Infect Dis Trop Med 2025; 11: e1766

# Is ChatGPT-4 accurate and complete when answering questions on tuberculosis? Results of the ChatGTB study

*A. De Vito[1], A. Colpani[1], D. Buonsenso[2,3], P.M.M. Candoli[4], E. Falbo[3,5], S. La Fauci[4], G. Madeddu[1], T. Masini[3,6], G. Misiano[4], C. Monari[3,7], A. Pontarelli[3,8], N. Riccardi[3,9], L. Saderi[3], F. Saluzzo[3,10], G. Sotgiu[3,11], M. Tadolini[3,12], G. Besozzi[3], A. Calcagno[3,13,14]*

[1]Unit of Infectious Diseases, Department of Medicine, Surgery, and Pharmacy, University of Sassari, Sassari, Italy
[2]Department of Woman and Child Health and Public Health, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy
[3]Stop TB Italy, Milan, Italy
[4]FIMMG Piemonte, RIMeG Research Group, Turin, Italy
[5]Centro Medicina del Viaggiatore e delle Migrazioni, ASP Catanzaro, Catanzaro, Italy
[6]Independent consultant, Lucca, Italy
[7]Infectious Diseases Unit, Azienda Ospedaliero-Universitaria L. Vanvitelli, Naples, Italy
[8]Unit of Respiratory Infectious Diseases, Cotugno Hospital, Azienda Ospedaliera dei Colli, Naples, Italy
[9]Infectious Diseases Unit, Azienda Ospedaliera Universitaria Pisana, Pisa, Italy
[10]Emerging Bacterial Pathogens Unit, Division of Immunology, Transplantation and Infectious Diseases, IRCCS San Raffaele Scientific Institute, Milan, Italy
[11]Clinical Epidemiology and Medical Statistics Unit, Department of Medicine, Surgery and Pharmacy, University of Sassari, Sassari, Italy
[12]Infectious Disease Unit, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy
[13]Unit of Infectious Diseases, Department of Medical Sciences, University of Turin, Turin, Italy
[14]Department of Translational Medicine, University of Western Piedmont, Novara, Italy

***ABSTRACT:***

— ***Objective:*** Artificial intelligence (AI), particularly large language models like ChatGPT, offers the potential to disseminate health information. This study aimed to assess the accuracy and completeness of ChatGPT-4's responses to TB-related questions.

— ***Materials and Methods:*** Ninety English-language TB questions based on official guidelines and clinical experience were formulated. ChatGPT-4o provided answers to these questions between February 1 and March 1, 2024. Three evaluation subgroups assessed the responses for accuracy (using a six-point Likert scale) and completeness (using a three-point Likert scale). Statistical analyses were performed using non-parametric tests.

— ***Results:*** The median accuracy score was 5 out of 6, with 88.9% of responses scoring at least 5, indicating high overall accuracy. However, only 34.4% achieved the highest score of 6, with diminished performance on medium and high level of expertise (LOE) questions. Low LOE questions had the highest

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License

Corresponding Authors: Andrea De Vito, MD; e-mail: andreadevitoaho@gmail.com;
Andrea Calcagno, MD; e-mail: andrea.calcagno@unito.it

accuracy, with 63.3% scoring 6. Completeness scores showed that 48.9% of responses were comprehensive (score of 3), particularly for low LOE questions (70% scored 3). In contrast, only 23.3% of high LOE questions achieved the highest completeness score. ChatGPT-4 often lacked specificity in complex topics, such as drug-resistant TB therapies, and provided outdated information not aligned with current World Health Organization guidelines.

— ***Conclusions:*** ChatGPT-4 effectively delivers accurate and comprehensive information for general TB inquiries, making it a valuable resource for the public and non-specialist clinicians. However, its performance declines with increasing question complexity, limiting its utility for advanced clinical decision-making in TB care. Continuous updates and enhancements are necessary to improve its accuracy and relevance in specialised medical contexts.

— ***Keywords:*** Tuberculosis, ChatGPT, Education, LLM, Prevention, Tuberculosis treatment.

