UNIVERSITY OF LEEDS

This is a repository copy of *Effectiveness of a comprehensive package based on electronic medication monitors at improving treatment outcomes among tuberculosis patients in Tibet: a multicentre randomised controlled trial*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/221827/

Version: Accepted Version

Article:
Wei, X., Hicks, J.P. orcid.org/0000-0002-0303-6207, Zhang, Z. et al. (10 more authors) (2024) Effectiveness of a comprehensive package based on electronic medication monitors at improving treatment outcomes among tuberculosis patients in Tibet: a multicentre randomised controlled trial. The Lancet, 403 (10430). pp. 913-923. ISSN 0140-6736

https://doi.org/10.1016/s0140-6736(23)02270-5

© 2024, Elsevier. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/. This is an author produced version of an article published in the Lancet. Uploaded in accordance with the publisher's self-archiving policy.

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Effectiveness of a comprehensive package based on electronic medication monitors at improving treatment outcomes among tuberculosis patients in Tibet: a multi-centre randomised controlled trial

Xiaolin Wei (PhD), Joseph Paul Hicks (PhD), Zhitong Zhang (MPH), Victoria Haldane (PhD), Pande Pasang (MD), Linhua Li (MD), Tingting Yin (MPH), Bei Zhang (MPH), Yinlong Li (PhD), Qiuyu Pan (PhD), Xiaoqiu Liu (PhD), John Walley (FFPH), Jun Hu (PhD)

Dalla Lana School of Public Health, University of Toronto, Toronto, Canada (Prof Xiaolin Wei [full professor], Zhitong Zhang); Nuffield Centre for International Health and Development, University of Leeds, Leeds, UK (Joseph P Hicks, Prof John Walley [full professor]); Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada (Prof Xiaolin Wei [full professor], Victoria Haldane); Shigatse Centre for Disease Control and Prevention, Shigatse, China (Pande Pasang, Linhua Li, Jun Hu); Weifang Medical College, Weifang, China (Tingting Yin, Bei Zhang); Jining Medical University, Jining, China (Yinlong Li); North Sichuan Medical College, Nanchong, China (Qiuyu Pan); National Center for TB control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China (Xiaoqiu Liu); Shandong University of Traditional Chinese Medicine, Jinan, China (Prof Jun Hu [full professor])

Correspondence to:

Xiaolin Wei and Jun Hu

Abstract: 433 words

Background

World Health Organisation recommends that electronic medication monitors (EMMs), a form of digital adherence technology, complement directly observed treatment (DOT) for tuberculosis (TB), as DOT is inconvenient and costly. However, existing evidence about their effectiveness is inconclusive. We evaluated the effectiveness of a comprehensive package based on EMMs among TB patients in Tibet.

Methods

We conducted this pragmatic, unblinded, multi-centre, individually-randomised, controlled, superiority trial in six counties in Shigatse, Tibet. Eligible participants were drug-susceptible TB patients aged ≥15 years starting standard TB treatment. TB doctors recruited patients from the

public TB dispensary in each county and randomised them to intervention or control (near 1:1 ratio), based on the predetermined randomised allocation sequence (using a permuted block design with variable block sizes of two/four/six, stratified by county). Intervention patients received an EMM box. This included audio medication-adherence reminders and recorded box-opening data, which were transmitted to a cloud-based server accessible to healthcare providers to allow remote adherence monitoring. A linked smartphone app enabled text/audio/video communication between patients and healthcare providers. We provided a free data plan to patients. We also trained family members to support patients with using the EMM and app. Control patients received usual care plus a deactivated EMM (this only recorded and transmitted box-opening data that was not accessible to healthcare providers), no access to the app, and we did not train family treatment supporters. Our primary outcome was a binary indicator of poor monthly adherence, defined as missing ≥20% of planned doses in the treatment month, measured using EMM opening data and verified by counting used medication blister packages during consultations. We recorded other secondary treatment outcomes based on national tuberculosis reporting data. We analysed the primary outcome based on the intention-to-treat population. Registered at ISRCTN: 52132803.

Findings

Between 17 November 2018 and 5 April 2021 we randomised 278 patients: 143 to the intervention and 135 to the control. The final patient ended treatment on 4 October 2021. In the intervention arm 10.2% (87/854) of patient treatment months showed poor adherence compared to 36.5% (290/795) in the control arm. The corresponding intervention versus control adjusted risk difference was -29.2 percentage points (95% CI: -35.3, -22.2; p≤0.001). Five out of six secondary treatment outcomes also demonstrated clear improvements including treatment success, which was 93.7% (133/142) in the intervention arm and 73.1% in the control arm (98/134), with an adjusted risk difference of 21 percentage points (95% CI: 12.4, 29.4); p≤0.001.

Interpretation

Our interventions were considerably effective at improving TB treatment adherence and outcomes, and our trial suggests that a comprehensive packaged involving EMMs may positively impact TB programmes in high-burden and low-resource settings.

Funding: TB REACH.

Introduction

Tuberculosis (TB) has long been the leading killer of infectious disease, with an estimated 9.9 million new cases and 1.3 million deaths in 2020 globally, only recently becoming the second leading infectious cause of death after COVID-19.[1] Tuberculosis treatment is lengthy, usually six months or longer, and is often difficult to complete due to medication side effects. Medication non-adherence poses a serious challenge to TB management efforts as it fuels the development of drug resistance, reduces treatment success, and contributes to disease transmission.[2] Recognising the need to support people with TB in completing their treatment, the World Health Organization (WHO) introduced directly observed treatment short course (DOTS) for TB treatment in 1990s. Directly observed treatment (DOT), a key strategy in TB programmes globally, requires people with TB to visit a health worker daily who observes them take their medication. However, delivering DOT also requires considerable resources and time commitments, particularly in low-resource settings, and imposes costs and the inconvenience of daily health facility visits on patients.[3]

Due to the lack of trial evidence for requiring patients to physically attend health facilities for DOT,[4] the WHO DOTS strategy has evolved towards more patient centred care, including recommending patient support (which may include DOT) be done by community health providers and lay health workers.[5] Yet, where DOT is still a requirement, as in China, in practice a significant proportion of TB patients self-administer therapy,[6] which is associated with non-adherence, lower rates of treatment success, and higher rates of developing drug resistance.[7] It

is therefore imperative to identify innovative ways to support TB patients to take their medications and complete their treatment.

Digital adherence technologies (DATs) are a promising tool to link TB patients to their care providers both for medication taking and for ongoing supportive care. These technologies include simple approaches such as short messenger services (SMS), as well as more advanced electronic medication monitors (EMM) that can record medication taking in real-time, and, via linked apps, enable communication with healthcare providers and provide health education.[8] However, systematic reviews reported mixed results in terms of their apparent ability to improve treatment adherence and other relevant outcomes.[8-11] In recent trials, DAT interventions using newer technologies, such as EMMs,[12,13] and video-observed therapy (VOT),[14] have shown promising effectiveness at improving medication adherence, but they have not demonstrated clear impacts on patients' treatment outcomes. Two studies employing 99DOTS in Uganda[15] and EMM in China[16] found that 20-50% of eligible participants did not enrol or use the new technologies, and reported that those who adopted the intervention appeared to reap the benefits, with non-adopters tending to be poorer, older, and having low smartphone literacy. To maximise use, DATs in future studies need to be tailored to the TB care delivery context, connect patients to their health care providers, and be versatile to accommodate varying patient circumstances.[8]

In this study we therefore evaluated a comprehensive package based on a new generation of EMMs for TB medication adherence among TB patients in the Tibet Autonomous Region

(hereafter Tibet), China. The monitors included features of accurate, real-time dose-recording (with the data accessible to healthcare providers), and audio voice reminders. We also provided an easy-to-use add-on for an existing popular smartphone app (WeChat) that aimed to improve communication between patients, their treatment supporters, and their health care providers. This setting has a high burden of TB and significant challenges with non-adherence given the typically long travel distances between communities and health facilities, the often-severe weather impeding access to care, and the severe shortage of human resources required to implement DOT.[17] We previously co-developed the intervention with partners in Tibet and piloted them for feasibility and acceptability testing. Here we report the results of a multi-centre, randomised, controlled trial comparing treatment adherence and other relevant outcomes between TB patients given an EMM, access to the associated app and a free data package, and a treatment supporter, with those who received usual care.

Methods

See the supplementary materials for a CONSORT checklist, further methodological details, and details of deviations from protocol-planned methods. We used *R* statistical software for randomisation, sample size calculation and analysis.[18]

Study design, setting, and participants

We conducted a pragmatic, multi-centre, individually-randomised, parallel-arm, controlled trial, to evaluate whether our intervention was superior to our control. The protocol was previously

published.[19] The intervention development was guided by implementation science and eHealth behavioural theory, specifically Unified Theory of Acceptance and Use of Technology.[20] We conducted the trial in six TB dispensaries, each located in a different county, in Shigatse, Tibet. This region is characterised by a very low population density with only four people per km^2. Most residents are ethnic Tibetans. Tuberculosis prevalence in Tibet was 758 per 100,000 people (0.76%) in 2014, twice the national average of China in 2010.[21] We originally planned to run the trial in two counties: Samzhubze and Sa'gya (recruitment started in November 2018). However, during the trial the COVID-19 pandemic caused periodic lockdowns in Tibet, and a change to the national TB care model interrupted treatment services, which substantially delayed recruitment. We therefore extended the trial to four counties based on their relatively high TB prevalence and TB patient numbers: Gyantse (recruitment started in June 2019), Ngamring and Tingri (recruitment started in June 2020), and Bainang (recruitment started in September 2020). Tuberculosis dispensaries are public facilities providing TB treatment on an outpatient basis, and they are the only designated health facilities providing TB treatment in these counties (one per county). All TB dispensaries in the six counties, except Gyantse, were originally located in the county Centres for Disease Control and Prevention (CDCs). All the TB dispensaries were transferred from CDCs to county general hospitals during the project period following a change in China's national TB care model.[22] Tuberculosis drug sensitivity analysis was not available in Shigatse during the trial. Therefore, patients who were assumed to have drug resistant TB were sent to Lhasa for diagnosis and treatment. Mobile connectivity in Tibet is functional but limited to residential areas. Most patients either had a smartphone or had access to one owned by a relative in their home.

During the recruitment phase all new pulmonary TB patients referred or self-presented to the TB dispensaries, who were diagnosed according to national and WHO guidelines of TB care,[23] aged ≥15 years, and started on the standard 6-month short-course chemotherapy on an outpatient basis were considered eligible. Patients were excluded if they had any serious communication impairment(s) (mental, visual, auditory or speech), or had any family members within their household who had already been enrolled into the trial to avoid contamination. Consequently, we also excluded retreatment patients because they received an 8-month treatment and were likely to be drug resistant, and any patients with multi-drug resistant TB because they received 18-month treatment based on national guidelines, and were often treated in Lhasa. We also excluded patients with ex-pulmonary TB as they had to receive individualised treatment in Lhasa. TB doctors in the dispensaries identified all eligible patients, and recruited patients by obtaining their written informed consent. We obtained ethical approval from the Office of Research Ethics at the University of Toronto (Ref: 36569) and the Ethics Review Committee of the Tibet Centre for Disease Control and Prevention (Ref: 006).

