J Clin Tuberc Other Mycobact Dis 11 (2018) 10–16

Contents lists available at ScienceDirect

J Clin Tuberc Other Mycobact Dis

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

Use of classification and regression tree (CART), to identify hemoglobin A1C (HbA_{1C}) cut-off thresholds predictive of poor tuberculosis treatment outcomes and associated risk factors

Josephine W. Mburu[a,b,*], Leonard Kingwara[a], Magiri Ester[b], Nyerere Andrew[b]

[a] National Reference Tuberculosis Laboratory, MOH, Kenya
[b] Jomo Kenyatta University of Agriculture and Technology (JKUAT), Kenya

ARTICLE INFO

Keywords:
Fractal geometry
Hollow fiber
Tuberculosis outcomes
Hyperglycemia

ABSTRACT

Background: Rifampin-based therapy potentially exacerbates glycemic control among TB patients who are already at high risk of hyperglycemia. This impacts negatively to the optimal care of TB- diabetes mellitus co-affected patients. Classification and regression tree (CART), a machine-learning algorithm impervious to statistical assumptions is one of the ideal tools for clinical decision-making that can be used to identify hemoglobin A1C (HbA_{1C}) cut-off thresholds predictive of poor TB treatment outcomes in such populations.

Methods: 340TB smear positive patients attending two peri-urban clinics were recruited and prospectively followed up for six months. Baseline HbA_{1C} and random blood glucose (RBG) levels were determined. CART was then used to identify cut-off thresholds and rank outcome predictors at end of therapy by determining Risk ratios (RR) and 95% confidence interval (CI) of each predictor threshold. Fractal geometry law explained effect of weight, while U-shaped curve explained effect of HbA_{1C} on these clinical outcomes.

Results: Of the 340 patients enrolled: 84%were cured, 7% completed therapy and 9% had unfavorable outcomes out of which 4% (n = 32) had microbiologic failure. Using CART HbA_{1C} identified thresholds were >2.95%, 2.95–4.55% and >4.55%, containing 8/11 (73%), 111/114 (97%) and 189/215 (88%) of patients who experienced favorable outcomes. RR for favorable outcome in patients with weight <53.25 Kg compared to >53.25 Kg was 0.61 (95% CI, 0.45–0.88) among patients with HbA1C >4.55%. Simulation of the CART model with 13 patients data failed therapy revealed that 8/11 (73%) of patients with HbA1C <2.95%, 111/114 (97%) with HbA_{1C} between 2.95% and 4.55% and 189/215 (88%) of patients with HbA1c >4.55% experienced microbiologic failure.

Conclusion: Using fractal geometry relationships to drug pharmacokinetics, low weight has profound influence on failure of anti-tuberculosis treatment among patients at risk for diabetes mellitus.

Introduction

Tuberculosis (TB), like other infections, can worsen glycemic control and complicate the clinical management of Diabetes mellitus (DM) [1–3]. The dual burden of TB and DM (gestational, type I and type II) have increased over the past decade with DM prevalence increasing in countries already afflicted with a high burden of TB [3]. The coexistence of two conditions presents a serious threat to global public health and patient clinical care resulting into worse clinical outcomes across the entire spectrum of either disease [2,4]. In one study, 7.5% (95% CI: 4.1%−11.5%) of TB incidence cases were attributable to hyperglycemia [5]. A south African cohort further demonstrated a correlation between active TB and the level of glycosylated hemoglobin (HbA1c) with a hazard ratio of 1.39 (95% CI: 1.18–.63) per unit increase [6,7]. In addition to the well-established contribution of DM to increased TB susceptibility, findings from most observational studies indicate that this co-morbidity is associated with delays in TB clearance during treatment, treatment failures, death, relapse, disease severity and re-infection [2,5]. Diabetic TB patients' management can play a critical role in understanding TB transmission dynamics [8]. The link between DM and TB and the implementation of the collaborative framework for care and control have thus widely been recommended as a means to potentially stimulate and strengthen the scale-up of non-communicable disease care and prevention programs, which may help in reducing not only the global burden of DM but also the global burden of TB [8–10].

* Corresponding author.
E-mail address: joewahogo@gmail.com (J.W. Mburu).

https://doi.org/10.1016/j.jctube.2018.01.002
Received 28 September 2017; Received in revised form 18 January 2018; Accepted 23 January 2018

In Kenya, the TB program recommends bidirectional screening, though there is dearth of data on performance of specific TB tests in individuals with DM, specific DM tests among TB patients, and screening and preventive therapy for latent TB infections in individuals with DM. Some of these can be attributed to the poor data management system in our health facilities. Studies indicate that TB patients with DM have a lower concentration of TB drugs and a higher risk of drug toxicity than those without diabetes [11]. In addition good glycaemic control, which reduces long-term DM complications and could also improve TB treatment outcomes, is hampered by chronic inflammation, drug-drug interactions, suboptimum adherence to drug treatments, and other factors [9,11,12]. Besides drug treatments for TB and DM, other interventions, such as education, intensive monitoring, and lifestyle interventions, might also be needed, especially for patients with newly diagnosed DM or those who need insulin. From a health systems point of view, delivery of optimum care and integration of services for tuberculosis and diabetes is a huge challenge in many countries more specifically resource limited setting.

DM screening among TB patients is now widely recommended, especially in setups with high DM prevalence [12]. A case example is India and the Pacific Islands region's TB control programs [13,14]. Despite these efforts, the best time and methods to diagnose for DM diagnosis among TB patients remains unclear [15–17]. In one of the observational studies, it was found that glycemic status was influenced by radiological manifestations of diabetic pulmonary TB [15,16]. DM pevalence rises steeply with age, but the most efficient age cut-off for screening is also unclear and varies among populations [18]. Blood glucose concentrations measurement at a single point in time might lead to a false diagnosis of DM in patients with TB because they could have intermittent hyperglycaemia through induction of insulin resistance, mediated by inflammation. TB [19,20]. Repeat testing could identify transient Hyperglycaemia concentration assessment is thus the only diabetes test that shows average glycaemia over time and in a single study [21,22], as it is more sensitive than fasting blood glucose when used as a screening test for newly diagnosed diabetes in patients with tuberculosis [22].

