Supplementary

Table S1 Distribution and chi-squared test results of clinical features for 782 DS-TB patients and 1,455 DR-TB patients. The null hypothesis of Pearson's chi-squared test is that clinical feature categories are independent from drug-sensitive or drug-resistant TB

Clinical features		No. (%) of cases		P value	Clinical features		No. (%) of cases		P value
		DS-TB	DR-TB				DS-TB	DR-TB	
Age	<15	5 (0.22)	3 (0.13)	1.40×10^{-3}	Patient type	New	707 (31.60)	916 (40.95)	5.64×10^{-46}
	15-24	105 (4.69)	182 (8.14)			Relapse	73 (3.26)	345 (15.42)	
	25-34	196 (8.76)	353 (15.78)			Failure	2 (0.09)	194 (8.67)	
	35-44	151 (6.75)	350 (15.65)						
	45-54	136 (6.08)	301 (13.46)		Gender	Female	257 (11.49)	468 (20.92)	0.74
	55-64	106 (4.74)	166 (7.42)			Male	525 (23.47)	987 (44.12)	
	65+	83 (3.71)	100 (4.47)						

https://dx.doi.org/10.21037/qims-21-290

Table S2 Distribution and chi-squared test results of 22 radiological features for 782 DS-TB patients and 1455 DR-TB patients with the null hypothesis that the number of affected sextants is independent from being DS-TB or DR-TB. GGD indicates ground glass density. The abbreviation "Multi" means multiple affected sextants

Radiological features	No. of affected sextants	No. (%) of cases		P value	Radiological features	No. of affected sextants	No. (%) of cases		P value
		DS-TB	DR-TB				DS-TB	DR-TB	
Small nodules	None	376 (16.81)	321 (14.35)	2.43×10^{-37}	Small cavities	None	628 (28.07)	988 (44.17)	8.72×10^{-12}
	Single	186 (8.31)	421 (18.82)			Single	122 (5.45)	292 (13.05)	
	Multi	220 (9.83)	713 (31.87)			Multi	32 (1.43)	175 (7.82)	
Middle nodules	None	553 (24.72)	594 (26.55)	1.88×10^{-43}	Medium cavities	None	692 (30.93)	1,207 (53.96)	9.24×10^{-6}
	Single	144 (6.44)	397 (17.75)			Single	80 (3.58)	173 (7.73)	
	Multi	85 (3.80)	464 (20.74)			Multi	10 (0.45)	75 (3.35)	
Large nodules	None	730 (32.6)	1,282 (57.31)	4.36×10^{-4}	Large cavities	None	759 (33.93)	1,329 (59.41)	2.14×10^{-7}
	Single	38 (1.70)	129 (5.77)			Single	21 (0.94)	80 (3.58)	
	Multi	14 (0.63)	44 (1.97)			Multi	2 (0.09)	46 (2.06)	
Huge nodules	None	776 (34.73)	1,405 (62.99)	5.20×10^{-4}	Multi-sextant cavities	None	764 (34.15)	1,304 (58.29)	3.93×10^{-12}
	Single	6 (0.27)	45 (2.01)			Single	16 (0.72)	69 (3.08)	
	Multi	0 (0)	5 (0.22)			Multi	2 (0.09)	82 (3.67)	
Multiple nodules	None	550 (24.59)	690 (30.84)	2.75×10^{-24}	Visible multiple cavities	None	729 (32.59)	1,214 (54.27)	1.96×10^{-10}
	Single	105 (4.69)	325 (14.53)			Single	35 (1.56)	127 (5.68)	
	Multi	127 (5.68)	440 (19.67)			Multi	18 (0.80)	114 (5.10)	
Calcified nodules	None	592 (26.46)	1,048 (46.85)	0.17	Any kind of cavities	None	553 (24.72)	870 (38.89)	1.34×10^{-10}
	Single	100 (4.47)	217 (9.70)			Single	166 (7.42)	320 (14.30)	
	Multi	90 (4.02)	190 (8.49)			Multi	63 (2.82)	265 (11.85)	
Noncalcified nodules	None	432 (19.31)	475 (21.23)	4.08×10^{-26}	Collapses	None	725 (32.41)	1,294 (57.85)	2.40×10^{-3}
	Single	175 (7.82)	391(17.48)			Single	42 (1.88)	93 (4.16)	
	Multi	175 (7.82)	589 (26.33)			Multi	15 (0.67)	68 (3.04)	
Clustered nodules	None	562 (25.12)	774 (34.60)	3.71×10^{-18}	Low GGD infiltrates	None	447 (19.98)	765 (34.20)	3.84×10^{-6}
	Single	126 (5.63)	315 (14.08)			Single	229 (10.24)	367 (16.41)	
	Multi	94 (4.20)	366 (16.36)			Multi	106 (4.74)	323 (14.44)	
Low GGD nodules	None	420 (18.78)	405 (18.10)	8.50×10^{-34}	Medium density infiltrates	None	401 (17.93)	781 (34.91)	0.23
	Single	174 (7.78)	420 (18.78)			Single	205 (9.16)	334 (14.93)	
	Multi	188 (8.40)	630 (28.16)			Multi	176 (7.87)	340 (15.20)	
Medium density nodules	None	432 (19.31)	373 (16.67)	2.55×10^{-45}	High density infiltrates	None	714 (31.92)	1,100 (49.17)	1.89×10^{-21}
	Single	179 (8.00)	428 (19.13)			Single	58 (2.59)	192 (8.58)	
	Multi	171 (7.64)	654 (29.24)			Multi	10 (0.45)	163 (7.29)	
High density nodules	None	682 (30.49)	1,164 (52.03)	1.12×10^{-6}	Any kind of nodules	None	345 (15.42)	258 (11.53)	1.47×10^{-42}
	Single	74 (3.31)	161 (7.20)			Single	103 (4.60)	184 (8.23)	
	Multi	26 (1.16)	130 (5.81)			Multi	334 (14.93)	1,013 (45.28)	