Randomisation and masking

We randomly assigned participants to either the intervention or control arm in a near 1:1 ratio using a computer-generated randomised permuted-block design, stratified by county. For each county the trial statistician (JPH) generated an allocation sequence with an initial unequally sized block (7: 3 intervention and 4 control) followed by randomly selected blocks of two, four or six (equal allocation), which further masked the allocation sequence. The lead statistician

(JPH) generated the allocation sequence and then passed it to the study team to print individually numbered cards with the allocation sequence. Study team members then sealed the cards in opaque envelopes and delivered them to study sites, where they were opened by TB doctors in each TB dispensary as each patient was recruited to decide patient assignment. Due to the nature of the intervention, it was not possible to mask patients or health care providers to treatment allocation.

Procedures

In both arms all patients received six months' standard TB care (or seven months' treatment if their sputum smear had not converted at the end of the second month of intensive treatment) based on WHO and Chinese national TB programme guidelines for the treatment of drug-susceptible TB patients, using daily fix-dose-combinations for the entire treatment period.[23] In Tibet, patients visit their TB dispensary every other month to collect their medication due to the particularly harsh travel conditions in the region. All patients were further supported by community-based health care providers known as village doctors, who acted as a treatment supervisors based on national TB programme guidelines.

Patients in the intervention arm also received an EMM box to store their medications in, which had two key features (Box 1). First, it reminded patients to take their medication on time using human-voice based recordings. Second, it recorded every time the box was opened and transmitted this information to a cloud-based server. This information could then be accessed by a patient's village doctor and their TB doctor could also access this information via a

computer web-based interface or a password protected WeChat-based smartphone app (a very popular Chinese multi-media messaging app similar WhatsApp) in real-time. Patients who refused to start/continue treatment, missed more than three consecutive doses, or where either doctor had concerns about their adherence, received daily real-time VOT/audio calls by the village doctor until they had taken more than three consecutive doses.

We also helped all intervention-arm patients select a treatment supporter, whose role was to help the patient adhere to their medication and use the electronic medication monitor and the WeChat app. Treatment supporters were typically a family member who lived with the patient and who had good smartphone literacy. Either TB doctors or family supporters then trained patients on how to use the monitor and the app during recruitment following their diagnosis, or alternatively village doctors trained them during their first home visit within the first week after diagnosis. The WeChat app also enables direct communication between the patient, their treatment supporter, and healthcare providers. We provided all patients with a free smartphone data plan, and although most patients used their own smartphones we also provided a few to patients who did not have access to one.

We gave patients in the control arm a deactivated EMM as a medication storage box. These boxes recorded each opening and transmitted the information to the cloud-based server, but patients, treatment supporters, and their healthcare providers could not access this information. Under usual care, the national guidelines recommend that village doctors visit patients at least once a week. However, they connected with patients through traditional

means, such as physical visits or phone calls. We also did not select or train family treatment supporters in this arm as per national guidelines.

Data collection

We recruited patients between 17 November 2018 and 5 April 2021 and followed the cohort until the final patient completed their treatment on 4 October 2021. This period overlapped with the COVID-19 pandemic. We collected routinely reported patient-level information from the national TB reporting database, regarding their name, age, sex, address, education level, profession, diagnoses, and standard TB treatment outcomes. We allocated a unique participant number to each patient to mask patient identifiable information to researchers. We collected patients' adherence data, recorded by their EMMs, from the cloud-based server.

We also asked patients to bring all used medication blister packages when replenishing their medications, and TB doctors conducted empty blister counts at each outpatient to allow us to correctly determine adherence given any discrepancies compared to the EMM data. Village doctors conducted blister counts by home visits for patients who were lost-to-follow-up. The main cause of discrepancies was when the EMM box was opened but did not record this on the server due to technical issues, such as a lack of power or offline.

Outcomes

During each of their six or seven months of treatment (each treatment month = 30 days) patients are expected to take 30 daily doses of their TB medication. The primary outcome is a

binary variable indicating if poor adherence to TB treatment medication occurred during each treatment month, and defined as occurring if the patient missed ≥20% of planned medication doses in the treatment month.[12] In a standard treatment month this was equal to missing six or more doses.

We also recorded the following secondary outcomes at the patient-level: 1) Total percentage of missed doses over the entire treatment period. 2) Overall poor adherence: a binary variable indicating if the patient missed ≥10% of all planned doses (based on the corresponding indicator from China's national TB programme). Then the following WHO standard TB outcomes.[24] 3) Treatment success: a binary variable indicating either patient cure or treatment completion, where cure indicates "a pulmonary TB patient with bacteriologically confirmed TB at the beginning of treatment who was smear or culture-negative in the last month of treatment and on at least one previous occasion", and treatment completion indicates "a TB patient who completed treatment without evidence of failure but with no record to show that sputum smear or culture results in the last month of treatment and on at least one previous occasion were negative, either because tests were not done or because results are unavailable." 4) Lost to follow-up: a binary variable indicating "a TB patient who did not start treatment or whose treatment was interrupted for two consecutive months or more." 5) Poor treatment outcome: a binary variable indicating a TB patient having either of the WHO standard TB outcomes[24] of death, treatment failure ("a TB patient whose sputum smear or culture is positive at month five or later during treatment") or loss to follow-up. 6) Sputum conversion at

the end of the second month of treatment: a binary variable indicating if a patient diagnosed as sputum positive tested sputum negative at the end of their second month of treatment.

We also conducted a complementary process evaluation to explore programme implementation using the Consolidated Framework for Implementation Research, as well as a cost-effectiveness study, which we will report elsewhere.

Statistical analysis

We based our sample size calculation on our primary outcome and we used a simulation-based approach. See our protocol for full details.[19] In summary, based on prior unpublished data we assumed 40% monthly poor adherence in the control arm and a moderate within-patient correlation between each patients' six monthly outcome values of 0.5,[12] a 1:1 intervention:control allocation ratio, and a target difference (based on stakeholder views) of an absolute reduction in monthly poor adherence of 12 percentage points (i.e. down to 28% in the intervention arm). Our simulations used generalised estimating equations (GEEs) with binomial errors, identity links, first-order autoregressive (AR[1]) correlation structures, and Wald-based, two-tailed p-values. They showed we needed 300 patients to have 81.5% power to detect the target difference or greater at the $p \leq 0.05$ level, assuming no loss of data.

We summarised study design, socio-demographic, and treatment related characteristics at baseline via frequencies and percentages or medians and interquartile ranges, as appropriate. We analysed all outcomes other than sputum conversion based on the intention-to-treat

population, which included all individuals recruited into the trial as per their original treatment allocations. Our analysis of the sputum conversion outcome analysed patients as per their original treatment allocation but was restricted to the subgroup of individuals who were sputum positive at diagnosis and could therefore potentially convert. However, our primary outcome, total percentage of missed doses outcome, and overall poor adherence outcome, are all based on dose adherence data, which was incomplete for patients who died. Therefore, for these outcomes we present complete case results below that use the dose adherence available from patients who died prior to their deaths, but in the supplementary materials we present results from sensitivity analyses of these outcomes based on imputing the missing, post-death, outcome data for patients who died under the two most extreme potential scenarios in each treatment arm (see supplementary materials for full details).[25] For all other secondary outcomes we have complete data.

As per our protocol, and as recommended by CONSORT,[26] we draw our conclusions about the impact of the intervention on our binary primary outcome primarily based on the (adjusted) risk difference, which we calculated via marginal standardisation.[27-29] This involved first fitting a GEE to the data with a binomial distribution and logit link, clustering at the patient-level assuming an AR(1) correlation structure (across treatment months), and covariates for county (the stratum and centre), age (years), sex (male or female), job (farmer or other), marriage status (married or single/divorced/widowed), treatment month (1-6/7, and treated as a "continuous" variable), treatment arm, and the interaction between treatment month and treatment arm. We then used the GEE coefficients to calculate monthly predicted outcome

probabilities (via inverse logit transformation) for all patients with their covariates at their observed values other than the treatment arm variable which for all patients was first set to the (possibly counterfactual) intervention arm. We then repeated the predictions while the treatment assignment variable was set to the control arm for all patients. We then computed the adjusted risk difference by contrasting (intervention minus control) the means of these two sets of counterfactual marginal probabilities. We calculated the associated 95% confidence intervals via a clustered (at the patient-level) bootstrapping approach using the bias-corrected and accelerated method,[30] and the associated two-tailed p-values via a clustered (at the patient-level) permutation approach (10,000 bootstraps and permutations).[31] This approach to estimating risk differences deviated from our original protocol.

As a sensitivity analysis we also calculated the crude risk difference using the same approach but while only adjusting for county and treatment month. To aid interpretation and usability of our results, we also calculated adjusted and crude risk ratio estimates by contrasting the two sets of probabilities on the ratio scale. We estimated the intervention's effect on all our (binary) secondary outcomes as adjusted and crude risk differences and risk ratios and their associated 95% confidence intervals and p-values via the same approach described above but while using logistic regression models instead of GEEs as all secondary outcomes were binary and recorded at the patient level.

We also explored whether there was any effect modification on the additive scale for our primary outcome between treatment month groups (months one to three and months four to

six or seven), and for our primary outcome and treatment success outcome between sexes (male and female) and age groups (≥56 and <56). We estimated the adjusted risk difference for these outcomes within each subgroup using the same approach described above, and then estimated the effect modification on the additive-scale as the difference between each subgroups' adjusted risk difference. We calculated all associated 95% confidence intervals and p-values using the same approaches described above.

Results

Between 16 November 2018 and 5 April 2021, 664 patients were diagnosed with TB across our six trial sites, when each was recruiting. Of these 250 were ineligible. Among the remaining 414 eligible patients 119 were diagnosed during periods when the trial was suspended due to either COVID-19 lockdowns or the TB service delivery model change (where TB dispensaries were transferred from CDCs to county hospitals), and a further 17 refused. Therefore, we ultimately recruited and randomised 278 patients (67% of eligible participants), slightly short of our target sample size. We subsequently removed one patient from each arm for ineligibility as they were both later diagnosed with multi-drug resistant TB (Figure 1). Our intention-to-treat population therefore included 142 patients in the intervention arm and 134 patients in the control arm. We followed-up all patients until the end of their planned treatment period or death, and the last patient ended treatment on October 4 2021.

Overall, 62% (170) of patients were male, the median age was 56 years old, 87% (239) only had an education level of primary or below, 77% (213) were farmers, 82% (226) were married, the median reported monthly income was US$154, and the median family size was five. At diagnosis 75% (206) were sputum negative. There were no sizable differences in the measured characteristics between the two treatment arms, and the distribution of recruitment dates appeared well balanced (Table 1). All health providers and patients were followed-up regularly based on either our interventions or the requirements of the national TB programme. In the intervention arm only 1.4% (2/142) of patients were triggered to receive VOT in month 1 but this rose to 31.2% (44/141) by month 6, while over 91% (93/102) of patients requiring VOT successfully received it, with the rest receiving audio calls via WeChat due to poor signal (Table S7). When comparing EMM adherence data to the prioritised empty blister counts 2.1% (3/142) of patients in the intervention arm and 8.2% (11/134) of patients in the control arm had some discrepancies.

