



Perspectives

Multi-modal deep learning based on multi-dimensional and multi-level temporal data can enhance the prognostic prediction for multi-drug resistant pulmonary tuberculosis patients



Zhen-Hui Lu, Ming Yang, Chen-Hui Pan, Pei-Yong Zheng, Shun-Xian Zhang*

Longhua Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai 200032, China

ARTICLE INFO

Keywords:

MDR-PTB
Multi-modal
Deep learning
Prognosis

ABSTRACT

Despite the advent of new diagnostics, drugs and regimens, multi-drug resistant pulmonary tuberculosis (MDR-PTB) remains a global health threat. It has a long treatment cycle, low cure rate and heavy disease burden. Factors such as demographics, disease characteristics, lung imaging, biomarkers, therapeutic schedule and adherence to medications are associated with MDR-PTB prognosis. However, thus far, the majority of existing studies have focused on predicting treatment outcomes through static single-scale or low dimensional information. Hence, multi-modal deep learning based on dynamic data for multiple dimensions can provide a deeper understanding of personalized treatment plans to aid in the clinical management of patients.

1. Background

Despite the emergence of new drugs, diagnosis and treatment schemes of multi-drug resistant pulmonary tuberculosis (MDR-PTB) still remains as one of the global health threats [1,2]. MDR-PTB seriously impedes the realization of “ending the tuberculosis (TB) epidemic” by 2035 [1]. China is one of the countries with a high burden of MDR-PTB: 56% of the MDR-PTB patients remain to be identified and second-line MDR-PTB treatments started only in 38% of the diagnosed patients [3]. Significant challenges for MDR-PTB control efforts have been due to the determination of MDR-PTB treatment success [3]. Current recommendations for MDR-PTB treatment surveillance depend on sputum and culture conversion. The low sensitivity and long-periodic time undermines their diagnostic value as predictors of clinical outcomes. Mathematical models based on multi-modal data can accurately and quickly predict the treatment outcome at prophase treatment stage [4,5]. It is crucial to prompt doctors to perform precise patient treatment and improve outcomes.

2. Factors causing MDR-PTB disease progression

2.1. Demographics, comorbidities, clinical scores, signs and symptoms

Abundant evidence shows that risk factors are associated with failure outcomes for MDR-PTB patients, including demographics (age, ethnicity, gender, occupation, etc.), comorbidities (malnutrition, diabetes mellitus,

hypertension, human immunodeficiency virus (HIV)/acquired immune deficiency syndrome (AIDS), etc.), drug exposure variables (regimens of various compositions and dosing, irregular treatments, etc.), clinical variables (such as history of TB treatment, pathological classification, etc.), disease severity (smear grade, sputum culture grade, etc.) and clinical signs and symptoms (weight, cough, lung function, etc.). It is necessary to assess the utility of clinical scoring systems and risk stratification to predict treatment outcomes in future studies.

2.2. Imageomics

Chest computed tomography (CT) is more sensitive and specific for MDR-PTB prognosis predictions. Changes in CT findings over time have been shown to correlate with treatment outcomes. Greater volume, thick-walled and multiple cavities on chest CT performed at prophase treatment stage had a high predictive value for poor treatment results. In addition, “caseating lung consolidation with cavitation” is an important indicator to delayed culture conversion and unfavorable treatment outcomes for MDR-PTB patients.

2.3. Biomarker

The MDR-PTB illness development is related to multi-level biological information, including genomics, transcriptome, epigenome, proteome, immunome, metabolome, *Mycobacterium tuberculosis* strain characteristics

* Corresponding author.

E-mail addresses: Dr_luzh@shutcm.edu.cn (Z.-H. Lu), yangpluszhu@sina.com (M. Yang), panchenhui06@163.com (C.-H. Pan), zpychina@sina.com (P.-Y. Zheng), zhangshunxian110@163.com (S.-X. Zhang).

