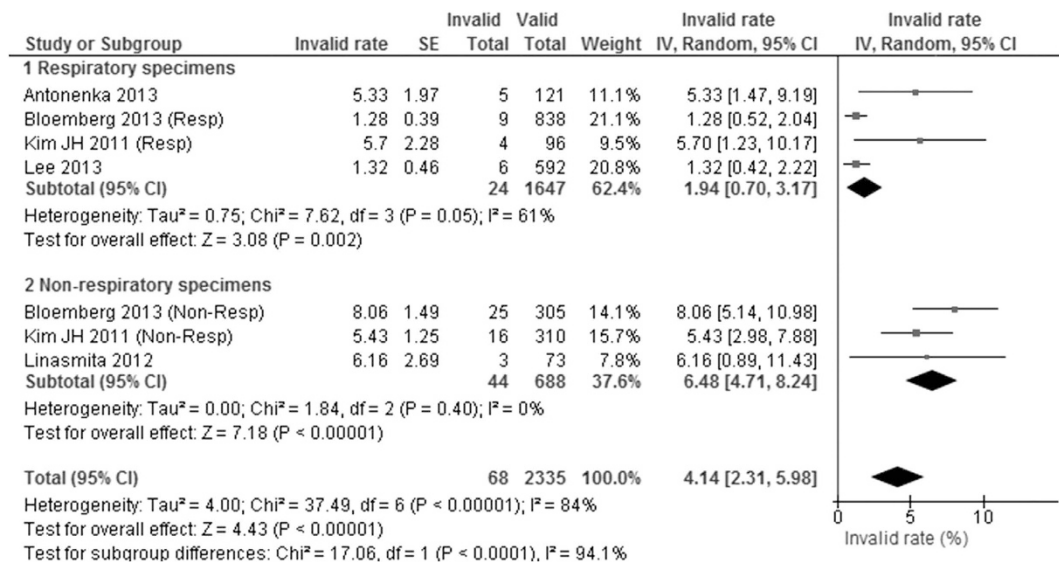
**Invalid rate.** Number of specimens with the invalid data was described in five original studies with 2403 specimens. The pooled invalid rate was 4.1% (95% CI 2.3–6.0%) with considerable heterogeneity of 84% (Fig. 4). In a subgroup analysis, the invalid rate was higher for the non-respiratory specimens (6.5%) than for the respiratory specimens (1.9%) ( $P < 0.001$ ).

## Discussion

To our knowledge, this is the first systematic review and meta-analysis for CTM diagnostic accuracy. According to the AUC and the DOR, diagnostic accuracy of CTM was excellent for the respiratory specimens (Table 2). The invalid rate was also as low as 1.9% for the respiratory specimens (Fig. 4). Even though both the diagnostic accuracy and the invalid rate were poorer for the non-respiratory specimens, these results were acceptable for practice (Table 2, Fig. 4). Worldwide, nearly 10 million people, including a million children, fell ill with TB every year. We strongly hope CTM, which can provide the quick accurate result, will contribute to early and accurate diagnosis of TB.

In the current study, we conducted a subgroup analysis comparing the smear positive and the negative respiratory specimens. Heterogeneity assessed by  $I^2$  statics between these subgroups was generally not strong (Table 2), which meant that CTM has the similar diagnostic accuracy for both subgroups. However, the summary estimates of the sensitivity and the specificity of CTM for the culture-proven TB clearly depended on the smear status. It is not surprising given the difference of bacterial load depending on the smear status. Positive sputum smear may reflect the abundant bacterial load probably above a detection threshold. The CTM summary estimate sensitivity for the smear-positive respiratory specimens was 0.952, which suggested that negative result in this category is reliable. Thus, a CTM negative for a smear-positive respiratory specimen strongly suggests the presence of non-TB mycobacterium or a false-positive smear that is sometimes obtained on Auramine-Rhodamine staining. The moderate CTM summary estimate sensitivity for the smear-negative specimen of 0.600 meant that CTM potentially have inability to detect TB in 40% of the smear-negative culture-positive respiratory specimens. The CTM summary estimate of specificity for smear-negative respiratory specimens of 0.989 meant that a CTM positive in this category result strongly suggests TB diagnosis. The summary estimate specificity for the smear-positive respiratory specimens of 0.916 was significantly lower than that for the smear-negative respiratory specimens, which meant that, based on definition of specificity, there are non-negligible number of CTM false positives for smear-positive specimens. However, this situation is caused by combinations of “smear positive, CTM positive, and culture negative,” or “smear positive, CTM positive, culture positive for non-TB Mycobacterium”. In such situations, we should consider the possibility of culture false negative, misclassification of TB into non-TB Mycobacterium, and dead TB bacilli. Cross reactivity for non-TB mycobacterium may be another concern. However, CTM cross reactivity for non-TB mycobacterium is believed very rare<sup>26</sup>.

