Online Appendix S1

Methods

Study design

Study subjects were recruited from the Black Lion University Hospital (Tikur Anbessa Specialized Hospital) in collaboration with eleven health centres, all located in Addis Ababa, Ethiopia: Tekelehaymanot, Lideta, Woreda, Yeka, Gulele, Kasanchise, Selam, Bole, Belatshachew, Sheromeda, and Arada health centres. The laboratory work of the study was performed in collaboration with the Armauer Hansen Research Institute (AHRI) and the International Clinical Laboratory (ICL) in Addis Ababa.

Ethical approvals

The clinical trial was conducted in agreement with the principles of the Declaration of Helsinki and the Guidelines for Good Clinical Practice. Ethical approval in Ethiopia was obtained both at national and regional levels (from the University hospital and research laboratory) including the National Research Ethics Review Committee (NRERC; 3.10/608/04) and the Institutional Review Boards (IRB; 008/11/IM) at Addis Ababa University and at the Armauer Hansen Research Institute (AHRI; P031/11) and also from the Food, Medicine and Health Care Administration and Control Authority of Ethiopia (Medical Product Agency, FMHACA; U2/6/22/12). Ethical approval for analyses of clinical trial samples in Sweden was obtained from the local ethical review board (EPN; 2011/1014-31/1), in Stockholm.

Patients

Sputum-smear microscopy and Mtb-culture negative clinical TB, defined according to the World Health Organization (WHO) 2006 criteria, was confirmed using clinical diagnosis based on: a) typical clinical symptoms for more than 2 weeks' duration, including persistent cough, general illness, fever, sweating, loss of appetite, and weight loss; b) chest X-ray revealing pulmonary infiltrates or lesions; and c) improved clinical and radiographic response to standard TB treatment (ie, resolution of clinical symptoms and X-ray findings). Patients with a diagnosis of clinical TB (n=31 placebo and n=35 vitD$_3$+PBA group) either had: a) both a negative sputum microcopy and Mtb sputum culture result (n=16 placebo and n=18 vitD$_3$+PBA group) or b) a negative microcopy result and no available Mtb-culture data (n=15 placebo and n=17 vitD$_3$+PBA group). Most (94%) of the clinical TB cases had a positive QuantiFERON (QFT) test at baseline and/or at follow-up ie. Mtb-specific IFN-γ production in whole blood samples. One QFT-negative patients discontinued treatment at week 4 (placebo n=1) while another patient discontinued treatment at week 0 and had no QFT test taken (vitD$_3$+PBA, n=1). Two QFT-negative patients who completed treatment per-protocol (vitD$_3$+PBA, n=2) responded to treatment according to the WHO criteria described above.

Interventions

Standard chemotherapy was provided for 6 months according to the Ethiopian guidelines for management of TB. The dosing scheme of vitD$_3$ and PBA was selected as follows: the dose of vitD$_3$ was based on a previous study from P. Bergman *et al.*, which showed that daily doses of vitD$_3$ (4000 IU/day) are required to enhance serum concentrations of 25(OH)D$_3$ to a level that is relevant to prevent respiratory tract infections [1]. The dose of PBA was chosen from a study conducted by M. Akhirunnesa *et al*, where 500 mg PBA given twice daily (together with 5000 IU vitD$_3$) to healthy volunteers proved to be the optimal oral dose to induce LL-37

and enhance mycobactericidal activity in monocyte-derived macrophages obtained from PBA-treated individuals [2].

Previously, we designed a 4-arm intervention trial in Dhaka, Bangladesh, including both microbiological and clinical primary endpoints [3]. While the Bangladeshi trial included sputum-smear positive patients allocated into four different treatment arms (placebo, vitD_3, PBA or vitD_3+PBA), this trial included both sputum-smear positive and negative TB patients allocated into two treatment arms (placebo and vitD_3+PBA). Designing a 2-arm intervention trial enabled an increased sample size in each arm (180 patients/arm=360 patients in total) compared to the Bangladeshi trial (72 patients/arm=288 patients in total), which was likely to increase the power of the primary analyses.

Randomization

Placebo tablets for vitD_3 and PBA had an identical appearance and taste as the corresponding active drug. Pharmacists at the Black Lion University Hospital prepared the study medication in polyethylene (PE)-bottles (Scandinavian Formulas) that were labeled with printed stickers (Merck Serono) including the numbers from the randomization list. The pharmacists were the only staff members with access to the randomization list: treatment allocation was concealed from patients, primary investigators, and other staff. Coded, sealed, opaque envelopes containing the key to the patient IDs were stored separately from the randomization list.

