U.PORTO

DOUTORAMENTO

CIÊNCIAS MÉDICAS

Tuberculosis in Portugal - evaluation of the surveillance system

Carlos Afonso de Carvalho

D

2024

Carlos Afonso de Carvalho.Tuberculosis in Portugal- evaluation of the surveillance system

D.ICBAS **2024**

Tuberculosis in Portugal- evaluation of the surveillance system

Carlos Filipe Afonso de Carvalho

Instituto de Ciências Biomédicas Abel Salazar

CARLOS FILIPE AFONSO DE CARVALHO

TUBERCULOSIS IN PORTUGAL
EVALUATION OF THE SURVEILLANCE SYSTEM

Tese de Candidatura ao grau de Doutor em
Ciências Médicas;
Programa Doutoral da Universidade do Porto (Instituto de Ciências Biomédicas de Abel Salazar)

Orientador:
Doutor António Guilherme de Almeida Gonçalves
Delegado de Saúde Coordenador
Unidade de Saúde Pública do ACeS Ave-Famalicão
Administração Regional de Saúde do Norte, I.P.

Co-orientadoras:
Doutora Raquel Duarte Bessa de Melo
Professora Associada
Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto

Dra. Ana Maria Azevedo e Vasconcelos Correia
Assistente Graduada Sénior de Saúde Pública
Instituto Nacional de Saúde Doutor Ricardo Jorge

ARTICLES PUBLISHED

This thesis includes results from the following published articles:

[1]**Carvalho C**, Alba S, Harris R, Abubakar I, Van Hest R, Correia AM, Gonçalves G, Duarte R. Completeness of TB notification in Portugal, 2015: an inventory and capture-recapture study. Int J Tuberc Lung Dis. 2020;24(11):1186–93.

Role/contribution: Conceived, designed, and coordinated the study; obtained access to country specific data sources and led data analysis and interpretation of results; led the manuscript writing and submission after approval from all co-authors.

[2]Straetemans M, Bakker MI, Alba S, Mergenthaler C, Rood E, Andersen PH, Schimmel H, Simunovic A, Svetina P, **Carvalho C**, Lyytikäinen O, Abubakar I, Harris RJ, Ködmön C, van der Werf MJ, van Hest R. Completeness of tuberculosis (TB) notification: Inventory studies and capture-recapture analyses, six European Union countries, 2014 to 2016. Eurosurveill. 2020;25(12):1–10.

Role/contribution: Developed country specific study protocol; facilitated access to the TB data sources; contributed to data analysis and interpretation of the study results; manuscript drafting and approval of the final manuscript.

[3]Rito T, Matos C, **Carvalho C**, Machado H, Rodrigues G, Oliveira O, et al. A complex scenario of tuberculosis transmission is revealed through genetic and epidemiological surveys in Porto. BMC Infect Dis. 2018;18(1):53.

Role/contribution: Collected and validated epidemiological and georeferencing data; reviewed and approved the final manuscript.

[4]Oliveira O, Gaio R, **Carvalho C**, Correia-Neves M, Duarte R, Rito T. A nationwide study of multidrug-resistant tuberculosis in Portugal 2014–2017 using epidemiological and molecular clustering analyses. BMC Infect Dis. 2019; 19:567.

Role/contribution: Contextualised the findings within a public health perspective and handled the database; critically reviewed the manuscript.

[5]Sentís A, Vasconcelos P, Machado RS, Caylà JA, Guxens M, Peixoto V, Duarte R, Carvalho I, **Carvalho C**. Failure to complete treatment for latent tuberculosis infection in Portugal, 2013–2017: geographic-, sociodemographic-, and medical-associated factors. Eur J Clin Microbiol Infect Dis. 2020;39:647–656.

Role/contribution: Corresponding author; co-designed the study; interpretation of results; critically reviewed the final manuscript.

[6]Sousa S, Rocha D, Silva JC, Ribeiro AI, Gonçalves G, Almeida Á, Correia AM, Duarte R, **Carvalho C**. Comparing the cost-effectiveness of two screening strategies for latent tuberculosis infection in Portugal. Pulmonology 2021;27(6):493–499.

Role/contribution: Co-designed the study; interpretation of results; critically reviewed the final manuscript.

[7]Sentís A, Prats-Uribe A, Peixoto VR, Caylà JA, Gomes MD, Sousa S, Duarte R, Carvalho I, **Carvalho C**. Decline of tuberculosis notification rate in different populations and regions in Portugal, 2010–2017. Pulmonology 2021 (published online ahead of print).

Role/contribution: Corresponding author; co-designed the study; interpretation of results; critically reviewed the final manuscript.

[8]Sousa S, Macedo R, Alves CM, **Carvalho C**, Goncalves G, Duarte R. Coffee shops, a hub for TB clusters? Pulmonology 2023 (published online ahead of print).

Role/contribution: Co-designed the study; interpretation of results; critically reviewed the final manuscript.

[9]Sousa S, Alves CM, Macedo R, **Carvalho C**, Goncalves G, Duarte R. An investigation of TB infection and reinfection among stone quarry workers. Pulmonology 2023 (published online ahead of print).

Role/contribution: Co-designed the study; interpretation of results; critically reviewed the final manuscript.

Disclosure:

Articles [3-9] have been or will be included in other theses, by the respective principal investigators. Those articles were included in this thesis as they are relevant in demonstrating the usefulness of the TB surveillance system.

To my mother

To my father

To my sister and to my brother

To my wife and to my kids

ACKNOWLEDGMENTS

This work would not have been possible without the support from many people to whom I would like to express my gratitude:

To my supervisor Guilherme Gonçalves, for his time, constant encouragement and unvaluable advice.

To my co-supervisor Raquel Duarte, for being permanently available to share her TB knowledge and experience, from programme management to field work, and for making me feel part of her team since the day we met.

To my co-supervisor Ana Maria Correia, for introducing me to “the TB world”, for trusting me, and for being such a role model to all Public Health professionals.

To all co-authors of the articles presented in this thesis, for their contributions and collaboration.

To all my colleagues and friends, who in any way shared or crossed my path and from whom I have learned so much.

To my family, who has always been there for me.

TABLE OF CONTENTS

ABSTRACT

Tuberculosis in Portugal – evaluation of the surveillance system

Tuberculosis (TB) is still an important public health problem worldwide, although steadily declining in incidence over the last decades. Portugal was the last West-European country to become a TB "low-incidence" country (under 20 cases per 100,000 population per year). In these circumstances, a well-implemented surveillance system is a key part of the strategy to prevent and control tuberculosis through swift and efficient public health measures.

Through the ongoing and systematic collection, analysis, interpretation and dissemination of health information, public health surveillance plays a critical role in informed decision-making and a real evidence-based public health action. Public health problems must be active and efficiently assessed and monitored, ensuring surveillance tasks support intervention that reduces morbidity and mortality, ultimately improving the population's health. To achieve this, surveillance systems must be routinely evaluated and improved.

When evaluating a public health surveillance system, different aspects must be considered to achieve the main goals of maximizing the system's utility and efficiency. Those aspects include the public health importance of the event under surveillance, the objectives, operation, and usefulness of the system, and qualitative attributes (simplicity, flexibility, acceptability), quantitative attributes (sensitivity, predictive positive value, representativeness, timeliness), and cost. Specifically designed studies are required in order to evaluate those attributes and generate the evidence that will be needed to support intervention and ultimately improve health.

The tuberculosis surveillance system in Portugal, designed and operated by the Directorate-General of Health (*Direção-Geral da Saúde – DGS*), has never been thoroughly evaluated. The National Tuberculosis Programme (*Programa Nacional para a Tuberculose – PNT*) routinely produces regular reports that provide support for public health action, but at this moment in which Portugal became a "Tuberculosis low-incidence country", an external in-

depth evaluation of the surveillance system is important to estimate the number of tuberculosis cases that are missed and, as importantly, that the collected data is promptly made available and analysed, triggering effective and timely public health intervention and supporting clinical management.

Given the complexity of performing an evaluation covering all surveillance systems' attributes, I prioritised and structured my thesis according to three of the most important: sensitivity (the ability of a system to detect the events under surveillance, also designated as completeness of notification or exhaustiveness), usefulness (the degree at which the surveillance and other public health objectives are met or supported by the surveillance system) and timeliness (the delay between the onset of a given disease and generation of evidence that will support intervention to minimise the impact of this disease in the community). These three dimensions of the tuberculosis surveillance system in Portugal were evaluated and constitute the chapter thread for this document. Nine articles, published in international peer-reviewed journals indexed to ISI Web of Knowledge (from which I was first author in one, last author in three, and co-author in five), and a research protocol with a positive review from the Ethical Board of the School of Medicine and Biomedical Sciences of the Porto University (*Instituto de Ciências Biomédicas de Abel Salazar - ICBAS-UP*), illustrate and support this thesis' conclusions.

In Chapter 2, I present two articles on sensitivity of the National Tuberculosis Programme (NTP) surveillance system (SVIG-TB), documenting and describing the current components and data sources but also including making recommendations for their integration in tuberculosis surveillance. Taking advantage of the availability of different national data sources for tuberculosis in Portugal, we used a capture-recapture methodology to estimate that completeness of reporting to SVIG-TB in 2015 might have been as low as 77%. These results have to be interpreted with caution, given the limitations of the method. Integration of different surveillance systems does not mean that one should replace the other. Although they might contribute to common surveillance objectives, such as estimating incidence of disease and identifying trends, it is clearly that some systems should be more oriented towards clinical aspects of the disease and others towards the immediate public health intervention aspects. For the first, which is illustrated by the NTP intrinsic surveillance system (SVIG-TB), specific objectives entail identification of individual risk factors and evaluation and improvement of clinical guidelines that ensure better treatment and outcomes. For the latter, which in Portugal is represented by the National Notifiable Diseases' surveillance system (SINAVE), timeliness is of paramount importance and a

special focus should be placed on the risk that cases pose to the rest of the (theoretically) healthy community.

In Chapter 3, I present seven articles that illustrate the usefulness of the tuberculosis surveillance system in Portugal, taking into account the objectives of the National Tuberculosis Programme (1) Monitoring of TB notification rates in the country, supporting regular production of regional, national, and international TB epidemiological reports; 2) Analysis of clinical data and production and dissemination of evidence in the areas of prevention, diagnosis and treatment of TB; 3) Supporting implementation and monitor TB-specific strategies, such as the Directly Observed Therapy Short-course (DOTS), drug-resistant TB prevention and control, roll-out and evaluation of newer diagnosis tests, molecular surveillance and monitoring of implementation of national guidelines on TB), but also international orientations such as the World Health Organization's End TB Strategy pillars and components. Through those seven articles, we demonstrated how TB surveillance data can be used to produce a better understanding of the epidemiology of active tuberculosis disease and latent tuberculosis infections, including specific aspects such as trends in notification rates, molecular characterisation in the context of outbreaks, drug-resistant TB features, epidemiology of TB and screening of latent tuberculosis infection (LTBI) in especially vulnerable populations such as the stone quarry workers, and factors associated with failure to complete treatment for LTBI. All the articles relied on data from the NTP surveillance system (SVIG-TB), exclusively or complemented by other public health service- or laboratory-based sources, reinforcing the need for data integration and the importance of case-base data to address surveillance objectives.

In Chapter 4, I present a research protocol for the study of timeliness of the tuberculosis surveillance system in Portugal, defined as the intervals between the onset of symptoms in a TB patient to the dates of first medical consultation and diagnosis, and then to the date at which a notification was accessible by the National Tuberculosis Programme. Although there were previous studies investigating delays to diagnosis and start of treatment, in both the component attributed to the patients and to healthcare services, this study will also provide an overview of the subsequent delay to notification, which reflects better the NTP and public health services capacity to monitor trends and detect outbreaks, as well as support timely implementation and evaluation of control measures.

The availability of independent systems or data sources enables a high-quality evaluation the disease surveillance system sensitivity, supporting a national strategy that involves multiple epidemiological surveillance systems for the same disease – provided that

redundancy is avoided and there is sufficient integration that allows program managers and decision makers to monitor the epidemiological situation and intervene, avoiding the exhaustion and distrust from the notifiers. Additional case-based data sources or surveillance systems, such as laboratory- or hospital-based, are of paramount importance and should be integrated in any tuberculosis surveillance system, through a process of data linkage allowing that data collected one single time can be incorporated in different systems with minimal effort. This entails complex data-protection issues, which should not be an obstacle or delay integration by the public health services, but on the other hand should be respected to ensure that individuals' rights and confidentiality are not compromised. The integration and the availability of improved data standards and electronic exchange of health data will support a more efficient response of public health services and the NTP to the threat that tuberculosis still represents in Portugal.

Evaluation is a continuous and cyclical process, that starts from an assessment and re-starts after the implementation of changes aiming at improving the system. Therefore, this thesis is not an end-point - just a first step in the process of improving the quality and efficiency of the tuberculosis surveillance system in Portugal that should be followed up and repeated regularly. As for the surveillance system itself, I hope that this thesis might be useful in the current context of digitalisation and integration of surveillance systems, specifically those for infectious diseases and other threats to public health.

RESUMO

Tuberculose em Portugal - avaliação do sistema de vigilância

A tuberculose (TB) mantém-se como um importante problema de saúde pública em todo o mundo, apesar da gradual redução da incidência nas últimas décadas. Portugal foi o último país da Europa Ocidental a tornar-se um país de "baixa incidência" de TB (menos de 20 casos por 100,000 habitantes por ano). Nestas circunstâncias, um sistema de vigilância robusto é uma parte fundamental da estratégia de prevenção e controlo da tuberculose através de medidas de saúde pública rápidas e eficientes.

Através da colheita, análise, interpretação e disseminação contínua e sistemática de informações de saúde, a vigilância em saúde pública desempenha um papel crítico na tomada de decisões informadas e na intervenção baseada em evidência. Os problemas de saúde pública devem ser monitorizados ativa e eficientemente, garantindo que as tarefas de vigilância apoiam a intervenção, reduzindo a morbidade e mortalidade e em última instância levam à melhoria do estado de saúde da população. Para alcançar isso, os sistemas de vigilância devem ser rotineiramente avaliados e melhorados.

Diferentes aspetos devem ser considerados na avaliação de um sistema de vigilância em saúde pública, no sentido de se maximizar a utilidade e eficiência do sistema. Esses aspetos incluem a importância para a saúde pública do evento sob vigilância, os objetivos, operação e utilidade do sistema, atributos qualitativos (simplicidade, flexibilidade, aceitabilidade), atributos quantitativos (sensibilidade, valor preditivo positivo, representatividade, rapidez), e custo. Para avaliar estes atributos e gerar as evidências necessárias para apoiar a intervenção em saúde pública e melhorar a saúde da população, são necessários estudos desenhados especificamente para o efeito.

O sistema de vigilância da tuberculose em Portugal, projetado e operado pela Direção-Geral da Saúde (DGS), nunca foi formalmente avaliado. O Programa Nacional para a Tuberculose (PNT) produz rotineiramente relatórios regulares que suportam a intervenção

em saúde pública, mas no atual contexto em que Portugal atingiu o estatuto de “baixa incidência” de tuberculose é importante levar a cabo uma avaliação externa do sistema de vigilância para estimar o número de casos perdidos e, de forma igualmente importante, garantir que os dados recolhidos são prontamente disponibilizados e analisados, suportando a intervenção eficaz e atempada por parte da saúde pública e apoiando a gestão clínica dos doentes com tuberculose.

Dada a complexidade de realizar uma avaliação que abrangesse todos os atributos do sistemas de vigilância, priorizei e estruturei minha tese em três dos mais importantes: sensibilidade (a capacidade de um sistema detetar os eventos sob vigilância, também designada como exaustividade), utilidade (a medida em que os objetivos da vigilância e outros objetivos em saúde pública são cumpridos ou apoiados pelo sistema) e rapidez (o atraso entre o início da doença e a produção da evidência que irá sustentar a intervenção para minimizar o impacto da doença na comunidade). Estas três dimensões do sistema de vigilância da tuberculose em Portugal foram avaliadas e constituem o fio condutor ao longo do qual se estruturam os capítulos deste documento. Nove artigos, publicados em revistas internacionais revisadas por pares indexadas (dos quais fui primeiro autor em um, último autor em três e coautor em cinco), e um protocolo de investigação com parecer positivo da Comissão de Ética do Instituto de Ciências Biomédicas de Abel Salazar (ICBAS-UP), ilustram e apoiam as conclusões desta tese nos capítulos respetivos.

No Capítulo 2 apresento dois artigos sobre a sensibilidade do sistema de vigilância intrínseco do PNT (SVIG-TB), documentando e descrevendo os seus componentes e fontes de informação, mas também incluindo recomendações para sua integração na vigilância da tuberculose. Aproveitando a disponibilidade de diferentes fontes de informação para tuberculose em Portugal, usámos uma metodologia de captura-recaptura para estimar que a sensibilidade do sistema (proporção de casos de tuberculose notificados ao SVIG-TB) em 2015 poderá ter sido tão baixa quanto 77%. A integração de diferentes sistemas de vigilância não significa que um deva anular o outro. Embora possam contribuir para objetivos comuns de vigilância, como estimar a incidência da doença e identificar tendências, é claro que um sistema com maior latência deve ser mais orientado para os aspetos clínicos da doença e outro para os aspetos imediatos da intervenção em saúde pública. Para o primeiro, ilustrado pelo SVIG-TB, objetivos específicos incluem a identificação de fatores de risco individuais e avaliação e melhoria das orientações clínicas que garantem melhores tratamento e resultados. Para o último, que em Portugal é representado pelo sistema nacional de vigilância epidemiológica (SINAVE), a rapidez é de

suma importância e o foco deve ser colocado no risco que os casos representam para o resto da comunidade, (teoricamente) saudável.

No Capítulo 3 apresento sete artigos que ilustram a utilidade do sistema de vigilância da tuberculose em Portugal, levando em consideração os objetivos do Programa Nacional para a Tuberculose (1) Monitorização das taxas de notificação de TB no país, apoiando a produção regular de relatórios epidemiológicos de tuberculose regionais, nacionais, e internacionais; 2) Análise de dados clínicos e produção e disseminação de evidência nas áreas de prevenção, diagnóstico e tratamento da TB; 3) Apoio à implementação e monitorização estratégias específicas para TB, como o Tratamento Observado Diretamente (TOD), prevenção e controlo da TB resistente a antibacilares, implementação e avaliação de novos testes de diagnóstico, vigilância molecular e monitorização das orientações nacionais de TB), mas também orientações internacionais como os pilares e componentes da Estratégia *End TB* da Organização Mundial da Saúde. Através desses sete artigos, demonstra-se como os dados de vigilância da TB são fundamentais para compreender a epidemiologia da tuberculose doença e infeção tuberculosa latente, incluindo aspetos específicos como evolução das taxas de notificação, caracterização molecular no contexto de surtos, características da TB resistente a medicamentos, rastreio de tuberculose latente em populações especialmente vulneráveis como os trabalhadores de pedreiras, e fatores associados ao abandono terapêutico. Todos os artigos foram suportados por dados recolhidos a partir do sistema de vigilância do PNT (SVIG-TB), de forma exclusiva ou complementada por outras fontes de informação da saúde pública, hospitalares ou laboratoriais, reforçando a necessidade de integração de sistemas e a importância dos dados de saúde individuais (e não agregados) para se atingirem os objetivos de vigilância.

O Capítulo 4 inclui um protocolo de investigação para o estudo da “rapidez” (*timeliness*) do sistema de vigilância da tuberculose em Portugal, definida como o conjunto de intervalos entre as datas de início dos sintomas, primeira consulta médica, diagnóstico de tuberculose, notificação para o SVIG e disponibilização ao Programa Nacional para a Tuberculose. Embora tenha havido estudos anteriores sobre os atrasos no diagnóstico e início de tratamento, tanto na componente atribuída aos pacientes quanto aos serviços de saúde, este estudo fornecerá também uma perspetiva sobre o atraso subsequente até à notificação, o que que reflete melhor a capacidade do PNT e dos serviços de saúde pública para monitorizar tendências e detetar surtos precocemente, bem como apoiar a implementação atempada de medidas de controle e fazer a sua avaliação.

A disponibilidade de sistemas ou fontes de dados independentes permite a avaliação da sensibilidade do sistema de vigilância, apoiando uma estratégia nacional de manutenção de sistemas em paralelo para a mesma doença – desde que a redundância seja evitada e haja integração suficiente que permita aos gestores de programas e decisores monitorizar a situação epidemiológica e intervir, sem desgaste excessivo e desconfiança por parte dos notificadores. Fontes de dados ou subsistemas de vigilância adicionais de base individual, como laboratoriais ou hospitalares, são de suma importância e devem ser integrados em qualquer sistema de vigilância da tuberculose, utilizando para isso técnicas de ligação de registos que permitam que os dados recolhidos num determinado contexto possam ser incorporados em diferentes sistemas com mínimo esforço. Este processo implica questões complexas de proteção de dados, que não devem ser um obstáculo ou atrasar a integração pelos serviços públicos de saúde, mas que devem ser respeitadas para garantir os direitos e confidencialidade dos indivíduos. A integração e o desenvolvimento de padrões e a troca eletrónica de dados de saúde irão permitir uma vigilância epidemiológica mais eficiente, e consequentemente uma resposta mais eficiente por parte dos serviços saúde pública e do PNT à ameaça que a tuberculose ainda representa em Portugal.

A avaliação é um processo contínuo e cíclico, que começa com uma avaliação e recomeça após a implementação de mudanças visando melhorar o sistema. Portanto, esta tese não é um ponto final – é antes um primeiro passo no processo de melhoria da qualidade e eficiência do sistema de vigilância da tuberculose em Portugal, que deve ser repetido regularmente. Quanto ao próprio sistema de vigilância, espero que esta tese possa ser útil no contexto atual de digitalização e integração dos sistemas de vigilância, especificamente aqueles para doenças infeciosas e outras ameaças à saúde pública.