When considering the complete case analyses for our primary outcome the percentage of patients' treatment months with poor adherence was 10.2% (87/854) in the intervention arm and 36.5% (290/795) in the control arm, and the corresponding adjusted risk difference was -29.2 percentage points (95% CI: -35.3, -22.2; $p<0.001$). The crude risk difference was very similar. The sensitivity analyses exploring how imputing the missing primary outcome data (due to patient death) under the two most extreme potential scenarios showed very little differences from the complete case results (Table S1). Figure 2 shows that levels of poor adherence increased in both intervention and control arms across patient treatment months,

with the gap between the two arms appearing to remain broadly constant. This was reflected by the effect modification analysis of the primary outcome comparing treatment periods, which indicated a small but clear reduction, on the additive scale, in the adjusted risk difference during treatment months four to six or seven compared to during treatment months one to three, under the complete case and more plausible intervention-best control-worst scenario, but not the intervention-worst control-best scenario (Table S2). There was no evidence for any additive-scale effect modification of the primary outcome due to sex or age (Table S3 and S4). All the effect modification analyses had very low power though and should be treated as exploratory.

The intervention appeared to beneficially impact all secondary outcomes other than sputum conversion (Table 2). Treatment success was 93.7% (133/142) in the intervention arm and 73.1% (98/134) in the control arm, with an adjusted risk difference of 21 percentage points (95% CI: 12.4, 29.4; p<0.001) and a similar crude risk difference. There was no clear evidence of any effect modification on this outcome by sex (Table S5) or age group (Table S6). The total percentage of missed doses (complete case data) was 7.7% (1976/25594) in the intervention arm and 29.1% (6937/23872) in the control arm, with an adjusted risk difference of -22.4 percentage points (95% CI: -28.3, -16.9; p<0.001) and a similar crude risk difference. Overall poor adherence (complete case data) was 22.5% (32/142) in the intervention arm and 53.7% (72/134) in the control arm, with an adjusted risk difference of -32.7 percentage points (95% CI: -43.2, -21.8; p<0.001) and a similar crude risk difference. For the total percentage of missed doses and overall poor adherence outcomes the sensitivity analyses exploring how imputing

the missing outcome data due to patient death under the two most extreme potential scenarios again showed very minimal differences from the complete case results (Table S1). Lost to follow-up was 2.8% (4/142) in the intervention arm and 21.6% (29/134) in the control arm, with an adjusted risk difference of -19.3 percentage points (95% CI: -26.7, -11.8; p<0.001) and a similar crude risk difference. Poor treatment was 7% (10/142) in the intervention arm and 29.1% (39/134) in the control arm, with an adjusted risk difference of -23.1 percentage points (95% CI: -31.6, -14.1; p<0.001) and a similar crude risk difference. Finally, among patients who were sputum positive at diagnosis, sputum conversion at the end of the second month of treatment was 90% (36/40) in the intervention arm and 73.1% (19/26) in the control arm, with an adjusted risk difference of 13.2 percentage points (95% CI: -2.5, 41.1; p=0.2). The crude risk difference was slightly higher (19.2 percentage points [95% CI: -3, 39.5; p=0.077), but neither estimate provided clear evidence for any treatment effect on this outcome.

Discussion

In this trial, we evaluated the effect of a comprehensive intervention package including an EMM, an associated app to facilitate adherence monitoring and communication between patients and healthcare workers, provision of a treatment supporter, and a free smartphone data package, on TB treatment adherence and outcomes. Our intervention substantially reduced the monthly-level percentage of poor adherence by -29 percentage points (95% CI: -35, -22), representing a very clear benefit for patients and public health. We also observed clear improvements in treatment success, the total percentage of missed doses, overall poor

adherence, lost to follow-up, and poor treatment outcomes. In all cases, the size of the improvements implied clear benefits for patient and public health.

Our adherence-related results are broadly consistent with previous studies employing DATs to improve TB treatment adherence. Among five published randomised controlled trials using DATs to support TB treatment,[12-15,32] four studies were conducted in LMIC settings,[12,13,15,32] but only two, which were both in China, looked at adherence measures.[12,13] Compared with the two trials Liu and colleagues conducted in China[12,13] that employed an earlier version of EMMs, we identified a larger reduction of monthly-level poor adherence (missing ≥20% of monthly doses), and compared to their earlier trials we also found a larger reduction in the more stringent measure of exceeding 10% non-adherence overall, which has been strongly linked with unfavorable treatment outcomes.[33] We would recommend future adherence trials also report this more stringent measures. The 99DOTS trial in Uganda reported no improvement regarding treatment outcomes, but improved treatment success for TB patients who used the DAT, whom accounted for 52% of the total during the intervention period.[15] To the best of our knowledge, our trial, which employed a comprehensive package including a DAT, is one of the first such trials to report clear and substantial improvements in TB treatment success. We also demonstrated a large reduction in poor treatment outcomes, though direct comparison with the Kenya trial[32] is prohibitive due to heterogeneity in trial design and reporting. The possible reasons for these findings are 1) the poor adherence rate in Tibet was much higher compared to that of the rest of China, 2) we employed a new generation of EMMs that can update healthcare providers on adherence in real time, 3) we employed a comprehensive package of

interventions that improved communications between patients and health providers at the village, township, and county levels, and 4) most patients in our intervention arm used their EMMs during treatment.

One of our strengths is that we employed a comprehensive intervention package based on EMMs with a wide range of components to support treatment adherence. These included voice reminders and, via a linked app, simultaneous messaging to doctors' smartphones to record medication-taking, which facilitated follow-up communications (often in the form of WeChat audio/video calls). Our interventions also included family treatment supporters, who helped patients managing the EMM and app, which was particularly beneficial for older patients who had low smartphone literacy. We asked patients to report any side effects to their TB doctors by using the app. Patients also received health education messages from the app. Other intervention components have emphasised the importance of equitable access to interventions relying on technology, as such we also provided a data package to every patient in the intervention arm. While most intervention-arm patients used their own smartphones we provided a limited number (10.6% [15/142]) to those who did not have access to one. All the intervention components were co-designed with partners in Tibet to ensure good understanding of the cultural context.[34] We followed the Medical Research Council (MRC)'s guidelines in designing complex interventions,[35] which helped us to identify the specific aim of improving the communication between patients and health providers. Indeed, systematic reviews suggest that DATs alone will not improve treatment unless they also improve communication between patients and their care teams.[7] We first piloted our intervention

package to test feasibility and acceptability, before revising and then evaluating it in this trial. Another strength of the trial is that it comprises a robust real-world evaluation of the impact of EMMs on TB treatment outcomes in a remote and low-resource setting, similar to other settings in high TB-burden LMICs. We implemented a hybrid pragmatic trial that tested both the intervention package and its implementation strategy. We embedded our interventions within the roles of existing care providers to ensure feasibility. We used minimal exclusion criteria and our findings can apply widely to new pulmonary TB patients, given that 99% (248/250) of ineligible patients were ineligible due to their (pre-defined) type of TB and only 6.1% (17/278) of patients meeting those criteria refused to participate (Figure 1). We also followed the MRC guideline[36] in conducting a process evaluation to understand key issues of reach, fidelity, feasibility, adaptation and sustainability, which will be reported elsewhere.

The trial has several limitations. First, the operational nature of the trial made it impossible to blind patients, TB doctors, village doctors, or treatment supporters to the intervention but the investigators were blinded during data preparation. Second, the trial recruitment experienced long suspension periods due to the COVID-19 pandemic and TB service delivery model transitions. Tibet implemented several strict lockdowns in response to the COVID-19 outbreaks in China, and patients were not able to travel to TB dispensaries for diagnosis or treatment during the lockdowns. During the TB model transfer periods there were staff shortages, meaning that only routine diagnosis and treatment services were provided and trial recruitment had to suspend. We found no clear difference in the distribution of sex or age characteristics between the 136 eligible TB patients registered in the national TB programme

during the recruitment period who were not approached compared with the 278 recruited patients (Table S8). We expanded the trial to another four counties to fulfill the recruitment target. Third, there may be some ascertainment bias because healthcare providers were not blinded during the blister counts, and we used EMMs to collect adherence data while EMMs were one of the intervention components. Though our process evaluation did not find any evidence of "pill-dumping", i.e. participants disposing of pills prior to prescheduled adherence evaluation visits, this was still a possibility. Ideally we would have used objective measures such as urine isoniazid testing, but this was not feasible in this setting. Fourth, the EMM with voice message may not be applicable to people who live in dormitories, such as students or workers, as seen from our refusal cases. Fifth, we did not include retreatment cases, non-pulmonary TB patients, and MDR-TB patients because their treatment is much longer, may involve injectable medication, and may require hospitalisation in Lhasa. Therefore, our results may not be generalisable to this patient population group, who may have greater need for EMMs. Consequently, specific treatment adherence-support guidance using EMMs should be developed and trialed in these patient populations in future research. Finally, we included the option of VOT but some patients opted to use WeChat audio calls with their health providers.

There is an urgent global need for more DATs that are patient-centred and act as a low-cost alternative to DOT to ensure effective adherence to TB treatment. This need became more urgent during the pandemic when lockdowns and travel disruptions were more frequent. A variety of DATs, including EMMs, VOT, and 99DOTs were included in the recently published WHO consolidated guidelines for treating drug-susceptible and drug-resistant TB.[37] Our

experience showed that, although promising, new DATs need to be adapted into the local context to improve patient communications with health providers so to maximise patient uptake and use, and these technologies need to be embedded into the routine work of health providers.

In conclusion, our trial has provided evidence that a comprehensive package of interventions involving EMMs can improve treatment adherence and treatment outcomes among drug-susceptible TB patients in Tibet. Based on our results and continuous engagement with policy makers, our intervention package, including the EMM and associated app, has been included in the national TB programme within Tibet, with the local government providing funding to ensure every TB patient can access an EMM and the other interventions in our package.[38] In settings where mobile connectivity functions, EMMs that are adapted to the local context can be a key tool towards reaching the ambitious goals of the global END-TB Strategy by 2035.

Contributors

XW, JPH, ZZ, XL, and JH contributed to the study design. JDW, ZZ, JH and PP developed the guideline and training. ZZ, LL, TY, BZ, YL, VH and QP collected and prepared the data. JPH and ZZ analysed the data. JPH provided substantial scientific input in statistical methods and interpretation of the results. XW, JH, PP, LL and ZZ implemented the study. XW took the lead in drafting the manuscript. JPH, JDW, and VH provided substantial comments to improve the draft. All authors contributed to the collection and/or interpretation of data, provided critical revisions to the manuscript, and approved the final draft. XW and JH are the guarantors of the study.

Declaration of interests: We declare no competing interests.

Transparency declaration: The lead authors affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Data sharing: According to the ethical agreements, the data is owned by the Shigatse CDC, and available for research purpose of the study team. Individual patient data cannot be posted and downloaded in a public data depository, and not be able to be transmitted out of China due to the National Privacy Regulation. Only aggregated data can be shared upon request to Shigatse CDC (shigatsecdc@163.com).