We compared the accuracies of CTM and other commercialized PCR methods. Cobas Amplicor was one of the most commonly used commercialized PCR system until the Cobas TaqMan system became available<sup>5–7</sup>. According to three large-scale studies that evaluated the diagnostic accuracy of Amplicor, the sensitivity ranged from 93% to 98% for the smear-positive specimens and from 40% to 62% for the smear-negative specimens; and the specificity ranged from 75% to 91% for the smear-positive specimens and from 99% to 99% the smear-negative specimens. According to these existing reports and the current meta-analysis, CTM has the better specificity for the smear-positive specimens. Xpert MTB/RIF is another currently available PCR technology, whose high performance is supported by heminested real-time PCR technology and single-use cartridge system with low risk of contamination<sup>13</sup>. Although a number of studies concerning the diagnostic accuracy of Xpert MTB/RIF have been reported since 2010, very few head-to-head comparative studies have been conducted for CTM and Xpert MTB/RIF<sup>13–15</sup>. Despite inconsistency, these head-to-head comparisons suggested that CTM and Xpert MTB/RIF had the generally equivalent diagnostic accuracy<sup>13–15</sup>. Therefore, key features for Xpert MTB/RIF compared to CTM are ability to detect resistance to rifampicin, simple procedure, and high cost. Thus, we believe Xpert MTB/RIF has good indication in middle/high income regions such as EU, where drug-resistant TB is prevalent. Loop-mediated



**Figure 4. The forest plot for invalid rate.** Valid: sum of positive and negative results.

isothermal amplification assay is also a recently developed gene amplification method<sup>27</sup>. In contrast to the PCR technology in which the reaction is carried out with a series of alternating temperature steps, isothermal amplification can be carried out at a constant temperature. The pooled sensitivity and the pooled specificity based on a meta-analysis published in 2014 were 80% (95% CI 78–83%), and 96% (95–97%), respectively<sup>27</sup>, which seems inferior to those by CTM in the current analysis. However, in the high-quality study subgroup, the pooled sensitivity and the pooled specificity was 90.0% (95% CI, 86.0–93.0%) and 99.0% (95% CI, 98.0–100.0%), respectively. Given the low cost and simplicity, these diagnostic accuracies of loop-mediated isothermal amplification assay are acceptable, especially in the developing countries. The real-time PCR, developed two decades earlier, measures PCR product accumulation by a dual-labeled fluorogenic TaqMan Probe. Unlike the conventional quantitative PCR methods, the real-time PCR does not demand post-PCR sample handling. This prevents potential PCR product carry-over contamination and results in a faster and higher throughput analyses. Therefore, the real-time PCR can provide the accurate and reproducible results<sup>28</sup>.

The primary limitations of the current meta-analyses stemmed from the study design of each original study. The non-blinded examinations and the non-prospective study designs usually impair the quality of a study for a diagnostic test accuracy<sup>18</sup>. However, these factors probably may not have considerably affected our results, as both the reference culture and the index CTM usually provide the inflexible non-subjective results. Another limitation was the possibility of the false negatives from the reference test, the TB culture<sup>2,3</sup>. One strategy to account for the culture false negative is introducing a composite reference standard combining the results of several imperfect references<sup>29</sup>. Although we acknowledged the merit of the composite reference standard for TB diagnosis, we could not use this for the current meta-analysis because most researchers frequently used the non-composite culture as the simple golden standard. The third possible limitation is that we did not consider the interaction between CTM and HIV status, which potentially affected the results of CTM. Reviewing nature of the study also might be a possible limitation.