ACRONYMS AND ABBREVIATIONS

AIC – Akaike Information Criteria

ARSN – *Administração Regional de Saúde do Norte* (Northern Regional Health Administration)

CDP – *Centro de Diagnóstico Pneumológico* (Tuberculosis outpatient centre)

CI – Confidence Interval

CRC – Capture-Recapture

Df – Degrees of freedom

DGS – *Direção-Geral da Saúde* (Directorate-General of Health Portugal)

DOT – Directly Observed Therapy

DST – Drug susceptibility testing

ECDC – European Centre for Disease Prevention and Control

EU/EEA – European Union/European Economic Area

GDH – *Grupos de Diagnósticos Homogéneos* (Hospital discharge classification system)

HIV – Human Immunodeficiency Virus

ICBAS-UP – *Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto* (School of Medicine and Biomedical Sciences, University of Porto)

IDU – Injecting Drug Users

ICER – Incremental Cost-Effectiveness Ratio

ICU – Intensive Care Unit

IGRA – Interferon-Gamma Release Assays

IQR – Inter-Quartile Range

IS – Inventory Studies

LTBI – Latent Tuberculosis Infection

LTV – Lisbon and Tagus Valley Health Region

MDR-TB – Multidrug-Resistant Tuberculosis

MIRU – Mycobacterial Interspersed Repetitive Units

MTC – *Mycobaterium tuberculosis* Complex

NTM – Non-Tuberculous Mycobacteria

NTP – National Tuberculosis Programme

OR – Odds Ratio

RR-TB – Rifampicin-resistant Tuberculosis

SDG – Sustainable Development Goals

SINAVE – Sistema Nacional de Vigilância Epidemiológica

SVIG-TB – *Sistema de vigilância intrínseco do Programa Nacional para a Tuberculose* (National Tuberculosis Programme surveillance system)

TB – Tuberculosis

TST – Tuberculin Skin Test

TESSy – The European Surveillance System

VNTR – Variable Number of Tandem Repeats

WHO – World Health Organization

1

INTRODUCTION

Tuberculosis (TB) has been an important cause of death and disability throughout history, and it remains an important public health problem around the world. After being declared a global emergency in April 1993,[10] significant resources have been allocated and efforts have been made to control tuberculosis worldwide. In 2014, the World Health Assembly adopted the “End TB strategy”, aiming at reducing TB incidence rate up to 90% and death up to 95% by 2035 as part of the Sustainable Development Goals.[11]

Despite the progress made, the situation is still of concern and requires intensified public health efforts by public health to decrease TB incidence and deaths and increase the number of persons receiving TB curative and preventive treatment.[12] The End TB strategy builds on three strategic pillars: 1) Integrated, patient-centred care and prevention (early diagnosis of tuberculosis including universal drug-susceptibility testing, and systematic screening of contacts and high-risk groups; treatment of all people with tuberculosis including drug-resistant tuberculosis, and patient support; collaborative tuberculosis/HIV activities, and management of co-morbidities; preventive treatment of persons at high risk, and vaccination against tuberculosis); 2) bold policies and supportive systems (political commitment with adequate resources for tuberculosis care and prevention; engagement of communities, civil society organizations, and public and private care providers; universal health coverage policy, and regulatory frameworks for case notification, vital registration, quality and rational use of medicines, and infection control; social protection, poverty alleviation and actions on other determinants of tuberculosis); and 3) Intensified research and innovation (through discovery, development and rapid uptake of new tools, interventions and strategies; and research to optimize implementation and impact, and promote innovations).[11]

1.1. EPIDEMIOLOGY OF TUBERCULOSIS

Tuberculosis is present in all countries and age groups, but the magnitude of the epidemic is geographically very heterogeneous, meaning that different goals and strategies are needed according to the context. [11] According to the World Health Organization (WHO) estimates, 10.6 million people developed TB and 1.6 million died from the disease in 2021.[13]

The WHO regions of South-East Asia and Africa are especially affected, accounting for 45% and 23% of all newly diagnosed cases of tuberculosis worldwide, corresponding to estimated incidence rates of 234 and 212 cases per 100,000, respectively, in 2021.[13] In

Africa, where the highest regional prevalence of HIV infection is recorded, coinfection with HIV is a significant factor in the TB epidemic (21% of patients with known HIV status are HIV-positive, and represent 27% of TB deaths).[13]

In the WHO European region, where the estimated incidence rates are much lower (25 cases per 100,000 in 2021), there is an increasing concern with resistance to antimicrobials – 26% of new cases in 2021 were resistant to either rifampicin (RR) or rifampicin + isoniazid (MDR), vs. 2.8% to 5.3% in the other WHO regions.[13] The antimicrobial resistance phenomenon is asymmetric in Europe, affecting predominantly the eastern part of the continent (27.4% in non-EU/EEA countries vs. 3.3% in EU/EEA, in 2021).[14]

The COVID-19 pandemic has had a significant impact on the provision of and access to essential tuberculosis (TB) services, the number of people diagnosed with TB, and TB disease burden (incidence and mortality).[15] According to recent reports, twelve months of COVID-19 had reversed years of global progress against tuberculosis, resulting in a decrease in diagnosis, treatment, and notification of TB around the world.[16,17]

The COVID-19 pandemic had a great impact in the number of diagnosed and reported cases of TB, falling from 7.1 million in 2019 to 5.8 million in 2020 (far short from the WHO estimates of 10 million diagnoses in 2020).[15] It remains to be seen if the reduction in number of cases reflects an actual decline in new infections, underdiagnosis, underreporting, or a combination of all these factors.[18,19]

Also in Europe, the progress made in reducing the burden of TB between 2015 and 2019 was hindered in 2020 and 2021 by the COVID-19 pandemic. The downward trend in TB mortality stopped and incidence reversed, and fewer people were tested for drug resistance and enrolled in treatment for drug-resistant TB.[14] This was due to delays or lack of TB diagnosis caused by disruptions to TB services, leading to increased severity of disease and an associated increase in deaths. The rate of successful treatment outcomes among new and relapse TB cases was lower in 2021 when compared to previous years, indicating that countries are facing increasing challenges in providing appropriate care to TB patients, and remains below regional and global targets.[14] The Tuberculosis Action Plan for the WHO European Region 2016–2020 ended in 2020 and its progress was assessed,[20] leading to the development of the new Tuberculosis Action Plan for the WHO European Region 2023–2030.[21] The updated plan is aligned with the Sustainable Development Goals (SDGs) and End TB Strategy, aiming at “ending the spread of drug-susceptible and drug-resistant TB

by achieving universal access to prevention, diagnosis and treatment in all Member States of the Region".[22]

Despite the steady decline of the incidence in Portugal in the last decades (averaging 4.3%/year from 2000 to 2019), Portugal was the last West-European country to become a TB low-incidence country (under 20 cases per 100,000 population per year).[14,23,24] If this decline was to remain constant, one could estimate that in 2035 the Portuguese TB incidence rate would be reduced by 59% when compared to 2015 (Figure 1), well above the target of a 90% reduction set at the End-TB strategy.[11]

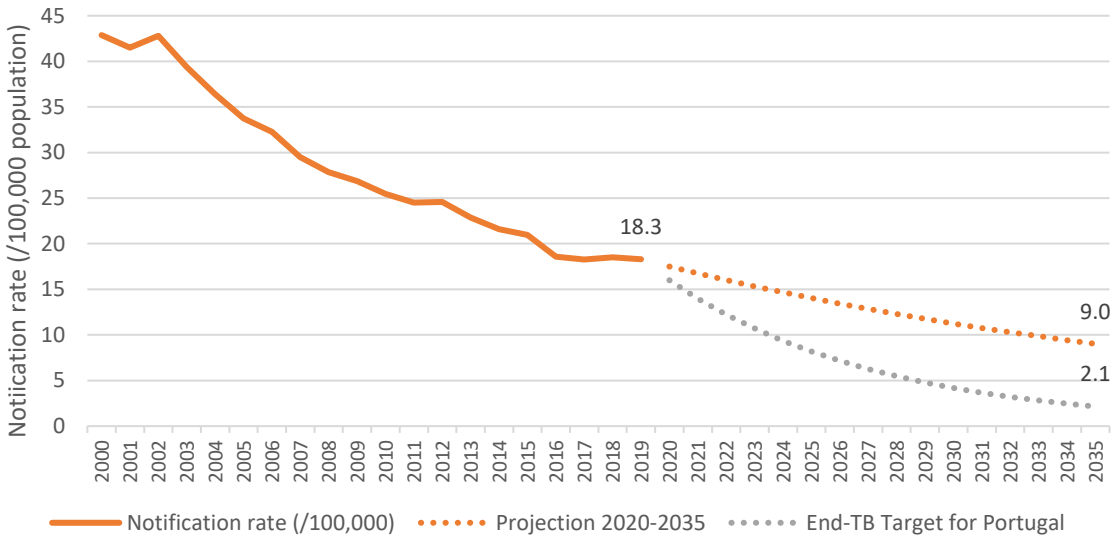


Figure 1. Reported and projected TB notification rates in Portugal, 2000-2035. Projection 2020-2035 calculated by applying the average 2000-2019 annual reduction (4.3%) to each subsequent year. The annual decline needed to achieve the End-TB strategy goal of 90% incidence reduction by 2035 would be 13%.

1.2. TUBERCULOSIS SURVEILLANCE SYSTEMS

Public health surveillance, through the ongoing and systematic collection, analysis, interpretation and dissemination of health information, plays a critical role in informed decision-making and a real evidence-based public health action.[25] In order to assist countries in implementing national TB surveillance systems, the World Health Organization (WHO) published in 1994 the WHO Framework for Effective Tuberculosis Control.[26] Surveillance was one of the five core components of the "TB control policy package": Countries should establish and maintain a monitoring system to be used both for programme supervision and evaluation, based on recording individual patient information in registers at district/county level and on regularly reporting from the same level.[26] This guidance underwent significant revisions in 2006[27] and 2013,[27] but a new WHO TB surveillance guide was recently published encompassing the entire screening, diagnosis, treatment, and care pathway for individuals with TB infection or disease and aiming at facilitating the implementation of digital, case-based, real-time TB surveillance systems.[28]

The new Tuberculosis Action Plan for the WHO European Region 2023–2030 also refers to the investment in surveillance systems, and the need to invest in real-time strategic information and digital health, as a priority action.[21] The most important aspect of a TB surveillance system should is that it serves a purpose, i.e. there are clear surveillance objectives and the data collected is kept to a minimum – while allowing the fulfilment of these objectives. The principles of personal data handling in terms of data minimization and data protection require that personal data collected and processed for public health surveillance of infectious diseases should be limited to what is necessary and relevant for the specific purpose. This means that only the minimum amount of personal data necessary to achieve the public health objective should be collected, and that data should be kept accurate, up-to-date, and secure.[29] Data protection measures should be implemented to safeguard personal data while ensuring that the system can still effectively fulfil its public health objectives, including measures such as anonymizing or pseudonymizing personal data, implementing strict access controls, and ensuring transparency and accountability in data processing, but without hampering the ability of those who need the information to take decisions.

For many infectious diseases, notification represents only the "tip of the iceberg", with "the iceberg" being the total number of people infected or ill (Figure 2).

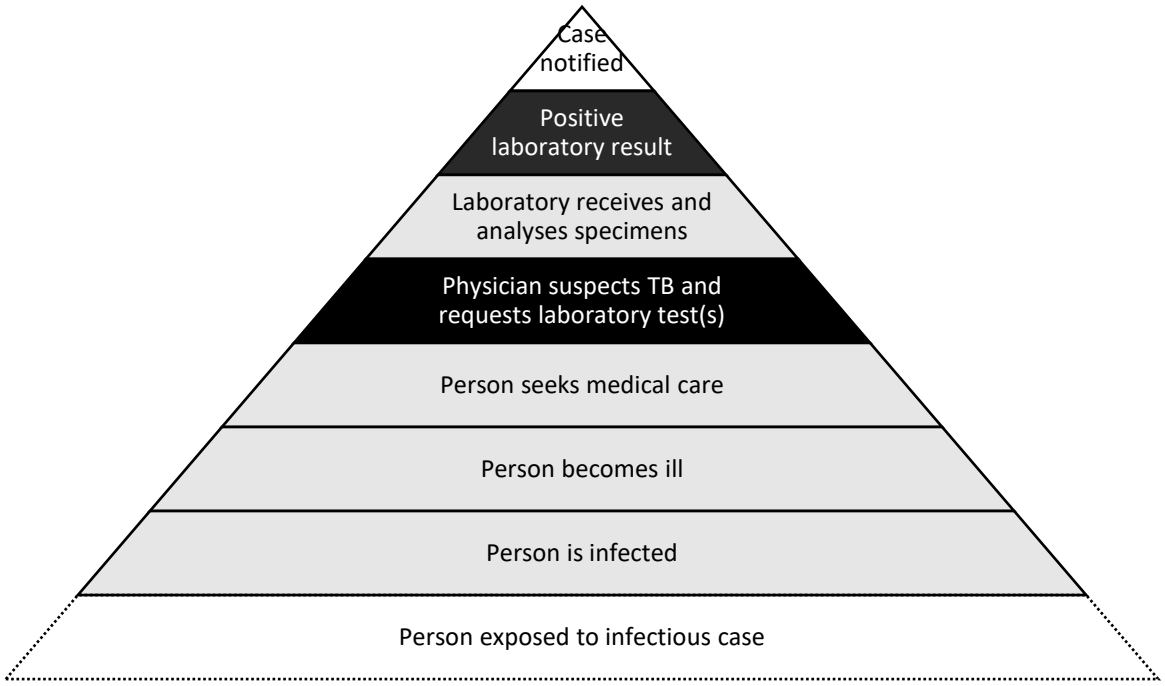


Figure 2. Surveillance pyramid for tuberculosis. Layers with black background refer to the "entry points" for the traditional surveillance systems based on notifications (clinical and laboratory)

In Portugal, surveillance of tuberculosis is supported by two main national notification systems: 1) a compulsory-notification one, designed to promptly inform public health authorities and trigger intervention; and 2) a clinical one, designed to keep record of diagnosis and follow-up of every patient. Compulsory notification, traditionally made using paper forms and including only cases of respiratory or disseminated TB, underwent significant changes with the introduction of an online notification platform, SINAVE (*Sistema Nacional de Vigilância Epidemiológica*) in June 2014.[30] Notification became an exclusively online-submission process, and all forms of active TB disease were then included. SINAVE is designed to quickly inform public health authorities and support intervention, covering the diseases under epidemiological surveillance according to the national and international legislation.[31,32] Laboratory notifications, clinical notifications, and epidemiological investigation reports are electronically submitted by laboratories, clinicians, and public health professionals, respectively, to SINAVE as mandated by law.

The second system is a clinical system, built in 1995 for the National TB Program, designed to maintain records of TB diagnoses and follow-up for each patient - SVIG-TB (*Sistema de Vigilância da Tuberculose*).[33] This system is intended for use by clinicians who are treating patients with active or latent tuberculosis and non-tuberculous mycobacterial infections in

specialized TB centres (*Centros de Diagnóstico Pneumológico* – CDPs). SVIG-TB has been a key component of the NTP, allowing for the monitoring of key indicators and targets at the local, regional, and national levels,[34,35] as well as the international reporting of data to the European Surveillance System – TESSy.[36]

For each TB patient (and also for treated cases of latent TB and non-tuberculous mycobacterial infections), a physician records clinical information in specific paper forms: F1 and F2,[33] that are subsequently computerised and stored at the regional and national levels in an anonymised way. Unlike SINAVE, the initial notification (F1 form) is progressively complemented with clinical data that comes available during treatment (F2 forms, which are filled whenever there is an update of the therapeutic regimen and upon completion of treatment). Given that the duration of treatment varies and can go up to 24 months (although usual treatment duration is 6 months), it is very important to ensure that the information system is set up in a "longitudinal way", i.e. all clinical information is stored in the patient's records – including the outcome, which is an important indicator to evaluate the performance of the NTP and status of TB control in the country.

In addition to active tuberculosis (clinical disease as a result of infection by bacteria from the *Mycobacterium tuberculosis* complex), Non-tuberculous Mycobacteria (NTM) Infections and cases of latent tuberculosis (LTBI) are also notified to SVIG-TB – as these cases are also treated at the TB-outpatient centres. Contacts of infectious cases of TB are also referred to these centres, where they are screened and followed up as appropriate – although these data are stored locally (the number of individuals screened for each case of active TB is reported in form F2, but no further information about each of them). As of 2022, there were approximately 30 offices for digitisation of the SVIG-TB paper forms in Portugal, being some of those responsible for multiple TB-outpatient centres in the respective region. These "digitisation offices" export the clinical data monthly, and from that point all the data becomes available for processing and analysis at the regional and national levels.

There are three main objectives associated with SVIG-TB: 1) Monitoring of TB notification rates in the country, supporting regular production of regional and national TB epidemiological reports; 2) Analysis of clinical data and production and dissemination of evidence in the areas of prevention, diagnosis and treatment of TB; 3) Supporting implementation and monitor TB-specific strategies, such as the Directly Observed Therapy Short-course (DOTS),[37] drug-resistant TB prevention and control, roll-out and evaluation of newer diagnosis tests, molecular surveillance and monitoring of implementation of national guidelines on TB.[33]

Despite recent progress on digitalisation of healthcare, the TB surveillance system in Portugal is still heavily dependent on active notification from physicians. This is done through the two pathways described above, with an associated risk of redundancy, but on the other hand enabling cross validation, evaluation, and integration of the different surveillance systems.

The Portuguese Directorate-General Health (DGS) routinely produces reports that provide support for public health action, but at this crucial moment in which Portugal is becoming a "TB low-incidence country", an external in-depth evaluation of the surveillance system is essential to ensure that no cases of TB are missed and, most importantly, that the information collected is triggering effective and timely intervention.[38]

1.3. EVALUATION OF THE TUBERCULOSIS SURVEILLANCE SYSTEM

In order to ensure that public health problems are efficiently and effectively monitored, and that information leads to an intervention that reduces morbidity and mortality, ultimately improving the population's health, surveillance systems must be routinely evaluated.[39] This will allow or improve the integration of surveillance and health information systems, the establishment of data standards and the electronic exchange of health data and facilitate the response of public health to emerging health threats (e.g., new diseases).[39]

When evaluating a public health surveillance system, different aspects must be considered to achieve the main goals of maximizing the system's utility and efficiency. Those aspects include the public health importance of the event under surveillance, the objectives and usefulness of the system, the operation, and qualitative attributes (simplicity, flexibility and acceptability), quantitative attributes (sensitivity, predictive positive value, representativeness and timeliness), and cost.[39,40]

The relative importance of the evaluation attributes may vary depending on the specific methods, scope, purpose, and objectives of the surveillance system being evaluated. A system should prioritize those characteristics that are most crucial for achieving its goals, understanding that improvement in one area may negatively impact others (for example, an increase in sensitivity may compromise simplicity or timeliness). As such, an evaluation

must take into consideration the characteristics that are most important for the given system and its goals.[39]

Given the complexity of the formal evaluation process, a multistep approach is usually recommended – focusing at each step on only one or two major attributes, such as sensitivity and timeliness.[41] Fit for purpose study designs are required to evaluate those attributes and generate the evidence that will be needed to support the intervention.[41]

Surveillance was one of the five core components in the original WHO Framework for Effective Tuberculosis Control (the WHO DOTS strategy) established in 1994.[26] WHO emphasizes the importance of monitoring and evaluating program performance[11,42] and the importance of quality surveillance towards the goal of measuring TB incidence directly from TB notifications and generate information that supports action.[43]

As a disease approaches elimination, the priorities of surveillance systems shift to focus on detecting and preventing the re-emergence of the disease,[44] requiring more robust surveillance measures, such as active case finding, enhanced laboratory capacity, use of advanced diagnostic tools, and increased need for physician training to recognise the disease and notify it to public health.[45]

In monitoring progress in TB control, the ultimate aim for all countries is to count, accurately and comprehensively, TB cases (incidence) and deaths (mortality) through routine surveillance and vital registration. To reach that goal, systematic assessments of surveillance systems and data should be carried out to determine how accurately case and death reports measure TB burden, nationally and sub-nationally.[46]

On the assumption that not all cases of tuberculosis are reported to the national surveillance systems, and that not all national surveillance systems perform at the same level, WHO estimates the burden of disease caused by TB in terms of incidence, prevalence and mortality using not only the information gathered directly from surveillance systems, but also special studies such as prevalence surveys, mortality surveys, inventory studies on under-reporting, in-depth analysis of surveillance and other data, expert opinion and consultations with countries.[46,47] The methodology used does not come without limitations and risks, as although it facilitates cross-country comparability it might be actually counterproductive in some settings, especially in low-income countries.[48]

In 2020 and 2021, due to the impact of the COVID-19 pandemic, TB incidences in some countries were estimated using dynamic models, to account for large absolute reductions

in the number of notifications relatively to pre-2020 trends.[47] For high-income countries, WHO estimates of TB incidence and mortality in 2020-2021 were produced using the same methodology, i.e. notification data with a standard adjustment for incidence and vital registration data for mortality.[12]

In the last decades many studies, in different countries worldwide, have been published assessing sensitivity of the tuberculosis notification systems (more frequently described as completeness of notification), following the recommendations from WHO.[49] In 2003, a literature review in the UK showed that reporting of tuberculosis was incomplete and could be missing 7 to 27% of cases;[50] and in 2012 another literature review including 46 studies in the UK showed a median reporting completeness of 73%.[51] In the Netherlands, in 1998, an inventory study using three data sources showed an under-notification of 13%, but after implementing a capture-recapture methodology (possibly flawed by the violation of assumptions underlying capture-recapture analysis) the under-notification was estimated at 37%.[52] In Germany, using an inventory and 2-source capture-recapture methodology, authors estimated the sensitivity of the TB surveillance system in 2013-2017 to be between 93 and 97%.[53] In Romania, the same methodology using three data sources yielded an estimated sensitivity of 82-85% in 2004.[54] In an inventory study in Greece in 2016, using three data sources, the authors found that 55% of TB cases were not notified to the national level;[55] a figure that was found to be even lower in a regional study in West Greece.[56] In China, a study covering nine counties showed that 19% of TB cases were not reported to the national TB surveillance system.[57] Better results were found in Republic of Korea, with sensitivities ranging from 90 to 94% in 2012-2014,[58] and Taiwan (sensitivity in 2005-2007 was estimated at 96.3%.[59] In other regions of the globe the situation regarding the sensitivity of the TB surveillance systems, and the methods used to estimate it, is heterogeneous: 45% in a province of Saudi Arabia in 2012,[60] 59-74% in Peru in 2017-2018,[61] 75-79% in South Africa in 2011,[62] 60% in a district in Cape Verde in 2006-2012.[63]

Any surveillance system should allow the constant monitorization and analysis of data, in order to produce information that supports intervention. In their updated guidelines for evaluating Public Health Surveillance Systems, the U.S. Centers for Disease Control and Prevention (CDC) state that a system might be useful if, among other things, it “stimulates research intended to lead to prevention and control, leads to improved practices and allows estimation of the magnitude of morbidity and mortality (including the identification of factors associated with the event)”.[39]

In this thesis, usefulness was assessed by describing how TB surveillance data was used to produce a better understanding of the epidemiology of tuberculosis, including specific aspects such as trends in notification rates, molecular characterisation in the context of outbreaks, drug-resistant TB features, screening of latent TB and factors associated with failure to complete treatment.

The usefulness of a surveillance system is more subjective and difficult to assess and interpret, as it usually links up with the surveillance objectives (which might differ according from region to region, according to the context), as well as with all other system's attributes. Usefulness of the surveillance data and the system should be evaluated in the context of the two key surveillance functions: early warning (detect outbreaks of diseases in a timely manner, inform appropriate and effective public health responses, determine the distribution and spread of disease, illustrate the epidemiology of new diseases, provide information to categorize the outbreak as of national or international importance, provide data to evaluate control measures) and routine programme monitoring (estimate disease burden, identify risk groups, determine incidence trends over time, measure outcomes and impacts of preventive and public health interventions, evaluate the overall control interventions).[64] In Ireland, in 2018, a study investigated usefulness of the National TB surveillance system after the introduction of a computerised electronic reporting system, showing that most stakeholders considered the system was available when needed for TB surveillance and included the variables necessary for effective surveillance.[65] In other regions of the globe, such as Afghanistan, a formal evaluation of the TB surveillance system in 2013 showed that two different components were useful, one more focusing on planning and monitoring and the other one more useful for detecting TB outbreaks, reporting to national and international stakeholders, and development and implementation of national and international studies.[66] In Yemen, a qualitative evaluation of the TB surveillance system was conducted in 2021, revealing that the system was useful to monitor TB trends over time, for procurement of anti-TB drugs and laboratory reagents, for identification of TB high-risk groups, and assessment of the impact of interventions, but not so useful to identify treatment failure or helping in resource planning, prevention, care, and control.[67] Although these are only two examples of formal evaluation of usefulness of the system, the amount of literature that has been published on TB using as data source the national TB surveillance systems, covering aspects that range from medical, microbiological, and radiological aspects to epidemiological reports and public health intervention, at all geographical levels, has been huge – and reflects well the importance and usefulness of well-established national TB surveillance systems.[13,14,68–74]

Timeliness can also be defined as the pace at which information progresses through various stages in a surveillance system, from a health-related event to its reporting.[75] Numerous factors influence the duration of this interval, including patients' awareness and recognition of symptoms, the patient's access to medical care, the attending physician's diagnosis or utilization of laboratory tests, the laboratory's timely reporting of test results to the physician and/or public health agency, and finally, the physician's prompt communication of the event to the public health agency.[59,76]

Timeliness in surveillance extends to the amount of time required for identifying trends, outbreaks, or the efficacy of implemented control and prevention measures. This identification process is influenced by several factors, including the severity and communicability of the health-related event, the staffing levels of the responsible public health agency, and the efficacy of communication between involved health agencies and organizations.[39]

For infectious diseases, and specifically for tuberculosis, the most frequently used time intervals include the time between exposure and infection, onset of symptoms and diagnosis, and diagnosis to notification to the surveillance system. Such comprehensive assessments of time intervals play a crucial role in optimizing the efficiency and effectiveness of public health surveillance systems.[77–79]

Development and integration of digital, case-based, real-time surveillance systems for TB have the potential to greatly improve timeliness. These systems have several advantages over traditional paper-based reporting of aggregated data, including enabling the use of automated data linkage and quality checks, as well as enabling close-to-real-time access to data from the programme managers and public health services.[79]

Timeliness is a crucial component of surveillance systems, closely linked to its usefulness, and should be evaluated frequently. It depends on several factors, including the intended use of the data and the nature of the disease being monitored (for infectious diseases, transmissibility represents a key factor).[39,76] For infectious diseases such as tuberculosis, timeliness is very important when it comes to implement immediate disease control and prevention activities, but also to recognise changes in epidemiology and update clinical recommendations. The timeliness requirements also depend on the geographical level of the public health system, as the intervention is different at the local, regional, national or supranational levels.[76] In notification-based surveillance systems, such as the ones for tuberculosis, information about cases is usually reported at a first stage by healthcare

providers and laboratories to the local public health services, and then it progresses hierarchically to the regional, national, and supranational levels, triggering different actions that contribute to reporting timeliness delays. There are not many published evaluations of timeliness of disease surveillance systems, but there is a plethora of publications on the delays that precede the disease diagnosis – both on the component attributable to the patient and the component attributable to healthcare services – which also contribute to the overall surveillance system timeliness and should be considered in a formal evaluation, especially for tuberculosis.