## INTRODUCTION

Tuberculosis (TB) is an important cause of morbidity and mortality despite the availability of drugs for prevention and therapy worldwide. The World Health Organization (WHO) reported 7.5 million new cases, with 1.3 million deaths in 2023[1]. While several ambitious targets have been set for 2030 and 2035, recent data suggest they will not be attained. The negative impact of COVID-19 on TB services has hindered the achievement of public health targets[1-3]. The adoption of WHO's End TB Strategy and the United Nations (UN) Sustainable Development Goals included large reductions in the TB incidence rate, mortality, and costs faced by those affected by TB and their households[4,5]. While the wide and appropriate use of antitubercular drugs for drug-sensitive and drug-resistant TB is a key element of the strategy, several other pillars and components have been identified. They include early diagnosis of TB, universal drug-susceptibility testing, identification, screening, and preventive treatment of contacts and high-risk groups, as well as community engagement. Additionally, the third pillar strongly emphasises the need for intensified research and innovation[5].

In this setting, artificial intelligence (AI) may be a relevant tool given the complexity of the disease and the need for new strategies. The use of computer-aided detection for TB diagnosis using digital chest radiography has been identified as a promising AI-based screening test[6,7]; the WHO[8] has recently suggested its adoption along with other more traditional diagnostic strategies. AI-based spatial data processing and analysis tools (such as MATCH) have been used for the integration of Geographic Information Systems with routine health surveillance data and other contextual data sources[9]: their application in identifying TB hotspots has been reported in India and Nigeria[10]. While other potential applications of this powerful technological tool may be applied to TB care and research, there is immediate potential for using it to inform individuals, patients, and healthcare workers. With the large availability of smartphones also in intermediate and limited-resource countries, many people can access the internet and look for information on TB symptoms, diagnosis, and treatment. The search for medical information and potential self-diagnosis on the internet is a common practice worldwide. A recent study[11] reported that in 2019, 72% of Germans searched online for health-related purposes. The provision of accurate, complete, and responsible information (although no computer-based self-diagnosis can substitute for patient-healthcare worker interactions) is paramount in a contagious and potentially deadly disease. Although people have a general idea of what TB is, several gaps in knowledge on transmission, treatment, and prevention may lead to diagnostic and treatment delays for people living with TB[12]. While patients' knowledge may be highly variable around the world, high-burden countries may be particularly at risk of poor health literacy[13]. Health literacy is associated with TB treatment adherence and care quality, thus identifying it as a potential target in the fight against TB[14]. Searching the word "tuberculosis" on Google Trends reports that most of the queries were performed in countries located in Central and Latin America and Africa and looking for information on the disease, its infectivity, and its treatment[15].

ChatGPT is an AI large language model implemented by OpenAI as an advanced conversational AI technology with more than 180 million monthly users. ChatGPT use in several medical disciplines has been recently explored and evaluated: while positive evaluations have been reported, some authors have raised concerns about legal issues, as well as the quality and nature of the training data, which can impact the accuracy of AI-generated text[16]. A recent article[17] in the Lancet Infectious Diseases recognised that "ChatGPT neglects a neglected disease" by forgetting to include Histoplasmosis, among other differential diagnoses. While the performance of ChatGPT on TB information has not been explored, there are several examples in other fields of medicine[18-20]. Papers have been published on infectious diseases, such as HIV and bacterial infections[21-26]. Our group previously reported that ChatGPT provided accurate information on HIV-related topics but lacked the socio-political context and inclusivity essential for effective health communication[21]. In this research, we aimed to study ChatGPT's role in providing answers to TB-related questions.

## MATERIALS AND METHODS

A study was carried out to assess the effectiveness of ChatGPT in providing accurate and complete information on TB to the general public and to medical professionals.

Fifteen medical doctors with different TB experience were recruited and divided into four subgroups based on their experience. The first team included two junior doctors working in the Infectious Diseases field and a senior doctor. Three additional teams comprised five doctors, specialists in Infectious Diseases, Pneumology, or general practitioners, with different experiences in the TB field.

### Questions Development and Answers Collection

One subgroup formulated questions in English based on official TB websites, clinical and consultancy experiences. They were grouped into three levels of expertise (LOE): "low" was a question that every medical doctor should be able to answer, "medium" was a question that an Infectious Diseases or Pneumology specialist should be able to address, "high" was a question addressed by TB experts. Answers from ChatGPT 4 were collected between February 1, 2024, and March 1, 2024, and recorded in an Excel spreadsheet.

### Questions and Answers Evaluation

The responses from ChatGPT were meticulously recorded by team one and shared with the other teams for evaluation. A designated team member (team one) oversaw the recording and transcription of all dialogues between ChatGPT and the staff members inputting questions, ensuring the fidelity of the information captured in a comprehensive file, which was then sent to the other teams. Three different subgroups analysed low, medium, and high LOE answers, respectively.