Though HbA_{1c} concentration assessment has been indicated as the most preferable test method for DM among TB infected patients, hyperglycemia levels predictive of poor clinical outcomes among patients with both DM and TB, for which clinicians must target for optimal care, is not well defined [18–20]. It has been shown that experimental hollow-fiber models of tuberculosis chemotherapy, and in TB patients, that pharmacokinetic variability directly leads to failure of TB therapy [23,24]. Moreover, rifampin is known to exacerbate hyperglycemia in patients starting TB therapy, suggesting that therapy in patients at risk of TB can be optimized and personalized [23]. To better inform TB programs, we sought to examine and rank factors predictive of clinical outcomes in routine low-resourced clinical settings using CART analysis. We also determined thresholds for those predictors that could be used for clinical decision-making. The focus of this study was TB patients at risk for DM. Previously; the hypothesis has been used to explain poor long-term outcomes in supposedly 'adequately' treated tuberculosis patients: wherein systemic tuberculosis infection produces early organ damage that leads to death in the long-term [25,26]. In this study we proposes the use of boosted classification and regression trees (CART) to determine baseline HbA_{1c} and random blood glucose (RBG) levels predictive of clinical outcomes in tuberculosis patients with/out diabetes mellitus. CART is an agnostic machine-learning algorithm more adept for data-driven pattern recognition, selecting predictors and their thresholds based on reproducible models.

Methods

Definition of terms

Clinical outcomes were defined using WHO and Kenyan tuberculosis program guidelines standard operational terms (http://nltp.co.ke/guidelines/) 27,28. The primary analyses compared favorable versus unfavorable outcomes at end of treatment. The secondary analyses only compared "cures" versus "failures" at similar time points as is the standard practice when examining chemotherapy efficacy.

Study design

We carried out a prospective cohort in two peri-/urban counties of Kenya (Kiambu and Nairobi) between February 2014 and August 2015. We enrolled patients aged >15 years who were TB smear positive and were not pregnant (Pregnant women are high risk group of Gestatinal DM) at time of diagnosis. TB culture test was used to confirm all the positives. Ethical approval was obtained from the Kenyatta National Hospital Ethical Research Committee. We collected one venous blood draw at baseline in two separate tubes (one for fasting or random blood glucose levels and the other for HbA1c levels). Each patient then had physical examination and questionnaire administered by trained healthcare personnel where detailed history, including signs and symptoms of diabetes mellitus, cigarette smoking and other life-style information were ascertained. Patients were followed prospectively at two, three, five, six months and at end of therapy. During each visits we assessed adherence treatment and treatment outcome with sputum microscopy examination. The initial sputum examinations were submitted for culture and pathogen identification. Patients were examined at each visit for both TB and DM.

Care and treatment

New tuberculosis patients were put on a six-months category I regimen comprising of 2 months of isoniazid, rifampin, pyrazinamide and ethambutol followed by four months of isoniazid and rifampin.. Previously treated patients, including those who had failed prior therapy were put on category II regimen which is similar to category I only that, streptomycin is additionally administered for the first two months, while the other four drugs is prolonged by one month and isoniazid, rifampin and ethambutol are given for a further five more months. Dosing was as per Daily fixed dose combinations formulations as per NTLD and WHO guidelines were given using directly observed Treatment, short course (DOTs) [28].

Glycosylated hemoglobin (HbA1c) measurements and laboratory quality assurance

HbA1c measurement was performed on Ethylenediamine-tetraacetic acid (EDTA) whole blood samples within 4–6 h of sample collection using cation exchange high performance liquid chromatography (HPLC). The assay values were standardized to the Diabetes Control and Complications Trial - Deflection Check: 0.5, Discreteness Check: 0.1 Abs, Sensitivity Check: 0.14, Blank Level Checks −0.1- −0.7. Laboratory procedures were performed as per manufacturer instructions. Commercial controls consisting of HbA1c Abnormal Low, Normal and Abnormal High and HbA1c calibrators were included during the assay procedure. HbA_{1C} and random blood glucose results were then issued to patients to receive appropriate further management after consulting with their physician.

Data analyses

Richardson-Mandelbrot log-log plots were used to examine fractal behavior [29,30]. We modified the equation to $f(t) \sim \left(\frac{t}{tp}\right)^{-q}$ to circumvent obvious concavity and inflection point estimate from the cumulative distribution plot. Observed variable value is denoted by t while tp denotes observed value at the inflection point. Since we were interested in physiological function in addition to morphological

dimensions, values between zero and one, one and two as well as two and three, were also examined. An estimate of the slope from the regression equation of the log-log plot was taken as the fractal dimension.

We used one sample *t*-test and the nonparametric Kruskall-Wallis test to test to assess the representativeness of the sample in representing a population of patients with diabetes mellitus, or pre-diabetes mellitus using cut-off points of 6.5% and 5.7%, respectively [31,32]. Stepwise multivariate logic regression analyses (validation of the CART Model findings and sensitivity analysis of the full model using the same variables used in CART for both development and validation samples with true classification using stepwise) was then performed and additional covariates added or removed based both on parsimony, Akaike information criteria (AIC), odds ratio, and 95% confidence interval (CI). Receiver operating characteristics (ROC) value was then reported in table format. Relative risk ration and the CI were also computed given the prospective nature of the study. Confidence intervals for proportions were computed using exact methods approximation.