Delay to diagnostic has been proven to be associated with sputum smear positivity and pulmonary cavities on lung X-ray, with the obvious public health implications.[80–89] In a study published in England, investigators found that one third of the TB cases had a diagnostic delay (defined as the time between symptom onset and treatment start date) of more than 4 months, with females, patients aged 45 years and older or residing outside of London, and extra-pulmonary TB cases, being especially affected.[90] In France, mean delay between first symptoms and diagnosis in 2010 was 97 days, equally distributed between patient delay (47 days between first symptoms and health care contact) and healthcare delay (48 days between health care contact and diagnosis).[91] In the Emilia-Romagna region of Italy, in 2003, the median delay to diagnosis of pulmonary tuberculosis was 65 days, with median patient delay and health care delays of 7 days and 36 days, respectively.[92] In a study on hospitalised TB patients in Serbia, in 2015, median delay to diagnostic was 118 days, with median patient delay and health care delays of 93 days and 19 days, respectively, and with reported excessive alcohol consumption being associated with a more prolonged time to seek medical help.[93] In Portugal, in 2010-2014, median delay to diagnostic was 68 days, with patient delay and health care delays of 33 and 17 days, respectively; in this study, bigger delays were observed in patients with extra-pulmonary TB and pulmonary comorbidities (such as lung cancer, sarcoidosis and chronic obstructive pulmonary disease), as well as patients residing further away from a healthcare services.[84] In a systematic review covering almost 19,000 cases from multiple low to middle income countries, from 2007 to 2005, conducted in 2017, the median diagnostic delay varied widely from 30 to 367 days, with 4 to 199 days attributable to the patient and 2 to 129 due to healthcare delays.[94] In a study from the Kerala region of India, mean time or diagnosis was 78 days, with 42 days being attributable to healthcare and 36 days attributable to the patient.[95]

In another study in the USA, timeliness of TB reporting was calculated using the number of days between date of TB diagnosis and date of report to the local or state health department: timeliness of reporting varied between sites from a median of 7 days to a median of 38 days, varying with infectiousness, type of provider, diagnosing provider, and reporting source.[96] In Republic of Korea, in 2012-2014, delay to notification of tuberculosis cases to the surveillance system was of 80-82% in the first 7 days after diagnosis, increasing to 90% within 30 days after diagnosis.[58] In Taiwan, in 2005-2007, 82% of cases had been notified within 7 days of start of treatment.[59]

In a study published in Canada, authors investigated timeliness and comprehensiveness of national tuberculosis surveillance reporting between 2000 and 2020, presenting another important timeliness indicator which was the delay to publication of annual surveillance data at the national level: this was on average 25 months.[97] In an evaluation study in Brazil, timeliness of the tuberculosis surveillance system in 2012-2014 was assessed as a composite indicator encompassing timeliness of notification, timeliness of investigation, timeliness of typing, timeliness of treatment and timeliness of case closure; the proportion of cases that were treated and notified in a timely fashion were 25 and 63%, respectively.[98]

1.4. THESIS AIMS AND LAYOUT

The aim of this thesis was to conduct an external evaluation of the tuberculosis surveillance system in Portugal, producing evidence that can be used to improve the system to better adapt to the changing epidemiology of TB in the country.

Specific objectives

1. To estimate the sensitivity (completeness of notification) of the National Tuberculosis Programme surveillance system (SVIG-TB) in 2015.

2. To evaluate the usefulness of SVIG-TB, in terms of research products that could be derived from this information system and complemented by data from other surveillance systems or data sources.

3. To assess the timeliness of SVIG-TB and SINAVE, in terms of delay from date of onset of symptoms, diagnosis, date of notification and information being made available to the National Tuberculosis Programme and to public health services.

The objectives and research products included in this thesis were defined taking into account the End TB strategy pillars and components,[11] as described in Table 1.

Table 1. Articles included in this thesis, according to specific End TB strategy pillars/components

End TB strategy pillar/component	**Articles included in this thesis**
Early diagnosis of tuberculosis including universal drug-susceptibility testing, and systematic screening of contacts and high-risk groups	3.2, 3.4, 3.7, 4.1.[4,6,9]
Treatment of all people with tuberculosis including drug-resistant tuberculosis, and patient support	3.2.[4]
Collaborative tuberculosis/HIV activities, and management of co-morbidities	3.3, 3.5.[5,7]
Preventive treatment of persons at high risk, and vaccination against tuberculosis	3.3, 3.4, 3.7.[5,6,9]
Engagement of communities, civil society organizations, and public and private care providers	3.6, 3.7.[8,9]
Discovery, development and rapid uptake of new tools, interventions and strategies	2.1, 2.2, 3.1, 3.2.[1–4]
Research to optimize implementation and impact and promote innovations	All articles.[1–9]

Thesis layout

This thesis is organised in five chapters, with the first one being this introduction and covering the epidemiology of TB, surveillance systems and evaluation principles. In the second, third and fourth chapters the results of the thesis are presented, according to the objectives on the surveillance system attributes sensitivity (two published articles), usefulness (seven published articles) and timeliness (one research protocol). The fifth and last chapter of the thesis covers a general discussion, providing an overview of the main findings in the state-of-the-art context, a reflection on strengths and limitations of this work, future perspectives, and recommendations for improvement of the TB surveillance system.

2

SENSITIVITY OF THE TUBERCULOSIS SURVEILLANCE SYSTEM

Sensitivity is one of the most important attributes of a surveillance system, referring to the ability of the system to detect the events under surveillance. The terms "case-ascertainment" and "completeness of notification" are often interchangeably used.

Two fundamental questions are raised when trying to estimate the proportion of cases that are correctly identified: 1) what is exactly a "true case" of tuberculosis? and 2) assuming that there is no such thing as a 100% sensitive surveillance system, how can we know what is the "true" number of cases of tuberculosis if not all of them are diagnosed or notified? The first question is answered by one of the first steps of any research: a specific case definition, based on uniform criteria that are built upon variables available in the surveillance system;[99] the answer to the second is much more complex, often relying on estimates produced by modern epidemiology methods.[49]

In order to estimate of sensitivity of the National Tuberculosis Programme surveillance system (SVIG-TB) in Portugal, we conducted an inventory study, coupled with a capture-recapture analysis, using two additional national data sources: SINAVE (compulsory-notification system), and hospital admission records (GDH).[1,2] This methodology is in line with the WHO recommendations to assess TB under-reporting.[49]

The two articles presented in this chapter were published in indexed international peer-reviewed journals and were supported by the European Centre for Disease Prevention and Control (ECDC/2016/004 – Framework Contract 'Assessment of tuberculosis underreporting through inventory studies').[100]

[1] **Carvalho C**, Alba S, Harris R, Abubakar I, Van Hest R, Correia AM, Gonçalves G, Duarte R. Completeness of TB notification in Portugal, 2015: an inventory and capture-recapture study. Int J Tuberc Lung Dis. 2020;24(11):1186–93. (Impact factor Web of Science: 2.4)

[2] Straetemans M, Bakker MI, Alba S, Mergenthaler C, Rood E, Andersen PH, Schimmel H, Simunovic A, Svetina P, **Carvalho C**, Lyytikäinen O, Abubakar I, Harris RJ, Ködmön C, van der Werf MJ, van Hest R. Completeness of tuberculosis (TB) notification: Inventory studies and capture-recapture analyses, six European Union countries, 2014 to 2016. Eurosurveill. 2020;25(12):1–10. (Impact factor Web of Science: 6.3)

2.1. COMPLETENESS OF NOTIFICATION OF TUBERCULOSIS IN PORTUGAL 2015: AN INVENTORY AND CAPTURE-RECAPTURE STUDY

INT J TUBERC LUNG DIS 24(11):000–000

http://dx.doi.org/10.5588/ijtld.19.0094

Completeness of notification of tuberculosis in Portugal 2015: an inventory and capture-recapture study

C. Carvalho,[1] S. Alba,[2] R. Harris,[3] I. Abubakar,[4] R. Van Hest,[5] A. M. Correia,[6] G. Gonçalves,[1] R. Duarte[7,8]

[1]Multidisciplinary Unit for Biomedical Research (UMIB), Institute of Biomedical Sciences Abel Salazar (ICBAS), University of Porto, Porto, Portugal; [2]Royal Tropical Institute, KIT Health, Amsterdam, The Netherlands; [3]National Infection Service, Public Health England, London, [4]Institute for Global Health, University College of London, London, UK; [5]Department of Tuberculosis Control, Regional Public Health Service (GGD) Groningen, Groningen, The Netherlands; [6]Braga Health Centre Group, Portuguese Northern Regional Health Administration Cávado I, Braga, [7]EPIUnit, Institute of Public Health, University of Porto (ISPUP), Porto, [8]Public Health Science and Medical Education Department, Faculty of Medicine, University of Porto, Porto, Portugal

SUMMARY

BACKGROUND: Despite the steady decline in the last few decades, Portugal remains the Western European country with the highest TB notification rates. The aim of this study was to estimate the completeness of notification to the National Tuberculosis Programme (NTP) Surveillance System (SVIG-TB) in 2015.

METHODS: We implemented an inventory study and a three-source log-linear capture-recapture analysis using two additional data sources that were deterministic and probabilistically lined: the national notifiable diseases surveillance system (*Sistema Nacional de Vigilância Epidemiológica* [SINAVE]) and the national hospital discharge database (*Grupos de Diagnósticos Homogéneos* [GDH]).

RESULTS: We identified 2328 unique probable/confirmed TB cases across the three data soources. We found a positive dependency between SVIG-TB and SINAVE (incidence rate ratio [IRR] 8.9, 95%CI 6.6–12.0) and between GDH and SINAVE (IRR 2.6, 95%CI 2.0–3.4). After adjusting for these dependencies, we estimated that 266 cases (95%CI 198–358) were not reported, indicating a notification (to SVIG-TB) completeness rate of 77%.

CONCLUSION: True incidence rate of TB in Portugal in 2015 could have been as high as 26.1 per 100 000. This could be an overestimation because of false-positive cases recorded in both SINAVE and GDH or due to a smaller scale, false non-matches. Studies aimed at validating potentially false-positive cases should be implemented to address these limitations.

KEY WORDS: TB; completeness of notification; record-linkage; capture-recapture; Portugal

Correspondence to: Carlos Carvalho, Instituto de Ciências Biomédicas de Abel Salazar da Universidade do Porto, Departamento de Estudo das Populações, Rua Jorge de Viterbo Ferreira 228, 4050-313 Porto, Portugal. e-mail: cfcarvalho@icbas.up.pt

Article submitted 20 February 2020. Final version accepted 27 April 2020.

Background

Notification rates form the basis of TB incidence estimates, which are the primary indicator to monitor the new global strategy and targets for TB prevention, care and control after 2015 (the End TB Strategy).[101] In Portugal, incidence rates of TB have been estimated and internationally reported by the National TB Programme (NTP) using as single data source the NTP surveillance system *Sistema de Vigilância da Tuberculose* (SVIG-TB).[33] TB notification rates have been steadily declining in the last decades in Portugal, reaching 20.1 notifications per 100 000 population in 2015 (close to the cut-off between intermediate- and low-incidence in a region).[102]

According to the 2017 European Centre for Disease Control and Prevention Surveillance Report, however, the true incidence rate in Portugal would be 15% higher (23/100 000 in 2015) due to under notification.[103] When multiple data sources are available, it is possible to estimate the completeness of notification using inventory studies, combined or not with capture-recapture (CRC) modelling.[49,104]

In the present study, we aimed at assessing the completeness of notification to the NTP Surveillance System and estimating the true incidence rate of TB in Portugal in 2015, implementing an inventory study and a CRC analysis with two additional national TB data sources.

Methods

Design and study population

We conducted an inventory study and CRC analysis including all records of diagnosed TB, retrieved from three national electronic, case-based databases for 2015 in mainland Portugal. Autonomous Island Regions of Azores and Madeira were excluded as there were no data on hospital discharges in these regions.

Case definition

We included all notified cases of active TB in mainland Portugal in 2015. Cases were considered 1) confirmed using a positive culture for Mycobacterium tuberculosis complex (gold standard) or polymerase chain reaction (PCR) and sputum smear; 2) probable, if clinical and laboratory results (PCR, microscopy or histopathological findings) were compatible with TB; or 3) possible, if diagnosis was made on clinical, radiological and/or epidemiological arguments.

Data sources

NTP surveillance system (SVIG-TB)

This system is based on the clinical notification by physicians from TB outpatient centres, which are part of the Portuguese National Health Service. Cases are notified and monitored using two different paper forms and sent to 'notification centres' for computer recording and subsequent national data aggregation.

At the national level, all data are anonymised, but information on date of birth, sex and place of residence are kept (as well as clinical information).

Patients with suspect TB are referred to TB outpatient centres from hospitals, general practitioners, private practitioners, public health services or by their own initiative. If suspicion is high enough treatment is started, triggering a first notification to SVIG-TB. When culture and drug susceptibility testing (DST) results are available, a second notification to SVIG-TB is triggered. There is a continuous update with serious adverse events, laboratory results and outcome (Figure 3). Any false diagnostic of TB (non-tuberculous mycobacterial infection or other diagnosis) is recorded in 'treatment outcome' and the case is excluded ('de-notified').

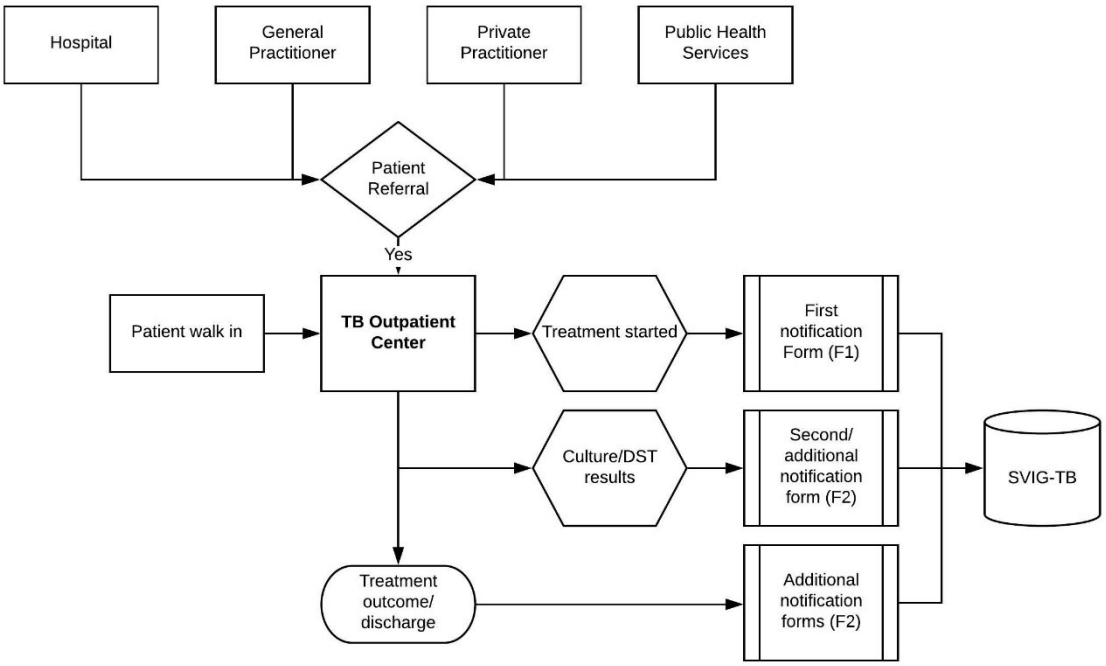


Figure 3. National Tuberculosis Programme (NTP) Surveillance System (SVIG-TB) Patient Referral and Data Flow, Portugal.

National notifiable diseases surveillance system (SINAVE)

TB is a statutory notifiable disease in Portugal since 1902.[105] In 2014 a new web-based notification platform, SINAVE (Sistema Nacional de Vigilancia - Epidemiológica), was introduced.[30] All physicians from the public and private sectors are legally required to notify to SINAVE whenever they suspect TB. The primary purpose of this surveillance system is triggering epidemiological investigation, risk assessment and community intervention by public health services. In theory, SINAVE should be updated with all available information before cases are validated as possible, probable, confirmed or non-cases (Figure 4). However, as culture results usually take some weeks, some cases are validated before there is enough information to accurately classify or exclude TB diagnosis.

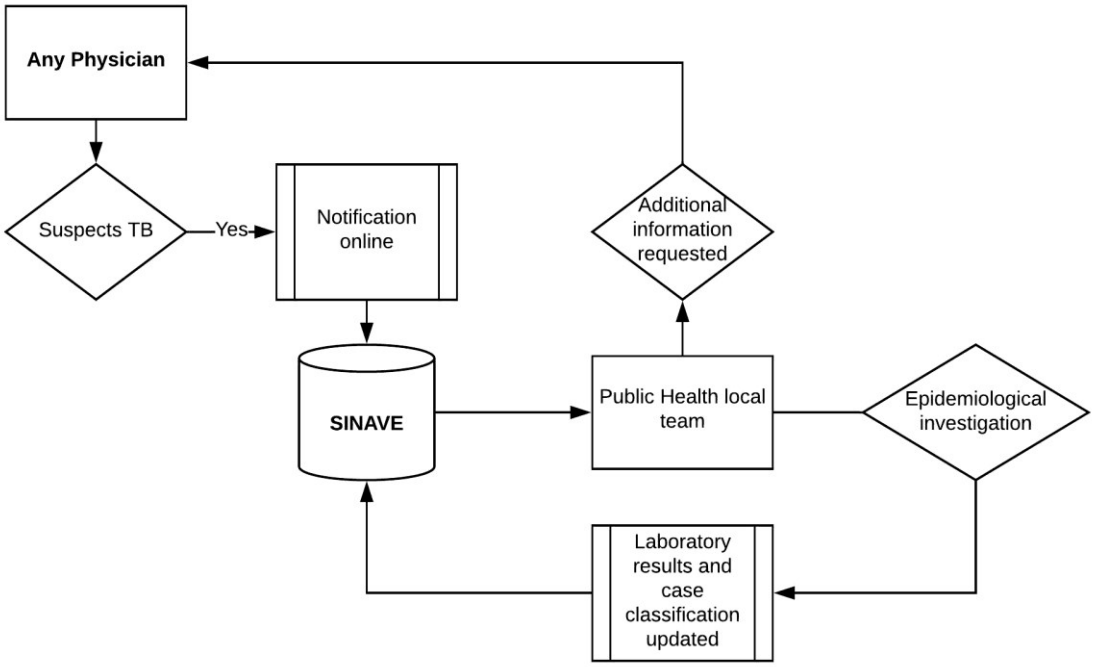


Figure 4. National Notifiable Diseases Surveillance System (SINAVE) Data Flow, Portugal.

Hospital discharge database (Grupos de Diagnósticos Homogéneos)

This database is managed by the Portuguese Central Administration of the Health System. All diagnoses made in the Portuguese National Health Service hospitals are coded upon discharge, using the International Statistical Classification of Diseases, Injuries and Causes of Death (ICD) version 9 before 2016 and version 10 since then. ICD9 codes referring to diagnoses of TB start with 010–018 (TB), 37105 (Phthisical cornea) or 6473 (TB in pregnancy).

Record linkage

All records of TB diagnosed in the study period were received in electronic spreadsheet format, including the variables date of birth, sex, place of residence (district, municipality and parish), date of diagnosis (estimated in Grupos de Diagnosticos Homogéneos [GDH], the hospital discharge database, as date of admission plus one third of hospitalisation days), site of disease and laboratory results (ICD9 codes termination digits in GDH), HIV

co-infection (ICD9 codes started by '042', '07953' or 'V08' in GDH) and vital status at discharge.

Data were abstracted from each data source for cases diagnosed from 1 July 2014 to 30 June 2016 to allow for correction of late notification. After record linkage, records from 1 July 2014 to 31 December 2014 and from 1 January to 30 June 2016 found in only one data source were excluded.

An initial step of deterministic linkage was performed,[49] matching exactly date of birth, sex and place of residence. These variables were expected to enable unique identification of a single case across the three databases.

To account for possible recording errors in the matching variables (records that in reality referred to the same case, although there were actually differences in the data recorded), probabilistic matching was performed using the Stata (StataCorp, College Station, TX, USA) user-written command *reclink*. [106] Agreement or disagreement weights were assigned to specific properties of each record (to accommodate different probability of error in different fields), which were then combined into a single probability score. Linked records were visually inspected, sorted by calculated probability scores, and different matching variables and minimum matching weights were used until most links were considered acceptable.

As a first step, the GDH data set was merged with SVIG-TB and thereafter the resulting dataset was merged with the SINAVE data set (minimum overall matching score of 0.87 to declare a match). The following relative weights were used: first digit of day of birth (10); second digit of day of birth (10); first digit of month of birth (10); second digit of month of birth (10); third digit of year of birth (15); fourth digit of year of birth (10); sex (15); district of residence (15); municipality of residence (5); parish of residence (3).

Cases were allocated to one of seven different strata (case existing in data sets A, B and C, A and B, A and C, B and C, only A, only B and only C) and assigned an unique case-classification depending on the information available: confirmed, if confirmed in at least one data source; probable, if not confirmed but probable in at least one data source; possible, if not confirmed nor probable in any of the three data sources.

Capture-recapture analysis

CRC analyses rely on Poisson regression models to estimate the total number of unreported cases in a data system using the information from all available data sources. The data consist of aggregated totals for each data set and overlaps between data sets (i.e., all different portions of the Venn Diagram in Figure 5). The models are built in a way that accounts for the fact that some data sets are ‘‘dependent’’ on each other, meaning that the probability of a case being recorded in one data set is related to the probability of being recorded in another data set. Such dependencies can result in biased estimates of the number of unreported cases if not accounted for. The incidence rate ratio (IRR) of the interaction term for belonging to two data sets is thus a measure of dependency between databases: if the IRR is greater than 1, the conclusion is that the two sources are positively associated.[107,108]

The models fitted in this study included three potential dependencies between pairs of datasets. As a result, eight possible models were considered, ranging from no dependencies (the base model) to all three (saturated model). The choice of model was

determined by balancing model fit with parsimony, based on Akaike Information Criterion (AIC) scores.[109] The chosen model was further expanded and analyses stratified by the covariates sex, region of residence, case classification, site of disease, HIV coinfection and vital status in turn.

Capture-recapture analysis was repeated after excluding possible cases. This was done in order to limit the impact of including potentially false-positive cases in the analysis.[110–112] All data management, including record linkage and CRC analysis, was performed using Stata v15.[113]

Ethics approval

This study received clearance by the Portuguese National Data Protection Committee (no 1043/2017) and was approved by the Northern Regional Health Administration Ethical Board (no 114/2017).

Results

From 1 January to 31 December 2015, 2086 cases were diagnosed with TB and subsequently notified to the National Tuberculosis Programme (NTP) surveillance system (SVIG-TB). In that same period, 1874 cases were recorded in the national notifiable diseases surveillance system (SINAVE) and 1358 in the hospital discharge database (GDH) (Table 2).