Patient Involvement: Patients were involved in our process evaluation and provided valuable insights regarding how the interventions were experienced from patients. Those will be published elsewhere.

Ethics approval: We obtained ethics approval from the Office of Research Ethics at the University of Toronto (Ref: 36569) and the Ethics Review Committee of the Tibet Centre for Disease Control and Prevention (Ref: 006).

Funding source: The study was funded by TB REACH, administrated from the STOP TB Partnership at the United Nations. The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Acknowledgements: We thank Dr. Lixia Wang who was the previous Director of the China Centre for Disease Control and Prevention in facilitating the trial. We thank the Beijing FlowCloud Data Technology Co., Ltd. who provided the system and 324 electronic monitors to the trial. We also thank colleagues from the Shigatse Centre for Disease Control and Prevention (CDC), county CDCs and county hospitals of Samzhubze, Sa'gya, Gyantse, Ngamring, Tingri and Bainang for their efforts in implementing the study, and Tibetan provincial CDC and Shigatse Health Commission for their policy support. We also thank all research assistants and postgraduate students from University of Toronto, Tibetan University, and Weifang Medical College who participated in the data collection and interventions, for their technical support. Finally we thank all clinicians and patients participated in the study.

References

1. WHO. Global Tuberculosis Report, 2021. Geneva: World Health Oranization, 2021.
2. WHO. Adherence to long-term therapies. Geneva: World Health Organization, 2003.
3. Karumbi J, Garner P. Directly observed therapy for treating tuberculosis. *Cochrane Database Syst Rev* 2015; (5): CD003343.

4. Walley JD, Khan MA, Newell JN, Khan MH. Effectiveness of the direct observation component of DOTS for tuberculosis: a randomised controlled trial in Pakistan. *Lancet* 2001; **357**(9257): 664-9.
5. WHO. Guidelines for treatment of drug-susceptible tuberculosis and patient care, 2017 update. Geneva: World Health Organization, 2017.
6. Hou WL, Song FJ, Zhang NX, et al. Implementation and community involvement in DOTS strategy: a systematic review of studies in China. *Int J Tuberc Lung Dis* 2012; **16**(11): 1433-40.
7. Alipanah N, Jarlsberg L, Miller C, et al. Adherence interventions and outcomes of tuberculosis treatment: A systematic review and meta-analysis of trials and observational studies. *PLoS Med* 2018; **15**(7): e1002595.
8. Subbaraman R, de Mondesert L, Musiimenta A, et al. Digital adherence technologies for the management of tuberculosis therapy: mapping the landscape and research priorities. *BMJ Glob Health* 2018; **3**(5): e001018.
9. Checchi KD, Huybrechts KF, Avorn J, Kesselheim AS. Electronic medication packaging devices and medication adherence: a systematic review. *JAMA* 2014; **312**(12): 1237-47.
10. Nglazi MD, Bekker LG, Wood R, Hussey GD, Wiysonge CS. Mobile phone text messaging for promoting adherence to anti-tuberculosis treatment: a systematic review. *BMC Infect Dis* 2013; **13**: 566.
11. Ngwatu BK, Nsengiyumva NP, Oxlade O, et al. The impact of digital health technologies on tuberculosis treatment: a systematic review. *Eur Respir J* 2018; **51**(1).
12. Liu X, Lewis JJ, Zhang H, et al. Effectiveness of Electronic Reminders to Improve Medication Adherence in Tuberculosis Patients: A Cluster-Randomised Trial. *PLoS Med* 2015; **12**(9): e1001876.
13. Liu X, Thompson J, Dong H, et al. Digital adherence technologies to improve tuberculosis treatment outcomes in China: a cluster-randomised superiority trial. *Lancet Glob Health* 2023; **11**(5): e693-e703.
14. Story A, Aldridge RW, Smith CM, et al. Smartphone-enabled video-observed versus directly observed treatment for tuberculosis: a multicentre, analyst-blinded, randomised, controlled superiority trial. *Lancet* 2019; **393**(10177): 1216-24.
15. Cattamanchi A, Crowder R, Kityamuwesi A, et al. Digital adherence technology for tuberculosis treatment supervision: A stepped-wedge cluster-randomized trial in Uganda. *PLoS Med* 2021; **18**(5): e1003628.
16. Wang N, Zhang H, Zhou Y, et al. Using electronic medication monitoring to guide differential management of tuberculosis patients at the community level in China. *BMC Infectious Diseases* 2019; **19**: 844.
17. Haldane V, Zhang Z, Ma Q, et al. A qualitative study of perspectives on access to tuberculosis health services in Xigaze, China. *Infect Dis Poverty* 2021; **10**(1): 120.
18. Team RC. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria* <*https://wwwR-projectorg/*> 2023.
19. Wei X, Hicks JP, Pasang P, et al. Protocol for a randomised controlled trial to evaluate the effectiveness of improving tuberculosis patients' treatment adherence via electronic monitors and an app versus usual care in Tibet. *Trials* 2019; **20**(1): 273.
20. Venkatesh V, Morris MG, Davis GB, Davis FD. User Acceptance of Information Technology: Toward a unified view. *MIS Quarterly* 2003; **27**(3): 425-78.

21. Wang L, Zhang H, Ruan Y, Chin D, Xia Y, Cheng S. Tuberculosis prevalence in China, 1990-2010; a longitudinal analysis of national survey data. *Lancet* 2014; **383**: 2057 - 64.
22. Wei X, Zou G, Walley J, et al. China tuberculosis policy at crucial crossroads: comparing the practice of different hospital and tuberculosis control collaboration models using survey data. *PloS one* 2014; **9**(3): e90596.
23. TB CARE I. International Standards for Tuberculosis Care, Edition 3. The Hagure: TB CARE I, 2014.
24. WHO. Definitions and reporting framework for tuberculosis- 2013 revision. Geneva: World Health Orgnaization, 2014.
25. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Medical Research Methodology* 2017; **17**(1): 162.
26. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; **340**: c869.
27. Kleinman LC, Norton EC. What's the Risk? A simple approach for estimating adjusted risk measures from nonlinear models including logistic regression. *Health Serv Res* 2009; **44**(1): 288-302.
28. Muller CJ, MacLehose RF. Estimating predicted probabilities from logistic regression: different methods correspond to different target populations. *Int J Epidemiol* 2014; **43**(3): 962-70.
29. Pedroza C, Truong VT. Performance of models for estimating absolute risk difference in multicenter trials with binary outcome. *BMC Med Res Methodol* 2016; **16**(1): 113.
30. Thomas JD, Bradley E. Bootstrap confidence intervals. *Statistical Science* 1996; **11**(3): 189-228.
31. Davison AC, Hinkley DV. Bootstrap Methods and their Application. Cambridge: Cambridge University Press; 1997.
32. Yoeli E, Rathauser J, Bhanot SP, et al. Digital Health Support in Treatment for Tuberculosis. *N Engl J Med* 2019; **381**(10): 986-7.
33. Imperial MZ, Nahid P, Phillips PPJ, et al. A patient-level pooled analysis of treatment-shortening regimens for drug-susceptible pulmonary tuberculosis. *Nat Med* 2018; **24**(11): 1708-15.
34. Haldane V, Li BP, Ge S, et al. Exploring the translation process for multilingual implementation research studies: a collaborative autoethnography. *BMJ Glob Health* 2022; **7**(5).
35. Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008; **337**: a1655.
36. Moore GF, Audrey S, Barker M, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ* 2015; **350**: h1258.
37. WHO. WHO operational handbook on tuberculosis: module 4: treatment: tuberculosis care and support. Geneva: World Health Organization, 2022.
38. Tibetan CDC. Notice to employ electronic medication monitors and associated app for management of patients with tuberculosis in Tibet Autonomous Region. In: Disease Control Department, editor. Lhasa, Tibet, China: Tibetan Centre for Disease Control and Provention; 2021.

Box 1. Detailed intervention and control arm procedures

Intervention components	Intervention arm	Control arm
Electronic medication monitors (EMMs)	1. At recruitment the patient's TB doctor assigned them an EMM box, linked it with the patient's medical record, and demonstrated how to use the EMM box. The treatment supervisor (village doctor) visited the patient's home within the first week of treatment and helped them solve any remaining problems they may have had when using the EMM. 2. The EMM reminded the patient to take their medicine on time everyday using human voice recordings. 3. The EMM transmitted medication adherence history, based on when the patient opened their EMM, to a cloud-based server that could be accessed by their TB doctor and treatment supervisor through the website or an app. These health staff were then able to use this information to monitor the patient's adherence history and outpatient visits, and could provide specific follow-up support accordingly.	1. At recruitment the patient's TB doctor assigned them an EMM box, linked it with the patient's medical record, and demonstrated how to use the EMM. The treatment supervisor (village doctor) visited the patient's home within the first week of treatment and helped them solve any remaining problems they may have had when using the EMM. 2. The EMM reminder function was disabled. 3. EMMs transmitted medication adherence history, based on when the patient opened their EMM, to a cloud-based server, but this information could not be accessed by their TB doctor or treatment supervisor. These health staff visited patients to provide follow-up support at their own discretion.
Treatment supporter (family member)	1. During recruitment or at the first home visit by the treatment supervisor a family member was chosen to act as a treatment supporter. The family member had to live in the same house as the patient, care about the patient, and be literate in using a smartphone. 2. When appointed, the family member was trained by the treatment supervisor about their responsibilities, which included an expectation of providing psychological support to patients and helping them to use the EMM and the WeChat app (see below).	Family treatment supporters were not selected.
Smartphone-enhanced communication	1. At recruitment or during the first home visit the TB doctor or treatment supervisor demonstrated to the patient how to install and use the WeChat app on their smartphone, or if the patient was not able to use a smartphone or WeChat they demonstrated to the patient's treatment supporter how to install and use the WeChat app on their smartphone. 2. This app enabled direct communication between the patient (or their treatment supporter) and their TB doctor and/or treatment supervisor about their medication, medication adherence history, follow-up issues, and/or specific health education.	Patients or treatment supporters were not able to use the WeChat app, and were not provided with free data plans or smartphones, and TB doctors and treatment supervisors contacted patients through traditional means, such as physical visits or phone calls.

	3. Free smartphone data plans were provided to all patients, and free smartphones were provided to patients or their treatment supporters who did not have one.	
Video-observed treatment (VOT)	1. TB treatment supervisors or TB doctors identified patients at high risk of being lost to follow-up (not starting treatment or missing all medication doses for two or more consecutive treatment-months) based on whether the patient a) refused to start or continue treatment, b) missed three consecutive medication doses (as recorded by the EMM data), or c) expressed serious concerns about maintaining their adherence to their medication. 2. Treatment supervisors were trained that if their patient was considered a high risk of becoming lost to follow-up they should explore with the patient their reasons for non-compliance and what help may be needed. The treatment supervisor or TB doctor then initiated VOT while trying to provide any help needed. 3. The treatment supervisor documented the implementation of VOT. Patients who accepted the invitation to begin VOT were given instructions either in the TB dispensary or at home during the visit of their treatment supervisor. Using the WeChat app VOT was done with the TB supervisor via a live video conversation or via recorded video/pictures showing them taking their medicine. VOT was provided until three consecutive doses were completed on time, after which the patient switched back to the normal approach using their EMMs. Patients who did not agree to receive VOT were called daily by treatment supervisors until three consecutive doses were completed.	Video-observed treatment was not provided to patients.