Classification and regression tree analysis (CART)

Patients' age, gender, weight, body mass index, blood glucose, creatinine, blood urea nitrogen and HbA1c levels at baseline, as well as HIV result and ever cigarette smoker status were put in a boosted CART model for binary outcome:. None of the patients enrolled reported a previous diagnosis of DM, so that variable was not included in model. Boosted CART modeling was performed using Tree Net in Salford software version 8, based on methodological approaches [33,34]. ROC curve then was used to compare models after 10-fold cross-validation.

Results

Of the 347 patients who agreed to participate in the study, 7 (2%) were excluded from further analyses because they grew Mycobacterium bovis on culture 285 (84%) attained microbiologic cure, 13 (4%) had microbiologic failure, while 23 (7%) completed treatment without attaining cure status. The remainder, 19 (5%) patients did not have clinical outcomes ascertained for the following reasons: 10 (53%) were lost to follow-up, while the remainder 9 (47%) transferred out of the two counties. There was no significance difference in demographic and clinical characteristics between patients whose outcomes were ascertained compared with those with an unknown outcome **(**Table 1**)** with exception to weight, smoking and regimen I TB therapy. Patients whose outcomes were unknown were therefore combined with failures to make a composite outcome category of unfavorable outcomes in all primary analyses.

We then examined whether the relationship between clinical outcomes and weight could be explained by complex nonlinear phenomenon and fractal geometry, since both have been used to describe metabolic rate in physiology and pharmacology [35,36]. First, we show in Fig. 1**A** that the distribution of HbA_{1C} was not normally distributed. Both its mean and median values of 5.76% and 5.28%, respectively, were significantly less than the threshold of 6.5% used to define DM, based on parametric and non-parametric sample tests, $p < 0.001$. However, the mean difference of 0.74% (95% CI: 0.50–0.98) was within the 5.7% cut-off used to define pre-DM**.** This implies that the study sample was derived from a population of patients at risk of DM, or pre-DM. Fig. 1 also shows that both the BMI and weight were very low for these patients given their mean age and height as indicated in Fig. 1**B** and **C**. The low weight was due to under nutrition from tuberculosis disease. The BMI ranged between 10.36 kg/m^2 and 49 kg/m^2 and was significantly correlated with weight, $r = 0.71$. Next, we show in Fig. 1**A** that the relationship between weight and predicted clinical outcomes was best described by ¾ fractal geometry law with $R^2 = 0.905$, indicative of a fractal with 0.75 dimension. Prediction (P_p) was given by the following equation: $P_p = P_{otyp} * \left(\frac{Wt}{Wtotyp}\right)^{3/4}$. Where Wt is observed weight, Wt_{otyp} is the typical weight and P_{otyp} is probability of the typical weight. The typical weight for this sample was then determined as 53.25 kg. This suggests that complex nonlinear pharmacokinetics relationships and variability mediated by changes in patients' weight was driving outcomes. The relationship shown in Fig. 1**B** of HbA_{1C} suggests an asymptotic fractal with a dimension.

To further examine nonlinear relationship between weight and other covariates with clinical outcomes we ran up to 200 boosted classification and regression trees (b CART) in Tree Net software for each binary outcome. The predictors are ranked and scored in Table 2 for favorable versus unfavorable as well as favorable versus microbiologic failure outcomes. Finally, we subjected the output for the fractal geometry analysis and CART to compute risk factors

Discussion

There is no significance difference in demographic and clinical characteristics between patients whose outcomes were ascertained compared with those with an unknown outcome with exception to weight, smoking and regimen I TB therapy. There was no correlation between blood glucose and HbA_{1C} levels at baseline: Spearman's rho = 0.035, p = 0.515 while patients who weigh less had unfavorable clinical outcomes compared to their heavier counterpart at p = 0.011.Cumulative probability for favorable outcomes or microbiologic cure decreased with increase in HbA_{1C}: and, the nonlinear trends are even more apparent before the 6.5% HbA_{1C} thresholds are reached as indicated. In our population, distribution of HbA_{1C} was not normally distributed. Both its mean and median values of 5.76% and 5.28%, respectively, were significantly less than the threshold of 6.5% used to define DM, based on parametric and non-parametric sample tests, $p < 0.001$. However, the mean difference of 0.74% (95% CI: 0.50–0.98) was within the 5.7% cut-off used to define pre-DM**.** This implies that the study sample was derived from a population of patients at risk of DM, or pre-DM. The low weight witnessed in our population was due to under nutrition from tuberculosis disease. The BMI of the study participants ranged between 10.36 kg/m^2 and 49 kg/m^2 and was significantly correlated with weight, $r = 0.71$. Relationship between weight and predicted clinical outcomes was best described by ¾ fractal geometry law with $R^2 = 0.905$, indicative of a fractal with 0.75 dimension. The typical weight for this sample was then determined as 53.25 kg. This suggests that complex nonlinear pharmacokinetics relationships and variability mediated by changes in patients' weight was driving outcomes [24]. 121/136 (89%) who had their weight < 53.25 Kg had favorable outcome compared to 162/172 (94%) who had their weight >53.25 Kg The relative risk of weighing >53.25 Kg was 0.71 (95% CI, 0.53–1.07)**.** Similarly, among the 215 patients with HbA1C > 4.55%, 75/92 (82%) had favorable outcomes with weight <53.25 Kg compared to 114/123 (93%) with weight >53.25 Kg. The relative risk of weighing above 53.25 Kg was 0.61 (95% CI, 0.45–0.88).