Accuracy of the answers was based on a six-point Likert scale, with 1 associated with an incorrect response, 2 with more incorrect than correct elements, 3 with equal correct and incorrect elements, 4 with more correct than incorrect elements, 5 with an almost full correct response, and 6 with a correct response. A three-point Likert scale was adopted to assess completeness of the answers: 1 was associated with an incomplete answer, 2 with an answer addressing all aspects of the question and providing the minimum information required, and 3 with a comprehensive response including additional information. Details on the evaluation are described in the **Supplementary File 1**.

### Statistical Analysis

Categorical variables were summarised with absolute and relative (percentage) frequencies. The Kruskal-Wallis test was adopted to assess accuracy and completeness score differences based on their non-parametric distribution. Chi-square or Fisher's exact tests were used for qualitative variables. A *p*-value <0.05 was considered statistically significant, and all analyses were conducted with STATA version 16.1 (College Station, TX, USA).

## RESULTS

Ninety questions were formulated to be submitted to ChatGPT version 4, and they are shown in **Supplementary File 2**. They concerned all aspects of TB prevention and care, including symptoms, consequences, contagiousness, diagnostic tests, and treatment: some were formulated as directly asked by individuals and patients, while others were from healthcare workers' points of view.

### Accuracy

The median (IQR) accuracy was 5 (5-6), with the majority (88.9%) of responses achieving at least a score of 5. Only two answers received a score of 3: (i) "What therapy is used for pre-eXtensively Drug-Resistant (pre-XDR) tuberculosis?" and (ii) "What therapy is used for Rifampicin-Resistant (RR) tuberculosis?" (Figure 1).

63.3% of "low" questions scored 6, while the percentage decreased to 20.0% for medium and high LOE questions, with a concomitant increase in scores of 4 and 5 (Table 1).

### Completeness

The overall median (IQR) completeness score was 2 (2-3), with 48.9% scoring 2 and 3. Only two answers scored 1: i) "At the end of the therapy, I was declared cured, but I still have dyspnea and often cough; is it possible that the tuberculosis has reactivated?"; ii) "What tests should I perform before prescribing therapy for tuberculosis?" (Figure 2).

Low LOE questions achieved higher completeness: about 70% scored 3. For medium-high LOE questions, although the median (IQR) score was 3 (2-3), the proportion achieving a score of 3 decreased to 53.4%. Only 23.3% of more complex questions scored 3, and 76.7% scored 2 (Table 2).

Important to note, ChatGPT recommends seeking medical advice in 58 out of 90 (64.4%) of the questions.

## DISCUSSION

The study assessed the effectiveness of ChatGPT-4 to provide accurate and complete information following questions on tuberculosis with different levels of complexity. ChatGPT-4 is effective, especially for low LOE questions. Nevertheless, the percentage of correct answers declined as the complexity of the questions increased.

Median accuracy scores were generally high, and 88.9% of responses scored at least a 5, showing that ChatGPT-4 is trustworthy in giving facts on tubercu-