Table 1. Characteristics of the 276 participants in the trial

Patient characteristics	Intervention group	Control group	Overall
Number of patients	142	134	276
County of residence			
Samzhubze	36 (25.4%)	34 (25.4%)	70 (25.4%)
Sa'gya	39 (27.5%)	37 (27.6%)	76 (27.5%)
Gyantse	44 (31.0%)	41 (30.6%)	85 (30.8%)
Ngamring	12 (8.5%)	12 (9.0%)	24 (8.7%)
Tingri	9 (6.3%)	5 (3.7%)	14 (5.1%)
Bainang	2 (1.4%)	5 (3.7%)	7 (2.5%)
Sex			
Male	89 (62.7%)	81 (60.4%)	170 (61.6%)
Female	53 (37.3%)	53 (39.6%)	106 (38.4%)
Age (years)	57.0 (40.2, 64.8)	55.0 (40.0, 62.8)	56.0 (40.0, 64.0)
Education level			
Primary school or below	127 (89.4%)	112 (83.6%)	239 (86.6%)
High school	12 (8.5%)	13 (9.7%)	25 (9.1%)
College or above	3 (2.1%)	9 (6.7%)	12 (4.3%)
Job			
Farmer	106 (74.6%)	107 (79.9%)	213 (77.2%)
Other	36 (25.4%)	27 (20.1%)	63 (22.8%)
Monthly income (U.S. Dollars*)	154.6 (0.0, 309.1)	185.5 (0.0, 463.7)	154.6 (0.0, 343.9)
Number of family members	5.0 (3.0, 7.0)	6.0 (4.0, 8.0)	5.0 (3.0, 7.0)
Marital status			
Married	116 (81.7%)	110 (82.1%)	226 (81.9%)
Single, divorced, or widowed	26 (18.3%)	24 (17.9)	50 (18.1%)
Sputum smear test at diagnosis			
Negative	102 (71.8%)	104 (77.6%)	206 (74.6%)
Positive	40 (28.2%)	30 (22.4%)	70 (25.4%)
Planned treatment length			
6 months	138 (97.2%)	128 (95.5%)	266 (96.4%)
7 months	4 (2.8%)	6 (4.5%)	10 (3.6%)
Recruitment date count†	519.5 (280.0, 711.5)	545.0 (305.2, 709.0)	527.0 (289.2, 710.5)

Data are frequency (%) for categorical variables and median (IQR) for numerical variables. There are no missing values for any variable. * Based on the exchange rate on 30 April 2021: 1 US dollar = 6.47 RMB. † Count of the number of days between each patients' recruitment date and the date the first patient was recruited.

Table 2. Intervention effects on all outcomes

	Summary values		Risk difference in percentage points (95% CI); p-value		Risk ratio (95% CI); p-value	
Outcome	**Intervention**	**Control**	**Adjusted**	**Crude**	**Adjusted**	**Crude**
Primary outcome						
Monthly poor treatment adherence (missing ≥20% of planned doses in the treatment month)	10.2% (87/854)	36.5% (290/795)	-29.2 (-35.3, -22.2); <0.001	-28.7 (-34.5, -21.8); <0.001	0.33 (0.23, 0.43); <0.001	0.34 (0.26, 0.45); <0.001
Secondary outcomes						
Treatment success (cured [negative at completion and ≥1 other month] or completed treatment [complete with no evidence of failure but results proving cured were not done/unavailable])	93.7% (133/142)	73.1% (98/134)	21 (12.4, 29.4); <0.001	21.4 (12.8, 29.8); <0.001	1.29 (1.16, 1.45); <0.001	1.3 (1.17, 1.46); <0.001
Total % of planned doses missed during treatment	7.7% (1976/25594)	29.1% (6937/23872)	-22.4 (-28.3, -16.9); <0.001	-22.2 (-27.9, -16.8); <0.001	0.25 (0.18, 0.35); <0.001	0.25 (0.18, 0.35); <0.001
Overall poor treatment adherence (missing ≥10% of all planned doses)	22.5% (32/142)	53.7% (72/134)	-32.7 (-43.2, -21.8); <0.001	-32.2 (-42.7, -21.4); <0.001	0.4 (0.28, 0.56); <0.001	0.41 (0.28, 0.57); <0.001
Lost to follow-up (never started treatment after diagnosis or missed ≥2 consecutive months of treatment)	2.8% (4/142)	21.6% (29/134)	-19.3 (-26.7, -11.8); <0.001	-19.7 (-27.1, -12.2); <0.001	0.12 (0.03, 0.3); <0.001	0.12 (0.03, 0.29); <0.001
Poor treatment (died, lost to follow-up, or treatment failure [positive at treatment month ≥5])	7.0% (10/142)	29.1% (39/134)	-23.1 (-31.6, -14.1); <0.001	-23.1 (-31.5, -14.3); <0.001	0.23 (0.11, 0.42); <0.001	0.23 (0.11, 0.41); <0.001
Sputum conversion (positive to negative) at end of 2nd month	90% (36/40)	73.1% (19/26)	13.2 (-2.5, 41.1); 0.2	19.2 (-3, 39.5); 0.077	1.18 (1, 2.74); 0.2	1.28 (1, 1.88); 0.077

Treatment arm summary values are % (number of events/total possible events). Denominators for the primary outcome represent the number of patient treatment-months per arm across all patients. Denominators for the total % of planned doses missed during treatment outcome represent the number of

planned doses per arm across all patients. All other outcome denominators represent the number of patients per arm. For the treatment effect results the binary primary outcome was derived and analysed at the patient treatment-month level (planned treatment length being six/seven months), and all other outcomes were derived and analysed at the patient level. All treatment effect results calculated via a marginal standardisation approach using bootstrapping and permutation methods to obtain the confidence intervals and p-values respectively (incorporating a clustered approach for the primary outcome). All adjusted results were adjusted for the covariates county (the stratum/centre), age (years), sex (male/female), job (farmer/other), marriage status (married/single or divorced or widowed), and the primary outcome alone was also adjusted for treatment month (treated as a “continuous” variable). All crude results were only adjusted for county (the stratum/centre), and the primary outcome alone was again also adjusted for treatment month. All outcomes other than sputum conversion included the intention-to-treat population. When analysing the primary outcome, the total % of planned doses missed during treatment outcome, and the overall poor adherence outcome, these analyses included outcome values for patients who died during treatment that were derived only from their dose adherence data available prior to death (see supplementary materials for full details and sensitivity analyses). Sputum conversion analyses included the subset of the intention-to-treat population who were sputum negative at diagnosis.

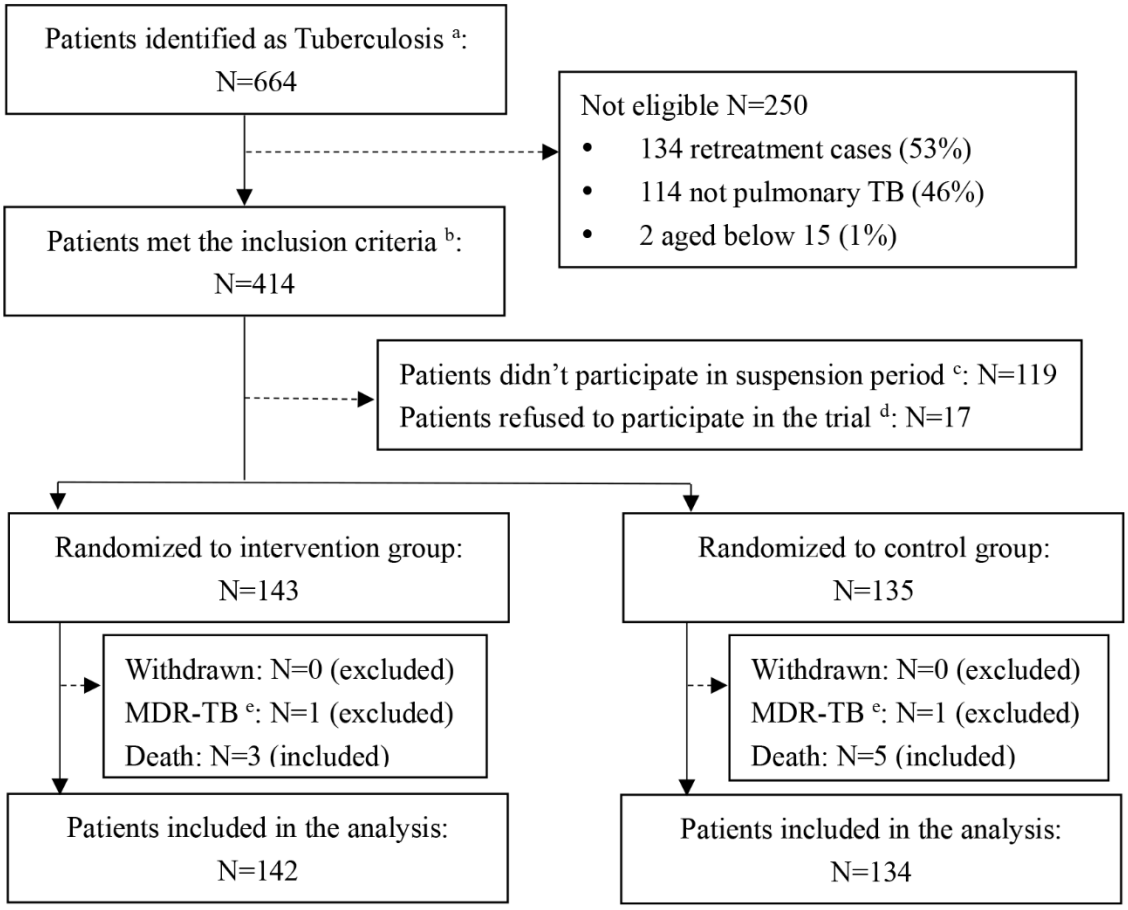

Figure 1. Trial profile

a. Patients were identified from 16 November 2018 to 31 March 2021 in six counties in Tibet, China.
b. Inclusion criteria were patients aged 15 years or older and newly diagnosed as active pulmonary TB starting outpatient (ambulatory) care on the standard 6-month treatment regimen.
c. During the recruitment period the trial was suspended for 9 months in Samzhubze, 7 months in Sa'gya, and 3 months in Gyantse in total due to either 1) impact of COVID-19 emergency response or 2) TB diagnosis and treatment functions were transferred from county CDC to county hospital according to the local policy change. During the suspension period, all patients received usual TB management but were not invited to participate in the trial. There were 50, 46, and 23 eligible patients identified during the suspension period in Samzhubze, Sa'gya and Gyantse respectively.
d. Among 17 patients who refused to participate in the trial, 6 patients (35.3%) were students who did not want to bring the electronic monitor into their dormitories, 6 patients (35.3%) were bus drivers or construction workers who did not have a place for the electronic monitor in their dormitory, and 5 patients (29.4%) simply refused to participate in the trial without providing clear reasons.
e. According to the local TB management policy, bacteriologically positive TB sample will be sent for drug susceptibility testing. There were two participants identified as MDR-TB later and thus no longer met the inclusion criteria, as such these participants were excluded for analysis.