<https://doi.org/10.1016/j.soh.2022.100004>

Received 1 October 2022; Accepted 19 November 2022

Available online 23 November 2022

2949-7043/© 2022 The Author(s). Published by Elsevier B.V. on behalf of Shanghai Jiao Tong University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

(protein structures, phenotypic and genotypic resistance, load and fitness of bacteria, etc.) [6]. Bionomic analyses have yielded signatures that are associated with disease outcomes. For instance, 22-gene ribonucleic acid (RNA) from blood can accurately predict individual treatment outcome [7]. Meanwhile, MDR-PTB patients usually have poor prognosis when *Mycobacterium tuberculosis* is resistant to second-line drugs (moxifloxacin, cycloserine, etc) or the strain contain genes. In addition, proinflammatory cytokines and immunocytes are also potential biomarkers for treatment outcome monitoring. Biomarker-guided model can predict treatment outcome and prompt doctors to implement individualized and precise therapeutic treatment plans.

3. New method for predicting MDR-PBT progression

3.1. Artificial intelligence (AI)

In this era of interdisciplinary science, many scientific achievements, especially in artificial intelligence (AI), have brought dramatic revolutions to human society. With the rapid development of AI, especially the emergence and rapid development of deep learning technology, AI has entered the stage of large-scale medical application. Concurrent advances in information technology infrastructure and mobile computing power in many low and middle-income countries (LMICs) have raised hopes that AI might help to address challenges unique to the field of global health and accelerate achievement of the health-related sustainable development goals [8]. Signal processing methods are often used together with machine learning to automate the diagnosis of communicable diseases. For instance, signal processing interventions focused specifically on the use of radiological data for tuberculosis and drug-resistant tuberculosis [8]. With the research of big models becoming the focus of AI, it is possible that the big model leads the technological transition in the next few years and brings a new industrial pattern. The big models play the basic role of 'intelligence producer', which can generate high-quality intelligence power under the support of large amount computing power and serve various AI applications. Through the development of big information model and big biomimetic model, the research process in the fields of electronic information and biomedicine can be accelerated. Meanwhile, the development of big models can help those innovative enterprises and individual developers construct high-intelligence applications, thus promoting the intellectual update of real world application.

AI has an overwhelming advantage in scientific exploration by effectively leveraging massive multi-dimensional data. AI-driven interventions can also focus on the diagnosis and prognosis of communicable diseases. AI based on imaging omics, using deep-learning features extracted from chest CT, performed significantly better in predictions as compared with clinical data-derived severity scores. Hence, AI can take advantage of imaging omics data to identify high-risk MDR-PTB patients sooner and improve outcomes.

The disease performances have continuously changed in the development of MDR-PTB illness. The cross-sectional data can only provide relatively limited information to understand the disease progression, while the multi-dimensional data in real world are generally longitudinal data (time-series data) [9]. Hence, mathematical models based on dynamic data have more accuracy to predict disease prognosis.

Deep learning has shown incomparable advantages in the analysis of time-series data [9]. The long short-term memory neural network has the biggest advantage in resolving the problem that recurrent neural network (RNN) cannot "remember" the information features of previous states. Combined with the high efficiency of feature extraction with convolutional neural network [9], long short term memory has been successfully applied to various fields of medical research involving time-series dates in recent years.

3.2. The 'One Health' approach

Over 2300 years ago, a Taoist philosopher by the name of Zhuang Zi argued that a human being is an integral part of nature [4,5,10]. The

individuals health is the result of the interaction between environment, time, location and self-factor. Over time, the combined effects of various factors have led to changes in health status. Many factors interact and change dynamically. Therefore, new methods are needed to evaluate the factors that affect the prognosis of disease. The 'One Health' approach aims to balance and optimize the health of the people and ecosystems in a sustainable way [4,5,10]. It has been suggested to address complex health problems at the human-environment interface, coupled with inter- and trans-factor involvement, which makes broader contributions to achieve optimal public health outcomes. It can provide service for formulating policies to promote the prevention and control of emerging infectious diseases, and improve the cure rate [4,5,10].

3.3. Multi-modal deep learning (DL)

Computer-aided prognosis (CAP) is a new and exciting complement to the field of computer-aided diagnosis (CAD) and involves developing and applying computerized image analysis and multi-modal data-fusion-algorithms to digitized patient data (e.g. imaging, tissue, genomic) for helping physicians to predict disease outcome and patient survival. While a number of data channels, ranging from the macro (e.g. CT) to the nano-scales (biological information, disease information etc.), are now being routinely acquired for disease characterization, one of the challenges in predicting patient outcome and treatment response has been in our inability to quantitatively fuse these disparate, heterogeneous data sources.