In conclusion, based on the meta-analysis using the bivariate model and the sufficient number of specimens, the diagnostic accuracy of CTM for the respiratory specimens was excellent. The summary estimates sensitivities and specificities using for the culture proven TB as reference test were 0.952 and 0.916 for the smear positive specimens and 0.600 and 0.989 for the smear negative specimens, respectively. The summary estimate sensitivity and specificity were 0.586 and 0.984 for the non-respiratory specimens, respectively. We believe that the information obtained from this study will aid the decision making of physicians who take care of patients with possible *M. tuberculosis* infection.

## References

1. World Health Organization. Global tuberculosis report 2014. Available from [http://www.who.int/tb/publications/global\\_report/en/](http://www.who.int/tb/publications/global_report/en/). Accessed on July 3<sup>rd</sup>, 2015.
2. The American Thoracic Society and the Centers for Disease Control and Prevention. Diagnostic standards and classification of tuberculosis in adults and children. *Am J Respir Crit Care Med.* **161**, 1376–1395 (2000).
3. American Thoracic Society; Centers for Disease Control and Prevention; Infectious Diseases Society of America. Controlling tuberculosis in the United States. *Am J Respir Crit Care Med.* **172**, 1169–1227 (2005).
4. Reischl, U., Lehn, N., Wolf, H. & Naumann, L. Clinical evaluation of the automated COBAS AMPLICOR MTB assay for testing respiratory and nonrespiratory specimens. *J Clin Microbiol.* **36**, 2853–2860 (1998).
5. Eing, B. R., Becker, A., Sohns, A. & Ringelmann, R. Comparison of Roche Cobas Amplicor Mycobacterium tuberculosis assay with in-house PCR and culture for detection of *M. tuberculosis*. *J Clin Microbiol.* **36**, 2023–2029 (1998).
6. D'Amato, R. F., Wallman, A. A., Hochstein, L. H., Colaninno, P. M., Scardamaglia, M. *et al.* Rapid diagnosis of pulmonary tuberculosis by using Roche AMPLICOR Mycobacterium tuberculosis PCR test. *J Clin Microbiol.* **33**, 1832–1834 (1995).