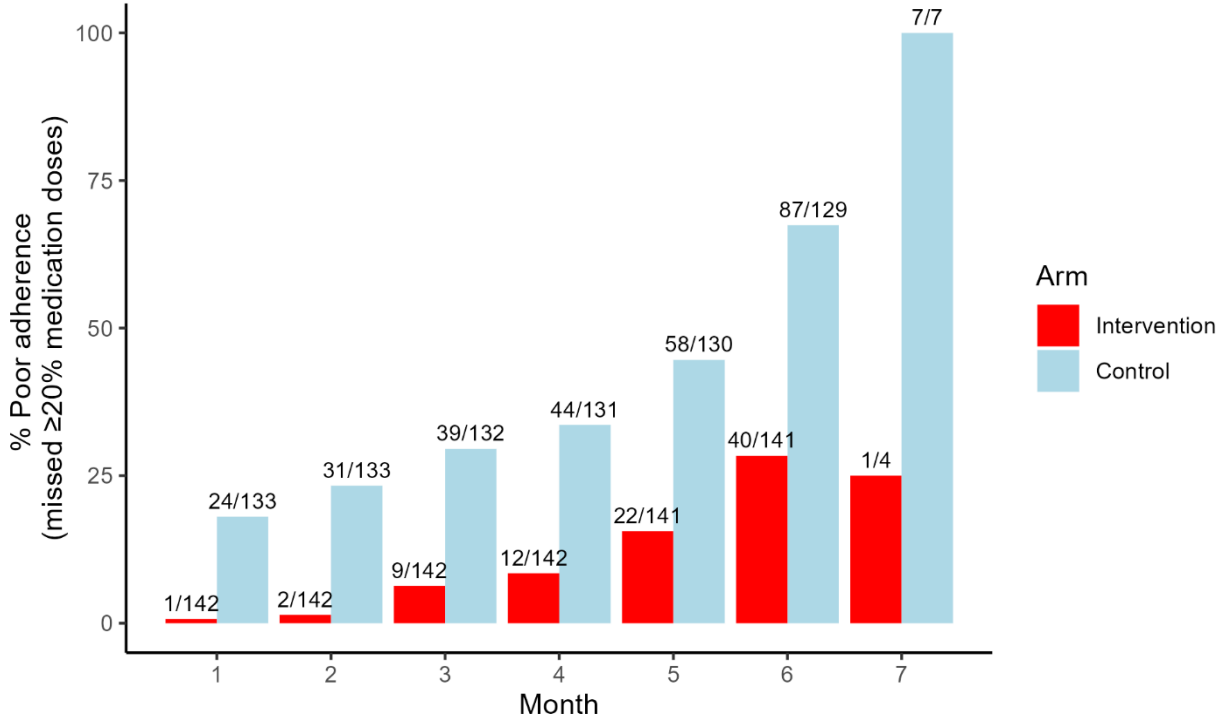

Figure 2. Percentage of patients missing ≥20% planned medication doses per treatment month

Patients are expected to take 30 medication doses per treatment month. Bar heights represent the percentage of patients who missed ≥20% of their 30 planned medication doses in the given treatment month by treatment arm. Numbers above bars represent the number of patients who missed ≥20% of their 30 planned medication doses in the given treatment month out of the number of patients receiving treatment that month. Patients who died are included in the figure's results for months up to and including the month of their death if it was possible to determine their poor/non-poor adherence status for that month, but otherwise excluded, and are also excluded from the results for any months subsequent to their death.

Supplementary Materials

Effectiveness of a comprehensive package based on electronic medication monitors at improving treatment outcomes among tuberculosis patients in Tibet: a multi-centre randomised controlled trial

Xiaolin Wei (PhD), Joseph Paul Hicks (PhD), Zhitong Zhang (MPH), Victoria Haldane (PhD), Pande Pasang (MD), Linhua Li (MD), Tingting Yin (MPH), Bei Zhang (MPH), Yinlong Li (PhD), Qiuyu Pan (PhD), Xiaoqiu Liu (PhD), John Walley (FFPH), Jun Hu (PhD)

Dalla Lana School of Public Health, University of Toronto, Toronto, Canada (Prof Xiaolin Wei [full professor], Zhitong Zhang); Nuffield Centre for International Health and Development, University of Leeds, Leeds, UK (Joseph P Hicks, Prof John Walley [full professor]); Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada (Prof Xiaolin Wei [full professor], Victoria Haldane); Shigatse Centre for Disease Control and Prevention, Shigatse, China (Pande Pasang, Linhua Li, Jun Hu); Weifang Medical College, Weifang, China (Tingting Yin, Bei Zhang); Jining Medical University, Jining, China (Yinlong Li); North Sichuan Medical College, Nanchong, China (Qiuyu Pan); National Center for TB control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China (Xiaoqiu Liu); Shandong University of Traditional Chinese Medicine, Jinan, China (Prof Jun Hu [full professor])

Correspondence to:

Xiaolin Wei and Jun Hu

Contents

1. Additional statistical analysis details

1.1. Software

We used the *blockrand*[1] package to create the randomisation scheme and cards. We used the *geepack*[2] package to fit the generalised estimating equations (GEEs), *base* R functions to fit the logistic regressions, the *rsample*[3] package to do the bootstrap resampling, the *coxed*[4] package to calculate the bias-corrected accelerated bootstrap confidence intervals, and the *permute*[5] package to do the permutations necessary to calculate the permutation-based p-values.

1.2. Missing data

1.2.1. Overview

For all patients we were able to collect full covariate data and full outcome data for our secondary outcomes of treatment success, lost to follow-up, poor treatment outcome, and sputum conversion at the end of the second month of treatment. Therefore, for these outcomes we had no missing data. Note that lost to follow-up is a WHO standard tuberculosis outcome, and is defined as a patient either not initiating treatment or missing treatment for two consecutive months or more. It therefore represents a patient being "lost" from treatment but necessarily not lost from the study, and in this study we were indeed able to track all patients and their outcomes for the duration of the study, either via direct contact or via contact with their treatment provider or treatment supporter. However, our primary outcome of monthly poor adherence and our secondary outcomes of the total percentage of doses missed and overall poor adherence were all computed from patient medication dose adherence data. Therefore, patients who died during treatment by definition did not have full dose adherence data for their entire treatment period (while all other patients did).

Our broad approach to deal with this missing data problem for these three outcomes was to present complete case analyses based on the available data in the main paper, and then present a set of sensitivity analyses here in the supplementary materials to explore how the results for these outcomes change when imputing the missing data under two contrasting extreme assumptions.

1.2.2. Complete case analyses

For our primary outcome complete case analyses for patients who died we set the binary monthly indicator of poor adherence (missing ≥20% of planned doses for that month) for the treatment month during which they died to missing unless there was sufficient data to impute a guaranteed outcome. Specifically, if the patient had lived long enough during that month and had either missed enough doses or taken enough doses such that that month's primary outcome could only be poor/good adherence respectively, we imputed the primary outcome for that month accordingly. For all patients who died prior to their final month of treatment we

set all primary outcome monthly indicators subsequent to the month in which they died to missing.

For our secondary outcome of the total percentage of doses missed for our complete case analyses for patients who died we simply used the dose adherence data that was available prior to their death. For our secondary outcome of overall poor adherence (≥10% of all doses missed) for our complete case analyses for patients who died we computed the binary indicator based on the available data.

1.2.3. Sensitivity analyses

Our sensitivity analyses used a "best-worst" "worst-best" approach to explore how imputing the missing data for the three outcomes under the contrasting two most extreme scenarios altered the results and conclusions. For both sets of sensitivity analyses we only imputed data for patients who died, on the assumption that they had lived for the full extent of their planned treatment duration (in all cases this was six months). For our best-worst analyses we assumed that intervention arm patients who died actually survived and had full adherence to their medication subsequent to their death, while control arm patients who died actually survived and had full non-adherence to their medication subsequent to their death. For our worst-best analyses we simply assumed the opposite scenario: intervention arm patients who died actually had survived and had full non-adherence to their medication subsequent to their death, while control arm patients who died actually survived and had full adherence to their medication subsequent to their death. To be clear, we did not change their adherence data prior to death.

For each scenario we then used the partially imputed dose-level adherence data to re-compute the three outcomes and re-analyse them using the methods previously described (i.e. producing crude and adjusted measures of treatment effect in the form of risk differences and risk ratios and their associated confidence intervals and p-values). We present the complete case, best-worst, and worst-best sets of results for these three outcomes together below in table S1 for easy comparison.

2. Deviations for protocol planned methods: justifications and further details not discussed in the main paper

2.1. Treatment completion secondary outcome switched to treatment success

We dropped treatment completion and switched to including treatment success as a secondary outcome prior to the data analysis because we realised we had made a mistake when planning to include treatment completion instead of treatment success. As per the WHO definitions and reporting framework for TB outcomes[6] treatment success is a more comprehensive indicator than treatment completion as it measures if a patient either completes treatment without evidence of cure or is show to be cured. It is therefore more

relevant as an indicator of a successful clinical outcome for a patient, following their treatment for TB.

2.2. Analysis methods

2.2.1. Primary outcome

We changed our protocol-planned approach to estimating the primary outcome's adjusted and crude risk differences and risk ratios directly via GEEs with binomial errors and identity links to using the marginal standardisation approach described in the main paper because we had universal model convergence issues with our originally planned approach (a known issue with the approach that, with hindsight, we should have planned for in the protocol). We then also switched to calculating the associated 95% confidence intervals and p-values via the bootstrap and permutation approaches described in the main paper, as a robust way to obtain these quantities. The other commonly used approach is to rely on delta-method based standard errors, but bootstrapping and permutation requires fewer assumptions to obtain valid measures. This change in approach therefore also applied to the effect modification analyses for the primary outcome.

2.2.2. Secondary outcomes

We already planned in the protocol to use marginal standardisation to obtain adjusted and crude risk differences for all secondary outcomes other than the total percentage of missed doses. However, for this secondary outcome we had previously planned to use linear regression to directly estimate the treatment effect, with the outcome being derived as the patient-level proportion of missed doses. With hindsight this was a poor choice as such an outcome was never likely to be approximately normally distributed. Therefore, for a more appropriate approach that was also consistent with how we analysed our other binary secondary outcomes we instead estimated the adjusted and crude risk differences and risk ratios for this outcome via marginal standardisation using logistic regression models, by treating the outcome as a true binomial outcome (with the number of successes being the number of missed doses per patient and the number of trials being the number of planned doses per patient). We did initially try converting the outcome to a binary, dose-level outcome, but the data became so large that it was computational not an efficient approach when having to run large numbers of bootstraps and permutations.