Procedures

Details on the primary (clinical TB score) and secondary (modified TB score, sputum-smear microscopy, sputum-culture and chest X-ray) outcomes, as well as adverse events (AEs) and statistical methods, are described below.

Clinical TB Score

In the primary TB score, a total of eleven clinical variables were recorded as present (1p) or absent (0p), resulting in a maximum TB score of 13 points (p). Note that BMI<16=1p and BMI<18=1p as well as MUAC<20=1p and MUAC<22=1p, and therefore a patient could obtain 2p if BMI or MUAC were very low [4]. To enhance the sensitivity of the primary score, we also generated a modified TB score that contained the same clinical variables as the primary score; however, the variables were recorded using a 3-point scale as absent (0p), improved (1p), or no change/worse (2p) (Table S1), resulting in a maximum TB score of 22p. In the primary TB score [4], anemia was assessed using conjunctiva pallor (palpebral conjunctival hue, assessing the paleness of conjunctivae at eye-examination). In the modified TB score, anemia was assessed using blood levels of hemoglobin (Hb), defining anemia in males with levels <13 mg/dl and in women <12 mg/dl.

Table S1 *Modified clinical TB score*

Variables	0 p	1 p	2 p
Cough	Absent	Improved	No change/worse
Night sweats	Absent	Improved	No change/worse
Chest pain	Absent	Improved	No change/worse
Anemia (Hb)	Hb>12-13	Hb=10-12	Hb<10
Hemoptysis	Absent	Improved	No change/worse
Dyspnea	Absent	Improved	No change/worse
Tachycardia	Pulse<80	Pulse=80-100	Pulse>100
Lung auscultations	Absent	Improved	No change/worse
Fever	Temp<37.5°C	Temp=37.5-39°C	Temp>39°C
BMI <18	>18	=16-18	<16
MUAC <220	>220 cm	=200-220 cm	<200 cm

Hb, hemoglobin mg/dl; BMI, Body Mass Index; MUAC, Mid-Upper-Arm-Circumference

Empirical validation of the primary TB score was performed in a cohort of 698 pulmonary TB patients from an epidemiological study in Guinea Bissau, assessing sensitivity to change, clinical improvement and the ability to predict mortality [4]. WHO guidelines were used to select the clinical signs or symptoms that were included in the score, which ensured content validity as all the symptoms had been described as relevant by an expert panel. Criterion validity was assessed by comparison with WHO outcomes (gold standard: Cure, Treatment

completed, Death, Failed, Defaulted, Transferred out, Not accounted for), also with the intention to determine if a high TB score could predict mortality and smear grade. The empirical validation showed that the TB score was sensitive to change during chemotherapy, showed good responsiveness and had predictive capacity. In addition, the score showed a strong association with mortality and grade of smear positivity. Another study performed in India, demonstrated that the TB score has good inter- and intra-observer variability in a resource-poor setting and is easy to learn and to monitor [5]. The TB score has also been prospectively evaluated during intensive phase anti-TB treatment in a cohort of 250 smear-positive and smear-negative pulmonary TB patients in Ethiopia [6]. This study confirmed that early changes in the TB score ie. between baseline and 2 weeks or 8 weeks, can predict treatment outcome recorded according to WHO definitions. An elevated TB score at week 8 (ie. severity class-III) was associated to a higher mortality and poor prognosis.

Sputum-microscopy and sputum-culture

Sputum-smear microscopy and Mtb-culture were the methods used to detect and monitor bacterial growth in this study [7]. For microscopy, acid-fast bacilli (AFB) were monitored weekly up to week 4 and at week 8. Ziehl-Neelsen staining (carbol-fuschin and methylene blue) was used to detect AFB by light microscopy (HumaScope Binocular laboratory LED illuminated microscope, Human Diagnostics, Wiesbaden, Germany). The density of AFB was graded according to WHO guidelines as no AFB (negative), scanty (0–1), +1, +2, or +3 AFB (positive) [8]. At weeks 0 and 8, part of the sputum sample was also used for Mtb culture. Sputum was processed using the N-acetyl-l-cysteine-sodium hydroxide (NALC-NaOH) method and cultured on Löwenstein-Jensen medium (BD) at 37°C for 8 weeks, with weekly read-outs of bacterial growth. Sputum samples that were positive for bacterial growth after 8 weeks of treatment (at the follow-up), were tested for drug susceptibility using GeneExpert

(Cepheid, Sunnyvale, CA) and a line probe assay. Occasionally, sputum microscopy or culture results were not available either due to technical reasons or the sputum sample was not enough or the sample was missing.