Table 2. Cases of TB Notified to SVIG-TB and SINAVE and Admitted to Hospital in Mainland Portugal (Date of Diagnosis 01 January to 31 December 2015)

Variable	**Categories**	**Cases in each of the data sources (proportion within data source)**		
		SVIG-TB	***SINAVE***	***GDH***
Region of residence	Northern Region (population 3.6M)	865 (41.5%)	805 (43%)	445 (32.8%)
	Lisbon and Tagus Valley (population 3.6M)	872 (41.8%)	750 (40.0%)	612 (45.1%)
	Central Region (population 1.7M)	179 (8.6%)	175 (9.3%)	181 (13.3%)
	Alentejo (population 0.48M)	79 (3.8%)	59 (3.1%)	52 (3.8%)
	Algarve (population 0.44M)	91 (4.4%)	85 (4.5%)	68 (5.0%)
Sex	Male	1391 (66.7%)	1246 (66.5%)	940 (69.2%)
	Female	695 (33.3%)	628 (33.5%)	418 (30.8%)
Age group (in years)	[0-5[	11 (0.5%)	9 (0.5%)	17 (1.3%)
	[5-15[	21 (1.0%)	21 (1.1%)	19 (1.4%)
	[15-25[	182 (8.7%)	154 (8.2%)	105 (7.7%)
	[25-35[	248 (11.9%)	217 (11.6%)	124 (9.1%)
	[35-45[	409 (19.6%)	374 (20%)	247 (18.2%)
	[45-55[	409 (19.6%)	383 (20.4%)	257 (18.9%)
	[55-65[	325 (15.6%)	287 (15.3%)	220 (16.2%)
	[65-75[	214 (10.3%)	183 (9.8%)	148 (10.9%)
	75+	267 (12.8%)	246 (13.1%)	221 (16.3%)
Site of disease	Pulmonary	1475 (70.7%)	1301 (69.4%)	881 (64.9%)
	Extra-pulmonary	602 (28.9%)	507 (27.1%)	477 (35.1%)
	Unknown	9 (0.4%)	66 (3.5%)	-
Case classification	Confirmed	1278 (61.3%)	1222 (65.2%)	147 (10.8%)
	Probable	466 (22.3%)	447 (23.9%)	784 (57.7%)
	Possible	342 (16.4%)	205 (10.9%)	427 (31.4%)
HIV co-infection	Yes	223 (10.7%)	184 (9.8%)	214 (15.8%)
	No/not recorded	1863 (89.3%)	1690 (90.2%)	1144 (84.2%)
Vital status [1]	Dead	153 (7.3%)	94 (5%)	152 (11.2%)
	Alive/not recorded	1933 (92.7%)	1780 (95%)	1206 (88.8%)
TOTAL		**2086 (100%)**	**1874 (100%)**	**1358 (100%)**

GDH: Hospital Discharge Database; SINAVE: national notifiable diseases surveillance system; SVIG-TB: National Tuberculosis Programme surveillance system. [1] Last updated at the end of treatment (SVIG-TB), recorded by Public Health services within 1-3 months after diagnosis/notification but might be updated by the regional or national level at a later stage (SINAVE), at the date of hospital discharge (GDH).

Identification of matches

Using deterministic record-linkage, 3561 unique records were identified across all three datasets. Most of the unmatched cases were observed in the hospital discharge database (61.9% of cases admitted to hospital) (Table 3).

Table 3. Deterministic Record-linkage of Possible, Probable and Confirmed Cases of Tuberculosis in Three Data Sources: SVIG-TB, SINAVE and GDH, Mainland Portugal, 2015.

Data sources [1]	**Cases recorded [2]**	**%**
SVIG-TB	682	19.2%
SVIG-TB and *SINAVE*	1046	29.4%
SVIG-TB and *GDH*	74	2.1%
SVIG-TB and *SINAVE* and *GDH*	342	9.6%
SINAVE	430	12.1%
SINAVE and *GDH*	118	3.3%
GDH	869	24.4%
TOTAL	**3561**	**100%**

GDH: Hospital Discharge Database; SINAVE: national notifiable diseases surveillance system; SVIG-TB: National Tuberculosis Program surveillance system. [1] Date of diagnosis from 1 January to 31 December 2015 recorded in at least one of the data sources; [2] Records deterministically matched on date of birth, sex and place of residence (district, municipality, parish).

The probabilistic record-linkage between the three data sets (matching scores shown on Table 4) greatly increased the overlap between data sources, allowing the identification of 2786 unique records (Table 5).

Table 4. Accepted Matching Scores (Probabilistic Record-linkage of Tuberculosis Datasets), Frequency and Description of Corresponding Mismatch Between Linking Variables, Mainland Portugal, 2015.

Overall matching score	**Proportion of cases matched (first record linkage, *GDH* with *SVIG-TB*) [1]**	**Proportion of cases matched (second record linkage, *GDH_SVIG-TB* with *SINAVE*) [2]**	**Mismatch between linking variables [3]**
1.0000	15.67%	42.00%	Perfect match
0.9709	18.70%	20.64%	Different parish
0.9223	1.54%	2.69%	Different municipality (and parish)
0.9029	1.05%	1.40%	One digit of date of birth
0.8738	2.59%	1.87%	One digit of date of birth and parish of residence
< 0.8700	60.45%	31.41%	Rejected match

GDH: Hospital Discharge Database; *SINAVE*: national notifiable diseases surveillance system; *SVIG-TB*: National Tuberculosis Programme surveillance system. [1] Denominator includes all non-matched records from SVIG-TB and GDH plus matched records (N=2668); [2] Denominator includes all non-matched records from the database generated in [1] (SVIG-TB_GDH) and SINAVE plus matched records (N=2786); [3] Variables used for record-linkage: date of birth (day, month, year), sex and place of residence (district, municipality, parish).

Table 5. Probabilistic Record-linkage of Possible, Probable and Confirmed Cases of Tuberculosis in Three Data Sources: SVIG-TB, SINAVE and GDH, Mainland Portugal, 2015.

Data sources [1]	**Cases recorded [2]**	**%**
SVIG-TB	323	11.6%
SVIG-TB and *SINAVE*	850	30.5%
SVIG-TB and *GDH*	127	4.6%
SVIG-TB and *SINAVE* and *GDH*	928	33.3%
SINAVE	118	4.2%
SINAVE and *GDH*	133	4.8%
GDH	307	11.0%
TOTAL	**2786**	**100%**

GDH: Hospital Discharge Database; *SINAVE*: national notifiable diseases surveillance system; *SVIG-TB*: National Tuberculosis Program surveillance system. [1] Date of diagnosis from 1 January to 31 December 2015 recorded in at least one of the data sources; [2] Records probabilistically matched on date of birth, sex and place of residence (district, municipality, parish).

The pattern of overlap between data sources varied across regions and case classifications (Table 6, Table 7).

Table 6. Probabilistic Record-linkage of Probable and Confirmed Cases of Tuberculosis in Three Data Sources: SVIG-TB, SINAVE and GDH, Stratified by Region of Residence of Case, Mainland Portugal, 2015.

Data sources/Region [1]	**Northern Region**	**Central Region**	**Lisbon and Tagus Valley**	**Alentejo**	**Algarve**
SVIG-TB	73 (7,9%)	33 (13,1%)	103 (10,6%)	17 (20,2%)	6 (6,1%)
SVIG-TB and *SINAVE*	430 (46,6%)	48 (19,1%)	250 (25,7%)	18 (21,4%)	33 (33,7%)
SVIG-TB and *GDH*	21 (2,3%)	16 (6,4%)	55 (5,7%)	5 (6%)	4 (4,1%)
SVIG-TB and *SINAVE* and *GDH*	312 (33,8%)	74 (29,5%)	426 (43,8%)	32 (38,1%)	41 (41,8%)
SINAVE	33 (3,6%)	18 (7,2%)	35 (3,6%)	6 (7,1%)	6 (6,1%)
SINAVE and *GDH*	27 (2,9%)	35 (13,9%)	51 (5,2%)	1 (1,2%)	3 (3,1%)
GDH	27 (2,9%)	27 (10,8%)	52 (5,3%)	5 (6%)	5 (5,1%)
TOTAL	**923 (100%)**	**251 (100%)**	**972 (100%)**	**84 (100%)**	**98 (100%)**

GDH: Hospital Discharge Database; *SINAVE*: national notifiable diseases surveillance system; *SVIG-TB*: National Tuberculosis Program surveillance system. [1] Cases diagnosed from 1 January to 31 December 2015, recorded as probable or confirmed in at least one of the data source.

Table 7. Probabilistic Record-linkage of Cases of Tuberculosis in Three Data Sources: SVIG-TB, SINAVE and GDH, Stratified by Case Classification, Mainland Portugal, 2015.

Data sources/Case classification [1]	Confirmed [2]	Probable [3]	Possible [4]
SVIG-TB	155 (8.9%)	77 (13.3%)	91 (19.9%)
SVIG-TB and *SINAVE*	648 (37.0%)	131 (22.7%)	71 (15.5%)
SVIG-TB and *GDH*	62 (3.5%)	39 (6.7%)	26 (5.7%)
SVIG-TB and *SINAVE* and *GDH*	726 (41.5%)	159 (27.5%)	43 (9.4%)
SINAVE	68 (3.9%)	30 (5.2%)	20 (4.4%)
SINAVE and *GDH*	70 (4.0%)	47 (8.1%)	16 (3.5%)
GDH	21 (1.2%)	95 (16.4%)	191 (41.7%)
TOTAL	**1750 (100%)**	**578 (100%)**	**458 (100%)**

GDH: Hospital Discharge Database; *SINAVE*: national notifiable diseases surveillance system; *SVIG-TB*: National Tuberculosis Program surveillance system. [1] Cases diagnosed from 1 January to 31 December 2015; [2] Positive culture for Mycobacterium tuberculosis complex or [Polymerase Chain Reaction (PCR) and sputum smear] in at least one of the data sources; [3] Cases not confirmed but with PCR, microscopy or histopathological findings compatible with TB in at least one of the data sources; [4] all other cases.

Considering only probable and confirmed cases, 2328 unique records were identified (Figure 5).

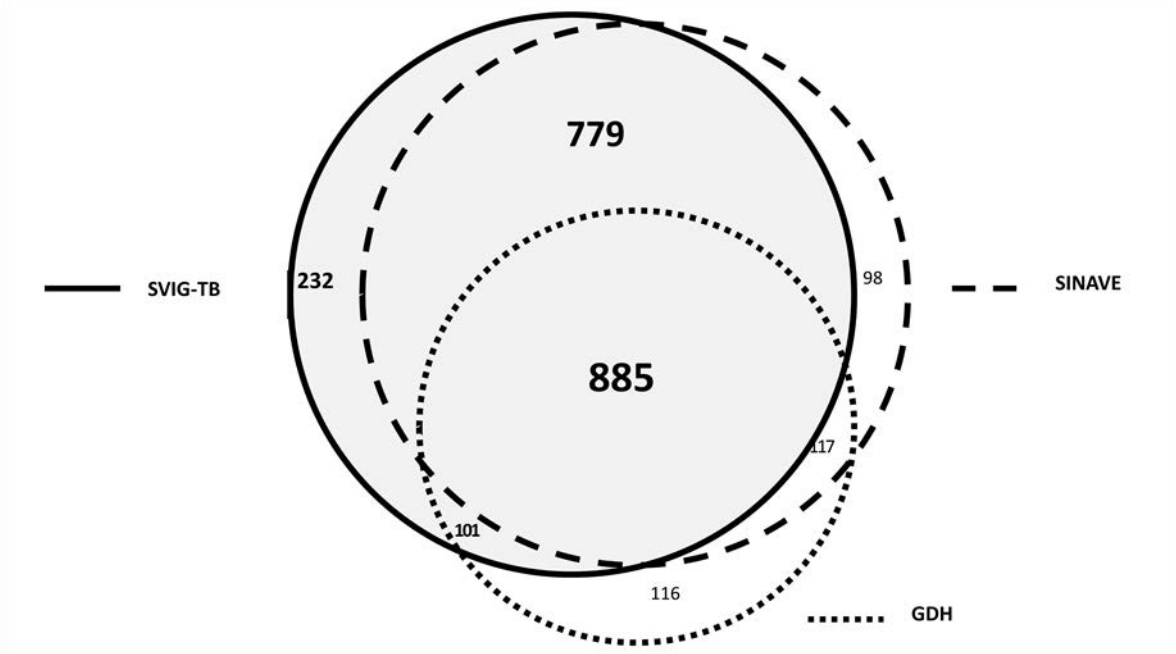


Figure 5. Probabilistic Linkage of Probable and Confirmed Cases of Tuberculosis Between Three Data Sources: SVIG-TB, SINAVE and GDH, Mainland Portugal, 2015.
GDH: hospital discharge database; *SINAVE*: national notifiable diseases surveillance system; *SVIG-TB*: National Tuberculosis Programme surveillance system. Numbers represent the number of cases observed in each one, two or three data sources (inventory study, N=2328).

Capture-recapture analysis

The model with better fit included two-way interactions between SVIG-TB and SINAVE, as well as between GDH and SINAVE (Table 8). Both indicated a positive dependence between sources (probabilistically linked data set, including only probable and confirmed cases—SVIG-TB:SINAVE, IRR 8.9 (95%CI 6.6–12.0); SINAVE:GDH IRR 2.6 (95%CI 2.0–3.4).

Table 8. Model Selection Statistics and Estimated Number of Unreported Cases of Tuberculosis in the Probabilistically Record-linked Dataset, Mainland Portugal, 2015.

Model (Interaction terms)	Df	Possible, probable and confirmed cases			Probable/confirmed cases		
		AIC	Unreported cases [1]	95% CI	AIC	Unreported cases [1]	95% CI
M1 (No interactions)	4	653.3	92.2	81.9-103.8	263.2	36.6	31.4-42.5
M2 (SVIG-TB:SINAVE)	5	150.8	333.6	291.0-382.5	121.9	116.6	95.5-142.5
M3 (SVIG-TB:GDH)	5	549.4	46.7	38.3-57.0	236.9	24.7	19.8-30.9
M4 (GDH:SINAVE)	5	654.9	94.6	81.9-109.3	252.8	43.5	36.6-51.8
M5 (SVIG-TB:SINAVE; SVIG-TB:GDH)	6	149.9	272.4	207.5-357.5	121.9	97.2	70.3-134.4
M6 (SVIG-TB:SINAVE; GDH:SINAVE)	6	64.5	780.8	618.0-986.4	62.6	266.5	198.2-358.3
M7 (SVIG-TB:GDH; GDH:SINAVE)	6	550.6	44.8	35.9-55.9	229.8	29.2	22.8-37.3
M8 (SVIG-TB:SINAVE; SVIG-TB:GDH; GDH:SINAVE)	7	66.5	756.3	531.3-1076.7	64.5	253.6	168.1-382.5

AIC: Akaike Information Criteria; CI: Confidence Interval; Df: Degrees of freedom used by the model (i.e. number of parameters); *GDH*: Hospital Discharge Database; *SINAVE*: national notifiable diseases surveillance system; *SVIG-TB*: National Tuberculosis Programme surveillance system. [1] Estimated number of unreported cases as exponentiated intercept coefficient.

The estimated number of unreported cases differed substantially across different models—but was comparable in models with similar AIC statistics. The model with all three two-way interactions (M8) had similar AIC to the chosen model (M6), and a similar number of unreported cases. This suggests that M6 and M8 have similar performance and that the choice of a more parsimonious model does not affect accuracy (Table 8).

The inclusion of covariates in the model made little difference in the overall estimated number of unreported cases, but those covariates affected capture probabilities and overlap patterns. Unreported cases were more likely to be possible cases (as opposed to being confirmed or probable), HIV-negative and alive (Table 9).

Table 9. Estimated Number of Unreported Cases of Tuberculosis by Selected Covariates in the Probabilistic Record-linked Dataset, Mainland Portugal, 2015.

Covariates	**Unreported cases [1] (95% CI)**	
	Possible, probable and confirmed cases	**Probable/confirmed cases**
None	780.8 (618.0-986.4)	266.5 (198.2-358.3)
Sex		
Male	410.5 (309.3-544.7)	149.4 (103.8-215.0)
Female	420.5 (274.5-644.0)	121.6 (71.8-206.0)
Geographical area		
Northern region	381.5 (239.2-608.5)	93.9 (50.8-173.5)
Lisbon and Tagus Valley	243.6 (173.8-341.5)	97.4 (63.6-149.0)
Rest of the country	222.9 (139.1-357.2)	82.9 (46.8-146.7)
Case classification		
Confirmed	52.5 (31.2-88.2)	52.5 (31.2-88.2)
Probable	187.6 (121.5-289.6)	187.6 (121.5-289.6)
Possible	668.5 (422.7-1057.2)	-
Site of disease		
Pulmonary	476.1 (349.2-649.0)	172.6 (118.8-250.9)
Extra-pulmonary	288.3 (201.2-413.1)	84.7 (51.6-139.2)
HIV co-infection		
HIV	52.6 (29.1-94.8)	22.3 (10.7-46.6)
No-HIV	765.6 (591.9-990.3)	252.4 (182.3-349.4)
Vital status		
Dead	23.5 (11.0-50.0)	9.6 (4.0-22.6)
Alive	783.5 (608.4-1008.9)	267.0 (192.7-369.9)

CI: Confidence interval; *GDH*: Hospital Discharge Database; *SINAVE*: national notifiable diseases surveillance system; *SVIG-TB*: National Tuberculosis Programme surveillance system. [1] Estimated number of unreported cases as exponentiated intercept coefficient, according to Model M6.

Completeness of notification (sensitivity)

Taking into account the figures above, the estimated sensitivity of SVIG-TB for probable/confirmed cases of TB was 77.0% [1997/(2328+266)] (Table 10).

Table 10. Estimated Completeness of Notification (Sensitivity) of SVIG-TB, SINAVE and GDH, According to Case Definition, Mainland Portugal, 2015.

Data source	Sensitivity (%) [95% CI]	
	Possible, probable and confirmed cases	**Probable/confirmed cases**
SVIG-TB	62.5% [59.1% - 65.5%]	77.0% [74.3% - 79.1%]
SINAVE	56.9% [53.8% - 59.6%]	72.4% [70.0% - 74.4%]
GDH	41.9% [39.6% - 43.9%]	47.0% [45.4% - 48.3%]
Inventory study (*SVIG-TB+SINAVE+GDH*)	**78.1% [73.9% - 81.8%]**	**89.7% [86.7% - 92.2%]**

CI: Confidence Interval; *GDH*: Hospital Discharge Database; *SINAVE*: national notifiable diseases surveillance system; *SVIG-TB*: National Tuberculosis Program surveillance system.

Discussion

This study was a first attempt at evaluating the sensitivity of the surveillance systems for tuberculosis in Portugal. It included a comprehensive CRC analysis using log-linear regression, based on three data sources both deterministically and probabilistically linked using two different case definitions. We included in the analysis various covariates, which showed different capture probabilities and between-source dependencies but had little impact on the overall estimate of unreported cases. The high dependency between SVIG-TB and SINAVE is consistent with the process of notification, since all physicians are legally required to notify new TB cases to Public Health (through SINAVE). The positive dependency between GDH and SINAVE may also imply that hospitalised cases are more likely to be notified to SINAVE (rather than the reverse). Probable errors in data entry to the data sources prevented accurate deterministic data linkage and hence the capture-recapture analysis in this data set produced an unrealistically high estimate of the number of unreported TB cases. This is an important limitation of CRC methods[114] that we addressed using probabilistic record-linkage. Applying the CRC analysis to a probabilistically merged dataset led to the estimation of 781 unreported cases of TB (95%CI 618–986). This seemed to be still an overestimation of unreported cases and could be a

result of including potentially false-positive cases in the analysis, as observed in other studies.[110–112]

An important assumption for application of CRC methods is that all records should correspond to true cases.[114–117] We suspect that at least some cases observed in our study, especially those from the GDH, are false-positive—possibly non-tuberculous mycobacterial infections and/or diagnostic miscoding.[111,112] Unreported cases were more likely to have been possible cases and that most cases observed only in GDH were classified as possible, which supports that hypothesis. To address this limitation and produce a more reliable estimate of the completeness of notification, we applied the CRC analysis to a data set, including only cases that could be classified as probable or confirmed. Assuming this more specific case definition, completeness of notification to SVIG-TB in Portugal would be 77.0%, corresponding to an incidence rate of 26.1 TB cases/100 000 population in 2015 (30% above what was reported to WHO and ECDC, which was 20.1/100 000 population).

Although this is a more acceptable estimate, it could still not correspond to the truth. In a CRC study, all cases should have equal probability of being captured by any of the data sources,[116,118–121] which is not true for the GDH (as only the most severe cases of TB are admitted to hospital).[122] Furthermore, there should not be subgroups with very different probabilities of being observed in one data source and reobserved in another data source.[118,123,124] In this respect, it is reassuring that the estimated number of unreported cases did not differ after the inclusion of a number of covariates that could be potential sources of heterogeneity.

Our study brings important considerations for TB surveillance in Portugal. CRC methods provide feasible way of correcting incidence rates for undernotification,[125,126] but this should not replace the effort to improve the sensitivity of the NTP surveillance system and continuously monitor the quality of the data. CRC analysis would produce the best estimates of the true incidence of TB and could be the recommended method to produce national statistics if data sources were 100% specific. Regarding sensitivity, previous studies suggested possible reasons for under-notification of TB.[57] This was not addressed in our study but will be a research priority. From an operational research perspective, another study should be conducted to clarify ‘potential false positive cases’, to make sure that they were true cases of tuberculosis and their matching variables were recorded correctly. In the meantime, as SVIG-TB combines a relatively high sensitivity with a higher specificity, it should remain as the main data source used for TB surveillance.