Human brains are complex systems that can process information from multiple sensory modalities. That is, they are capable of simultaneously handling language, image, video and other modal informations, which allow us to accurately and efficiently complete perception, understanding and decision-making tasks. To mimic these core cognitive abilities in humans, it is necessary and also promising for AI models to train on large-scale multi-modal data. The key challenge of training a big multi-modal model is to learn the correlations and the complementarity of multiple modalities due to the heterogeneity of multi-modal data [11,12]. The goal of big multi-modal models is to align data from different modalities and acquire the ability to transfer their learned knowledge to various downstream tasks, to ultimately approach strong AI. In general, current research on multi-modal models mainly focus on video, image, speech and language modals. Exploiting data from more modalities is also very meaningful, and it may make exciting and major discoveries, eventually leading AI models more similar to human brains [11-13].

Considering that redundant and irrelevant features in high-dimensional features usually influence learning accuracy, the least absolute shrinkage and selection operator (lasso) regression algorithm was performed to reduce feature dimensions by retaining high discriminative features [12]. Based on the selected features, the random forest classification model and multi-layer perceptron, which is frequently used in the field of supervised machine learning, was built for our task of prognostic diagnosis [12]. The performance of predictive models was assessed according to the commonly used accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and the area under the curve of the receiver operating characteristic, F1 score, Cohen's Kappa and balanced accuracy, etc [14]. In addition, information fusion can occur in a myriad of ways. In machine learning, early, intermediate and late fusion is typified if all the information flows into a single model (early), a step-wise fashion where outputs from one model become inputs for another (intermediate), and lastly, where all unique data types undergo separate modelling after which ensembling and/or voting occurs (late). Finally, the field application of the model is the most effective way to verify the feature fusion method and model performance.

4. Conclusion

Remarkable progress has been made for learning cross-lingual cross-modal representations. Multi-modal model is developed for predicting MDR-PTB treatment outcome. It is expected that the combination of

convolutional neural network identified deep-learning features and multi-dimension time-series data, such as demographic information, clinical variables, imaging data, lab measurements, therapeutic interventions are in time series. Multi-modal data modeling might be possible to take advantage of this readily available multi-dimension time series data to identify high-risk patients early and improve treatment outcomes.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The full study protocol and the datasets are available, following manuscript publication, upon request from the corresponding author (ZHANG Shun-Xian, zhangshunxian110@163.com).

Authors' contributions

S-XZ, Z-HL and P-YZ conceived the idea. C-HP and MY critically drafted the paper. All authors read and approved the final manuscript.

Funding

The study was supported by the fund of Medical Innovation Research Special Project of the Shanghai 2021 "Science and Technology Innovation Action Plan" (21Y11922500, 21Y11922400), the promotion and application of deep learning in traditional Chinese medicine to improve the ability and level in clinical research field (SHDC2022CRS039), and the talent fund of Longhua Hospital of (LH001.007). Funding sources had no role in the design and conduct of the study, collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

Declaration of competing interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

None.

Abbreviations

AI Artificial intelligence
AIDS Acquired immune deficiency syndrome

CAD Computer-aided diagnosis
CAP Computer-aided prognosis
COVID-19 Coronavirus disease 2019
CT Computed tomography
DL Deep learning
HIV Human immunodeficiency virus
LMICs Low and middle-income countries
LSTM Long short-term memory
MDR-PTB Multi-drug resistant pulmonary tuberculosis
RNA Ribonucleic acid
RNN Recurrent neural network
TB Tuberculosis