Chest X-ray

Chest X-ray assessments were performed at the Department of Radiology at the Black Lion Hospital in Addis Ababa, Ethiopia. Standard full-size posteroanterior chest X-rays were used to grade TB disease severity among the study subjects at the time of diagnosis into normal (grade 0), mild (grade 1), moderate to severe (grade 2), or advanced (grade 3) using the diagnostic standards and classifications of TB set forth by the American Thoracic Society, the National tuberculosis association of the United States. Radiographs were examined visually by an experienced radiologist to identify the presence and extent of infiltrates and other pathological features (eg, opacification, single or multiple cavitations, fibrosis and nodular or interstitial lesions). During follow-up examinations at week 4, 8, 16, and 24, each patient's disease status was rescored using a new X-ray as follows: normal/complete resolution or scarring (grade 0), >50% reduction or significant resolution (grade 1), <50% reduction or mild resolution (grade 2), and 0% reduction or unchanged/worsened (grade 3).

Adverse events (AE)

Solicited and unsolicited AEs were monitored during the complete study period including clinical examinations to record TB-specific clinical complications and laboratory tests to record liver function: serum alanine aminotransferase (ALT), aspartate aminotransferase (AST), total bilirubin and alkaline phosphatase (ALP); kidney function: serum creatinine and urea; and calcium/phosphate homeostasis: serum calcium, phosphate, and albumin. Blood chemistry analyses were conducted using kits from BD (ESR, bilirubin) and Abbott

Diagnostics (IL, USA). AEs were graded according to DAIDS AE grading Table (The Division of AIDS Table for Grading the Severity of Adult and Pediatric Adverse Events) [9]. Both clinical and laboratory AEs were graded as mild (grade 1), moderate (grade 2), severe (grade 3) and potentially life-threatening (grade 4). The clinical and laboratory safety data was reviewed by an external data and safety monitoring board (DSMB) appointed at the Black Lion University Hospital in Adds Ababa, Ethiopia.

Statistical analysis

Following the intention-to-treat (ITT)-concept, subjects who discontinued treatment but provided follow-up values were analyzed as though they were still part of the treatment group. In this study, we used a modified ITT (mITT) analysis, generally defined as a subset of the ITT population that allows exclusion of some randomized subjects in a justified way. Modified ITT is commonly used when the disease diagnosis is not immediately available at randomization or at start of treatment. Therefore, reasons for patients to be excluded from ITT could be that tests after randomization show that the patient is misdiagnosed and/or ineligible. mITT is commonly used in antimicrobial/anti-infective trials, when test results are often obtained after randomization. Accordingly, it usually takes days or weeks to receive a confirmatory Mtb diagnosis or results from drug sensitivity testing. The randomization and the start of treatment cannot wait until the confirmatory diagnosis results are available. In the mITT analyses, multiple imputation [10], as opposed to single imputation or last observation carried forward, ensures that standard errors and *P*-values are not artificially deflated [11].

Patients with negative sputum-smear or culture results at baseline were excluded from sputum microscopy and culture conversion analyses, respectively. Sputum conversion was defined as the earliest week of recurrent sputum negativity. Drop-outs with prior conversion were

assumed to be recurrent, while drop-outs without prior conversion were counted as censored. Persons with intermittent missing values were still considered to be in the risk set. A sensitivity analysis incorporating interval censoring was also made [12], but results did not differ noticeably from the data presented in the paper.

Tests for heterogeneity of treatment effect with regard to the primary outcome were performed for the following baseline variables: sputum-smear status (smear-negative vs. smear-positive), TB score severity (TB score<5=severity class I vs. TB score>5=severity class II-III), vitD_3 status (25(OH)D_3≤50 nmol/l vs. 25(OH)D_3>50nmol/l), vitD_3 status and TB score severity (25(OH)D_3≤50 nmol/l+TB score>5 vs. 25(OH)D_3>50 nmol/l+TB score<5), and vitD_3 response (25(OH)D_3≤50 nmol/l at baseline and >75 nmol/l at week 16).