The ROC score for each train sample was >0.95 and misclassification rates <0.02 indicating that the model described the data well. On the other hand, ROC scores on test samples after cross validation was >0.65 for each of the model examined which is reassuring in prediction on future similar sample. Simulation on how to interpret the classification tree indicated that ten (77%) out of 13 patients who experienced microbiologic failure had $HbA_{1C} > 4.55\%$ while 8 (62%) of these had weight <53.25 kg. The three HbA_{1C} cut-off levels depicting U-shaped pattern that interacted with both weight and BMI were identified by CART. When the model was applied to the entire cohort it revealed that 8/11 (73%) of patients with $HbA_{1C} < 2.95\%$, 111/114 (97%) with HbA_{1C} between 2.95% and 4.55% and 189/215 (88%) of patients with $HbA_{1c} > 4.55\%$ experienced microbiologic failure. HbA_{1c} thresholds were also statistically significant for outcomes observed at end of treatment based and the other known risk factors for either tuberculosis or diabetes mellitus were not associated with these

Table 1
Difference in demographic and clinical characteristics in patients with/out clinical outcomes ascertained.

Variable	Level	All patients n = 340 (%)	Outcomes Unknown n = 19 (n)	Outcomes Ascertained n = 321 (n)	p-value
Demographic	Characteristics				
County	Kiambu	56 (17)	4 (21)	52 (16)	0.584
	Nairobi	283 (83)	268 (79)	268 (83)	
	Missing data	1 ()		1 (1)	
Gender	Female	96 (28)	5 (26)	91 (28)	0.848
	Male	244 (72)	14 (74)	230 (72)	
Cigarettes	Ever-smoker	97 (29)	1 (5)	96 (30)	0.021
Age (years)	Mean (SD)	32.11 (8.80)	29 (5.16)	32.29 (8.94)	0.172
Clinical	Characteristics				
HIV test	Negative	238 (70)	11 (58)	227 (71)	0.312
	Positive	77 (23)	7 (37)	70 (22)	
	Not Done	25 (7)	1 (5)	24 (7)	
BMI (kg/m^2)	Mean (SD)	19.55 (3.91)	17.96 (2.35)	19.65 (3.97)	0.055
Underweight	<18.5	133 (39)	11 (58)	122 (38)	0.156
Ideal	18.5–25	184 (54)	8 (42)	176 (55)	
Overweight	>25	23 (7)	0	23 (7)	
Weight (KG)	Mean (SD)	56 (61)	49.93 (5.61)	54.88 (10.46)	0.023
Blood glucose (mmol/dL)	Mean (SD)	3.61 (1.19)	3.66 (0.69)	3.60 (1.22)	0.548
Glycosylated hemoglobin (%)	Mean (SD)	5.75 (2.22)	6.27 (2.09)	5.72 (2.23)	0.145
Creatinine (mmol/dL)	Mean (SD)	89.39 (20.33)	83.79 (22.78)	89.72 (20.17)	0.087
BUN (mmol/dL)	Mean (SD)	3.76 (1.16)	3.51 (1.09)	3.78 (1.17)	0.378
Treatment	Regimen				
Regimen I	2HRZE/4HR	308 (91)	14 (74)	294 (92)	0.009
Regimen II (retreatment)	2HRZES/1HRZE/5HRE	32 (9)	5 (26)	27 (8)	

There is no significance difference in demographic and clinical characteristics between patients whose outcomes were ascertained compared with those with an unknown outcome with exception to weight, smoking and regimen I TB therapy.