2.2.3. Covariate adjustment choice

In our protocol we planned for our adjusted and crude analyses of the primary outcome to adjust for treatment month as a categorical variable. However, after examining partial residual plots from the adjusted and crude GEEs there was no clear indication of a non-linear relationship between time and the primary outcome. Therefore, we decided instead to adjust for treatment month as a continuous variable (in practice it was recorded as a discrete count) as this provided an adequate fit to the data while requiring far fewer model

parameters, particularly when it came to the effect modification analyses. We also planned to adjust the adjusted analyses of the primary outcome and all secondary outcomes for age as a categorical variable (15-49 or 50+), marriage (married or never married/not currently married), and employment status (farmer, other, or unemployed/retired). However, the partial residual plots again showed no clear there was no clear non-linear relationship between age and any outcomes, so we instead adjusted for age as a continuous (in reality discrete) variable for the same reason. Then, due to data sparsity, we instead adjusted all adjusted outcome analyses for marriage status as married or other (i.e. single/divorced/widowed), and employment status as farmer or other.

2.2.4. Effect modification: treatment completion, travel time, and treatment period

We originally planned to analyse effect modification for the originally-planned secondary outcome of treatment completion. However, as we changed this outcome to treatment success we analysed effect modification for this outcome instead. In our protocol we also planned to analyse effect modification for our primary outcome and treatment completion due to travel time to TB dispensaries, but we were unable to collect the required data to do these analyses and so they were dropped. Then lastly, subsequent to the protocol we decided it would be important to understand whether the effect of the intervention on the primary outcome and treatment success differed between the earlier and latter phases of treatment. We therefore added effect modification analyses for these two outcome where we looked at the intervention effect during the first three treatment months and during the final three treatment months (or during the final four months of treatment for patients who had seven-month treatment plans).

Table S1. Intervention effect on primary and relevant secondary outcomes under complete case, intervention-best control-worst, and intervention-worst control-best imputation scenarios for death-related missing dose adherence data

	Summary values		Risk difference in percentage points (95% CI); p-value		Risk ratio (95% CI); p-value	
Outcome	**Intervention**	**Control**	**Adjusted**	**Crude**	**Adjusted**	**Crude**
Complete case analysis						
Primary outcome						
Monthly poor adherence	10.2% (87/854)	36.5% (290/795)	-29.2 (-35.3, -22.2); <0.001	-28.7 (-34.5, -21.8); <0.001	0.33 (0.23, 0.43); <0.001	0.34 (0.26, 0.45); <0.001
Secondary outcomes						
Total missed doses	7.7% (1976/25594)	29.1% (6937/23872)	-22.4 (-28.3, -16.9); <0.001	-22.2 (-27.9, -16.8); <0.001	0.25 (0.18, 0.35); <0.001	0.25 (0.18, 0.35); <0.001
Overall poor adherence	22.5% (32/142)	53.7% (72/134)	-32.7 (-43.2, -21.8); <0.001	-32.2 (-42.7, -21.4); <0.001	0.4 (0.28, 0.56); <0.001	0.41 (0.28, 0.57); <0.001
Intervention-best control-worst imputation scenario						
Primary outcome						
Monthly poor adherence	9.9% (85/856)	37.7% (306/811)	-30.3 (-36.5, -23.5); <0.001	-30 (-35.9, -23.1); <0.001	0.32 (0.24, 0.42); <0.001	0.32 (0.24, 0.43); <0.001
Secondary outcomes						
Total missed doses	7.7% (1976/25680)	30.4% (7395/24330)	-23.8 (-29.7, -18.1); <0.001	-23.5 (-29.2, -18); <0.001	0.24 (0.17, 0.33); <0.001	0.24 (0.17, 0.33); <0.001
Overall poor adherence	22.5% (32/142)	55.2% (74/134)	-34.3 (-44.8, -23.4); <0.001	-33.7 (-43.9, -22.8); <0.001	0.39 (0.27, 0.54); <0.001	0.4 (0.28, 0.54); <0.001
Intervention-worst control-best imputation scenario						

Primary outcome						
Monthly poor adherence	10.5% (90/856)	35.6% (289/811)	-27.1 (-33.1, -20.2); 0	-26.8 (-32.6, -19.9); 0	0.36 (0.27, 0.47); <0.001	0.36 (0.25, 0.46); <0.001
Secondary outcomes						
Total missed doses	8.0% (2062/25680)	28.5% (6937/24330)	-21.5 (-27.2, -16); <0.001	-21.3 (-26.9, -15.8); <0.001	0.27 (0.19, 0.37); <0.001	0.27 (0.19, 0.37); <0.001
Overall poor adherence	22.5% (32/142)	53.7% (72/134)	-32.7 (-43.2, -21.8); <0.001	-32.2 (-42.7, -21.4); <0.001	0.4 (0.28, 0.56); <0.001	0.41 (0.28, 0.56); <0.001

Treatment arm summary values are % (number of events/total possible events). Denominators for the primary outcome represent the number of patient treatment months per arm across all patients. Denominators for the total % of planned doses missed during treatment outcome represent the number of planned doses per arm across all patients. Denominators for overall poor treatment represent the number of patients per arm. For the treatment effect results the binary primary outcome was derived and analysed at the patient treatment-month level (planned treatment length being six/seven months), and all other outcomes were derived and analysed at the patient level. All treatment effect results calculated via a marginal standardisation approach using bootstrapping and permutation methods to obtain the confidence intervals and p-values respectively (incorporating a clustered approach for the primary outcome). All adjusted results were adjusted for the covariates county (the stratum/centre), age (years), sex (male/female), job (farmer/other), marriage status (married/single or divorced or widowed), and the primary outcome alone was also adjusted for treatment month (treated as a “continuous” variable) and the interaction between treatment arm and treatment month. All crude results were only adjusted for county (the stratum/centre), and the primary outcome alone was again also adjusted for treatment month and the interaction between treatment arm and treatment month. All outcomes included the intention-to-treat population. Complete case analyses included outcome values for patients who died during treatment that were derived only from their dose adherence data available prior to death. Intervention-best control-worst analyses included outcome values for patients who died during treatment that were derived from their dose adherence data following imputation of the missing, post-death, dose adherence data, until the end of their planned treatment, assuming full dose adherence after death for death-patients in the intervention arm and full dose non-adherence after death for death-patients in the control arm. Intervention-worst control-best analyses followed the same imputation approach but assuming non-full adherence for death-patients in the intervention arm and full adherence for death-patients in the control arm following death.

Table S2. Primary outcome effect modification by treatment period (months 1-3 and 4-6/7) under complete case, intervention-best control-worst, and intervention-worst control-best imputation scenarios for death-related missing dose adherence data

Summary value/intervention effect estimated	Estimate
Complete case analysis	
Month 1-3 intervention outcome summary (% [n/N])	2.8% (12/426)
Month 1-3 control outcome summary (% [n/N])	23.6% (94/398)
Month 1-3 intervention effect: ARD (95% CI); p-value	-24.6 (-29.6, -16.6); <0.001
Month 4-6/7 intervention outcome summary (% [n/N])	17.5% (75/428)
Month 4-6/7 control outcome summary (% [n/N])	49.4% (196/397)
Month 4-6/7 intervention effect: ARD (95% CI); p-value	-33.3 (-42, -26.6); <0.001
Effect modification: month 4-6/7 ARD - month 1-3 ARD (Percentage point [95% CI]; p-value)	-8.7 (-18.7, -4); 0.011
Intervention-best control-worst scenario	
Month 1-3 intervention outcome summary (% [n/N])	2.8% (12/426)
Month 1-3 control outcome summary (% [n/N])	24.4% (98/402)
Month 1-3 intervention effect: ARD (95% CI); p-value	-25.2 (-30.5, -17.3); <0.001
Month 4-6/7 intervention outcome summary (% [n/N])	17% (73/430)
Month 4-6/7 control outcome summary (% [n/N])	50.9% (208/409)
Month 4-6/7 intervention effect: ARD (95% CI); p-value	-35 (-43.5, -28.2); <0.001
Effect modification: month 4-6/7 ARD - month 1-3 ARD (Percentage point [95% CI]; p-value)	-9.7 (-19.5, -4.8); 0.004
Intervention-worst control-best scenario	
Month 1-3 intervention outcome summary (% [n/N])	2.8% (12/426)
Month 1-3 control outcome summary (% [n/N])	23.4% (94/402)
Month 1-3 intervention effect: ARD (95% CI); p-value	-24 (-29.3, -16.2); <0.001
Month 4-6/7 intervention outcome summary (% [n/N])	18.1% (78/430)
Month 4-6/7 control outcome summary (% [n/N])	47.7% (195/409)
Month 4-6/7 intervention effect: ARD (95% CI); p-value	-30.5 (-39.1, -23.6); <0.001
Effect modification: month 4-6/7 ARD - month 1-3 ARD (Percentage point [95% CI]; p-value)	-6.5 (-16.2, -1); 0.076

ARD = adjusted risk difference. Treatment arm summary values are % (number of events/total possible events). Denominators represent the number of patient treatment-months per arm across all patients. For the adjusted risk difference results the binary primary outcome was derived and analysed at the patient treatment-month level (planned treatment length being six/seven months). Effect modification adjusted risk difference results calculated on the additive scale as the difference between the adjusted risk difference of each subgroup as indicated. All adjusted risk difference results were calculated via a marginal standardisation approach using bootstrapping and permutation methods (clustered within patient) to obtain the confidence intervals and p-values respectively. All adjusted results were adjusted for the covariates county (the stratum/centre), age (years), sex (male/female), job (farmer/other), marriage status (married/single or divorced or widowed), treatment month (treated as a “continuous” variable), and the interaction between treatment arm and treatment month and treatment period (month 1-3/6 or 7). All analyses included the intention-to-treat population. Complete case analyses included outcome values for patients who died during treatment that were derived only from their dose adherence data available prior to death. Intervention-best control-worst analyses included outcome values for patients who died during treatment that were derived from their dose adherence data following imputation of the missing, post-death, dose adherence data, until the end of their planned treatment, assuming full dose adherence after death for death-patients in the intervention arm and full dose non-adherence after death for death-patients in the control arm. Intervention-worst control-best analyses followed the same imputation approach but assuming non-full adherence for death-patients in the intervention arm and full adherence for death-patients in the control arm following death.