Results

Enrolment and baseline characteristics

In total 894 patients were screened for eligibility based on a high clinical suspicion of pulmonary TB. However, more than half did not fulfil the pre-defined inclusion/exclusion criteria, primarily due to TB/HIV co-infection. Furthermore, 42 patients out of the 390 who were initially screened and randomized for treatment, later received a laboratory diagnosis confirming other diseases than TB (n=18, pulmonary fibrosis, cancer and pulmonary thromboembolism), other concomitant diseases (n=8, HIV infection, liver or renal disease) or MDR-TB (n=16). Accordingly, these 42 patients did not fulfil the inclusion/exclusion criteria for enrolment in the study. As a confirmed diagnosis for these patients were based on laboratory tests (mainly chest X-ray, sputum and blood tests) obtained after randomization, a modified ITT cohort (n=348) was generated excluding these patients from the ITT analyses. Please, note that the 42 patients who did not fulfil the pre-defined inclusion/exclusion criteria were evenly distributed between the placebo (n=20) and vitD_3+PBA treatment group (n=22).

Primary endpoint: Clinical TB score

The distribution of the primary TB score in the vitD_3+PBA treatment and placebo groups at the various follow-up time points is illustrated in Fig S1. From the baseline data in the clinical TB scores, it was evident that the proportion of TB patients with anemia was higher when using hemoglobin levels (32,4% anemic patients) compared with conjunctiva pallor (12,6% anemic patients). Perhaps this could be explained by reduced sensitivity of the palpebral conjunctival hue, as a previous study found a significant association of palpebral conjunctival hue with anemia only for hemoglobin levels <10 mg/dl [13].

Secondary endpoint: Chest X-ray

Chest X-rays were monitored throughout the study as a secondary endpoint. Although the X-ray results improved from moderate-severe disease to normal over the trial period, the difference in radiological findings was not statistically significant between the treatment and placebo groups (Fig. S2).

Subgroup analyses

We conducted subgroup analyses of the clinical response to vitD$_3$+PBA treatment in TB patients with 25(OH)D$_3$ levels ≤50 nmol/l [14] or a TB score>5 [4] at baseline. The differences in TB scores and 95% CI are shown in Table S2. Per-protocol analyses revealed a significant reduction in the primary as well as the modified TB score both in patients with 25(OH)D$_3$ levels ≤50 nmol/l (P = 0.015 and P =0.009) and in patients with a TB score>5 (P = 0.018 and P = 0.013) at week 8. In addition, per-protocol analysis that included patients with 25(OH)D$_3$ levels ≤50 nmol/l at baseline and >75 nmol/l at week 16 (i.e, vitD$_3$ responders) revealed a significant decrease in both primary and modified TB scores at week 8 (P = 0.034 and P = 0.014) and week 16 (P = 0.021 and P = 0.023).

Table S2 *Subgroup analyses: Clinical TB score in vitD$_3$+PBA versus placebo*

			Crude			Adjusted[a]		
Endpoint	Week	n	Difference	95% CI	p-value	Difference	95% CI	p-value
Patients with 25(OH)D$_3$ ≤50 nmol/l (per-protocol)								
TB score	8		-0.49	(-1.02 to 0.04)	0.072	-0.55	(-0.99 to -0.11)	0.015
Modified TB score	8	252	-0.65	(-1.42 to 0.13)	0.101	-0.72	(-1.27 to -0·18)	0.009
Patients with TB score >5 (per-protocol)								
TB score	8	150	-0.90	(-1.61 to -0.19)	0.013	-0.84	(-1.54 to -0.15)	0.018
Modified TB score	8	150	-1.12	(-2.02 to -0.23)	0.015	-1.07	(-1.90 to -0.23)	0.013
Patients with 25(OH)D$_3$ ≤50 nmol/l at baseline and >75 nmol/l at week 16 (per-protocol)								
TB score	8	238	-0.39	(-0.91 to 0.14)	0.145	-0.48	(-0.92 to -0.04)	0.034
TB score	16	238	-0.27	(-0.80 to 0.26)	0.323	-0.39	(-0.73 to -0.06)	0.021
Modified TB score	8	238	-0.56	(-1.33 to 0.21)	0.155	-0.70	(-1.26 to -0.14)	0.014
Modified TB score	16	238	-0.35	(-1.15 to 0.46)	0.400	-0.53	(-0.98 to -0.07)	0.023

CI, confidence interval; 25(OH)D$_3$, 25-hydroxyvitamin D

[a] Data are adjusted for gender, age, and TB score and sputum-smear positivity at baseline.