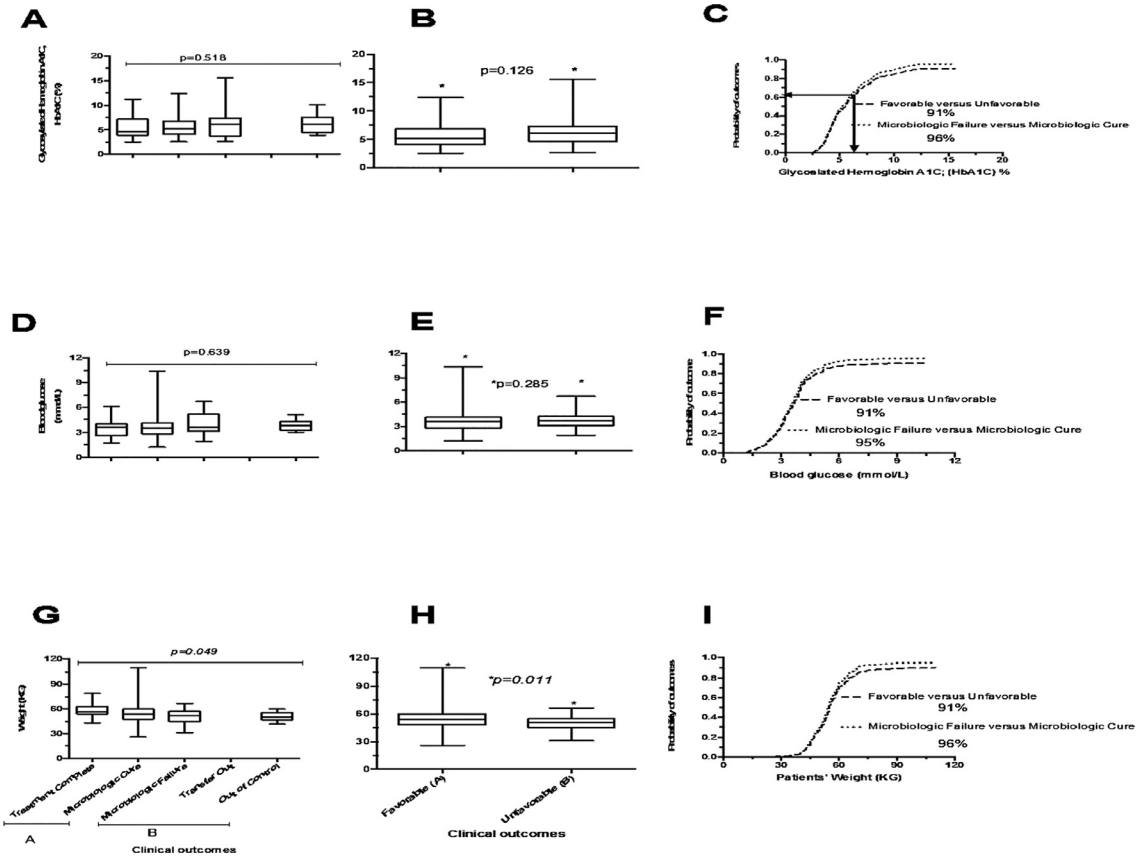

Fig. 1. Distribution of HbA_{1C}, blood glucose and weight across clinical outcomes.
There was no correlation between blood glucose and HbA_{1C} levels at baseline: Spearman's rho = 0.035, p = 0.515. None of the enrolled patients reported a history or received prior treatment for DM. Weight had a borderline significance for transfers (Fig. 1 G). In general patients who weigh less had unfavorable clinical outcomes compared to their heavier counterpart at p = 0.011(Fig. 1**H**). Cumulative probability for favorable outcomes or microbiologic cure decreased with increase in HbA_{1C}: and, the nonlinear trends are even more apparent before the 6.5% HbA_{1C} thresholds are reached as indicated in Fig. 1C (as indicated by the arrows in).

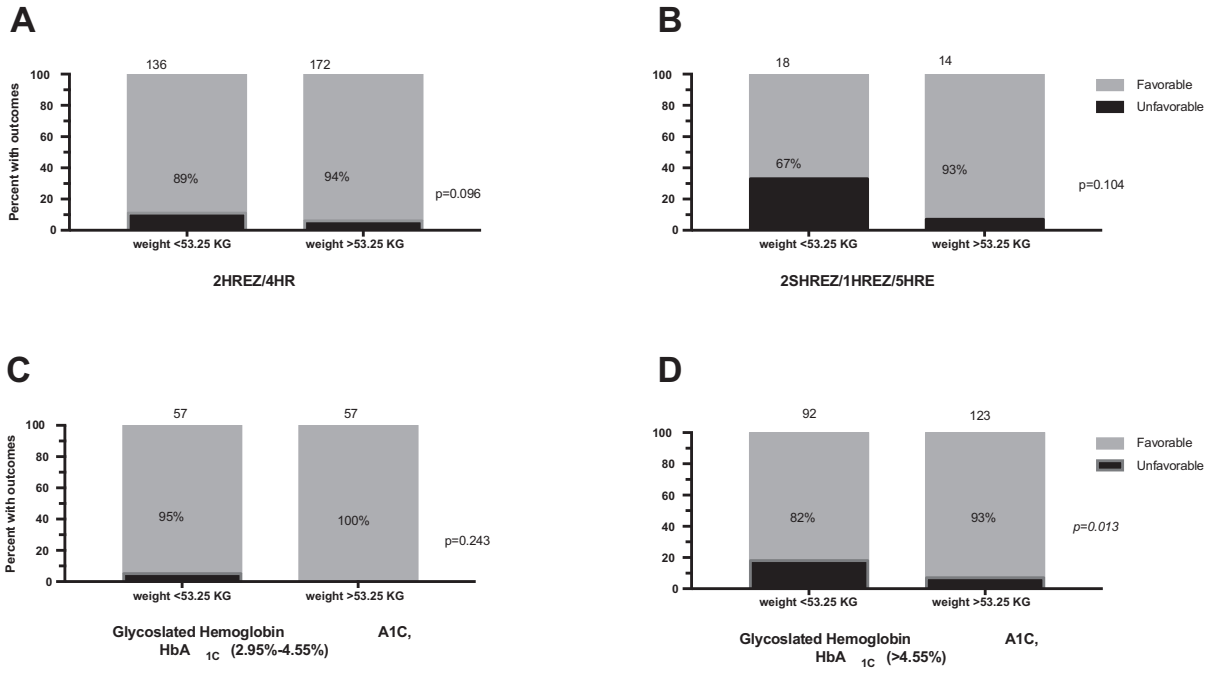

Fig. 2. Frequency distribution of glycated hemoglobin, blood glucose and weight.
Fig. 2 indicates significant differential effect of weight on outcomes when stratified by treatment regimen and HbA_{1C}. In 2A 121/136 (89%) who had their weight <53.25 Kg had favorable outcome compared to 162/172 (94%) who had their weight >53.25 Kg The relative risk of weighing >53.25 Kg was 0.71 (95% CI, 0.53–1.07). Similarly, among the 215 patients with HbA1C >4.55%, 75/92 (82%) had favorable outcomes with weight <53.25 Kg compared to 114/123 (93%) with weight >53.25 Kg. The relative risk of weighing above 53.25 Kg was 0.61 (95% CI, 0.45–0.88).

thresholds.

Our findings demonstrated the complex and nonlinear effects of weight on clinical outcomes. The effect of weight is probably mediated via pharmacokinetic variability of the different anti-tuberculosis drugs and impact of inadequate dosing [37]. In this study we show that when change in weight is adjusted for, the odds for favorable outcome increasingly gets worse with each unit increase in HbA1C above 4.95%. The odds ratio was 0.18 (0.05–0.61) and remained robust even when the outcomes were restricted to treatment failure. Indeed this is similar to currently used criteria to distinguish individuals at higher risk for premature mortality from micro-vascular (retinopathy) or macro vascular (heart attacks) complications based on mostly on observational studies which have established association between those diabetes-related complications and glycosylated hemoglobin (HbA1c) with/out general poor glycemic control [38–41].