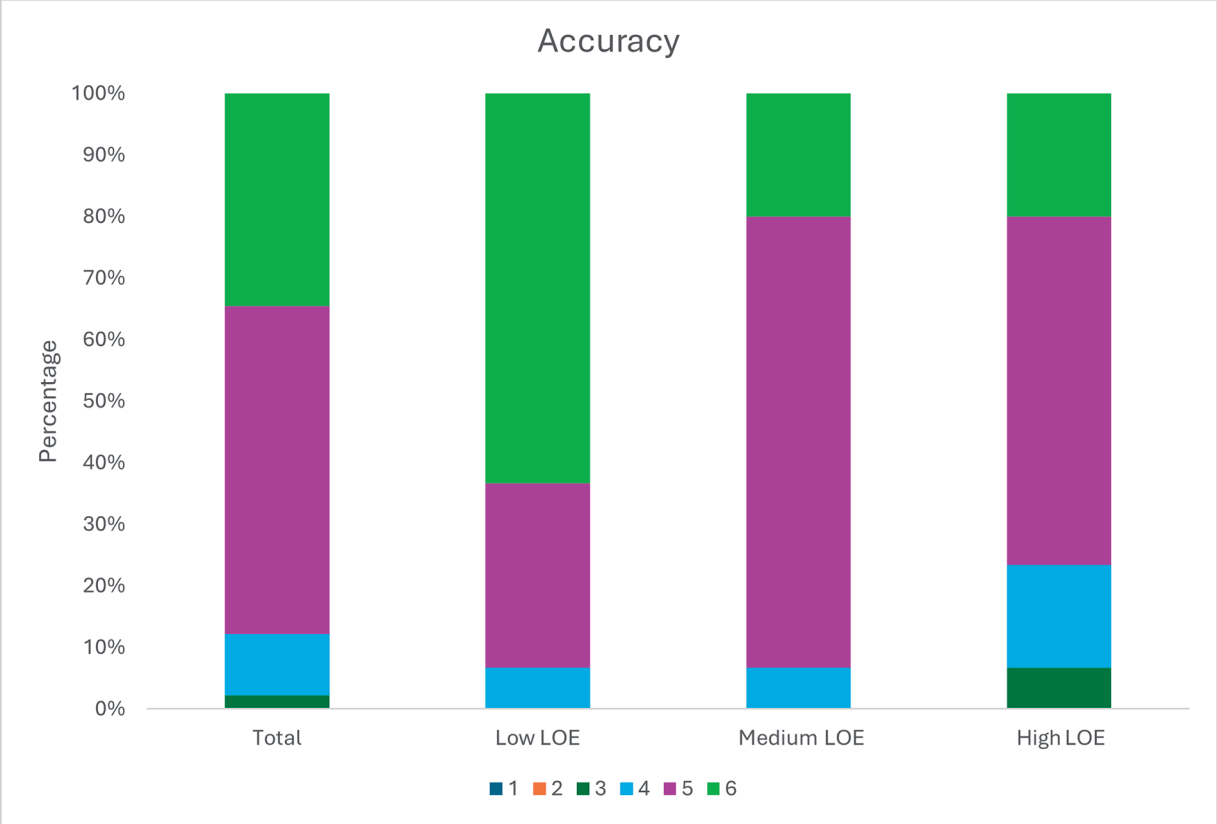


**Figure 1.** Accuracy scoring of ChatGPT-4 in answering open-ended questions on tuberculosis, divided according to the Level of Expertise (LOE).

losis. Moreover, a satisfactory accuracy was found for low LOE questions, with usefulness for the lay public and non-specialist clinicians. However, the accuracy decreased for medium and high LOE questions, with only 20.0% receiving the highest accuracy score. The model scored 3 in accuracy for two questions related to drug-resistant tuberculosis therapy, suggesting its poor specificity.

Regarding completeness, similar trends were observed. Low LOE questions had a higher percentage of comprehensive answers, with a median score of 2 out of 3. About 70% of low LOE questions received a completeness score of 3. For high LOE questions, only 23.3% achieved the highest score, lacking details for complex medical inquiries.

One of the main issues about ChatGPT responses is that the model ChatGPT does not include other pathogenic species like *M. bovis* or *M. africanum*, and less incident tuberculosis agents[27,28]. In addition, the model provided outdated responses in the context of drug-resistant TB; answers did not align with the latest World Health Organization (WHO) guidelines for therapies of multi-drug-resistant (MDR), rifampicin-resistant (RR), pre-extensively drug-resistant (pre-XDR), and extensively drug-resistant (XDR) tuberculosis[29]. Regarding this specific topic, the model also did not recommend consulting updated guidelines or seeking expert medical advice, which could result in suboptimal treatment decisions.

**Table 1.** Accuracy scoring and median of ChatGPT-4 in answering open-ended questions on tuberculosis, divided according to the Level of Expertise (LOE).

| | Total | Low LOE | Medium LOE | High LOE | *p*-value |
|---|---|---|---|---|---|
| Median, (IQR) | 5 (5-6) | 6 (5-6) | 5 (5-5) | 5 (5-5) | <0.001 |
| Accuracy, 6 points | 31 (34.4) | 19 (63.3) | 6 (20.0) | 6 (20.0) | <0.001 |
| Accuracy, 5 points | 48 (53.3) | 9 (30.0) | 22 (73.3) | 17 (56.7) | |
| Accuracy, 4 points | 9 (10.0) | 2 (6.7) | 2 (6.7) | 5 (16.7) | |
| Accuracy, 3 points | 2 (2.2) | 0 | 0 | 2 (6.6) | |
| Accuracy, 2 points | 0 | 0 | 0 | 0 | |
| Accuracy, 1 point | 0 | 0 | 0 | 0 | |