A statistical test for interaction revealed a significant interaction between vitD$_3$+PBA treatment and the effect on the primary TB score in TB patients with a vitD$_3$ insufficiency and a TB score>5 (-1.00, 95% CI -1.80, -0.19; P = 0.016). There was no significant differences in the effect of treatment in patients with smear-positive or smear-negative TB (P = 0.185) or in patients with a vitD$_3$ insufficiency or adequate levels of 25(OH)D$_3$ (P = 0.373) or in patients with a TB score<5 (mild disease) vs a TB score>5 (moderate-to-severe disease) (P = 0.060) or in vitD$_3$ responders ie. patients with 25(OH)D$_3$≤50 nmol/l at baseline and >75 nmol/l at week 16 vs non-responders (P = 0.096). Similarly, no significant interaction were found between vitD$_3$+PBA treatment and the effect on sputum conversion in TB patients with 25(OH)D$_3$≤50 nmol/l (P = 0.279), a TB score >5 (P = 0.576) or 25(OH)D$_3$≤50 nmol/l + TB score>5 (P = 0.102).

Adverse events

Patients' clinical symptoms were closely followed and monitored through patient exams. Only new symptoms i.e. TB-specific complications that emerged during treatment were documented, while mild-to-moderate symptoms typically related to the anti-TB drugs were not routinely recorded. The physiological state of each patient was monitored by the laboratory analysis shown in Tables S3 and S4. Four grade 4 laboratory AEs were detected in the study cohort; two in the vitD$_3$+PBA group (two patients with high ALT) and two in the placebo group (one patient with high bilirubin and one patient with high ALT). Most grade 1–3 laboratory AEs improved from the start of the study to the last visit; however, some patients maintained grade 1–2 AEs through week 16. Two patients maintained grade 3 AEs (low phosphate and calcium levels, respectively) through week 16. Calcium levels were typically higher in TB patients at the time of baseline measurements, and thus a calcium level of 12 mg/dl (3 nmol/l) was considered the upper limit of the normal range for this trial. Patients

with elevated laboratory results had an unscheduled visit for a clinical examination and extra blood chemistry analysis. None of the AEs had any connection to the study drugs (vitD_3 or PBA), as reviewed by the DSMB.

Table S3 *Laboratory adverse events*

Variables	Grade[a]	$VitD_3$+PBA (n=175)	Placebo (n=173)
Hemoglobin	1 (10-10.9)	6 (3.4%)	9 (5.2%)
(g/dl)	2 (9.0-9.9)	2 (1.1%)	3 (1.7%)
	3 (7-8.9)	2 (1.1%)	2 (1.2%)
	4 (<7)	0	0
Albumin	1 (3-3.5)	23 (13%)	24 (18.9%)
(g/l)	2 (2-2.9)	6 (3.4%)	4 (2.3%)
	3 (<2)	2 (1.1%)	0
	4 NA	0	0
Low calcium	1 (7.8-8.4)	46 (26.2%)	55 (31.8%)
corrected albumin	2 (7-7.7)	14 (8%)	17 (9.8%)
(mg/dl)	3 (6.1-6.9)	6 (3.4%)	4 (2.3%)
	4 (<6.1)	0	0
High calcium	1 (10.6-11.5)	12 (6.8%)	11 (6.3%)
corrected albumin	2 (11.6-12.5)	6 (3.4%)	5 (2.9%)
(mg/dl)	3 (12.6-13.5)	1 (0.5%)	1 (0.6%)
	4 (>13.5)	0	0
Phosphate	1 (2.5)	0	0
(mg/dl)	2 (2-2.4)	3 (1.7%)	3 (1.7%)
	3 (1-1.9)	2 (1.1%)	4 (2.3%)
	4 (<1)	0	0
Creatinine	1 (1.32-1.56)	1 (0.5%)	0
(mg/dl)	2 (1.68-2.56)	0	0
	3 (2.28-4.08)	0	0
	4 (>4.2)	0	0
Bilirubin (total)	1 (1.32-1.8)	5 (2.8%)	8 (4.6%)
(mg/dl)	2 (1.9-3)	2 (1.1%)	1 (0.6%)
	3 (3.1-6)	0	1 (0.6%)
	4 (>6)	0	1 (0.6%)
ALT	1 (50-100)	10 (5.7%)	14 (8.1%)
(IU/l)	2 (100-200)	2 (1.1%)	2 (1.2%)
	3 (200-400)	1 (0.5%)	1 (0.6%)
	4 (>400)	2 (1.1%)	1 (0.6%)
AST	1 (50-100)	4 (2.3%)	12 (6.9%)
(IU/l)	2 (100-200)	2 (1.1%)	4 (2.3%)
	3 (200-400)	2 (1.1%)	0
	4 (>400)	0	0
ALP	1 (382-765)	1 (0.5%)	1 (0.6%)
(mg/dl)	2 (795-1530)	0	0
	3 (1560-3060)	0	0
	4 (>3060)	0	0

NA, not applicable; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase

[a] Division of AIDS for grading the severity of adult AEs for laboratory results.