Higher HBA_{1C} thresholds were predictive of unfavorable outcomes in treated tuberculosis patients. Several studies including meta-analyses have shown that patients with diabetes mellitus and tuberculosis have worse outcomes for either disease, but there are no international guidelines on joint management and control of both diseases [19,20,22,39]. One recent Mexican study revealed that, among diabetes mellitus patients who received standard tuberculosis treatment, 81% of recurrences and 77% of relapses, were infections caused by the same Mycobacterium tuberculosis isolate based on IS6110-band restriction

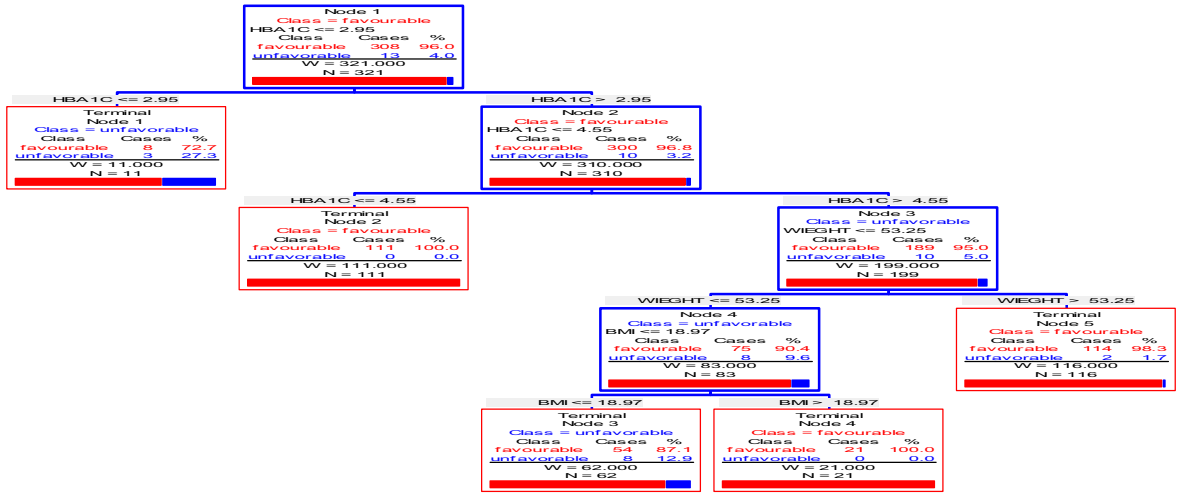

Figure 3. Classification and Regression Tree (CART), U shaped curve and cut-off.
Simulation on how to interpret the classification tree is by tracking the 13 patients who failed therapy. Ten (77%) out of 13 patients who experienced microbiologic failure had HbA_{1C} >4.55% while 8 (62%) of these had weight <53.25 kg. The three HbA_{1C} cut-off levels depicting U-shaped pattern that interacted with both weight and BMI were identified by CART. When the model was applied to the entire cohort it revealed that 8/11 (73%) of patients with HbA1C <2.95%, 111/114 (97%) with HbA_{1C} between 2.95% and 4.55% and 189/215 (88%) of patients with HbA_{1c} >4.55% experienced microbiologic failure.

Table 2
Ranking of most important variables predictive of clinical outcomes based on boosted classification and regression models.

Model I N = 340 Favorable n = 308 (91%) Unfavorable n = 32 (9%) Ranked Variable	Relative Score	Model II N = 321 Favorable n = 308 (96%) Microbiologic failure n = 13 (4%) Ranked Variable	Relative Score
1. HBA_{1C}	100	1. HBA_{1C}	100.00
2. Regimen	79.38	2. Regimen	72.61
3. Creatinine	70.09	3. Age	69.19
4. Body mass index	63.93	4. Weight	55.39
5. Blood urea nitrogen	62.47	5. Random blood glucose	49.87
6. Weight	62.63	6. Body mass index	48.74
7. Age	61.63	7. BUN	48.18
8. Random blood glucose	55.62	8. HIV positive results	46.51
9. HIV positive result	39.21	9. Ever smoker	39.69
10. Male gender	34.48	10. Creatinine	37.69
Model properties	$ROC_{learn} = 0.96 \pm 0.01$	Model properties	$ROC_{learn} = 0.97 \pm 0.01$
	$ROC_{test} = 0.65 \pm 0.06$		$ROC_{test} = 0.56 \pm 0.07$

The ROC score for each train sample was >0.95 and misclassification rates <0.02 indicating that the model described the data well. On the other hand, ROC scores on test samples after cross validation was >0.65 for each of the model examined which is reassuring in prediction on future similar sample. We depict the pruned back optimal tree for the favorable versus microbiologic failure outcome that examined 321 patients in Fig. 3.

Table 3
Glycosylated hemoglobin A_{1C} thresholds levels predictive clinical outcomes in patients treated for pulmonary Mycobacterium tuberculosis.

Variable	Total N = 340 (%)	HbA_{1c} (%) levels			p-value
Clinical Outcomes		2.95–4.55 n = 114 (%)	<2.95 n = 11 (%)	>4.55 n = 215	
At end of treatment					
Microbiologic failure	13 (4)	0	3 (27)	10 (5)	<0.001
Microbiologic cure	285 (84)	102 (89)	6 (55)	177 (82)	
Treatment complete	23 (7)	9 (8)	2 (18)	12 (6)	
Unknown (TO/OOC)	19 (6)	3 (3)	0	16 (7)	
Favorable outcome	308 (91)	111 (97)	8 (73)	189 (88)	0.002
Unfavorable outcome	32 (9)	3 (3)	3 (27)	26 (12)	
Modifiable factors					
Weight (mean [SD]) KG	54.61 (10.32)	53.64 (9.29)	50.82 (9.82)	55.32 (10.80)	0.173
BMI (mean [SD]) KG/m^2	19.55 (3.91)	19.35 (4.24)	18.36 (3.34)	19.72 (3.75)	0.417
Blood glucose (mean [SD]) mmol/L	3.61 (1.19)	3.52 (0.95)	3.22 (0.63)	3.67 (1.32)	0.301
Regimen I (2HRZE)	308 (91)	99 (87)	10 (91)	199 (93)	0.24
Regimen II (2SHRZE/1HRZE/5HRE)	32 (9)	15 (13)	1 (9	16 (7)	
Other factors					
Female	96 (28)	30 (26)	3 (27)	63 (29)	0.847
Male	244 (72)	84 (74)	8 (73)	152 (71)	
HIV test –positive	77 (23)	26 (23)	4 (36)	47 (22)	0.518
-negative	238 (70)	82 (72)	7 (64)	149 (69)	
-not done	25 (7)	6 (5)	0	19 (9)	

Table 3 above confirms that these HbA_{1c} thresholds were also statistically significant for outcomes observed at end of treatment based. The other known risk factors for either tuberculosis or diabetes mellitus were not associated with these thresholds.