2.2. COMPLETENESS OF TUBERCULOSIS (TB) NOTIFICATION: INVENTORY STUDIES AND CAPTURE-RECAPTURE ANALYSES, SIX EUROPEAN UNION COUNTRIES, 2014 TO 2016

RESEARCH

Completeness of tuberculosis (TB) notification: inventory studies and capture-recapture analyses, six European Union countries, 2014 to 2016

Masja Straetemans[1] , Mirjam I Bakker[1] , Sandra Alba[1] , Christina Mergenthaler[1] , Ente Rood[1] , Peter H Andersen[2] , Henrieke Schimmel[3] , Aleksandar Simunovic[4] , Petra Svetina[5] , Carlos Carvalho[6,7] , Outi Lyytikäinen[8] , Ibrahim Abubakar[9] , Ross J Harris[9,10] , Csaba Ködmön[11] , Marieke J van der Werf[11] , Rob van Hest[12,13]

1. KIT Royal Tropical Institute, Health Unit, Amsterdam, the Netherlands
2. Statens Serum Institute, (National Institute for Public Health), Department of Infectious Disease Epidemiology and Prevention, Copenhagen, Denmark
3. National Institute for Public Health and the Environment (RIVM), Centre for Infectious Disease Control, Bilthoven, the Netherlands
4. Croatian Institute of Public Health, Infectious Disease Epidemiology Service, Zagreb, Croatia
5. University Clinic of Pulmonary Diseases and Allergy Golnik, Department of Tuberculosis, Golnik, Slovenia
6. University of Porto, Institute of Biomedical Sciences Abel Salazar (ICBAS), Multidisciplinary Unit for Biomedical Research (UMIB), Porto, Portugal
7. Portuguese Northern Regional Health Administration, Public Health Department, Porto, Portugal
8. National Institute for Health and Welfare (THL), Helsinki, Finland
9. Institute for Global Health, University College of London (UCL), London, United Kingdom
10. Public Health England (PHE), Statistics Unit, London, United Kingdom
11. European Centre for Disease Prevention and Control (ECDC), Stockholm, Sweden
12. Department of Tuberculosis Control, Regional Public Health Service Groningen and Fryslân (GGD), Groningen, the Netherlands
13. Department of Pulmonology and Tuberculosis, University Medical Centre Groningen (UMCG), Groningen, the Netherlands

Correspondence: Masja Straetemans (m.straetemans@kit.nl)

Citation style for this article:
Straetemans Masja , Bakker Mirjam I , Alba Sandra , Mergenthaler Christina , Rood Ente , Andersen Peter H , Schimmel Henrieke , Simunovic Aleksandar , Svetina Petra , Carvalho Carlos , Lyytikäinen Outi , Abubakar Ibrahim , Harris Ross J , Ködmön Csaba , van der Werf Marieke J , van Hest Rob . Completeness of tuberculosis (TB) notification: inventory studies and capture-recapture analyses, six European Union countries, 2014 to 2016. Euro Surveill. 2020;25(12):pii=1900568. https://doi.org/10.2807/1560-7917.ES.2020.25.12.1900568

Article submitted on 11 Sep 2019 / accepted on 27 Jan 2020 / published on 26 Mar 2020

Background: Progress towards the World Health Organization's End TB Strategy is monitored by assessing tuberculosis (TB) incidence, often derived from TB notification, assuming complete case detection and reporting. This assumption is unlikely to hold in many settings, including European Union (EU) countries.**Aim:** We aimed to assess observed and estimated completeness of TB notification through inventory studies and capture–recapture (CRC) methodology in six EU countries: Croatia, Denmark, Finland, the Netherlands, Portugal Slovenia. **Methods:** We performed record linkage, case ascertainment and CRC analyses of data collected retrospectively from at least three national TB-related registers in each country between 2014 and 2016. **Results:** Observed completeness of TB notification by inventory studies was 73.9% in Croatia, 98.7% in Denmark, 83.6% in Finland, 81.6% in the Netherlands, 85.8% in Portugal and 100% in Slovenia. Subsequent CRC analysis estimated completeness of TB notification to be 98.4% in Denmark, 76.5% in Finland and 77.0% in Portugal. In Croatia, CRC analyses produced implausible results while in the Netherlands and Slovenia, it was methodologically considered not meaningful. **Conclusion:** Inventory studies and CRC methodology suggest a TB notification completeness between 73.9% and 100% in the six EU countries. Mandatory reporting by clinicians and laboratories, and cross-checking of registers, strongly contributes to accurate notification rates, but hospital episode registers likely contain a considerable proportion of false-positive TB records and are thus less useful. Further strengthening routine surveillance to count TB cases, i.e. incidence, accurately by employing record-linkage of high-quality TB registers should make CRC studies obsolete in EU countries.

Background

In 2014, the World Health Organization (WHO) published the Action Framework towards TB elimination in low-incidence countries, and in 2016, the WHO Regional Office for Europe (WHO/Europe) published the Roadmap to implement the tuberculosis action plan for the WHO European Region 2016-2020: Towards ending TB and multidrug-resistant tuberculosis. They outline blueprints to carry out the WHO's global End TB Strategy in Europe and to reach the sustainable development goal (SDG) target for tuberculosis (TB).[42,127,128] Key strategic targets include reduction of global TB incidence by 80% in 2030 and 90% in 2035, compared with 2015.[101,129] In the WHO European Region, the targets include a 25% reduction in TB incidence rate by the year 2020 compared with 2016.[42] TB incidence can be derived from TB notification rates, assuming complete case detection and reporting. This is considered a strong assumption unlikely to hold in many settings, including European Union (EU) countries. Several studies across the EU/EEA, e.g. those from France, Spain, Italy and Romania, have revealed high rates of under-reporting.[54,104]

Inventory studies (IS) are a widely accepted methodology to study the level of under-reporting of TB cases.[49] To determine completeness of TB notification, TB IS compare the number of cases meeting standard case definitions and recorded in multiple TB-related registers, such as laboratory registers or hospital episode registers, with the cases notified to local and national authorities.[130,131] For this comparison, record-linkage methodologies are used. Subsequently, through capture–recapture (CRC) analysis, the number of cases unknown to all registers can be estimated, thereby providing an estimate of the completeness of TB notification.[52]

At the time of the study and writing this article, the United Kingdom (UK) was still part of the EU and only two EU countries, the Netherlands and the UK, had assessed the level of TB under-reporting using IS/CRC studies on a national basis.[52,112] In 2016, the European Centre for Disease Prevention and Control (ECDC) commissioned an IS/CRC project to estimate TB under-notification in six to nine EU countries and expand the evidence base of the methodology.

The objective of this study was to assess observed and estimated TB notification completeness in EU countries, selected after an eligibility and feasibility appraisal, and determine whether data collected at national level reflect the real TB situation in these countries.

Methods

This study used inventory studies and capture-recapture analyses as per the methodology outlined by Van Hest and the WHO.[49,104]

Country selection process

To select the six most appropriate and representative EU and European Economic Area (EEA) countries for the study, we followed a predefined selection procedure that was developed by the consortium commissioned to perform this work and agreed upon with ECDC during the project kick-off meeting in June 2016. First, eligibility and feasibility for IS and more specifically CRC studies was assessed by an 11-item tick-box questionnaire that was distributed to the TB operational contact points for epidemiology at ECDC and WHO/Europe TB surveillance network meeting in Bratislava in June 2016 (Table 11).

Table 11. Eligibility and feasibility appraisal questionnaire with simple scoring system based on essential issues, issues "important (not) to have" (2 points) and issues "nice to have" (1 point).

	Question:	Score
Q1	Does your country have an electronic case-based TB notification database?	Essential
Q2	Are these the data you report to ECDC/TESSy?	1
Q3	Is TB notification mandatory in your country?	2
Q4	Only mandatory for the doctor or also for other, e.g. laboratory?	2
Q5	Does your country have other electronic case-based TB registers available?	Essential
Q6	Are standard TB case definitions used in these registers?	2
Q7	Do these registers contain variables that can be used for record linkage?	Essential
Q8	Does the NTP already routinely link certain TB registers?	1
Q9	Does the NTP have professionals that can assist in this study (e.g. data managers, statistician)?	2
Q10	Does your country have legal issues hindering capture-recapture studies?	Essential
Q11	Would your country be interested to participate?	Essential

TB: tuberculosis; ECDC: European Centre for Disease Prevention and Control; TESSy: The European Surveillance System; NTP: National TB Programme

Non-responders and countries not present at this meeting were invited by email to return the questionnaire by 1 August 2016. Countries who had not responded by 29 July received a reminder email from ECDC to fill out the form, but this did not yield any additional responses. Twenty forms were received from 31 countries (28 EU and 3 EEA): Belgium, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Malta, the Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, the UK and Norway. Four countries, France, Malta, the UK and Norway, indicated they did not want to participate. Germany and Ireland indicated they did not have a third case-based register or national data, which are essential requirements.

To obtain detailed information regarding TB-related registers, data availability, data access and data quality, the 14 countries identified in this rapid eligibility and feasibility assessment were contacted by 1 hour telephone or Skype calls, with a follow-up call or email for some countries. Subsequently, a simple scoring system was applied to ensure a consistent and systematic selection of countries (Table 11). A matrix of geographical location, economic status, TB incidence and proportion of foreign-born TB patients was also prepared to ensure diverse representation of countries. After the eligibility and feasibility appraisal, six countries, Croatia, Denmark, Finland, the Netherlands, Portugal and Slovenia, were selected for the IS and CRC analyses, and country-specific protocols were developed.

Case definition and inclusion criteria

This study included all cases recorded as active pulmonary or extra pulmonary TB, new as well as previously treated or diagnosed cases, in the various registers of the six EU countries in the selected study year. Double entries in each of the registers, cases caused by non-TB mycobacteria (NTM) or cases later diagnosed as not having TB were excluded. TB diagnosis was either: (i) bacteriologically confirmed by a positive culture for Mycobacterium tuberculosis complex (gold standard), (ii) bacteriologically confirmed by a positive auramine or Ziehl-Neelsen stained respiratory or non-respiratory sample and/or a positive PCR for M. tuberculosis complex, or (iii) based on clinical, radiological, histopathological or epidemiological findings with an intention to treat. Hospital episode records were included when coded with International Code of Disease (ICD)-9 code 010–018 or ICD-10 code A15-A19.[132,133]

Inclusion and period of data collection

TB cases with a date of notification, mycobacteriological confirmation, first day of hospital admission or death between 1 January and 31 December of the same study year were included. Study years varied in the countries between 2014 and 2016 (Table 12). To allow for correction of late positive results from laboratory diagnostics or late notification, the period of data collection was from 6 months before the year of study until 3 months after (6 months after in Portugal).

Table 12. Demographics of the 6 European Union Member States and main findings from inventory studies and capture-recapture analyses.

Country (study year)	**Croatia (2015)**	**Denmark (2015)**	**Finland (2014)**	**Portugal (2015)**	**Netherlands (2014)**	**Slovenia (2016)**
***n* inhabitants** [a, b]	4,225,316	5,659,715	5,451,270	10,374,822	16,829,289	2,064,188
Region	East	North	North	South	West	East
Gross Domestic Product per capita [a, b]	10,600€	48,000€	37,600€	17,400€	39,800€	19,500€
TB notifications TESSy [a] **/100K (*n* [c])**	11.6 (486)	6.3 (357)	4.8 (263)	21.0 (2,178)	4.8 (814)	5.7 (118)
TB notification national TB register	489	379	260	2,182	814	118
Estimated TB incidence [a] **/100K (n[c])**	13.0 (560)	6.5 (370)	5.3 (290)	23.0 (2,400)	5.8 (980)	6.5 (140)
Foreign-born (%)/unknown (%)[a]	15.2 / 32.9	67.8 / 0.0	33.1 /1.5	16.7 / 0.1	73.8 / 0.0	36.4 0.0
Geographic coverage	National	National	National	Islands excluded	National	National
National registers used	1. TB Notification register (NTR) 2. Mycobacterial Reference Laboratory database (NRL) 3. Hospital discharge register	1. TB notification register (MIS2) [d] 2. Mycobacterial Reference laboratory & Danish Microbiology Database (MiBa) 3. Hospital discharge register (National Patient Register)	1. National infectious disease Register (NIDR/TTR) 2. Hospital discharge register (Hilmo) 3. Primary Health Center Discharge (AvoHilmo) 4. Death register	1. TB register (SVIG TB) [i] 2. Notifiable disease [i] surveillance system (SINAVE) 3. Hospital discharge register (GDH) [i]	1. Netherlands Tuberculosis Register (NTR) 2. Mycobacteriology Reference Laboratory 3. Hospital Discharge Register	1. TB Notification Register 2. Laboratory Register 3. Hospital Register 4. Mortality Register
Record linkage	Probabilistic: First name, family name (Jaro-Wrinkler distance algorithm)	Deterministic: National Identity Code	Deterministic: National Identity Code	Probabilistic: Date of birth, sex, place of residence	Deterministic: (relaxed) combination of major and minor proxy identifiers [k]	Deterministic: Full name of patient
Statistical models	Poisson log-linear regression model 3 source data	Poisson log-linear regression model 3 source data	Poisson log-linear regression model 3 source data and 4 source data	Poisson log-linear regression model 3 source data	CRC not feasible due to complete overlap of 2 registers and 1 considered poor data quality	CRC not feasible due to almost complete overlap

Results of IS: observed completeness of notification	73.9% (489/662)	98.7% (379/384)	83.6% (260/311) [e],	85.8% (1,997/2,328)	81.6% (814/998)	100% (118/118)
Results of CRC: estimated unobserved cases (95% confidence interval)	Considered implausible result	*1 (95% CI: 0 – 3)*	117 (95% CI: 26-515)[f] / 20 (95% CI: 2-234) [g]	266 (95% CI 198 - 358)[j]	Not assessed	Not assessed
Estimated completeness of notification in %	Not assessed	98.4% (95% CI 97.9-98.7)	76.5% (95% CI 63.7-81.3) [h]	77.0% (95% CI 74.3% to 79.1%) [j]	Not assessed	Not assessed
Main challenge identified	81% (140/173) unnotified TB cases were only available in hospital register Possibly presence of false-positive TB cases in hospital discharge register Possible unmeasurable strong 3-way dependency	Direct referral of cases from one source to another conflicts with general notion of pairwise dependence that can be handled by log-linear models.	Many of the 313 (573-260) Hilmo and Avohilmo cases not notified are likely to be false positive	Probably false-positive TB cases in public health register and hospital discharge register Admission to hospital is likely related to severity of disease, potentially violating the assumption of homogeneous capture probabilities.	A check with non-TB mycobacteria register and latent TB infection data filtered 14 hospital-only cases out; the remaining 184 cases were registered in hospital discharge only which is expected to contain many false-positive records.	A priori notification and laboratory records expected to be highly interdependent, but other registers also appeared to be interdependent due to 1) Majority of TB patients starting treatment in hospital; 2) Regular mortality audits with hospital register's cause of death data
Conclusion	Impossible to interpret CRC analyses without following up on clinically diagnosed TB patients only availably in hospital register.	Systematic validation of notification and laboratory register led to improved accuracy of data.	Primary health center discharge data were found to be unreliable and follow-up of false positives is needed.	Number of unobserved cases is likely to be higher than previously thought but, due to likely presence of false-positive TB cases in hospital discharge registers, lower than estimated in this study.	Proportion of under-notification in the Netherlands in 2014 pending availability of time and resources necessary to investigate non-matching hospital-only TB cases.	Completeness of TB notification is high; only small percentage is culture negative

[a] For country-specific study year; [b] Source: Eurostat, [c] Source: Tuberculosis surveillance and monitoring in Europe 2018; absolute number as reported to TESSy; [d] MIS2 Meldesystemet for Infektions Sygdomme version 2, [e] after adjustment for 90% suspected false-positive TB cases; [f] Excluding death register and assuming that 10% of cases in primary health center and hospital discharge register are true TB cases; [g] Including death register, excluding primary health center discharge register, assumption 10% of Hilmo cases are true TB cases; [h] 4 source model under assumption that 10% of Hilmo and Avohilmo-only cases are true TB cases; [i] SVIG TB Sistema de Vigilância da Tuberculose; SINAVE Sistema Nacional de Vigilância Epidemiológica); GDH (Grupos de Diagnóstico Homogéneos);[j] After probabilistic matching and excluding possible cases; [k] Major proxy identifiers include: year or full date of birth; sex, 4 digit postal code, and minor proxy identifiers include date of notification, date of first bacteriology culture sample and date of hospital admission.

Sources used to identify tuberculosis cases

The national electronic databases with case-based records of registered TB cases used in the IS in the six EU countries are listed in Table 12.

Record linkage

Two different approaches were used to link records of the same individual in different databases.[49] Deterministic record linkage, identical identifiers across datasets for a match, was performed when a unique identifier was available (Denmark, Finland, Portugal and Slovenia). Relaxed deterministic record linkage was used in the Netherlands, allowing for slight differences between the combination of identifiers in the various registers. Probabilistic record linkage, based on a likelihood score of a match,[49] was used when a combination of proxy-identifiers, e.g. initials of the name, full date of birth, full postcode and sex of the patient, was available (Croatia, Portugal) (Table 12). In Portugal probabilistic record linkage was performed after deterministic record linkage provided unrealistic results due to errors in the data sources.

Analysis

Observed TB notification completeness was calculated as the proportion of TB cases notified among the total number of TB cases observed after record linkage through the retrospective ISs. Estimated TB notification completeness were obtained through CRC analysis. The preferred CRC method entails log-linear modelling of at least three possibly incomplete, partially overlapping and preferably, but not necessarily, independent registers. Three-source log-linear modelling is less compromised by violation of the underlying modelling assumptions (e.g. a closed and homogeneous population; independence of registers; absence of false-positive cases).[49,104,108,118,134–136] CRC aims to estimate the unobserved number of TB cases from the information provided by the overlap between the linked registers. The unobserved number of TB cases were modelled based on rates of presence in each register (main effects) and pairwise dependencies (interactions) between registers; the latter were estimated as incidence rate ratios (IRRs), where estimates > 1 indicated a positive dependence between registers. Three linked registers provide eight different possible models ranging from containing no dependencies (the base model), any possible combination of the three pairwise dependencies, to the saturated model,

containing all dependencies. The interaction between all three registers cannot be estimated and must be assumed absent. Selection of the preferred model (and thus the estimate of unobserved TB cases) was determined by balancing model fit with parsimony, based on Akaike Information Criterion (AIC) scores, coupled with knowledge of epidemiological plausibility, i.e. consistency with prior reports or estimates.[112] The estimated completeness of notification was the proportion of TB cases notified of the estimated total number of TB cases.[49,104,108,118,134–136] Poisson log-linear modelling of three independent registers was done in Croatia, Denmark and Portugal. For Finland, modelling was extended to four sources, incorporating proportions of false-positive TB cases, based on external evidence. CRC was considered not feasible in the Netherlands and Slovenia because of complete overlap of respectively two and three registers.

Study permission

National medical-ethical authority permission was obtained (Finland (Tietosuojavaltuutetun toimisto: THL/112/6.02.00/2017), the Netherlands (Data Protection Committee of the Netherlands Tuberculosis Register Number: 04-2017) and Portugal (Comissão Nacional de Protecção de Dados (CNPD) Processo no: 11543/2017)) or not required (Denmark, Croatia and Slovenia).

Results

In Croatia, record linkage of three national registers (notifications (489 cases), reference laboratory (334 cases), and hospital episodes (476 cases)) resulted in 662 registered TB cases, translating into an observed completeness of notification of 73.9% (489/662). Of the 173 TB cases not notified, 140 (81%) were only known to the hospital episode register (Figure 6).

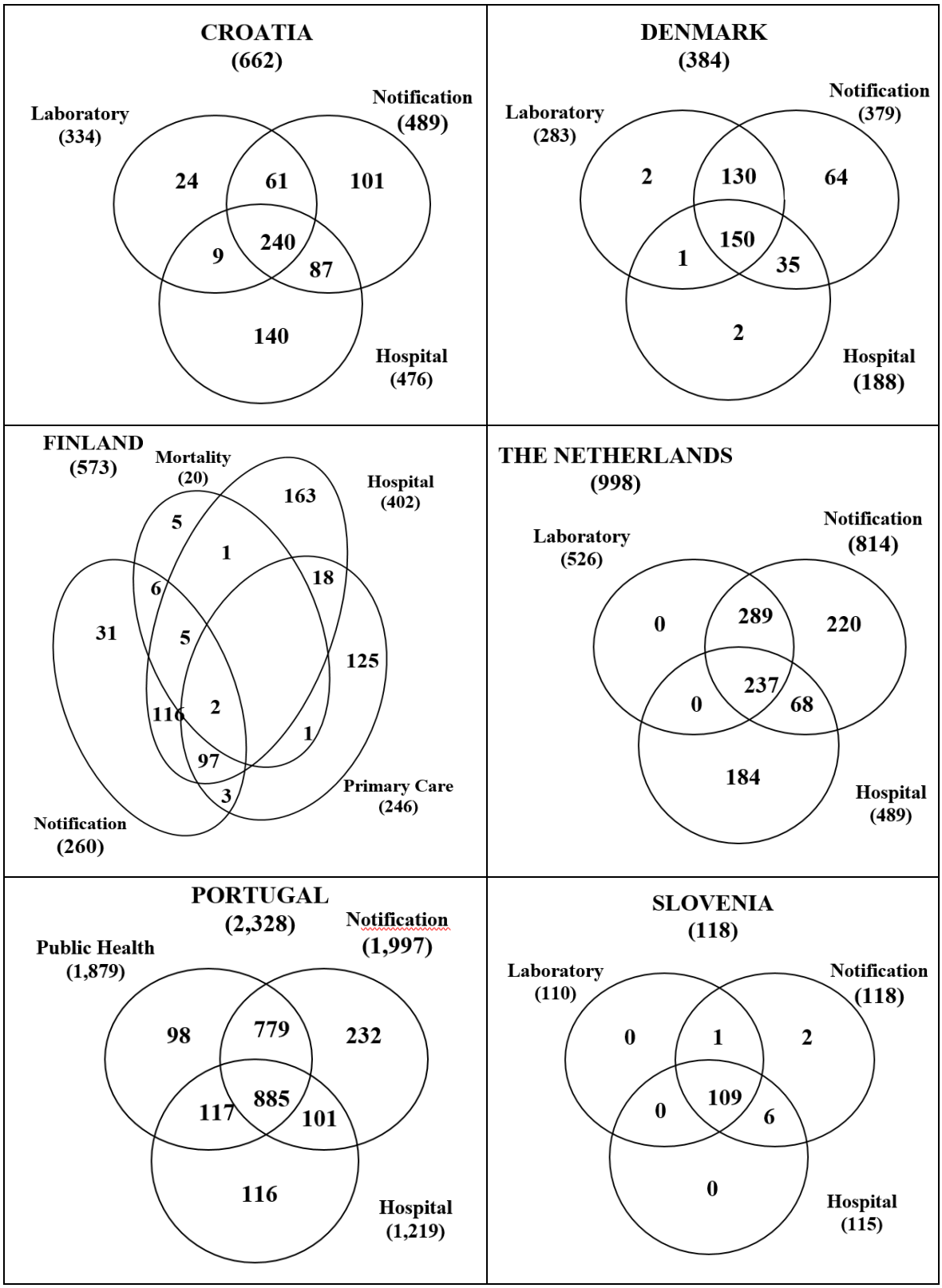


Figure 6. Schematic view of the registered number of TB cases in the six selected European Union Member States after record linkage of three TB-related registers.

Finland Venn diagram shows the number of TB cases based on the preferred four-source record-linkage adjusted for suspected 90% false positive TB cases. Portugal Venn diagram shows the number of TB cases based on probabilistic linkage and excluding possible TB cases. The circles show the number of patients in each registration and their overlap. For clarity, the Venn diagram is not projected to scale of the number of TB cases in each cell.

In CRC analysis, AIC scores favoured the saturated model, which estimated an implausibly high number of unobserved TB cases (1,705; 95% confidence interval (CI): 707–4,114).

In Denmark, record linkage of three national registers (notifications (379 cases), reference laboratory (283 cases), and hospital episodes (188 cases)) resulted in 384 registered TB cases, translating into an observed completeness of notification of 98.7% (379/384). Of the five TB cases not notified, two were only known to the hospital episode register, two to the laboratory register only and one to both registers. CRC analysis selected a model with one interaction (between the laboratory and hospital episode registers), estimating one unobserved TB case, and resulting in an estimated completeness of notification of 98.4% (379/385) (95% CI: 97.9–98.7).

In Finland, complete overlap between the reference laboratory and the notification registers was expected due to automatic registration of the laboratory results. Record linkage of the notification register (260 cases) with the hospital episode (402 cases) and primary care episode (246 cases) registers resulted in 569 registered TB cases, translating into an observed completeness of notification of 45.7% (260/569). After record linkage with a fourth register, the mortality register (20 cases), 573 TB cases were observed, translating into an observed completeness of notification of 45.4% (260/573). A large part of the TB cases not notified (573–260 = 313) appeared either only in the hospital register (52% (163/313) or in the primary care register (40% (125/313). Unpublished findings from a Finnish pilot study, brought to our attention during implementation of our research, in which hospital records of TB patients with ICD-10 codes A15-A19 were reviewed in detail, found that only 10% were true, notifiable, TB cases (personal communication, Hanna Soini, August 2019). Since it was not in the scope of our study to review clinical records, we assumed 90% of the hospital or primary care only cases to be false positive. After adjusting for these suspected false-positive TB cases, the observed completeness of notification was 83.6% (260/311). CRC analysis on the unadjusted data preferred the saturated model, which estimated an implausibly high number of unobserved TB cases in Finland in 2014, more than 15 times the number of observed TB cases. Several other models were considered, as described in detail elsewhere (country report for Finland available upon request from ECDC). The four-source CRC was the preferred analysis because it provided a very similar estimate to the three-source CRC, which included the mortality register, and it corroborated the assumption that many of the Hilmo-only or Avohilmo-only cases are false positive TB cases. The preferred four-source CRC analysis on the data adjusted for the 90% suspected false-positive TB cases estimated 29 unknown TB cases, translating into an estimated

completeness of notification of 76.5% (95% CI: 63.7–81.3%) (260 notified cases/340 estimated true TB cases).