Table S4 *Blood chemistry data*

Variables[a]	Week	$VitD_3$+PBA (n=175)	Placebo (n=173)
CD4 T cell counts	0	496.00 (379.00-667.00)	465.00 (324.00-636.00)
(cells/ml)	4	609.50 (447.50-772.00)	539.50 (405.50-694.00)
	8	578.50 (440.00-739.00)	556.00 (411.00-699.50)
	16	644.50 (466.00-818.00)	581.50 (443.00-728.00)
WBC	0	7.30 (5.30-9.60)	7.40 (5.90-9.20)
(SI units)	4	6.90 (5.30-8.30)	6.15 (4.95-7.29)
	8	5.70 (4.40-6.90)	5.40 (4.60-6.65)
	16	4.90 (3.90-6.30)	5.20 (4.25-6.30)
ESR	0	45.50 (30.00-56.00)	46.00 (30.00-54.00)
(mm/hour)	4	38.00 (20.00-50.00)	38.00 (20.00-48.00)
	8	24.00 (12.00-40.00)	25.50 (12.00-42.00)
	16	12.00 (8.00-20.00)	16.00 (8.00-26.00)
Hemoglobin	0	13.00 (11.70-14.00)	13.20 (11.90-14.20)
(mg/dl)	4	13.60 (12.50-14.70)	13.90 (13.00-14.95)
	8	14.10 (13.30-15.20)	14.20 (13.40-15.25)
	16	14.85 (13.60-15.80)	14.95 (14.00-15.95)
Albumin	0	3.60 (3.30-3.90)	3.60 (3.30-3.90)
(g/l)	4	3.80 (3.50-4.10)	3.80 (3.50-4.10)
	8	3.90 (3.70-4.20)	4.00 (3.70-4.20)
	16	4.00 (3.80-4.20)	4.00 (3.80-4.30)
Calcium	0	8.90 (8.30-9.50)	8.80 (8.20-9.40)
(g/dl)	4	9.10 (8.50-9.60)	8.80 (8.40-9.30)
	8	9.10 (8.60-9.60)	8.95 (8.40-9.50)
	16	9.00 (8.50-9.40)	8.90 (8.40-9.50)
Phosphate	0	3.70 (3.20-4.10)	3.60 (3.20-4.00)
(mg/dl)	4	3.70 (3.40-4.20)	3.65 (3.30-4.20)
	8	3.80 (3.40-4.20)	3.60 (3.30-4.10)
	16	3.70 (3.20-4.20)	3.60 (3.25-4.00)
Creatinine	0	0.80 (0.60-0.90)	0.79 (0.70-0.90)
(mg/dl)	4	0.80 (0.60-0.90)	0.70 (0.60-0.80)
	8	0.70 (0.60-0.90)	0.70 (0.60-0.80)
	16	0.70 (0.60-0.90)	0.70 (0.50-0.80)
Bilirubin (total)	0	0.81 (0.59-1.23)	0.80 (0.60-1.20)
(mg/dl)	4	0.60 (0.41-0.80)	0.60 (0.45-0.80)
	8	0.60 (0.48-0.78)	0.60 (0.50-0.73)
	16	0.65 (0.50-0.82)	0.62 (0.44-0.73)
Urea	0	20.00 (15.00-25.00)	18.00 (15.00-24.00)
(mg/dl)	4	17.00 (13.00-22.00)	17.00 (12.40-20.50)
	8	17.55 (13.10-22.80)	16.35 (12.75-20.95)
	16	19.00 (15.00-25.00)	19.00 (14.00-24.00)
ALT	0	16.00 (12.00-24.00)	17.00 (12.00-25.00)
(IU/l)	4	14.00 (10.00-22.00)	17.00 (11.00-26.00)
	8	15.00 (11.00-22.00)	16.00 (12.00-23.00)
	16	14.00 (10.00-20.00)	16.00 (12.50-23.00)
AST	0	20.00 (16.00-25.00)	21.00 (17.00-28.00)
(IU/l)	4	20.50 (16.00-30.00)	22.00 (16.00-31.00)
	8	21.00 (18.00-29.00)	24.00 (18.00-29.00)
	16	21.00 (17.50-25.80)	23.50 (19.00-31.00)
ALP	0	91.00 (79.00-114.00)	96.00 (75.00-125.00)
(mg/dl)	4	85.50 (70.00-102.00)	91.00 (73.00-111.00)
	8	79.50 (66.00-103.00)	92.50 (72.50-108.50)
	16	75.00 (62.50-95.00)	88.00 (73.00-111.00)

WBC, white blood cell; ESR, erythrocyte sedimentation rate;
ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase

[a] All data is presented as median and 25th percentile-75th percentile.