Table 4
Multivariate logistic regression analyses of associations between favorable outcome at end of therapy and glycosylated hemoglobin A_{1C} (HbA_{1C}) levels and weight at baseline.

Variable	Level	Odds ratio (95% CI)	p-value
HbA_{IC} (%)	2.95% − 4.55%	Referent	
	<2.95%	0.06 (0.01–0.39)	0.003
	>4.55%	0.18 (0.05–0.61)	0.006
Weight (Kg)	>53.25	Referent	
	<53.25	0.36 (0.16–0.76)	0.009
Constant		24.93 (7.74–80.30)	<0.001

fragment length polymorphism and spoligotyping fingerprint analyses [42,43]. The proportion with treatment failure in our study was 4.0 (95% CI, 2.2–6.8), similar to those seen in the larger Mexican study [44,45]. Indeed, the proportion with tuberculosis recurrence will be significantly larger if the lower threshold is used.

Attempt to show the key role of machine-learning algorithms in answering non-specific but pertinent decision-making questions commonly asked in the clinic indicated how output from CART can be the used to generate risk ratios and adjusted odds ratio from multivariate logistic regression for comparison with existing literature. Some machines prediction methods are not universally consistent [46,47]. A case example is random forest (RF) that is not consistent if splits are performed to purity [48,49]. Therefore, some impurity within the machine generated tree is required for consistency. Comparatively, the performance of CART during simulation returned consistency well above most machine- learning approaches for classification and probability estimation.

Our study has some notable limitations and strengths. First, HbA_{1C}

and blood glucose measurement were only taken at baseline, and thus do not fully reflect glucose control across the 6–8 months of tuberculosis therapy. Relatedly, the contribution of other key events, which might have occurred, during the course of therapy, including decisions to add DM treatment was not assessed. However, since we only included patients with no prior history of DM, we were able to assess impact of impaired glucose tolerance and standard tuberculosis treatment on patients at risk for DM in a routine program setting. Thus, our findings contain practical lessons informative to both clinicians and programs in similar settings.

We concluded that fractal geometry relationships to drug pharmacokinetics, low weight has profound influence on failure of anti-tuberculosis treatment among patients at risk for diabetes mellitus and the three HbA_{1C} cut-off levels,depicting U-shaped pattern that interacted with both weight and BMI, predictive of unfavorable outcomes among TB-DM co-affected patients can be identified by CART

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi: 10.1016/j.jctube.2018.01.002.