**Table 2.** Completeness scoring and median of ChatGPT-4 in answering open-ended questions on tuberculosis, divided according to the Level of Expertise (LOE).

| | Total | Low LOE | Medium LOE | High LOE | *p*-value |
|---|---|---|---|---|---|
| Median, (IQR) | 2 (2-3) | 3 (2-3) | 3 (2-3) | 2 (2-2) | 0.003 |
| Completeness, 3 points | 44 (48.9) | 21 (70) | 16 (53.4) | 7 (23.3) | 0.001 |
| Completeness, 2 points | 44 (48.9) | 8 (26.7) | 13 (43.3) | 23 (76.7) | |
| Completeness, 1 point | 2 (2.2) | 1 (3.3) | 1 (3.3) | 0 | |

Another issue is the missing discussion on subclinical or incipient TB[30,31]. Additionally, generic questions on TB are assumed to be related to pulmonary TB without mentioning other clinical types of disease. On the other hand, for all questions about the risk of transmission, the assumption that we were asking about pulmonary TB resulted in clear information about all the different measures to prevent the spread of the infection.

These results differ from those reported in previous studies on ChatGPT's performance with infectious diseases. For example, Tunçer and Kadir[32] assessed ChatGPT's ability to answer questions from social media, experts, and guidelines on various infectious diseases, including TB. They posed 19 tuberculosis-related questions to ChatGPT, and 90% of the answers were accurate. In contrast, only 34.4% of our answers received the highest accuracy score. This discrepancy may be due to the higher difficulty level of our questions.

When comparing ChatGPT's performance on TB with its performance on other infectious diseases, the model performed worse on TB than on HIV-related questions. A similar study[21] on HIV found that 41.5% of answers achieved the highest score, compared to 34.4% in our study. Additionally, higher completeness was reported for HIV, with 76.1% of answers receiving a completeness score of 3 out of 3, whereas only 48.9% achieved this level in our study. However, ChatGPT suggested seeking medical advice in a higher percentage of answers compared to the study on HIV (64.4% *vs.* 39.2%).

On the contrary, ChatGPT performed better on TB questions compared to other bacterial infections, where only 25% of answers received a perfect score, and 50% achieved the highest completeness score[22].

There are multiple reasons why ChatGPT-4 performed worse on TB questions compared to other infectious diseases investigated in other studies. The present study included more complex and specialised questions, particularly at medium and high LOEs, which the model may have struggled with. In addition, the model's knowledge cutoff likely missed the most recent guidelines or nuanced clinical practices essential for specialised fields like drug-resistant TB.