In the Netherlands, record linkage of three national registers (notifications (814 cases), reference laboratory (526 cases), and hospital episodes (489 cases)) resulted in 998 registered TB cases in 2014, translating into an observed completeness of notification of 81.6% (814/998). All culture-confirmed laboratory TB cases were notified (observed completeness of notification of 100%). Because of this complete overlap two linked registers remained, with all TB cases not notified (n = 184; after excluding NTM (10 cases) and latent TB infections (4 cases)) only known to the hospital episode register, indicating that 49.0% (489/998) of the TB cases in the Netherlands in 2014 were (initially) hospitalised. Scrutinising the individual patient records of the cases only available in the hospital episode registry to assess whether they had signs, symptoms or diagnostic test results suggestive for TB, or whether treatment was initiated and completed, was not possible for confidentiality, logistical and financial reasons. A CRC analysis was methodologically not considered meaningful due to the complete overlap of the notification this would be a two-source CRC (see Discussion).[137]

In Portugal, three TB registers were identified in 2015: the national TB programme register (2,086 cases), the public health register (1,874 cases) and the hospital episode register (1,358 cases). Case ascertainment was 3,561 (deterministic linkage; all TB cases); 2,914 (deterministic linkage; excluding 'possible' TB cases); 2,786 (probabilistic linkage; all TB cases); and 2,328 TB cases (probabilistic linkage; excluding 'possible' TB cases); with an observed completeness of notification of 60.2% (2,144/3,561), 64.6% (1,883/2,914), 80.0% (2,228/2,786), and 85.8% (1,997/2,328), respectively. For all datasets, a parsimonious CRC model with two (positive) pairwise interactions (national TB register*public health register and public health register*hospital episode register) was selected. The preferred model estimated completeness of notification at 77.0% (95% CI: 74.3%–79.1%), after probabilistic record linkage and excluding the 'possible' TB cases (1,997/(2,328 observed TB cases and 266 unobserved TB cases)).

In Slovenia, a strong positive interaction between the notification (118 cases), reference laboratory (111 cases) and hospital episode (115 TB cases) registers was expected, as all these services are mainly concentrated in one university hospital. Record linkage of these registers with the mortality (19 cases) register resulted in a case ascertainment of 119 TB cases. Excluding one person with a positive culture of M. tuberculosis starting treatment abroad and assumed notified there, 118 TB cases were registered in Slovenia in 2016,

translating into a completeness of notification of 100%. Due to the almost complete overlap of the registers used, CRC analysis was methodologically considered not meaningful.

Discussion

Observed completeness of TB notification was calculated to be 73.9% in Croatia, 98.7% in Denmark, 83.6% in Finland, 81.6% in the Netherlands, 85.8% in Portugal, and 100% in Slovenia. Subsequent CRC analyses estimated completeness of TB notification to be 98.4% in Denmark, 76.5% in Finland and 77.0% in Portugal. In the other three countries, CRC analysis gave implausibly high estimates of unobserved TB cases (Croatia) or was not performed because it was methodologically not considered meaningful (the Netherlands, Slovenia). In the Netherlands and Slovenia only two databases could be used and using two-source capture–recapture models for epidemiological data often violates the underlying capture–recapture assumptions, resulting in biased estimates and is thus not preferred.[137] According to the latest ECDC/WHO Europe TB surveillance and monitoring report, completeness of notification was estimated as 88.5%, 94.3%, 90.8%, 88.9%, 87.4%, and 84.3% for Croatia, Denmark, Finland, the Netherlands, Portugal and Slovenia, respectively.[138]

In monitoring progress in TB control, the ultimate aim is to count TB cases (incidence) accurately through routine surveillance. When routine surveillance is not robust, alternative methods can be used to assess completeness of notification, as described elsewhere.[43] One such method is a prevalence survey (active case finding).[46] However, prevalence surveys are costly and laborious and can only be justified in countries where many cases are thought to be missed and the expected smear positive TB prevalence is high enough (≥ 100 per 100,000 population) to obtain an accurate estimate with a reasonable sample size.[46,139]

Surveillance by notification may result in inaccurate incidence rates because of under-ascertainment of the true number of cases or over-ascertainment due to false-positive cases.[139] Standardised means are needed to evaluate and adjust incidence rates.[43] IS and CRC analysis are such methods but can be difficult to implement, even in resource-rich countries.[104] We encountered various challenges when assessing completeness of TB notification in the six EU countries through retrospective IS and CRC methodology: (i) high interdependency between notification and laboratory registers (as a result of improved surveillance, sometimes by mandatory reporting, automatic record linkage, and routine

cross-checking of registers); (ii) false-positive TB cases in hospital episode (and related) registers; (iii) privacy issues and aggregated data collection, preventing the use of prescriptions and health insurance registers; and (iv) selection of the CRC model. A comprehensive discussion of violation of underlying CRC assumptions can be found elsewhere.[104,118,136]

Previous regional and national IS and CRC studies in EU countries often used the notification, reference laboratory and hospital episode registers, assuming absence of major interdependencies.[52,104,110,112,140] In recent years, many EU countries made it mandatory for the laboratory to report positive bacteriological TB test results, in addition to notification by the diagnosing clinician. This can result in a (near) complete overlap between these registers, as observed in Denmark, the Netherlands and Slovenia. Often, notification and reference laboratory registers are maintained within the same institute, with routine cross-checks between registers and sometimes with (in)direct access to other TB-related registers as well. The assumption that notification and reference laboratory registers in EU countries are relatively independent should be judged critically. In some countries, such as Denmark and Slovenia, where, by convention, almost all TB patients are (initially) hospitalised, a (near) complete overlap and interdependence between the hospital and notification registers was observed. Likewise, negative dependencies are expected when the hospital database includes clinically diagnosed cases, which by definition do not appear in the laboratory register, such as in Croatia. Violation of the assumption of homogenous capture probabilities can lead to biased estimates.

The two previous national IS and CRC studies in EU countries suggested a considerable proportion of false-positive TB cases in hospital episode registers, possibly coding a differential diagnosis upon admission or a presumptive diagnosis upon discharge.[52,112] The diagnosis of TB, a disease with a sometimes prolonged diagnostic pathway, can be withdrawn later, e.g. due to absence of positive laboratory results for TB, results indicating NTM, or a final diagnosis other than TB. However, hospital episode registers are not dynamic, with continuous updates and corrections of the ICD codes. A retrospective study conducted in Turku and Tampere university hospitals in Finland (2014–2016), scrutinising the local hospital episode registers to assess whether patients were correctly coded with an ICD code for active TB, found 90% of the records to be false-positive (i.e. patients ultimately not diagnosed with and offered complete treatment for TB) (personal communication, Hanna Soini, August 2019). Based on this knowledge, a correction for (assumed) false-positive TB cases could be made, resulting in a dramatically reduced and more realistic CRC estimate of the unregistered TB cases in Finland. In Portugal, similar outcomes were

obtained after correction for 'possible' TB cases. When (nearly) all records only known to the hospital episode register in the Netherlands were considered false-positive, the observed completeness of notification was (nearly) 100% and, as a result of this correction, 37.5% of the TB cases in the Netherlands were (initially) hospitalised in 2014. This is in line with historical observations that around one third of the TB patients in the Netherlands are hospitalised for more than one week, 247 cases (30%) reported through routine surveillance in 2014.[141] In Croatia and the Netherlands, comparison of the hospital episode records with clinical information was not possible. Routine review of national hospital episode registers against patient records, in consultation with the clinicians, e.g. half a year after entry, could possibly increase the proportion of true-positive TB cases, possibly reducing over-estimation of the number of TB cases.

Alternative registers that could be used in some EU countries are mortality registers, when not nested in other registers. Other registers include health insurance registers and prescription registers for anti-TB medication, most specifically pyrazinamide, since other first-line drugs are also prescribed for different conditions, such as latent TB infection and treatment of NTM. Possible problems with these registers could be collection of only aggregate data, unsuitable for record linkage, or unavailability of data due to privacy regulations, such as the EU General Data Protection Regulation (GDPR),[142] which came into effect on 25 May 2018. Prescription registers could still be used for triangulation of IS and CRC results through pharmaco-epidemiology by estimating the number of TB cases based on the daily dose equivalent of pyrazinamide.[143] In our study, a limited number of registers were available for linkage: mortality registers only contained few TB-related records and prescription and health insurance registers were not suitable in existing formats or under current regulations.

In some studies, the saturated CRC model, i.e. the model incorporating all pairwise interdependencies but also the highest order interdependency between all registers which cannot be controlled, was preferred based on the lowest AIC.[135] The literature generally recommends extreme caution when this model is chosen as one or more assumptions underlying the CRC analysis are likely violated, making statistical inferences unreliable.[118] Further work would be beneficial, including alternative structural models for the data-generating process that do not rely on log-linear assumptions,[144] Bayesian models incorporating prior uncertainty on proportions of false-positive cases,[145] data from multiple years to account for individuals observed in different years, or the incorporation of covariate data,[146] as attempted in Portugal.

Given the intrinsically complex nature of TB epidemiology, variations in diagnoses (e.g. bacteriologically or clinically confirmed), socioeconomic background and disease severity, the assumption of homogenous probabilities of observations is unlikely to hold. Log-linear models may partially incorporate population heterogeneity, especially if relevant covariate information explaining variation is available, but the problem is likely persistent to some extent and a limitation of CRC.[118] The limited number of TB registers available and the absence of data to assess variations in individual detection probabilities in different countries did not allow to identify how heterogeneity affects population estimates across different settings. Hence, the accuracy of the final population estimates could only be assessed based on expert opinion, critical scrutiny of the nature of various sources and the quality of the data used in the analyses.

Another limitation of this study is that the six countries were selected after an eligibility and feasibility appraisal, with likely selection bias, limiting extrapolation of the results to other EU countries. The six countries selected were predominantly countries with a small population, a limited number of TB cases and a well-organised TB notification and control system.

Conclusion

Our analysis suggests that only in Denmark and Slovenia the assumption that TB notification reflects TB incidence is likely to be true. In these countries, the estimated number of unobserved TB cases was minimal (Denmark) or expected to be minimal because of complete overlap of TB cases between registers (Slovenia). In the other countries, observations and estimates were more difficult to interpret.

Mandatory reporting by both clinicians and laboratories, and cross-checking of registers, strongly contributes to accurate notification rates but hospital episode registers likely contain a considerable proportion of false-positive TB records and more scrutiny is needed. Further strengthening routine (computerised) infectious disease surveillance systems to count TB cases (incidence) accurately, by structurally employing electronic record linkage of high-quality TB registers, should make CRC studies obsolete in EU countries.

3

USEFULNESS OF THE TUBERCULOSIS SURVEILLANCE SYSTEM

A surveillance system is useful if it generates a public health response that leads to improved control and prevention of adverse health events or to a better understanding of the process that leads to adverse outcomes.[147]

[39]The seven articles presented in this chapter, published in indexed international peer-reviewed journals, illustrate the usefulness of the Portuguese National Tuberculosis Programme surveillance system while covering its surveillance objectives:[33] 1) Monitoring of TB notification rates in the country, supporting regular production of regional and national TB epidemiological reports; 2) Analysis of clinical data and production and dissemination of evidence in the areas of prevention, diagnosis and treatment of TB; 3) Supporting implementation and monitor TB-specific strategies, such as the Directly Observed Therapy Short-course (DOTS), drug-resistant TB prevention and control, roll-out and evaluation of newer diagnosis tests, molecular surveillance and monitoring of implementation of national guidelines on TB.

[3] Rito T, Matos C, **Carvalho C**, Machado H, Rodrigues G, Oliveira O, et al. A complex scenario of tuberculosis transmission is revealed through genetic and epidemiological surveys in Porto. BMC Infect Dis. 2018;18(1):53. (Impact factor Web of Science: 3.1)

[4] Oliveira O, Gaio R, **Carvalho C**, Correia-Neves M, Duarte R, Rito T. A nationwide study of multidrug-resistant tuberculosis in Portugal 2014–2017 using epidemiological and molecular clustering analyses. BMC Infect Dis. 2019; 19:567. (Impact factor Web of Science: 3.1)

[5] Sentís A, Vasconcelos P, Machado RS, Caylà JA, Guxens M, Peixoto V, Duarte R, Carvalho I, **Carvalho C**. Failure to complete treatment for latent tuberculosis infection in Portugal, 2013–2017: geographic-, sociodemographic-, and medical-associated factors. Eur J Clin Microbiol Infect Dis. 2020;39:647–656. (Impact factor Web of Science: 3.3)

[6] Sousa S, Rocha D, Silva JC, Ribeiro AI, Gonçalves G, Almeida Á, Correia AM, Duarte R, **Carvalho C**. Comparing the cost-effectiveness of two screening strategies for latent tuberculosis infection in Portugal. Pulmonology 2021;27(6):493–499. (Impact factor Web of Science: 3.6)

[7] Sentís A, Prats-Uribe A, Peixoto VR, Caylà JA, Gomes MD, Sousa S, Duarte R, Carvalho I, **Carvalho C**. Decline of tuberculosis notification rate in different populations and regions in Portugal, 2010–2017. Pulmonology 2021 (published online ahead of print). (Impact factor Web of Science: 3.6)

[8] Sousa S, Macedo R, Alves CM, **Carvalho C**, Goncalves G, Duarte R. Coffe shops, a hub for TB clusters?. Pulmonology 2023 (published online ahead of print). (Impact factor Web of Science: 9.2)

[9] Sousa S, Alves CM, Macedo R, **Carvalho C**, Goncalves G, Duarte R. An investigation of TB infection and reinfection among stone quarry workers. Pulmonology 2023 (published online ahead of print). (Impact factor Web of Science: 9.2)

3.1. A COMPLEX SCENARIO OF TUBERCULOSIS TRANSMISSION IS REVEALED THROUGH GENETIC AND EPIDEMIOLOGICAL SURVEYS IN PORTO

Rito *et al. BMC Infectious Diseases* (2018) 18:53
DOI 10.1186/s12879-018-2968-1

BMC Infectious Diseases

RESEARCH ARTICLE Open Access

A complex scenario of tuberculosis transmission is revealed through genetic and epidemiological surveys in Porto

Teresa Rito[1,2*], Carlos Matos[3], Carlos Carvalho[4,5], Henrique Machado[1,2], Gabriela Rodrigues[4], Olena Oliveira[6], Eduarda Ferreira[3], Jorge Gonçalves[7], Lurdes Maio[3], Clara Morais[8], Helena Ramos[9], João Tiago Guimarães[6,10,11], Catarina L. Santos[1,2], Raquel Duarte[6,12,13] and Margarida Correia-Neves[1,2]

Abstract

Background: Tuberculosis (TB) incidence is decreasing worldwide and eradication is becoming plausible. In low-incidence countries, intervention on migrant populations is considered one of the most important strategies for elimination. However, such measures are inappropriate in European areas where TB is largely endemic, such as Porto in Portugal. We aim to understand transmission chains in Porto through a genetic characterization of *Mycobacterium tuberculosis* strains and through a detailed epidemiological evaluation of cases.

Methods: We genotyped the *M. tuberculosis* strains using the MIRU-VNTR system. We performed an evolutionary reconstruction of the genotypes with median networks, used in this context for the first time. TB cases from a period of two years were evaluated combining genetic, epidemiological and georeferencing information.

Results: The data reveal a unique complex scenario in Porto where the autochthonous population acts as a genetic reservoir of *M. tuberculosis* diversity with discreet episodes of transmission, mostly undetected using classical epidemiology alone.

Conclusions: Although control policies have been successful in decreasing incidence in Porto, the discerned complexity suggests that, for elimination to be a realistic goal, strategies need to be adjusted and coupled with a continuous genetic characterization of strains and detailed epidemiological evaluation, in order to successfully identify and interrupt transmission chains.

Keywords: *Mycobacterium tuberculosis*, Genotyping techniques, Epidemiology of tuberculosis, Phylogenetic analysis, Public health

* Correspondence: teresarito@med.uminho.pt
[1]Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Gualtar Campus, 4710-057 Braga, Portugal
[2]ICVS/3B's, PT Government Associate Laboratory, Braga/4805-017, 4710-057 Guimarães, Portugal
Full list of author information is available at the end of the article

Background

Today, tuberculosis (TB) still ranks as one of the deadliest infectious diseases worldwide.[148] Although it has been consistently declining in Europe, some urban areas still display a high incidence.[103] This is the case for the city of Porto in Portugal where, in 2014, the reported incidence rate was 46.3 cases per 100,000 inhabitants.[149] This higher incidence in European urban areas has been associated with disproportionately affected sub-populations, including immigrants from high TB incidence countries, people who are HIV-infected, drug users, high alcohol consumers, prisoners and the homeless.[150] Current guidelines suggest that control and management efforts should target these particular groups, which are not only more likely to develop TB but also to transmit it to the remaining population.[151] In Porto, common risk factors such as HIV-infection and immigration are less frequent amongst TB patients than elsewhere,[152,153] but it remains necessary to understand their relevance within the overall transmission scenario.

In the TB-endemic Porto urban area, there has been a continuous reduction of TB incidence from 81 to 38 cases per 100,000 people in one decade (2002–2012)[153] leading to a very recent shift from relatively high to intermediate TB incidence. However, while current strategies and efforts have therefore been effective in decreasing incidence,[149,154] the reduction rate has been slowing down and renewed efforts are necessary to achieve disease elimination. Illustrating the difficulty in establishing such strategies, the World Health Organization does not foresee an elimination of TB in low incidence European countries before 2050.[151] In this scenario, the Porto urban area could potentially be a useful model with which to study and develop control strategies in European endemic urban areas. Disruption of transmission chains is essential for TB control.[155] These chains can be identified by classical epidemiological analysis complemented by detailed genetic characterization. It is essential that we understand whether new TB cases originate from small-scale outbreaks of a few strains of the causative agent Mycobacterium tuberculosis or through constant reappearance and low-level transmission, with Porto's population acting as a large genetic reservoir of M. tuberculosis strains.

Such an understanding is unachievable without a detailed genetic characterization of M. tuberculosis with a high level of lineage discrimination, which has yet to be done. The major Portugal-based genetic study, which included Porto, focused only on the classification of each strain into a branch of the geographically-clustered global M. tuberculosis tree[156] through a set of lineage defining SNPs[157], which were largely uninformative in detecting transmission chains.

One of the most commonly used genotyping systems for molecular epidemiology is the Mycobacterial Interspersed Repetitive Units Variable Number of Tandem Repeats (MIRU-VNTR 24 loci). MIRU provides sufficient detail to discriminate isolates through 24 fast-evolving markers of size variation,[158,159] whilst also providing an approximate strain classification in the global M. tuberculosis tree.[155,156] 24-loci MIRU-VNTR also has the advantages of genotyping simplicity and low cost, which renders it the current gold standard for large population-based TB transmission studies and molecular epidemiological studies worldwide.[160,161]

In this study, integrated within a larger public health initiative that aims to develop strategies for eliminating TB in the Porto urban area, we obtained for the first time a detailed genetic characterisation of M. tuberculosis in Porto over a two-year period. In order to detect probable clusters of transmission, we genotyped the M. tuberculosis isolates in the Porto urban area corresponding to laboratory-confirmed TB cases in 2014–2015. We introduced an evolutionary algorithm based on median networks, commonly employed in evolutionary research but used here for the first time in the context of 24-loci MIRU-VNTR, which has several advantages over previously used clustering algorithms. Following this approach, we should be able to address the following points: a) does genotypic data support current routine epidemiological investigations for detecting transmission chains in Porto; b) can molecular epidemiology improve our knowledge on the transmission scenario in Porto; and c) can the newly generated knowledge be used to improve TB control strategies in Porto?

Methods

Specimen collection and epidemiological interviews

Samples correspond to TB cases reported in 2014 and 2015 for the three health institutions in Porto where TB is diagnosed, treated and reported: São João Hospital Centre, Porto Hospital Centre and Porto TB Outpatient Centre. A total of 144 isolates, corresponding to the cases confirmed by laboratory growth, were genotyped, 89 from patients living in the city of Porto and the remaining 55 living in the suburban area of Porto.

No personal identifiable information was used. To ensure confidentiality, each case was anonymized by the assignment of a random identification number that can only be accessed by authorized public health professionals. This work was carried out in accordance with the recommendations by the Ethics Sub-commission of Life and Health Sciences (SECVS) from

the University of Minho (SECVS 135/2015), by the Health Ethics Committee of the ARSN (Northern Region Health Administration) (68/2014) and the Ethics Committee for Health of the São João Hospital Centre (CES-305/15), with written informed consent from all subjects. All procedures were in accordance with the ethical standards of the responsible committees and with the Helsinki Declaration, as revised in 2008. Information collected included age, place of residence, workplace and commonly frequented places, HIV infection status, whether or not homeless and migratory status. This information is routinely collected by public health teams whenever a case of TB is notified, and electronically recorded in the National Epidemiological Surveillance System. To each case, a specific GPS coordinate was attributed, which was never cross-checked for a concrete address but used solely to determine distances between cases to categorise putative transmission events.

Genotyping

We performed DNA extraction following Supply and colleagues.[158] All samples were genotyped using the MIRU-VNTR 24 loci kit that includes 24 size-variation markers. MIRU locus designations were MIRU_154, MIRU_424, MIRU_577, MIRU_580, MIRU_802, MIRU_960, MIRU_1644, MIRU_1955, MIRU_2059, MIRU_2165, MIRU_2347, MIRU_2401, MIRU_2461, MIRU_2531, MIRU_2687, MIRU_2996, MIRU_3007, MIRU_3171, MIRU_3192, MIRU_3690, MIRU_4052, MIRU_4156, MIRU_4348 and MIRU_2163b. This genotyping was performed by the company Genoscreen, France.

The genotypic profiles were reported as a series of 24 numbers corresponding to the number of alleles at each of the 24 loci.

Evolutionary analysis

We established the evolution of lineages in the 24 markers under a character-based evolutionary model. We employed median networks, commonly used in evolutionary studies but never applied to MIRU-VNTR data, which allow a most-parsimonious reconstruction of evolutionary relationships between genotypes and inference of ancestral or unsampled strains,[162] in contrast with a minimum-spanning tree.[163] We used two algorithms consecutively implemented in the Network 5 software (freely available at http://www.fluxus-engineering.com), the reduced-median algorithm[164] followed by the median-joining

algorithm[162] as suggested by the developers for size-variation data. We used the default values of the network software for both algorithms (threshold of 2 for the reduced median and epsilon of zero for the median joining algorithm). This hybrid approach allows for the median joining algorithm to be performed on the previously established reduced median matrix allowing a simplification of the data. For a more accurate and reliable reconstruction, we weighted the 24 loci according to their allelic diversity in a large dataset of 494 M. tuberculosis isolates.[158] We modified the initial standard weight of 10 for all loci in the Network software to a scale between 5 and 15 through direct comparison with the allelic diversity.[158] The lowest diversity in MIRU_4052 was considered zero and that diversity was subtracted to the diversity of the other markers.

This diversity was linearized from 0 (MIRU_4052) to 10 in the less diverse marker (MIRU_0154). Values were then rounded to the closest integer and inverted as 0 became 10 and 10 became 0. A value of 5 was added to place the weighting around the default value of 10, leading to a weight of 15 in the low-diversity marker MIRU_0154 and a weight of 5 in markers like MIRU_4052 and MIRU_2163b, as they are likely the less phylogenetically informative (Figure 7).