References

[1] Bailey SL, Grant P. 'The tubercular diabetic': The impact of diabetes mellitus on tuberculosis and its threat to global tuberculosis control. Clin Med J R Coll Phys Lon 2011;11:344–7.
[2] Dooley KE, Chaisson RE. Tuberculosis and diabetes mellitus: convergence of two epidemics. Lancet Infect Dis. 2009;9:737–46.
[3] Ruslami R, Aarnoutse RE, Alisjahbana B, Van Der Ven AJAM, Van Crevel R. Implications of the global increase of diabetes for tuberculosis control and patient care. Trop Med Int Heal 2010;15:1289–99.
[4] Lane MA, et al. The Roy Walford legacy: diet restriction from molecules to mice to monkeys to man and onto mimetics. Exp Gerontol 2004;39:897–902.
[5] Lee P-H, et al. Glycemic Control and the risk of tuberculosis: a cohort study. PLOS Med 2016;13:e1002072.
[6] Selvin E, et al. Meta-analysis: glycosylated hemoglobin and cardiovascular disease in diabetes mellitus. Ann Intern Med 2004;141:12. 421–I21.
[7] Zhang Y, Hu G, Yuan Z, Chen L. Glycosylated hemoglobin in relationship to cardiovascular outcomes and death in patients with type 2 diabetes: A systematic review and meta-analysis. PLoS One 2012;7.
[8] Harries AD, et al. Epidemiology and interaction of diabetes mellitus and tuberculosis and challenges for care: a review [Review article]. Public Heal Action 2013;3:3–9.
[9] Geneva S. World Health Organisation. Guidelines for surveillance of drug resistance in tuberculosis. WHO Doc 2009;94:1–24. **WHO/TB/**.
[10] Raghuraman S, Vasudevan KP, Govindarajan S, Chinnakali P, Panigrahi KC. Prevalence of diabetes mellitus among tuberculosis patients in urban Puducherry. N Am J Med Sci 2014;6:30–4.
[11] Jeon CY, Murray MB. Diabetes mellitus increases the risk of active tuberculosis: A systematic review of 13 observational studies. PLoS Med 2008;5:1091–101.
[12] Alsultan A, Peloquin CA. Therapeutic drug monitoring in the treatment of tuberculosis: An update. Drugs 2014;74:839–54.
[13] Azman AS, Golub JE, Dowdy DW. How much is tuberculosis screening worth? Estimating the value of active case finding for tuberculosis in South Africa, China, and India. BMC Med 2014;12:216.
[14] Sandhu G. Tuberculosis: Current situation, challenges and overview of its control programs in India. J Glob Infect Dis 2011;3:143.
[15] Balakrishnan S, et al. High diabetes prevalence among tuberculosis cases in Kerala, India. PLoS One 2012;7.
[16] Restrepo BI, Schlesinger LS. Host-pathogen interactions in tuberculosis patients with type 2 diabetes mellitus. Tuberculosis 2013;93.
[17] Bruins M, et al. Diagnosis of active tuberculosis by e-nose analysis of exhaled air. Tuberculosis 2013;93:232–8.
[18] Barba C, et al. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet 2004;363:157–63.
[19] Matthews DR, et al. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 1985;28:412–9.
[20] Kaona FA, Tuba M, Siziya S, Sikaona L. An assessment of factors contributing to treatment adherence and knowledge of TB transmission among patients on TB treatment. BMC Public Health 2004;4:68.
[21] Stratton IM. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. BMJ 2000;321:405–12.
[22] Borg R, et al. HbA1cand mean blood glucose show stronger associations with cardiovascular disease risk factors than do postprandial glycaemia or glucose variability in persons with diabetes: the A1C-Derived Average Glucose (ADAG) study. Diabetologia 2011;54:69–72.
[23] Srivastava S, Pasipanodya JG, Meek C, Leff R, Gumbo T. Multidrug-resistant tuberculosis not due to noncompliance but to between-patient pharmacokinetic variability. J Infec Dis 2011;204:1951–9.
[24] Srivastava S, Sherman C, Meek C, Leff R, Gumbo T. Pharmacokinetic mismatch does not lead to emergence of isoniazidor rifampin-resistant Mycobacterium tuberculosis but to better antimicrobial effect: a new paradigm for antituberculosis drug scheduling. Antimicrob Agents Chemother 2011;55:5085–9.
[25] Pusch T, Pasipanodya JG, Hall RG, Gumbo T. Therapy duration and long-term outcomes in extra-pulmonary tuberculosis. BMC Infect Dis 2014;14:115.
[26] Shaw JET, Pasipanodya JG, Gumbo T. Meningeal tuberculosis: high long-term mortality despite standard therapy. Medicine (Baltimore) 2010;89:189–95.
[27] Agweyu A, et al. Prevalence and correlates of treatment failure among Kenyan children hospitalised with severe community-acquired pneumonia: a prospective study of the clinical effectiveness of WHO pneumonia case management guidelines. Trop Med Int Health 2014;19:1310–20.
[28] WHO. Global Tuberculosis report 2014. WHO Rep 2014;2014:171.
[29] Ntsekhe M, Mayosi BM, Gumbo T. Quantification of echodensities in tuberculous pericardial effusion using fractal geometry: a proof of concept study. Cardiovasc Ultrasound 2012;10:30.
[30] Rigaut JP, Schoëvaërt-Brossault D, Downs AM, Landini G. Asymptotic fractals in the context of grey-scale images. J. Microsc. 1998;189:57–63.
[31] Drouin P, et al. Diagnosis and classification of diabetes mellitus. Diabetes Care 2009;32:S62–7.
[32] American Diabetes association. Classification and diagnosis of diabetes. Diabetes Care 2017;40:S11–24.
[33] Pasipanodya JG, et al. Serum drug concentrations predictive of pulmonary tuberculosis outcomes. J Infect Dis 2013;208:1464–73.
[34] Swaminathan S, et al. Drug concentration thresholds predictive of therapy failure and death in children with tuberculosis: bread crumb trails in random forests. Clin Infect Dis 2016;63:S63–74.
[35] Havlin S, et al. Fractals in biology and medicine. Chaos Solitons Fractals 1995;6:171–201.
[36] Higgins JP. Nonlinear systems in medicine. Yale J Biol Med 2002;75:247–60.
[37] Pasipanodya JG, Gumbo T. A new evolutionary and pharmacokinetic-pharmacodynamic scenario for rapid emergence of resistance to single and multiple anti-tuberculosis drugs. Curr Opin Pharmacol 2011;11:457–63.
[38] Selvin E, et al. Glycated hemoglobin, diabetes, and cardiovascular risk in nondiabetic adults. N Engl J Med 2010;362:800–11.
[39] Elizabeth S. Glycated hemoglobin, diabetes, and glycated hemoglobin, diabetes, and cardiovascular risk in nondiabetic adults. N Engl J Med 2010;362:800–11.
[40] Meigs JB, Nathan DM, Cupples LA, Wilson PWF, Singer DE. Tracking of glycated hemoglobin in the original cohort of the Framingham heart study. J Clin Epidemiol 1996;49:411–7.
[41] Chen G, et al. Contributions of the framingham heart study to the epidemiology of coronary heart disease. JAMA Cardiol 2016;41:279–81.
[42] Niemann S, et al. Stability of IS6110 restriction fragment length polymorphism patterns of Mycobacterium tuberculosis strains in actual chains of transmission. J Clin Microbiol 2000;38:2563–7.
[43] Goyal M, Saunders Na, van Embden JD, Young DB, Shaw RJ. Differentiation of Mycobacterium tuberculosis isolates by spoligotyping and IS6110 restriction fragment length polymorphism. J Clin Microbiol 1997;35:647–51.
[44] Poss JE. The meanings of tuberculosis for Mexican migrant farmworkers in the United States. Soc Sci Med 1998;47:195–202.
[45] de Jong BC, et al. Progression to active tuberculosis, but not transmission, varies by Mycobacterium tuberculosis lineage in The Gambia. J Infect Dis 2008;198:1037–43.
[46] Malley JD, Kruppa J, Dasgupta A, Malley KG, Ziegler A. Probability machines: consistent probability estimation using nonparametric learning machines. Methods Inf Med 2012;51:74–81.
[47] Wu T-F, Lin C-J, Weng RC. Probability estimates for multi-class classification by pairwise coupling. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit 2010;2:2301–11.
[48] Nicodemus KK, Malley JD, Strobl C, Ziegler A. The behaviour of random forest permutation-based variable importance measures under predictor correlation. BMC Bioinf 2010;11:110.
[49] Touw WG, et al. Data mining in the life science swith random forest: a walk in the park or lost in the jungle? Brief Bioinform 2013;14:315–26.