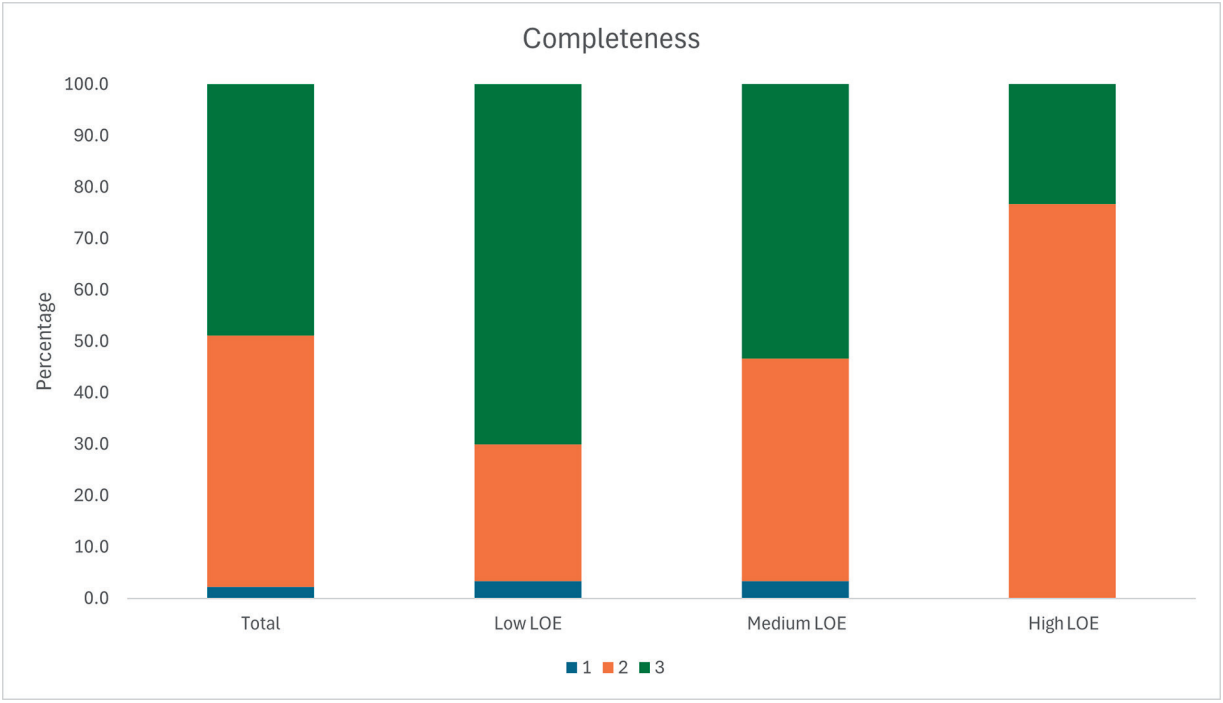


**Figure 2.** Completeness scoring of ChatGPT-4 in answering open-ended questions on tuberculosis, divided according to the Level of Expertise (LOE).

These comparisons suggest that ChatGPT-4's performance varies across different infectious diseases, potentially reflecting the complexity of the subject matter and the availability of relevant training data. Tuberculosis, particularly drug-resistant forms, involves complex diagnostic and treatment protocols that may not be as well-represented in the model's training data as other topics, like HIV.

One limitation of this study is the potential for evaluator bias. The assessment of accuracy and completeness was based on Likert scales, which involve subjective judgment. While the evaluators were medical professionals with varying levels of experience in TB, personal interpretations may have influenced scoring. Additionally, the study did not compare ChatGPT-4's performance with other information sources or previous versions of the model, which would have provided a more comprehensive understanding of its effectiveness.

## CONCLUSIONS

ChatGPT-4 delivers the best performance in providing comprehensive and accurate information about tuberculosis when asked general queries, although its ability declines as questions become more complex. It shows limitations in addressing specialised tuberculosis-related inquiries, especially when its performance is compared with that of other infectious diseases. While it can be useful for basic information, caution is needed when using it for advanced clinical decision-making. Ongoing updates and improvements are necessary to enhance its utility in specialised medical contexts.

**Fundings:**
No funding has been received for this manuscript.

**Authors' Contribution:**
Conceptualisation: ADV, AnCa; methodology: ADV, GS, AnCa; formal analysis: ADV, GS; investigation: AgCo, DB, PMMC, EF, SF, GiMa, TM, GiMi, CM, AP, NR, LS, FS, MT, GB; data curation: AgCo, DB, PMMC, EF, SF, GiMa, TM, GiMi, CM, AP, NR, LS, FS, MT, GB; writing – original draft preparation: ADV, AgCo, AnCa; writing – review and editing: DB, PMMC, EF, SF, GiMa, TM, GiMi, CM, AP, NR, LS, FS, MT, GS, GB; visualisation: ADV, AgCo; supervision: GB, AnCa.

**Conflict of Interest:**
The authors declare no conflicts of interest.

**Informed Consent:**
Informed consent was not required as no human participants were directly involved in this study.

**AI Disclosure:**
ChatGPT-4 (OpenAI) was used in this study to generate answers to tuberculosis-related questions, which were subsequently evaluated for accuracy and completeness by the research team. The model did not contribute to the writing or editing of the manuscript.

**Orcid ID:**
Andrea De Vito: 0000-0002-8265-5400
Agnese Colpani: 0000-0002-8959-1403
Danilo Buonsenso: 0000-0001-8567-2639
Elisabetta Falbo: 0009-0009-0680-3078
Giordano Madeddu: 0000-0001-6099-227
Tiziana Masini: 0000-0003-0400-1919
Caterina Monari: 0000-0003-2622-0573
Niccolò Riccardi: 0000-0002-2947-4707
Laura Saderi: 0000-0002-5131-9367
Francesca Saluzzo: 0000-0001-6705-1716
Giovanni Sotgiu: 0000-0002-1600-4474
Marina Tadolini: 0000-0002-1005-7532
Giorgio Besozzi: 0000-0002-2041-8493
Andrea Calcagno: 0000-0003-2534-8815

**Ethics Approval:**
The ethical review and approval requirement was waived because the study did not include any analysis of humans or animals.