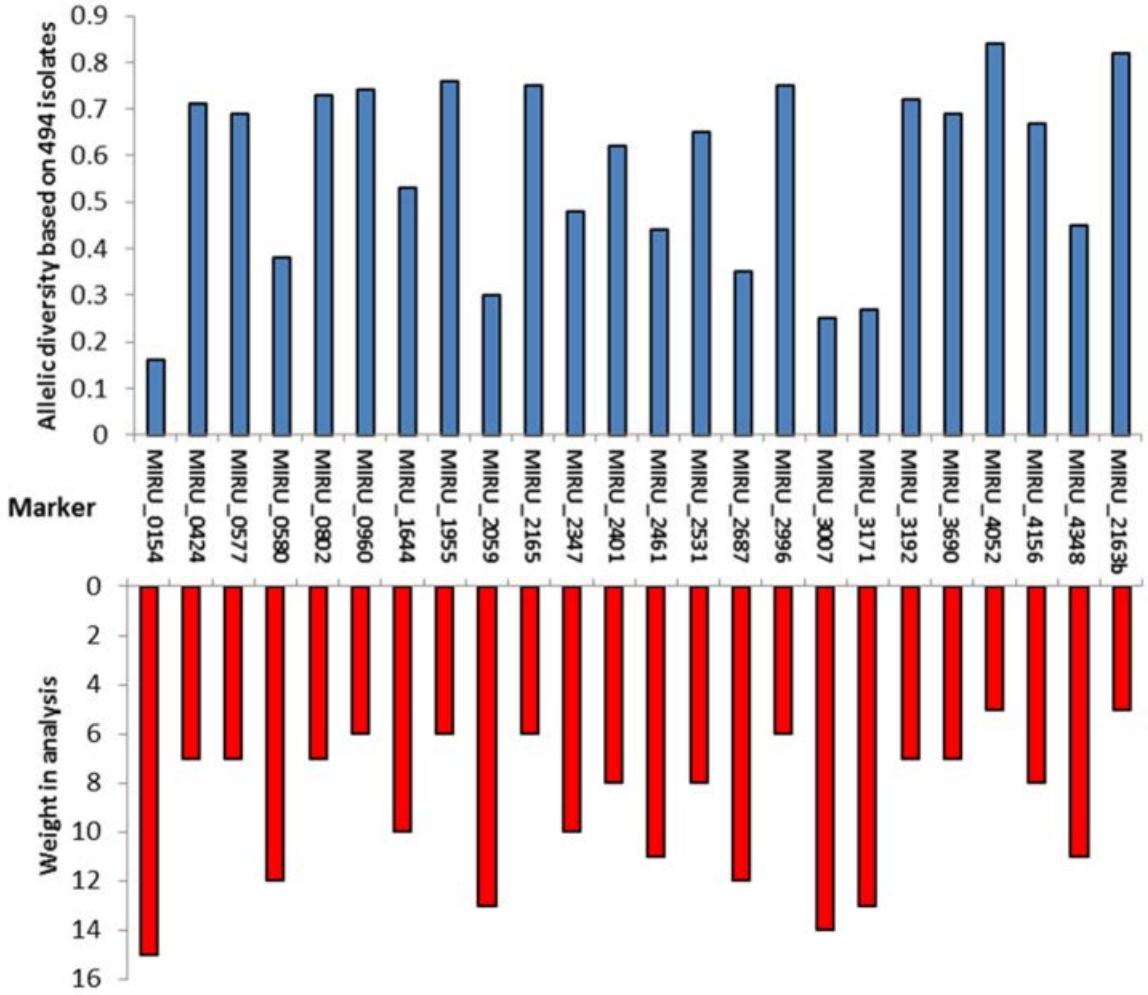


Figure 7. Conversion of allelic diversity of markers in MIRU-VNTR 24 loci kit into weight for the median vectors' analysis.
Top graphic (in blue) corresponds to the allelic diversity in each VNTR as calculated by Supply et al., 2006 and the lower graphic (in red) corresponds to the linear conversion of this diversity into weights from 5 to 15 in the network software.

We determined the main geographical lineages for the M. tuberculosis isolates by comparing the genotyped samples and the MIRU-VNTR reference strains whose lineage is well determined in the MIRU-VNTRplus website http://www.miru-vntrplus.org.

Statistical analysis

Isolates that share common genetic ancestry are referred to as genotypically clustered, a potential marker for episodes of direct TB transmission. Following the initial network that displayed a reasonably stable evolution of the 24 loci, we opted for defining clusters as

groups of samples that differed up to two mutations relative to the putative ancestor (root type) genotype of that cluster, as used by Walker and colleagues [165]. This allows us to conservatively account for random mutations occurring between transmission episodes, considering that VNTRs are fast evolving markers that could render real connections to be excluded. In addition, we wanted to account for hypothetical occasional mistyping errors. Although generally only similar genotypes are more meaningful to detect direct transmissions (for example as in [166]), we also opted for this more conserved approach as a few decades-old connections can be representative of local evolution and transmissions taking place within specific communities/localities in TB-endemic Porto which is highly relevant for public health.

A pairwise genetic distance (defined as the number of mutations observed in the tree between two lineages) was established for every two possible pairs of genotypes in the analysis. It allows for all the data to be analysed in an unbiased way in the sense that the pairwise genetic distance between genotypes is a direct measure on their probability of relatedness with zero differences (same genotype) representing the higher probability of representing a transmission event.

A geographical distance between every combination of two cases was established using georeferencing data for place of residence, work and common movements. One of the calculated distances was established between places of residence between each pair in the

sense that a geographical distance of zero Km, especially coupled with a genetic distance of zero in the genotypes would correspond likely to a household transmission. A second measure was established considered the lowest possible geographic distance between each pair considering the provided information by the patient. The analysis of the georeferencing and epidemiological data was performed by authorized public health professionals.

Results

Phylogenetic reconstruction of M. Tuberculosis MIRU-VNTR genotypes displays both close and deeper evolutionary relationships

The collection of M. tuberculosis isolates from the Porto urban area during 2014–2015 yielded 144 samples. The basic characterization of the infected population is depicted on Table 13.

Table 13. Basic sociodemographic information data on the 144 cases with tuberculosis.

		Study population
Sample size (n)		144
Males (n)		96
Age (years old, average)		49.0
HIV positivity (n)		9 (2 NA)
Homeless (n)		11 (2 NA)
Immigrants (n)		10 (1 NA)
Porto city (n)		89
Localities within Porto city (n)	Aldoar, Foz and Nevogilde	9
	Bonfim	10
	Campanhã	15
	Cedofeita, S. Ildefonso, Sé, S. Nicolau, Miragaia and Vitória	23
	Lordelo Ouro and Massarelos	6
	Paranhos	19
	Ramalde	7
Porto suburban area (n)		55

NA: epidemiological data not available

In summary, 61.8% of the samples collected were isolated from individuals registered as living in the city of Porto, whereas 38.2% were from individuals registered as living in the surroundings of the city (here called the Porto suburban area).

We genotyped all samples with the 24-loci MIRU-VNTR approach (Table 14), and submitted the 144 resulting M. tuberculosis genotypes to a character based phylogenetic tree (Figure 8).

Table 14. MIRU-24 loci VNTR profiles of the 144 Mycobacterium tuberculosis case isolates

Nr	Laboratory sample ID	Sample (n)	Cluster in analysis	Mtb global sub-lineage	MIRU-VNTR *Loci* 0154,0424,0577,0580,0802,0960,1644, 1955, 2059,2165,2347,2401,2461,2531, 2687,2996, 3007,3171,3192,3690,4052, 4156,4348,2163b
1	**1G9**	1	Unique	EAI	2,2,2,5,3,4,3,11,2,9,3,2,4,6,2,2,3,3,5,7,3,1,1,2
2	**3F6**	1	Unique	EAI	2,2,4,5,3,4,3,9,1,6,3,2,4,6,2,2,3,1,5,2,5,1,1,9
3	**2G3**	1	XXVII	ND	2,2,4,2,3,3,3,2,2,3,4,2,2,6,1,5,3,3,3,3,5,2,2,2
4	**3D6**	1	XXVII	ND	2,1,4,2,3,3,3,1,2,3,4,2,2,6,1,5,3,3,3,3,5,2,2,3
5	**3A1**	1	Unique	ND	2,2,4,2,1,3,3,2,2,4,4,2,2,5,1,5,3,3,3,3,6,1,2,3
6	**3B6**	1	Unique	ND	2,2,4,2,4,3,2,2,2,3,4,2,2,6,1,5,3,3,3,3,5,2,2,3
7	**3H4**	1	Unique	ND	2,2,4,2,4,3,1,2,2,3,4,2,2,5,1,5,3,3,3,7,5,2,2,4
8	**3D3**	1	Unique	ND	2,2,6,2,2,3,4,2,2,2,4,2,2,5,1,5,3,3,3,3,5,2,2,3
9	**2G8, 3E2**	2	XXII	LAM	2,4,4,2,1,3,2,3,2,2,4,1,1,6,1,4,3,3,3,2,8,2,2,4
10	**1B2, 2B7, 2F5, 2G9**	4	XXIII	LAM	2,4,4,2,1,3,2,3,2,2,4,1,1,6,1,4,3,5,3,2,8,2,2,3
11	**1E4**	1	XXIII	LAM	2,4,3,2,1,3,2,3,2,2,4,1,1,6,1,4,3,5,3,2,8,2,2,3
12	**2H2, 2H6, 3C1, 3C6, 3F3, 3G4**	6	XXIII	LAM	2,4,4,2,1,3,2,3,2,2,4,1,1,6,1,4,3,5,3,2,8,2,2,4
13	**1I8**	1	XXIII	LAM	2,4,4,2,1,3,2,3,2,2,4,1,1,6,1,4,3,5,3,2,9,2,2,4
14	**2C1**	1	Unique	LAM	2,4,4,2,1,3,2,3,2,2,4,1,1,6,1,4,3,5,3,2,8,0,2,4
15	**3E1**	1	XXIV	LAM	2,4,4,2,1,3,2,3,2,2,4,1,1,6,1,4,3,5,3,2,5,2,2,4
16	**2E2**	1	XXIV	LAM	2,4,4,2,1,3,2,3,2,2,4,1,1,6,1,4,3,5,3,2,4,2,2,4
17	**2E5**	1	Unique	LAM	2,4,4,2,1,3,1,3,2,2,4,1,1,6,1,5,3,5,3,2,5,2,2,3
18	**3D5**	1	Unique	LAM	2,4,4,2,1,3,1,3,2,2,4,1,1,6,1,4,2,5,3,2,6,2,2,4
19	**1A9, 3G2**	2	XXVI	LAM	2,4,5,2,1,3,1,3,2,2,4,1,1,4,1,3,2,5,3,2,3,2,2,2
20	**2C9**	1	Unique	LAM	2,4,4,2,1,2,1,3,2,2,4,1,1,4,1,4,2,5,3,2,7,2,2,4
21	**3F9**	1	Unique	LAM	2,4,4,2,1,3,1,3,2,2,4,1,1,6,1,4,2,6,3,2,8,2,2,2
22	**2G6, 3F2**	2	XXV	LAM	2,4,4,2,1,4,1,3,2,2,4,1,1,6,1,5,2,5,3,2,12,2,2,4
23	**3G9**	1	XXV	LAM	2,4,4,2,1,3,1,3,2,2,4,1,1,6,1,5,2,6,3,2,12,2,2,3
24	**2B9**	1	Unique	LAM	2,4,4,2,1,4,2,3,2,2,4,1,1,6,1,4,3,4,3,2,5,2,2,4
25	**3D8**	1	Unique	LAM	2,5,4,4,1,4,2,3,2,2,4,1,1,7,1,5,3,5,3,2,7,2,2,3
26	**2E9, 2I8**	2	XXI	LAM	2,5,4,2,1,3,2,3,2,2,4,1,2,7,1,6,3,3,3,2,7,2,2,2
27	**3C7**	1	Unique	LAM	2,4,4,2,1,3,2,3,2,2,5,1,2,6,1,2,3,3,3,2,7,2,2,2
28	**3E8**	1	Unique	LAM	3,3,4,2,1,3,2,3,2,2,4,1,2,6,1,2,3,3,2,2,5,2,2,4
29	**2D1**	1	XX	LAM	2,2,4,2,1,3,2,3,2,2,5,1,2,6,1,5,3,3,3,2,4,2,2,2

Nr	Laboratory sample ID	Sample (n)	Cluster in analysis	Mtb global sub-lineage	MIRU-VNTR *Loci* 0154,0424,0577,0580,0802,0960,1644, 1955, 2059,2165,2347,2401,2461,2531, 2687,2996, 3007,3171,3192,3690,4052, 4156,4348,2163b
30	**1A3, 2D4, 2D9, 2F2, 3D4, 3G6**	6	XX	LAM	2,1,4,2,1,3,2,3,2,2,5,1,2,6,1,5,3,3,3,2,4,2,2,2
31	**1D7, 1G3, 2H9, 3C4, 3H5**	5	XX	LAM	2,1,4,2,1,3,2,3,2,2,5,1,2,6,1,4,3,3,3,2,4,2,2,2
32	**3E9**	1	Unique	LAM	2,4,4,2,1,3,2,3,1,2,5,1,2,6,1,7,3,3,3,2,4,2,2,2
33	**3B4**	1	XIX	LAM	1,3,6,2,4,4,3,3,2,2,4,1,2,6,1,5,3,3,2,2,8,2,2,2
34	**1I5, 2E8, 2I5, 2I7, 3C8, 3D9**	6	XIX	LAM	1,3,6,2,4,4,3,3,2,2,4,1,2,6,1,5,3,3,2,2,6,2,2,2
35	**1I1**	1	Unique	LAM	1,3,6,3,4,4,2,3,2,2,4,1,2,6,1,4,3,3,2,2,7,2,2,2
36	**1E8, 1G1**	2	XVIII	LAM	1,3,4,2,5,2,3,3,2,2,4,1,2,6,1,5,3,3,2,2,8,2,2,2
37	**1A8**	1	XVIII	LAM	1,3,4,2,5,2,3,3,2,2,4,1,2,6,1,5,3,3,2,2,7,2,2,3
38	**1I7**	1	XVIII	LAM	1,3,4,2,5,2,3,3,2,2,4,1,2,6,1,5,3,3,2,2,6,2,2,2
39	**3F1**	1	XVII	LAM	1,3,4,2,6,4,3,3,2,2,4,1,2,6,1,5,3,3,2,2,6,2,2,2
40	**3D2**	1	XVII	LAM	1,4,4,2,5,4,3,3,2,2,4,1,2,6,1,5,3,3,2,2,7,2,2,2
41	**2C4**	1	Unique	LAM	1,3,4,2,4,7,3,3,2,2,4,1,2,6,1,4,3,3,2,2,7,2,2,2
42	**3A5**	1	Unique	LAM	1,3,3,2,4,4,3,3,2,2,4,1,2,7,1,5,3,3,2,2,7,2,2,2
43	**3A9, 3C9, 3G7**	3	XVI	LAM	1,3,3,2,2,4,3,3,2,2,4,1,2,6,1,5,3,3,3,1,7,3,2,2
44	**1A5**	1	XVI	LAM	1,3,3,1,2,4,3,3,2,2,4,1,2,6,1,5,3,3,3,1,7,3,2,2
45	**3H3**	1	XV	LAM	2,5,4,2,6,4,3,1,1,2,4,1,2,6,1,4,3,3,2,2,6,2,2,2
46	**2F4, 3H7**	2	XV	LAM	2,4,4,2,8,4,3,1,1,2,4,1,2,6,1,4,3,3,2,2,6,2,2,2
47	**1G2**	1	Unique	LAM	2,3,4,2,3,4,3,3,2,2,4,2,2,6,1,5,4,1,3,1,7,2,2,3
48	**2E1**	1	Unique	LAM	2,3,4,2,3,4,3,2,2,2,4,2,2,6,1,3,3,1,3,1,7,2,2,4
49	**2H8**	1	Unique	LAM	2,3,4,2,4,4,3,3,2,2,4,2,2,6,1,5,3,1,3,1,5,2,2,3
50	**3H1**	1	Unique	LAM	2,2,4,2,3,3,3,3,2,2,4,2,2,6,1,6,3,1,3,1,7,2,2,2
51	**1H3, 2B2**	2	XI	LAM	2,3,4,2,3,4,3,3,2,2,4,2,2,6,1,5,3,1,3,1,11,2,2,2
52	**3G5**	1	XIV	LAM	2,3,2,2,3,4,3,3,2,2,4,2,2,6,1,5,3,1,3,1,3,2,2,4
53	**1B3, 1E6, 2I3**	3	XIV	LAM	2,3,2,2,3,4,3,3,2,2,4,1,2,6,1,5,3,1,3,1,4,2,2,3
54	**3E3**	1	XIII	LAM	2,4,2,2,3,4,2,3,2,2,3,2,2,6,1,5,3,1,3,1,9,2,2,4
55	**1E1, 2B1, 2B4, 3B9, 3H2**	5	XIII	LAM	2,4,2,2,3,4,2,3,2,2,3,2,2,6,1,5,3,1,3,1,10,2,2,4
56	**1A1, 1B8, 2F7, 3B2**	4	XII	LAM	2,3,2,2,3,4,3,3,2,2,4,2,2,6,1,5,3,1,3,1,8,2,2,2
57	**3E5**	1	Unique	Beijing	2,4,5,2,3,3,3,5,2,4,4,4,2,5,1,5,3,3,5,3,8,2,3,6
58	**1B1**	1	I	Beijing	2,4,4,2,3,3,3,5,2,4,4,4,2,5,1,7,3,3,5,3,8,2,3,6
59	**1E9**	1	I	Beijing	2,4,4,2,3,3,3,5,2,4,4,4,2,5,1,7,3,3,4,3,9,2,3,6
60	**1F5**	1	Unique	Haarlem	2,2,3,2,3,5,3,3,2,3,4,4,2,5,1,5,3,3,3,2,3,3,2,5

Nr	Laboratory sample ID	Sample (n)	Cluster in analysis	Mtb global sub-lineage	MIRU-VNTR *Loci* 0154,0424,0577,0580,0802,0960,1644, 1955, 2059,2165,2347,2401,2461,2531, 2687,2996, 3007,3171,3192,3690,4052, 4156,4348,2163b
61	**2B5, 2F8**	2	II	Haarlem	2,2,3,2,1,2,3,3,2,3,4,4,2,5,1,4,3,3,2,2,6,3,2,2
62	**2G7, 3B7**	2	IV	Haarlem	2,2,3,2,4,5,3,3,2,3,2,3,2,3,1,5,3,3,3,3,7,3,2,3
63	**1C3, 1C4, 1C6, 1F8, 2F9**	5	III	Haarlem	2,2,3,2,3,5,3,3,1,3,2,4,2,3,1,5,3,3,2,3,7,3,2,5
64	**3E4**	1	Unique	Haarlem	2,2,3,2,3,5,3,2,1,3,2,4,2,3,1,5,3,3,3,3,5,2,2,2
65	**3F7**	1	VI	Haarlem	2,2,3,2,3,5,2,3,2,3,2,4,2,3,1,5,3,3,3,3,7,2,2,5
66	**2I1**	1	VI	Haarlem	2,2,3,2,4,5,2,3,2,3,2,4,2,3,1,5,3,3,3,3,7,2,2,5
67	**3B5**	1	V	Haarlem	2,4,3,2,4,4,3,3,1,3,4,4,2,4,1,5,3,3,3,4,3,3,2,2
68	**2D2**	1	V	Haarlem	2,4,3,2,4,4,2,3,1,3,4,4,2,4,1,5,3,3,3,3,3,3,2,2
69	**1B6, 1C2, 2B3, 2C5, 2D6, 2F1, 2G2, 2G5, 3C3, 3E7, 3F5, 3G3**	12	VII	Haarlem	2,2,3,2,4,4,3,2,2,3,4,4,2,4,1,5,3,3,3,3,3,3,2,2
70	**1F6**	1	Unique	Haarlem	2,2,3,2,3,5,3,2,2,3,4,4,2,5,1,5,4,3,3,3,5,3,2,9
71	**2D7**	1	Unique	Haarlem	2,3,2,2,3,5,3,2,2,3,4,4,2,5,1,5,4,3,3,3,6,3,2,4
72	**3D1**	1	IX	X	2,2,3,1,5,4,3,4,2,3,4,4,2,5,1,5,3,3,3,3,9,1,2,3
73	**1F2, 2B8, 2E6**	3	IX	X	2,2,3,1,5,4,3,4,2,3,4,4,2,5,1,5,3,3,3,3,9,1,2,4
74	**3F8**	1	Unique	X	2,1,4,2,3,4,2,4,2,3,4,4,2,5,1,5,3,3,2,7,9,3,2,4
75	**1G4**	1	VIII	X	2,1,4,2,3,4,3,4,2,3,4,4,2,5,1,5,3,3,3,6,9,3,2,4
76	**1I4, 2A1, 2H5**	3	VIII	X	2,1,4,2,3,4,3,4,2,3,4,4,2,5,1,5,3,3,3,6,9,3,2,3
77	**3C5**	1	Unique	X	2,1,4,2,3,4,3,4,2,3,4,4,2,5,1,5,3,3,3,7,8,3,2,3
78	**2A4**	1	Unique	X	2,1,4,2,3,4,3,4,2,3,4,4,2,5,1,5,3,3,3,6,6,3,2,4
79	**3F4**	1	Unique	X	2,2,3,2,3,9,2,6,2,3,4,4,2,5,1,6,3,2,4,4,7,3,2,2
80	**3E6**	1	Unique	X	2,2,3,2,3,4,3,4,2,3,4,4,2,5,1,4,3,3,3,3,8,3,2,5
81	**3G1**	1	Unique	ND	2,10,2,2,2,3,2,4,2,3,4,2,2,5,1,6,3,3,5,3,7,4,3,2
82	**2H1, 3G8**	2	X	ND	2,3,3,1,3,4,3,1,2,3,4,2,2,5,1,5,3,3,3,4,9,2,2,4

The table indicates the laboratory sample code, the cluster they are contained for the phylogenetic analysis and the classification from the global *M. tuberculosis* lineages. ND refers to non-determined *M. tuberculosis* sub-lineages.

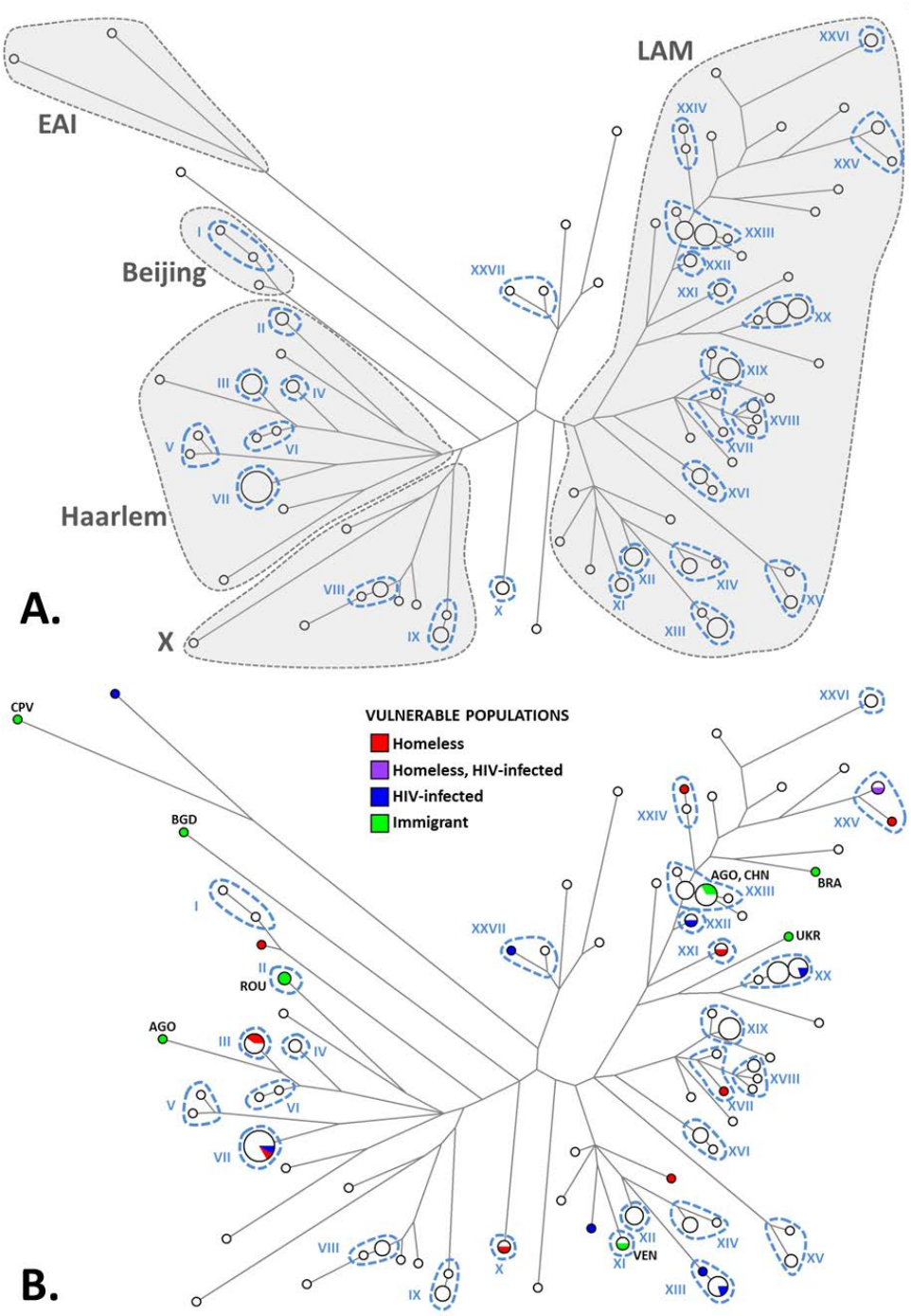


Figure 8. Median network of 144 Mycobacterium tuberculosis isolates.
The size of the circles is proportional to the number of cases of each genotype. A) Clades representing determined lineages from the global M. tuberculosis tree are indicated in the figure (Lineage 1: EAI, Lineage 2: Beijing, Lineage 4: LAM, X and Haarlem). The 27 hypothetical clusters of transmission are also indicated and numbered in dashed lines (I-XXVII). B) Vulnerable populations in the dataset (homeless, HIV infected and immigrants). The 27 hypothetical clusters of transmission are indicated in dashed lines. Countries of origin of immigrants are indicated as CPV (Cape-Verde), ROU (Romania), AGO (Angola), VEN (Venezuela), UKR (Ukraine), BRA (Brasil), BGD (Bangladesh) and CHN (China).