https://dx.doi.org/10.21037/qims-21-290

Table S3 Distribution and chi-squared test results of four radiological features for 782 DS-TB patients and 1,455 DR-TB patients based on the null hypothesis that the lesion volume or the presence of lesion is independent from DS-TB or DR-TB

Radiological features		No. (%) of cases		P value	Radiological features		No. (%) of cases		P value
		DS-TB	DR-TB				DS-TB	DR-TB	
Overall percentage of abnormal volume	0	17 (0.76)	8 (0.36)	1.50×10^{-3}	Pleural effusion percentage of hemithorax involved	0	175 (7.82)	461 (20.61)	1.98×10^{-5}
	≤50%	653 (29.19)	1,213 (54.22)			≤50%	603 (26.96)	987 (44.12)	
	>50%	112 (5.01)	234 (10.46)			>50%	4 (0.18)	7 (0.31)	
Presence of other non-TB abnormalities	No	686 (30.67)	1,299 (58.07)	0.27	Presence of mediastinal lymphadenopathy	No	720 (32.19)	1,301 (58.16)	0.04
	Yes	96 (4.29)	156 (6.97)			Yes	62 (2.77)	154 (6.88)	

Table S4 Detailed information of SVM model for classification between DS-TB and DR-TB

Environment	Intel Xeon CPU E3-1275 V6 @ 3.80 GHz 3.79 GHz	
Software	Matlab 2020a	
SVM model	Function	fitcsvm(trainData,trainlabel,'Standardize',true, ClassNames', {'sensitive', 'resistant'}, 'KernelFunction','RBF','KernelScale','auto');
	'Standardize'	Standardize the predictors before training using their corresponding weighted means and weighted standard deviations: $x_i^* = (x_i - \mu_i^*)/\sigma_i^*$, (1) where $\mu_i^* = \frac{1}{\sum_k w_k^*}\sum_k w_k^* x_{ik}$, (2) $(\sigma_i^*)^2 = \frac{v_1}{v_1^2 - v_2}\sum_k w_k^*(x_{ik} - \mu_i^*)^2$, (3) $v_1 = \sum_i w_i^*$, (4) $v_2 = \sum_i (w_i^*)^2$. (5) x_{ik} is observation k (row) of predictor i (column) and w_k^* is the weight of observation k (row). w_k^* is set to 1 in our work.
	Kernel function	RBF (Radial Basis Function) $h(x_i, x_j) = \exp(-\|x_i - x_j\|^2)$, (6)
	Optimization solver	SMO (Sequential Minimal Optimization)
	Class names	'sensitive'; 'resistant'
	No of predictors	25
	No of observations (1st round)	2618
	Weight matrix (1st round)	2618×1
	Support Vectors (1st round)	1934×25
	Support vector labels (1st round)	1934×1

Note: 25-feature-based training data and testing data for the 1st round of the ten-fold cross validation, as well as SVM testing code can be downloaded here: https://github.com/fyang5/ClassificationDRvsDS_SVM.git.

https://dx.doi.org/10.21037/qims-21-290