**Data Availability:**
The datasets generated and/or analysed during the current study are available from the corresponding author upon reasonable request.

## References

1. World Health Organization. Global Tuberculosis Report 2023. WHO; 2023. Available at: https://www.who.int/publications/i/item/9789240083851
2. Cioboata R, Biciusca V, Olteanu M, Vasile CM. COVID-19 and Tuberculosis: Unveiling the Dual Threat and Shared Solutions Perspective. J Clin Med 2023; 12: 4784.
3. Kyu HH, Ledesma JR. What is the impact of the COVID-19 pandemic on tuberculosis? Lancet Glob Health 2023; 11: e1323-e1324.
4. Health - United Nations Sustainable Development - Goal 3: Ensure healthy lives and promote well-being for all at all ages. https://www.un.org/sustainabledevelopment/health/ (accessed October 31, 2024).
5. World Health Organization. The End TB Strategy. Available at: https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/the-end-tb-strategy
6. Qin ZZ, Ahmed S, Sarker MS, Paul K, Adel ASS, Naheyan T, Barrett R, Banu S, Creswell J. Tuberculosis detection from chest x-rays for triaging in a high tuberculosis-burden setting: an evaluation of five artificial intelligence algorithms. Lancet Digit Health 2021; 3: e543-e554.
7. Geric C, Qin ZZ, Denkinger CM, Kik SV, Marais B, Anjos A, David PM, Khan FA, Trajman A. The rise of artificial intelligence reading of chest X-rays for enhanced TB diagnosis and elimination. The Int J Tuberc Lung Dis 2023; 27: 367-372.