7. Bergmann, J. S. & Woods, G. L. Clinical evaluation of the Roche AMPLICOR PCR Mycobacterium tuberculosis test for detection of M. tuberculosis in respiratory specimens. *J Clin Microbiol.* **34**, 1083–1085 (1996).
8. Navarro, E., Serrano-Heras, G., Castaño, M. J. & Solera, J. Real-time PCR detection chemistry. *Clin Chim Acta.* **439**, 231–250 (2015).
9. Huh, H. J., Koh, W. J., Song, D. J., Ki, C. S. & Lee, N. Y. Evaluation of the Cobas TaqMan MTB test for the detection of Mycobacterium tuberculosis complex according to acid-fast-bacillus smear grades in respiratory specimens. *J Clin Microbiol.* **53**, 696–698 (2015).
10. Jönsson, B., Lönnemark, E. & Ridell, M. Evaluation of the Cobas TaqMan MTB test for detection of Mycobacterium tuberculosis complex. *Infect Dis.* **47**, 231–236 (2015).
11. Kim, J. H., Kim, Y. J., Ki, C. S., Kim, J. Y. & Lee, N. Y. Evaluation of Cobas TaqMan MTB PCR for detection of Mycobacterium tuberculosis. *J Clin Microbiol.* **49**, 173–176 (2011).
12. Linasmita, P., Srisangkaew, S., Wongsuk, T., Bhongmakapat, T. & Watcharananan, S. P. Evaluation of real-time polymerase chain reaction for detection of the 16S ribosomal RNA gene of Mycobacterium tuberculosis and the diagnosis of cervical tuberculous lymphadenitis in a country with a high tuberculosis incidence. *Clin Infect Dis.* **55**, 313–321 (2012).
13. Causse, M., Ruiz, P., Gutiérrez-Aroca, J. B. & Casal, M. Comparison of two molecular methods for rapid diagnosis of extrapulmonary tuberculosis. *J Clin Microbiol.* **49**, 3065–3067 (2011).
14. Antonenka, U., Hofmann-Thiel, S., Turaev, L., Esenalieva, A., Abdulloeva, M. *et al.* Comparison of Xpert MTB/RIF with ProbeTec ET DTB and COBAS TaqMan MTB for direct detection of M. tuberculosis complex in respiratory specimens. *BMC Infect Dis.* **13**, 280 (2013).
15. Park, K. S., Kim, J. Y., Lee, J. W., Hwang, Y. Y., Jeon, K., *et al.* Comparison of the Xpert MTB/RIF and Cobas TaqMan MTB assays for detection of Mycobacterium tuberculosis in respiratory specimens. *J Clin Microbiol.* **51**, 3225–3227 (2013).
16. Roche Diagnostics GmbH. COBAS TaqMan MTB Test. Available from: <http://tbevidence.org/documents/rescentre/sop/MTB%20TaqMan%20PI.pdf#search='cobas+taqman+mtb'>. Accessed on July 3rd, 2015.
17. Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. G., & PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med.* **151**, 264–269 (2009).
18. Leeflang, M. M., Deeks, J. J., Takwoingi, Y. & Macaskill, P. Cochrane diagnostic test accuracy reviews. *Syst Rev.* **2**, 82 (2013).
19. Whiting, P. F., Rutjes, A. W., Westwood, M. E., Mallett, S., Deeks, J. J. *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* **155**, 529–536 (2011).
20. Leeflang, M. M. G., Deeks, J. J., Gatsonis, C. & Bossuyt, P. M. M. Systematic reviews of diagnostic test accuracy. *Ann Intern Med.* **149**, 889–897 (2008).
21. Higgins, J. P. & Green, S. Cochrane handbook for systematic reviews of interventions. Available from <http://handbook.cochrane.org/>. Accessed on July 3rd, 2015.
22. Doebler, P., & Holling, H. Meta-analysis of diagnostic accuracy with mada. Available from: <https://cran.r-project.org/web/packages/mada/vignettes/mada.pdf#search='mada+r+systematic+review+diagnostic+accuracy'>. Accessed on November 5th, 2015.
23. Doebler, P. Package 'mada'. Available from: <https://cran.r-project.org/web/packages/mada/mada.pdf#search='mada+r'>. Accessed on November 5th, 2015.
24. Huh, H. J., Kwon, H. J., Ki, C. S. & Lee, N. Y. Comparison of the genedia MTB detection kit and the cobas TaqMan MTB assay for detection of Mycobacterium tuberculosis in respiratory specimens. *J Clin Microbiol.* **53**, 1012–1014 (2015).
25. Deggim, V., Somoskovi, A., Voit, A., Böttger, E. C. & Bloemberg, G. V. Integrating the Xpert MTB/RIF assay into a diagnostic workflow for rapid detection of Mycobacterium tuberculosis in a low-prevalence area. *J Clin Microbiol.* **51**, 2396–2399 (2013).
26. Kim, J. H., Kim, Y. J., Ki, C. S., Kim, J. Y. & Lee, N. Y. Evaluation of Cobas TaqMan MTB PCR for detection of Mycobacterium tuberculosis. *J Clin Microbiol.* **49**, 173–176 (2011).
27. Yuan, L. Y., Li, Y., Wang, M., Ke, Z. Q. & Xu, W. Z. Rapid and effective diagnosis of pulmonary tuberculosis with novel and sensitive loop-mediated isothermal amplification (LAMP) assay in clinical samples: a meta-analysis. *J Infect Chemother.* **20**, 86–92 (2014).
28. Heid, C. A., Stevens, J., Livak, K. J. & Williams, P. M. Real time quantitative PCR. *Genome Res.* **6**, 986–994 (1996).
29. Alonzo, T. A. & Pepe, M. S. Using a combination of reference tests to assess the accuracy of a new diagnostic test. *Stat Med.* **18**, 2987–3003 (1999).

## Author Contributions

H.N. contributed for the study design, the study search, the analysis, and the drafting the first manuscript as a principal investigator. Y.M. a PCR and T.B. specialist revised the manuscript. N.K. carried out the study search, the study quality evaluation, and the data extraction. S.T., T.T., N.H., T.K., S.Y., W.H., N.K., U.R. and I.M. interpreted the data and revised the manuscript. S.K. contributed as a statistical advisor. Y.T. provided advice as a TB specialist and K.T. contributed to the general management. All authors reviewed and finally approved the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Horita, N. *et al.* Sensitivity and specificity of Cobas TaqMan MTB real-time polymerase chain reaction for culture-proven Mycobacterium tuberculosis: meta-analysis of 26999 specimens from 17 Studies. *Sci. Rep.* **5**, 18113; doi: 10.1038/srep18113 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>