References

1. Bergman P, Norlin AC, Hansen S *et al*. Vitamin D3 supplementation in patients with frequent respiratory tract infections: a randomised and double-blind intervention study. *BMJ open* 2012; **2:** e001663.
2. Mily A, Rekha RS, Kamal SM *et al*. Oral intake of phenylbutyrate with or without vitamin D3 upregulates the cathelicidin LL-37 in human macrophages: a dose finding study for treatment of tuberculosis. *BMC Pulm Med* 2013; **13:** 23.
3. Mily A, Rekha RS, Kamal SM *et al*. Significant effects of oral phenylbutyrate and vitamin D3 adjunctive therapy in pulmonary tuberculosis: A randomized controlled trial. *PloS one* 2015; **10:** e0138340.
4. Wejse C, Gustafson P, Nielsen J *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:** 111-20.
5. Raaby L, Bendix-Struve M, Nielsen J, Wejse C. Inter-observer variation of the Bandim TB-score. *Scand J Infect Dis* 2009; **41:** 220-3.
6. Janols H, Abate E, Idh J *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:** 828-34.
7. Keeler E, Perkins MD, Small P *et al*. Reducing the global burden of tuberculosis: the contribution of improved diagnostics. *Nature* 2006; **444 Suppl 1:** 49-57.
8. Weyer K. Laboratory services in tuberculosis control: Microscopy Part II. *Geneva, Switzerland: World Health Organization (WHO)* 1998
9. National Institute of Allergy and Infectious Diseases (NIAID). Division of AIDS Table for Grading the Severity of Adult and Pediatric Adverse Events. December, 2004.

10. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Me*d 2011; **30:** 377-99.

11. Sterne JA, White IR, Carlin JB *et al*. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009; **338:** b2393.

12. Turnbull BW. Nonparametric estimation of a survivorship function with doubly censored data. *J Am Stat Assoc* 1974; **69:** 169-73.

13. Kent AR, Elsing SH, Hebert RL, Conjunctival vasculature in the assessment of anemia. *Ophthalmology* 2000; **107:** 274-277.

14. Holick MF, Binkley NC, Bischoff-Ferrari HA *et al*. Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 2011; **96:** 1911-30.

Supplementary figure captions

Fig. S1 *Distribution of the primary TB score. (a) The absolute TB score in the vitD₃+PBA treatment and placebo groups at baseline and at weeks 4, 8, 16, and 24 after initiation of anti-TB chemotherapy. Adjunct vitD₃+PBA treatment was provided during the first 16 weeks of standard care. (b) The change in baseline TB score in vitD₃+PBA compared with placebo treatment at weeks 4, 8, 16, and 24. Placebo (white bars) vs vitD₃+PBA treatment (green bars) is shown in the graphs.*

Fig. S2 *Radiological findings. Secondary analysis involving chest X-ray grading among the study subjects in the vitD₃+PBA treatment and placebo groups at baseline compared with weeks 4, 8, 16 and 24 after initiation of anti-TB chemotherapy. Chest X-ray findings was graded as normal (grade 0), mild (grade 1), moderate-to-severe (grade 2), or advanced (grade 3). Data are shown in a bar graph with a colour scale from severity grade 0 (red) to grade 3 (blue).*