8. WHO consolidated guidelines on tuberculosis: Module 2: screening – systematic screening for tuberculosis disease [Internet]. Geneva: World Health Organization; 2021.
9. MATCH: Mapping and Analysis for Tailored disease Control and Health system strengthening. KIT Royal Tropical Institute 2019. Available at: https://www.kit.nl/institute/publication/match-mapping-and-analysis-for-tailored-disease-control-and-health-system-strengthening/
10. Alege A, Hashmi S, Eneogu R, Meurrens V, Budts AL, Pedro M, Daniel O, Idogho O, Ihesie A, Potgieter MG, Akaniro OC, Oyelaran O, Charles MO, Agbaje A. Effectiveness of Using AI-Driven Hotspot Mapping for Active Case Finding of Tuberculosis in Southwestern Nigeria. Trop Med Infect Dis 2024; 9: 99.
11. Link E, Baumann E. Use of health information on the internet: personal and motivational influencing factors. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2020; 63: 681-689.
12. Craciun OM, Torres MDR, Llanes AB, Romay-Barja M. Tuberculosis Knowledge, Attitudes, and Practice in Middle- and Low-Income Countries: A Systematic Review. J Trop Med 2023; 2023: 1014666.
13. Barańska A, Kłak A. Recent Trends in Health Literacy Research, Health Status of the Population and Disease Prevention: An Editorial. Int J Environ Res Public Health 2022; 19: 8436.
14. Chauhan A, Parmar M, Dash GC, Chauhan S, Sahoo KC, Samantaray K, Sharma J, Mahapatra P, Pati S. Health literacy and tuberculosis control: systematic review and meta-analysis. Bull World Health Organ 2024; 102: 421-431.
15. Tuberculosis - Google Trends n.d. https://trends.google.com/trends/explore?date=all&q=tuberculosis&hl=it (accessed October 31, 2024).
16. Dave T, Athaluri SA, Singh S. ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. Front Artif Intell 2023; 6: 1169595.
17. Nacher M, Françoise U, Adenis A. ChatGPT neglects a neglected disease. Lancet Infect Dis 2024; 24: e76.
18. Lechien JR, Naunheim MR, Maniaci A, Radulesco T, Saibene AM, Chiesa-Estomba CM, Vaira LA. Performance and Consistency of ChatGPT-4 Versus Otolaryngologists: A Clinical Case Series. Otolaryngol Head Neck Surg 2024; 170: 1519-1526.
19. Grünebaum A, Chervenak J, Pollet SL, Katz A, Chervenak FA. The exciting potential for ChatGPT in obstetrics and gynecology. Am J Obstet Gynecol 2023; 228: 696-705.
20. Cascella M, Montomoli J, Bellini V, Bignami E. Evaluating the Feasibility of ChatGPT in Healthcare: An Analysis of Multiple Clinical and Research Scenarios. J Med Syst 2023; 47: 33.
21. De Vito A, Colpani A, Moi G, Babudieri S, Calcagno A, Calvino V, Ceccarelli M, Colpani G, d'Ettorre G, Di Biagio A, Farinella M, Falaguasta M, Focà E, Giupponi G, Habed AJ, Isenia WJ, Lo Caputo S, Marchetti G, Modesti L, Mussini C, Nunnari G, Rusconi S, Russo D, Saracino A, Serra PA, Madeddu G. Assessing ChatGPT's Potential in HIV Prevention Communication: A Comprehensive Evaluation of Accuracy, Completeness, and Inclusivity. AIDS Behav 2024; 28: 2746-2754.
22. De Vito A, Geremia N, Marino A, Bavaro DF, Caruana G, Meschiari M, Colpani A, Mazzitelli M, Scaglione V, Venanzi Rullo E, Fiore V, Fois M, Campanella E, Pistarà E, Faltoni M, Nunnari G, Cattelan A, Mussini C, Bartoletti M, Vaira LA, Madeddu G. Assessing ChatGPT's theoretical knowledge and prescriptive accuracy in bacterial infections: a comparative study with infectious diseases residents and specialists. Infection 2025; 53: 873-881.
23. Sarink MJ, Bakker IL, Anas AA, Yusuf E. A study on the performance of ChatGPT in infectious diseases clinical consultation. Clin Microbiol Infect 2023; 29: 1088-1089.
24. Maillard A, Micheli G, Lefevre L, Guyonnet C, Poyart C, Canouï E, Belan M, Charlier C. Can Chatbot Artificial Intelligence Replace Infectious Diseases Physicians in the Management of Bloodstream Infections? A Prospective Cohort Study. Clin Infect Dis 2024; 78: 825832.
25. Rabaan AA, Bakhrebah MA, Alotaibi J, Natto ZS, Alkhaibari RS, Alawad E, Alshammari HM, Alwarthan S, Alhajri M, Almogbel MS, Aljohani MH, Alofi FS, Alharbi N, Al-Adsani W, Alsulaiman AM, Aldali J, Ibrahim F Al, Almaghrabi RS, Al-Omari A, Garout M. Unleashing the power of artificial intelligence for diagnosing and treating infectious diseases: A comprehensive review. J Infect Public Health 2023; 16: 1837-1847.
26. De Vito A, Geremia N, Bavaro DF, Seo SK, Laracy J, Mazzitelli M, Marino A, Maraolo AE, Russo A, Colpani A, Bartoletti M, Cattelan AM, Mussini C, Parisi SG, Vaira LA, Nunnari G, Madeddu G. Comparing large language models for antibiotic prescribing in different clinical scenarios: which performs better? Clin Microbiol Infect 2025; 31: 1336-1342.
27. Taye H, Alemu K, Mihret A, Wood JLN, Shkedy Z, Berg S, Aseffa A. Global prevalence of Mycobacterium bovis infections among human tuberculosis cases: Systematic review and meta-analysis. Zoonoses Public Health 2021; 68: 704-718.
28. Sharma A, Bloss E, Heilig CM, Click ES. Tuberculosis Caused by Mycobacterium africanum, United States, 2004-2013. Emerg Infect Dis 2016; 22: 396-403.
29. WHO consolidated guidelines on tuberculosis: Module 4: Treatment - Drug-susceptible tuberculosis treatment [Internet]. Geneva: World Health Organization; 2022.
30. Rotundo S, Tassone MT, Serapide F, Russo A, Trecarichi EM. Incipient tuberculosis: a comprehensive overview. Infection 2024; 52: 1215-1222.
31. Achkar JM, Jenny-Avital ER. Incipient and Subclinical Tuberculosis: Defining Early Disease States in the Context of Host Immune Response. J Infect Dis 2011; 204: S1179-S1186.
32. Tuncer G, Güçlü KG. How Reliable is ChatGPT as a Novel Consultant in Infectious Diseases and Clinical Microbiology? Infect Dis Clin Microbiol 2024; 6: 55-59.