References

- [1] X.Y. Ding, W.H. Mao, W. Lu, H. Yu, Q. Liu, P. Lu, et al., Impact of multiple policy interventions on the screening and diagnosis of drug-resistant tuberculosis patients: a cascade analysis on six prefectures in China, *Infect. Dis. Pov.* 10 (1) (2021 Jan 19) 8, <https://doi.org/10.1186/s40249-021-00793-9>.
- [2] S.X. Zhang, L. Qiu, C. Li, W. Zhou, L.M. Tian, H.Y. Zhang, et al., Efficacy of integrating short-course chemotherapy with Chinese herbs to treat multi-drug resistant pulmonary tuberculosis in students - China, 2015-2019, *China CDC Wkly* 3 (2021 Nov 6) 131, <https://doi.org/10.1186/s40249-021-00913-5>.
- [3] W. Su, Y. Ruan, T. Li, X. Du, J. Jiang, Y. He, et al., Epidemiological characteristics of rifampicin-resistant tuberculosis in students - China, 2015-2019, *China CDC Wkly* 3 (26) (2021 Jun 25) 549-552, <https://doi.org/10.46234/ccdcw2021.142>.
- [4] X.X. Zhang, J.S. Liu, L.F. Han, G. Simm, et al., One Health: new evaluation framework launched, *Nature* 604 (7907) (2022 Apr) 625, <https://doi.org/10.1038/d41586-022-01108-0>.
- [5] X.X. Zhang, J.S. Liu, L.F. Han, S. Xia, S.Z. Li, O.Y. Li, et al., Towards a global One Health index: a potential assessment tool for One Health performance, *Infect. Dis. Pov.* 11 (1) (2022 May 22) 57, <https://doi.org/10.1186/s40249-022-00979-9>.
- [6] J.G. Greener, S.M. Kandathil, L. Moffat, D.T. Jones, A guide to machine learning for biologists, *Nat. Rev. Mol. Cell Biol.* 23 (1) (2022 Jan) 40-55, <https://doi.org/10.1038/s41580-021-00407-0>.
- [7] J. Heyckendorf, S. Marwitz, M. Reimann, K. Avsar, A.R. DiNardo, G. Günther, et al., Prediction of anti-tuberculosis treatment duration based on a 22-gene transcriptomic model, *Eur. Respir. J.* 58 (3) (2021 Sep 2), 2003492, <https://doi.org/10.1183/13993003.03492-2020>.
- [8] N. Schwalbe, B. Wahl, Artificial intelligence and the future of global health, *Lancet* 395 (10236) (2020 May 16) 1579-1586, [https://doi.org/10.1016/S0140-6736\(20\)30226-9](https://doi.org/10.1016/S0140-6736(20)30226-9).
- [9] L. Wu, G. Huang, X. Yu, M. Ye, L. Liu, Y. Ling, et al., Deep learning networks accurately detect ST-segment elevation myocardial infarction and culprit vessel, *Front. Cardiovasc. Med.* 9 (2022 Mar 10), 797207, <https://doi.org/10.3389/fcvm.2022.797207>.
- [10] H.Q. Zhao, S.W. Fei, J.X. Yin, Q. Li, T.G. Jiang, Z.Y. Guo, et al., Assessment of performance for a key indicator of One Health: evidence based on One Health index for zoonoses in Sub-Saharan Africa, *Infect. Dis. Pov.* 11 (1) (2022 Oct 22) 109, <https://doi.org/10.1186/s40249-022-01020-9>.
- [11] A. Nair, M. Singh, S.R. Aglyamov, K.V. Larin, Multimodal heartbeat and compression optical coherence elastography for mapping corneal biomechanics, *Front. Med.* 9 (2022 Apr 5), 833597, <https://doi.org/10.3389/fmed.2022.833597>.
- [12] S. Guo, L. Wang, Q. Chen, L. Wang, J. Zhang, Y. Zhu, Multimodal MRI image decision fusion-based network for glioma classification, *Front. Oncol.* 12 (2022 Feb 24), 819673, <https://doi.org/10.3389/fonc.2022.819673>.
- [13] D. Hao, Q. Li, Q.X. Feng, L. Qi, X.S. Liu, D. Arefan, et al., Identifying prognostic markers from clinical, radiomics, and deep learning imaging features for gastric cancer survival prediction, *Front. Oncol.* 11 (2022 Feb 2), 725889, <https://doi.org/10.3389/fonc.2021.725889>.
- [14] K. Tan, W. Huang, X. Liu, J. Hu, S. Dong, A multi-modal fusion framework based on multi-task correlation learning for cancer prognosis prediction, *Artif. Intell. Med.* 126 (2022 Apr), 102260, <https://doi.org/10.1016/j.artmed.2022.102260>.