Fig. S1

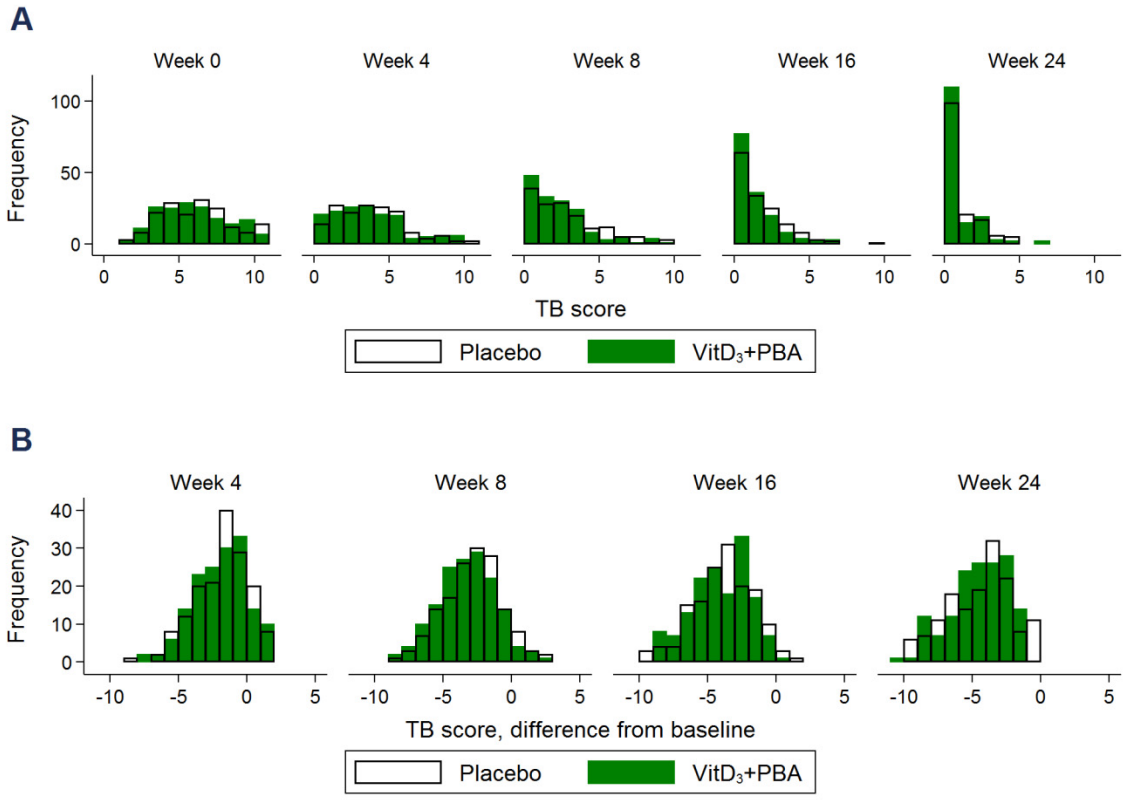
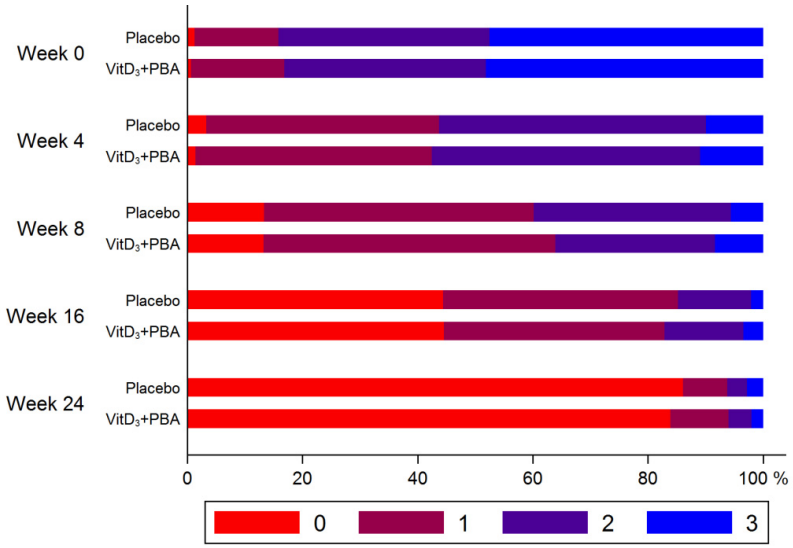

A
Week 0
Week 4
Week 8
Week 16
Week 24
Frequency
100
50
0
0
5
10
TB score
Placebo
VitD3+PBA
B
Week 4
Week 8
Week 16
Week 24
Frequency
40
30
20
10
0
-10
-5
0
5
TB score, difference from baseline
Placebo
VitD3+PBA

Fig. S2


Week 0
Placebo
VitD3+PBA
Week 4
Placebo
VitD3+PBA
Week 8
Placebo
VitD3+PBA
Week 16
Placebo
VitD3+PBA
Week 24
Placebo
VitD3+PBA
0
20
40
60
80
100 %
0
1
2
3