Table S3. Primary outcome effect modification by sex (male and female) under complete case, intervention-best control-worst, and intervention-worst control-best imputation scenarios for death-related missing dose adherence data

Summary value/intervention effect estimated	Estimate
Complete case analysis	
Male intervention outcome summary (% [n/N])	10.5% (56/534)
Male control outcome summary (% [n/N])	39.3% (188/478)
Male intervention effect: ARD (95% CI); p-value	-26.4 (-36, -15.3); <0.001
Female intervention outcome summary (% [n/N])	9.7% (31/320)
Female control outcome summary (% [n/N])	32.2% (102/317)
Female intervention effect: ARD (95% CI); p-value	-30.8 (-38.9, -21.8); <0.001
Effect modification: male – female ARD (Percentage point [95% CI]; p-value)	-4.3 (-18.4, 9.1); 0.570
Intervention-best control-worst scenario	
Male intervention outcome summary (% [n/N])	10.3% (55/534)
Male control outcome summary (% [n/N])	41.1% (202/492)
Male intervention effect: ARD (95% CI); p-value	-27.2 (-36.8, -16.1); <0.001
Female intervention outcome summary (% [n/N])	9.3% (30/322)
Female control outcome summary (% [n/N])	32.6% (104/319)
Female intervention effect: ARD (95% CI); p-value	-32.2 (-40.5, -23.3); <0.001
Effect modification: male – female ARD (Percentage point [95% CI]; p-value)	-5.1 (-19.3, 8.4); 0.507
Intervention-worst control-best scenario	
Male intervention outcome summary (% [n/N])	10.7% (57/534)
Male control outcome summary (% [n/N])	38.2% (188/492)
Male intervention effect: ARD (95% CI); p-value	-25 (-34.4, -14); <0.001
Female intervention outcome summary (% [n/N])	10.2% (33/322)
Female control outcome summary (% [n/N])	31.7% (101/319)
Female intervention effect: ARD (95% CI); p-value	-28.3 (-36.3, -19.5); <0.001
Effect modification: male – female ARD (Percentage point [95% CI]; p-value)	-3.3 (-17.4, 9.9); 0.663

ARD = adjusted risk difference. Treatment arm summary values are % (number of events/total possible events). Denominators represent the number of patient treatment-months per arm across all patients. For the adjusted risk difference results the binary primary outcome was derived and analysed at the patient treatment-month level (planned treatment length being six/seven months). Effect modification adjusted risk difference results calculated on the additive scale as the difference between the adjusted risk difference of each subgroup as indicated. All adjusted risk difference results were calculated via a marginal standardisation approach using bootstrapping and permutation methods (clustered within patient) to obtain the confidence intervals and p-values respectively. All adjusted results were adjusted for the covariates county (the stratum/centre), age (years), job (farmer/other), marriage status (married/single or divorced or widowed), and treatment month (treated as a “continuous” variable), and the interaction between treatment arm and treatment month and sex (male/female). All analyses included the intention-to-treat population. Complete case analyses included outcome values for patients who died during treatment that were derived only from their dose adherence data available prior to death. Intervention-best control-worst analyses included outcome values for patients who died during treatment that were derived from their dose adherence data following imputation of the missing, post-death, dose adherence data, until the end of their planned treatment, assuming full dose adherence after death for death-patients in the intervention arm and full dose non-adherence after death for death-patients in the control arm. Intervention-worst control-best analyses followed the same imputation approach but assuming non-full adherence for death-patients in the intervention arm and full adherence for death-patients in the control arm following death.

Table S4. Primary outcome effect modification by age (<56 and ≥56) under complete case, intervention-best control-worst, and intervention-worst control-best imputation scenarios for death-related missing dose adherence data

Summary value/intervention effect estimated	Estimate
Complete case analysis	
<56 intervention outcome summary (% [n/N])	8.7% (35/404)
<56 control outcome summary (% [n/N])	35.5% (147/414)
<56 intervention effect: ARD (95% CI); p-value	-29.1 (-36.7, -20.4); <0.001
≥56 intervention outcome summary (% [n/N])	11.6% (52/450)
≥56 control outcome summary (% [n/N])	37.5% (143/381)
≥56 intervention effect: ARD (95% CI); p-value	-28.7 (-38.5, -18.1); <0.001
Effect modification: <56 - ≥56 ARD (Percentage point [95% CI]; p-value)	0.3 (-12.6, 13.3); 0.966
Intervention-best control-worst scenario	
<56 intervention outcome summary (% [n/N])	8.7% (35/404)
<56 control outcome summary (% [n/N])	37.2% (158/425)
<56 intervention effect: ARD (95% CI); p-value	-30.2 (-37.8, -21.5); <0.001
≥56 intervention outcome summary (% [n/N])	11.1% (50/452)
≥56 control outcome summary (% [n/N])	38.3% (148/386)
≥56 intervention effect: ARD (95% CI); p-value	-30.1 (-39.7, -19.5); <0.001
Effect modification: <56 - ≥56 ARD (Percentage point [95% CI]; p-value)	0.1 (-12.8, 13.3); 0.989
Intervention-worst control-best scenario	
<56 intervention outcome summary (% [n/N])	8.9% (36/404)
<56 control outcome summary (% [n/N])	34.6% (147/425)
<56 intervention effect: ARD (95% CI); p-value	-26.8 (-34.5, -18.3); <0.001
≥56 intervention outcome summary (% [n/N])	11.9% (54/452)
≥56 control outcome summary (% [n/N])	36.8% (142/386)
≥56 intervention effect: ARD (95% CI); p-value	-27 (-36.6, -16.3); <0.001
Effect modification: <56 - ≥56 ARD (Percentage point [95% CI]; p-value)	-0.3 (-13.2, 13); 0.976

ARD = adjusted risk difference. Treatment arm summary values are % (number of events/total possible events). Denominators represent the number of patient treatment-months per arm across all patients. For the adjusted risk difference results the binary primary outcome was derived and analysed at the patient treatment-month level (planned treatment length being six/seven months). Effect modification adjusted risk difference results calculated on the additive scale as the difference between the adjusted risk difference of each subgroup as indicated. All adjusted risk difference results were calculated via a marginal standardisation approach using bootstrapping and permutation methods (clustered within patient) to obtain the confidence intervals and p-values respectively. All adjusted results were adjusted for the covariates county (the stratum/centre), sex (male/female), job (farmer/other), marriage status (married/single or divorced or widowed), and treatment month (treated as a “continuous” variable), and the interaction between treatment arm and treatment month and age group (<56/≥56). All analyses included the intention-to-treat population. Complete case analyses included outcome values for patients who died during treatment that were derived only from their dose adherence data available prior to death. Intervention-best control-worst analyses included outcome values for patients who died during treatment that were derived from their dose adherence data following imputation of the missing, post-death, dose adherence data, until the end of their planned treatment, assuming full dose adherence after death for death-patients in the intervention arm and full dose non-adherence after death for death-patients in the control arm. Intervention-worst control-best analyses followed the same imputation approach but assuming non-full adherence for death-patients in the intervention arm and full adherence for death-patients in the control arm following death.

Table S5. Treatment success effect modification by sex (male and female)

Summary value/intervention effect estimated	Estimate
Male intervention outcome summary (% [n/N])	94.4% (84/89)
Male control outcome summary (% [n/N])	71.6% (58/81)
Male intervention effect: ARD (95% CI); p-value	19.7 (5.2, 33.8); 0.007
Female intervention outcome summary (% [n/N])	92.5% (49/53)
Female control outcome summary (% [n/N])	75.5% (40/53)
Female intervention effect: ARD (95% CI); p-value	21.8 (11.5, 32.7); <0.001
Effect modification: male – female ARD (Percentage point [95% CI]; p-value)	2.2 (-15, 19.9); 0.818

ARD = adjusted risk difference. Treatment arm summary values are % (number of events/total possible events). Denominators represent the number of patients in that subgroup per arm. For the adjusted risk difference results the binary outcome was derived and analysed at the patient level. Effect modification adjusted risk difference results calculated on the additive scale as the difference between the adjusted risk difference of each subgroup as indicated. All adjusted risk difference results calculated via a marginal standardisation approach using bootstrapping and permutation methods to obtain the confidence intervals and p-values respectively. All adjusted results were adjusted for the covariates county (the stratum/centre), age (years), job (farmer/other), marriage status (married/single or divorced or widowed), and the interaction between treatment arm and sex (male/female). All analyses included the intention-to-treat population.

Table S6. Treatment success effect modification by age (<56 and ≥56)

Summary value/intervention effect estimated	Estimate
<56 intervention outcome summary (% [n/N])	97% (65/67)
<56 control outcome summary (% [n/N])	74.3% (52/70)
<56 intervention effect: ARD (95% CI); p-value	23.2 (12.3, 34.1); <0.001
≥56 intervention outcome summary (% [n/N])	90.7% (68/75)
≥56 control outcome summary (% [n/N])	71.9% (46/64)
≥56 intervention effect: ARD (95% CI); p-value	19 (5.8, 32); 0.003
Effect modification: <56 - ≥56 ARD (Percentage point [95% CI]; p-value)	-4.3 (-21.5, 12.6); 0.634

ARD = adjusted risk difference. Treatment arm summary values are % (number of events/total possible events). Denominators represent the number of patients in that subgroup per arm. For the adjusted risk difference results the binary outcome was derived and analysed at the patient level. Effect modification adjusted risk difference results calculated on the additive scale as the difference between the adjusted risk difference of each subgroup as indicated. All adjusted risk difference results calculated via a marginal standardisation approach using bootstrapping and permutation methods to obtain the confidence intervals and p-values respectively. All adjusted results were adjusted for the covariates county (the stratum/centre), age (years), sex (male/female), job (farmer/other), marriage status (married/single or divorced or widowed), and the interaction between treatment arm and age group (<56/≥56). All analyses included the intention-to-treat population.

Table S7. Intervention-arm patients requiring VOT and frequency of successful implementation of VOT for those patients

	Month							Total
	1	2	3	4	5	6	7	
Patients requiring VOT, % (n/total)	1.4% (2/142)	1.4% (2/142)	8.5% (12/142)	12.7% (18/142)	16.3% (23/141)	31.2% (44/141)	25% (1/4)	11.9% (102/854)
Successfully implemented VOT, % (n/total)	100% (2/2)	100% (2/2)	91.7% (11/12)	94.4% (17/18)	91.3% (21/23)	88.6% (39/44)	100% (1/1)	91.2% (93/102)

VOT = video-observed therapy.

Table S8. Comparison of basic patient characteristics (sex and age) between recruited patients and non-participating patients also undergoing treatment in the National Tuberculosis Programme during the trial period

	Recruited patients*	All new NTP patients*	P-value†
Number of patients	278	136	
Sex (% [n])			0.734
Male	61.5% (171)	63.2% (86)	
Female	38.5% (107)	36.8% (50)	
Age (Median [IQR])	56 (40, 64)	53 (24.5, 63)	0.132

IQR = interquartile range. * Recruited patients are all patients who were recruited into the trial. All new NTP patients are all patients who were registered with the Chinese National Tuberculosis Program for treatment for tuberculosis during the trial recruitment period but who were not recruited into the trial. † P-value for sex obtained from a chi-square test of independence comparing the distribution of the number of males and females between recruited patients and all new NTP patients, and the p-value for age obtained from a Mann-Whitney U test comparing the distribution of ages (in years) between recruited patients and all new NTP patients.

References

1. Snow G. blockrand: Randomization for Block Random Clinical Trials. 1.5 ed; 2020.
2. Højsgaard S, Halekoh U, Yan J. The R Package geepack for Estimating Equations. *Journal of Statistical Software* 2006; **15**(2): 1-11.
3. Frick H, Chow F, Kuhn M, Mahoney M, Silge J, Wickham H. General Resampling Infrastructure. 1.1.1 ed; 2022.
4. Kropko J, Harden JJ. coxed: Duration-Based Quantities of Interest for the Cox Proportional Hazards Model. 0.3.3 ed; 2020.
5. Simpson G. permute: Functions for Generating Restricted Permutations of Data. 0.9-7 ed; 2022.
6. World Health Organization. Definitions and reporting framework for tuberculosis (2013 revision, updated December 2014 and January 2020), 2020.