This tree, obtained directly from the median networks, displays a total of 82 distinct genotypes (59 of which were unique) organized into a set of inferred evolutionary clades with instances of genotypes displaying equal or closely linked genotypes (clusters) that probably contain related episodes of transmission.

The phylogenetic reconstruction also displays probable deeper evolutionary relationships between genotypes (50 inferred ancestral nodes) that can be of interest in terms of strain genetic characterization. By comparing our genotypes with MIRU-VNTR genotypes reference, the isolates were classified into lineages 1, 2 and 4 from the global geographically-defined M. tuberculosis lineages defined by complete M. tuberculosis genomes phylogenies.[156,167] Clades are labelled as groups “Latin American-Mediterranean” (LAM) (61.1%), “Haarlem” (20.1%), and “X” (9.0%) – which are all part of lineage 4, “Beijing” (2.1%) from lineage 2, “East-African Indian” (EAI) (1.4%) from lineage 1, and 9 unclassifiable samples (6.3%). This distribution is clearly reflected in the phylogenetic reconstruction, where these groups form monophyletic clades (Figure 8A), corroborating the power of the analysis to discern deeper as well as recent evolutionary relationships.

At deeper evolutionary levels homoplasy and longbranch attraction make evolutionary connections dubious. One example is the clustering of X and Haarlem lineages (lineage 4) together with Beijing (lineage 2) instead of LAM (lineage 4). Nevertheless, the phylogenetic reconstruction is reliable for clades that were probably established thousands of years ago[167] making it certainly reliable for recent clusters as the ones analysed in this work.

Great *M. tuberculosis* diversity within Porto’s autochthonous population

Conservatively, we considered as clusters samples whose genotypes differed by up to two differences across the 24 loci (as defined in [165]). Under this assumption, we obtained 27 clusters that include two or more isolates, each of which could represent transmission events within Porto (Figure 8A). 36 of the 54 unique haplotypes were located outside any cluster. The clustering frequency was 76.4%, considering all clusters identified, and 59.7% if we excluded those that contained only two isolates. Interestingly, there was no predominant chain of transmission of one or a limited number of specific genotypes. Instead, it seems that the Porto urban area functions as an M. tuberculosis genetic reservoir, where multiple genotypes circulate within the city. We would need to consider at least the 11 most frequent clusters to account for 50% of the cases, which highlights the high diversity

of M. tuberculosis strains circulating within Porto. There are, however, three clusters (clusters VII, XX and XXIII) that represent about 8.3% each, encompassing 25% of all genotyped cases, suggesting that a limited number of strains are more actively being transmitted.

Strain genotypes in vulnerable individuals are spread across the clusters

Taking into account the epidemiological information of each patient, we considered three risk factors: HIV infection, immigration and homelessness (Figure 8B). There were nine HIV-infected individuals and apart from two genotypes within cluster XIII (Figure 8B) their genotypes were unrelated, and they were quite spread across the tree. Seven of the nine HIV-infected individuals were clustered with individuals from the general population (clusters VII, XI, XII, XX, XXII, XV). There were 10 immigrants (Figure 8B): one, from Cape Verde, fell on the most diverged branch in the tree (the EAI clade); another from Bangladesh corresponded to one divergent unclassified lineage; two individuals from Romania formed an individual two-sample cluster (cluster II); one from Angola was a unique genotype while another from Angola fell within one of the largest clusters within the dataset (cluster XIII), which also included a Chinese individual with ten autochthonous individuals; one case from Venezuela formed a two-sample cluster with the strain detected in an autochthonous individual (cluster XI); finally, one case from a Brazilian and one from one Ukrainian were isolated. Eleven cases of TB in homeless people were reported during the study period: 10 from Porto city and one from the Porto suburban area. The cases are highlighted in Figure 8B where we note that only two pairs are located within the same clusters (cluster III and XXV), with 11 cases located within 10 unrelated genotypes. Six of the 11 cases are clustered with M. tuberculosis isolates from the general population (clusters III, VII, X, XVII, XXI, XXIV).

Hypothetical transmission episodes are not concordant with geographic proximity

The city of Porto is administratively and geographically divided into seven localities or parishes (Figure 9A).

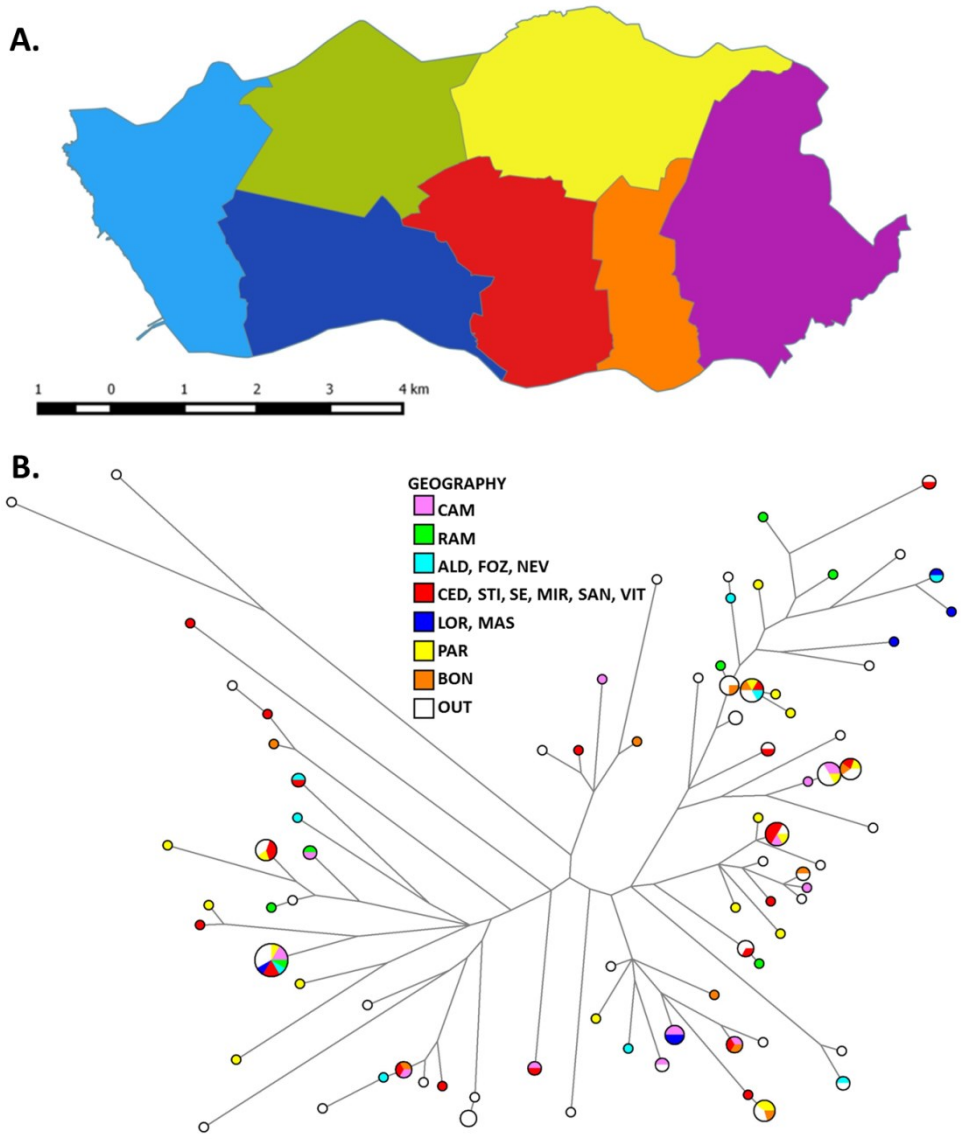


Figure 9. Geographic distribution of isolates according to Porto's localities or parishes.
A.: Map of Porto highlighting the localities. Indication of the geography of the samples in the network. The colour code represents different regions of Porto urban area: ALD (Aldoar), FOZ (Foz do Douro), NEV (Nevogilde), RAM (Ramalde), CAM (Campanhã), CED (Cedofeita), STI (Santo Ildefonso), SE (Sé), MIR (Miragaia), SAN (S. Nicolau), VIT (Vitória), LOR (Lordelo do Ouro), MAS (Massarelos), PAR (Paranhos), BOM (Bonfim) and OUT (Porto suburban area)

The tree displaying the localities for each case is shown in Figure 9B. Every two cases within a cluster can represent a transmission event between them. For example, a cluster with only two samples would only have one hypothetical transmission pair, while a cluster with three samples would have three possible pairs and a cluster with four samples would have six possible transmission pairs. Considering all pairs within the clusters of Porto residents, only 17.4% of them reside in the same locality.

To account for the fact that different related cases can be located in different localities of the city but actually separated by only a small distance (for example, cases close to the border of neighbour locations) or at extremes within the same locality, we performed a cross-validation analysis between all available pairs in the database, comparing the genetic distance in the tree and the geographic distance between coordinates of place of residence (Figure 10).

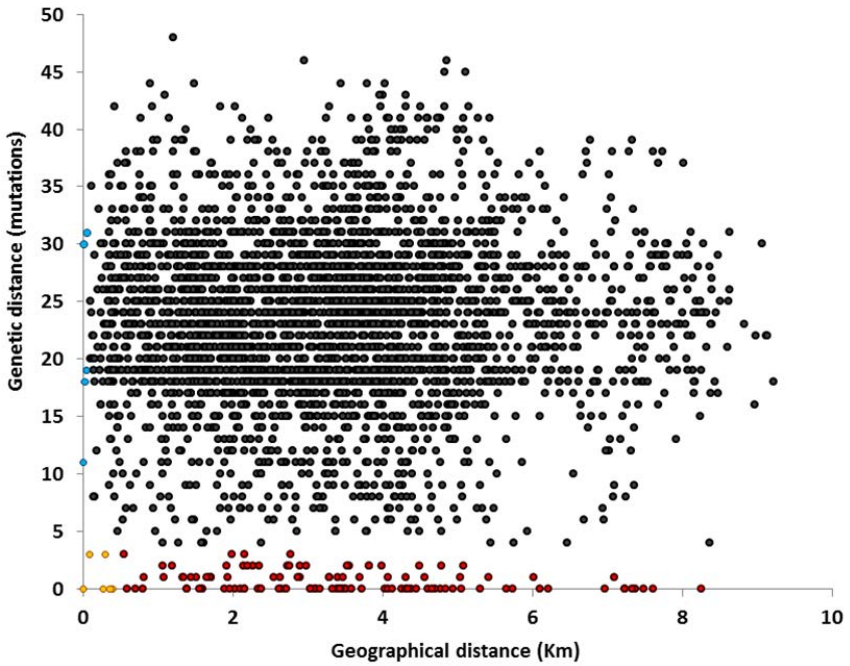


Figure 10. Graphical representation of genetic distance against the geographical distance between different tuberculosis cases.
The genetic differences were estimated using the genotypes and phylogenetic reconstruction and the geographical distance between different cases was calculated through the use of georeferencing coordinates. The 3486 datapoints were obtained comparing all possible combinations of case pairs. Data points are highlighted as: red in cases related genotypically but not displaying direct geographic connection; yellow for genotypically related cases in close proximity; in blue household cases with unrelated strains; and grey for cases not related either genotypically or geographically.

If structuring of M. tuberculosis diversity within Porto was to be observed, pairs with lower genetic distance would be expected to display overall lower distances between them. However, we did not observe any such correlation. Only five pairs from the same cluster were from individuals residing less than 0.5 Km apart (represented as yellow circles in Figure 10), and only a single case displayed evident household transmission. The values were not altered when we considered the lowest possible distance between two cases, varying the geographical parameters used (workplace, residence and frequented places).

The complete lack of sub-structuring of M. tuberculosis diversity, as well as the non-existence of connections between clusters and geospatial data were reflected in a comparison between clustered and genetically unrelated pairs. There was no difference in the geographic distance between genetically related pairs (genetic distance lower than two mutations – red and yellow data points in Figure 10) and non-related pairs (that should be random in this respect – black and blue points in Figure 10, respectively 3.37 and 3.31 Km (p-value = 0.719). At one extreme, two pairs representing household members displayed very different unrelated genotypes (blue data points in Figure 10).

A first epidemiological investigation of transmissions proposed using the genetic reconstruction

Considering this new genetic evidence, public health professionals reviewed epidemiological surveillance data from all TB cases who had reported contact with other TB patients 2 years prior to TB diagnosis. Their findings identified up to five links which had not been recognized at the time of the epidemiological investigation. If genetic data had been available in “real-time”, it would have been an important factor in directing subsequent epidemiological investigation in these cases. A surprising finding was a pair of TB cases with a reported epidemiological connection, who presented different genotypes.

In this situation, genetic data would have been an important factor in the decision to widen contact screening, had it been available during that investigation.

Given the relatively short duration of this study and the complex dynamics of TB transmission, it is plausible that such pattern would be identified more often in a setting of routine TB genotyping in Porto. Based on the results of this study, a workgroup was created which is currently investigating the largest genetic clusters that were identified.

Discussion

In this work, we initiated the process of genetic characterization of M. tuberculosis strains in Porto using a MIRU-VNTR approach over a 2-year period. The analysis shown here, based on a careful phylogenetic reconstruction, allowed us to discern both deep relationships and more refined connections (an essential feature for transmission detection and characterisation), allowing the discrimination of 82 genotypes in 144 samples and 48 inferred ancestral nodes between genotypes. The analysis showed a high proportion of unique cases and the lack of clear transmission links. These data reflect the high degree of M. tuberculosis diversity circulating within the autochthonous populations of Porto.

The reliability and accuracy of the methodology is demonstrated by the monophyletic status of groups of strains classified within the same clade of the global *M. tuberculosis* tree through comparison with MIRU reference strains. Additionally, the frequencies found in the Porto urban area were similar to those obtained by Lopes and colleagues for Portugal[157] using lineage defining SNPs.[168] The predominance of lineage 4 amongst circulating M. tuberculosis strains is corroborated by other studies in Europe.[169,170]

The data allowed us to discern 27 possible transmission clusters within Porto, but also to reject transmission links suggested by classical epidemiology, such as the cases of householders that were infected by different strains. Furthermore, epidemiological investigation data, including geospatial information (address, workplace and frequently visited places during the contagious period), was not enough to clarify clear transmission processes and chains of transmission, evoking an urgent need to re-establish epidemiological parameters.

Since TB transmission is a complex process and workplace and frequently visited places are not systematically recorded, the lack of correlation between strain characterisation and geospatial data should be interpreted with caution. Nevertheless, the inability to establish epidemiological links between cases that belong to the same genetic cluster is surprising, especially when compared with other studies, where often up to 50%[171,172] and even nearly 90% of the transmission links[173] are explained from this type of epidemiological data (such as residence and place of work). However, in most European urban areas, cases result from episodes of recent introduction of M. tuberculosis strains, which is not the case in Porto, where a wide diversity of strains is present. There is a high proportion of unique genotypes outside clusters (25%) that are often assumed to represent reactivation of latent

infection[171] or introductions by immigrants from high incidence countries as often reported from other Western Europe cities.[170,174]

From our dataset, only 7 genotypes in 82 and 8 cases in 144 (5.6%) could correspond to introduction through immigrants, placing this value close to the countries with lower proportion of foreign-origin TB like Bulgaria, Poland and Romania.[103] Two of the cases, one from an Angolan and one from a Chinese individual (cluster XIII) were contained within one of the three major transmission clusters. Interestingly this cluster is already closely linked with several other genotypes in Porto, strongly suggesting an autochthonous status.

As most samples were obtained from Portuguese Porto residents, the first scenario of reactivation is far more likely, which further enforces the notion that Porto urban area is a significant genetic reservoir of M. tuberculosis strains. In fact, from 36 unique genotypes in the tree only six were related with immigrants. Again, here the epidemiologically unlinked clusters involved autochthonous individuals, with the single exception from a two-sample cluster that involved individuals from Romania. This also alerts us to the possibility that an additional layer of genetic diversity might exist in the latent form of M. tuberculosis within the autochthonous Porto population, which can only be revealed with a continuous monitoring of M. tuberculosis strains in the coming years that could also reveal further transmission links. Although two years is assumed to be an adequate time period for epidemiological studies, reflecting the so-called cluster windows (the maximized probability of infected individuals to develop active TB in that time frame),[175] considering the scenario described here, a longer longitudinal study is essential.

Another possible improvement on further studies is the possibility of analysing *M. tuberculosis* genomes. Although genomic analysis would improve the resolution and point out possible false positives in the current study, it would not improve the traceability of transmission events involving clusters, which is the most pertinent aspect of the present study in Porto.

Ultimately the scenario we describe here reflects the endemic status of TB in Porto, where the emergence of new TB cases is not immediately traceable between cases but instead they might be related with more subtle transmission routes within the suburban area of Porto.

This poses an enormous challenge for health institutions and stakeholders aiming at the control of TB following a successful policy that has already led to a significant decline of TB cases.

The success of previous initiatives and future challenges in TB control strategies are mirrored in the homeless cases. In Portugal, TB incidence among the homeless has been estimated to be 122 per 100,000 homeless, five times higher than in the rest of the population.[176] While one of the common sources of transmission of TB within this vulnerable group was highlighted as co-transmission during time spent in homeless shelters,[177] measures were taken in Porto that minimize this transmission, including a mandatory

test for TB before entrance is granted to the shelters. The success of this measure is reflected in the fact that only two pairs of the 11 cases are related, and even one of these does not display the same exact genotype (Figure 8A), suggesting that this route of transmission has been largely blocked. This is corroborated by Nunes and Taylor, who reported that in Porto TB is significantly more rapidly diagnosed amongst homeless people than in others.[152] Further investigation should address whether the presence of risk factors (homelessness and HIV-infection) in most of the clusters simply reflects the expected high tendency to develop TB from circulating genotypes or if they play an active role in transmitting the disease as a source of the transmission clusters.

The Porto urban area has recently undergone a sustained decrease of TB incidence. Transient periods between incidence statuses are key moments to re-evaluate strategies as they contribute to define changes in conditions. While Porto might represent an ideal case study for TB control in urban areas, a full monitoring of TB transmission patterns is required to improve control strategies, including a detailed and continuous genetic characterization of M. tuberculosis strains. Epidemiological investigation needs to be enhanced, ideally guided by the established putative transmission links in molecular epidemiology, as suggested by our preliminary epidemiological investigations. If TB elimination is to be achieved, health professionals should exhaustively investigate and document potential sites of transmission for all TB cases. As previously suggested in the literature, continuous monitoring of social determinants that represent risk factors could greatly improve TB elimination programmes.[178] This task may prove difficult, since health care workers can only know what the patient provides voluntarily, which is often incomplete, incongruent with previous patient information and sometimes misleading. This might explain the lack of correlations obtained here. Some vulnerable populations, namely alcoholics and drug users, may be especially reluctant to report all contacts.[179] The absence of such

optimal and detailed epidemiological platforms poses constrains in terms of the limited information available regarding contact cases exposed to index cases.

Conclusions

MIRU-VNTR typing will likely be superseded by M. tuberculosis genomic data in a global scale. The greater resolution of complete genomes and the continuously decreasing prices caused this shift to have already occurred in several countries. However, there is over a decade of valuable MIRU genotyping data that when analysed with other tools (as the one described here) could provide further insights into modelling and understanding of transmissibility and outbreaks. Adding to that, MIRU-VNTR is still being continuously used in various countries and despite its resolution, a scenario where genomic data reaches this level of data availability is still remote.

MIRU-VNTR genotyping, coupled with an efficient phylogenetic reconstruction shows sufficient discrimination for us to recommend that its continuous implementation should become an essential tool for supporting TB control strategy implementation and policy-making in the coming years, with costs that could be sustainable for Public Health in Portugal and potentially in other countries. An ideal scenario would be the genotyping of the *M. tuberculosis* strain immediately following diagnosis. This would allow public health teams to identify chains of transmission that are unnoticeable using the present epidemiological questionnaire. This refined detection would lead to a secondary questionnaire, personalised for the cases, promoting an improved detection of contacts.

Overall this would allow the early detection of infected individuals and the characterization of chains of transmission currently unidentified, allowing the refinement of policies to control TB within Porto as it is in practice in several European cities.

A continuous follow-up on the genetic characterization of M. tuberculosis, combined with enhanced epidemiological field investigation, will prompt emerging strategic shifts aiming at TB eradication in scenarios of endemic low incidence of TB (that were likely high-incidence areas before). Porto could become a model for other TB endemic European cities.[103]

3.2. A NATIONWIDE STUDY OF MULTIDRUG-RESISTANT TUBERCULOSIS IN PORTUGAL 2014–2017 USING EPIDEMIOLOGICAL AND MOLECULAR CLUSTERING ANALYSES

Oliveira *et al. BMC Infectious Diseases* (2019) 19:567
https://doi.org/10.1186/s12879-019-4189-7

BMC Infectious Diseases

RESEARCH ARTICLE Open Access

A nationwide study of multidrug-resistant tuberculosis in Portugal 2014–2017 using epidemiological and molecular clustering analyses

Olena Oliveira[1,2,3], Rita Gaio[4,5], Carlos Carvalho[6,7], Margarida Correia-Neves[1,2], Raquel Duarte[3,8,9] and Teresa Rito[1,2*]

Abstract

Background: Increasing multidrug-resistant tuberculosis (MDR-TB) incidence is a major threat against TB eradication worldwide. We aim to conduct a detailed MDR-TB study in Portugal, an European country with endemic TB, combining genetic analysis and epidemiological data, in order to assess the efficiency of public health containment of MRD-TB in the country.

Methods: We used published MIRU-VNTR data, that we reanalysed using a phylogenetic analysis to better describe MDR-TB cases transmission occurring in Portugal from 2014 to 2017, further enriched with epidemiological data of these cases.

Results: We show an MDR-TB transmission scenario, where MDR strains likely arose and are transmitted within local chains. 63% of strains were clustered, suggesting high primary transmission (estimated as 50% using MIRU-VNTR data and 15% considering epidemiological links). These values are higher than those observed across Europe and even for sensitive strains in Portugal using similar methodologies. MDR-TB cases are associated with individuals born in Portugal and evolutionary analysis suggests a local evolution of strains. Consistently the sublineage LAM, the most common in sensitive strains in Europe, is the more frequent in Portugal in contrast with the remaining European MDR-TB picture where immigrant-associated Beijing strains are more common.

Conclusions: Despite efforts to track and contain MDR-TB strains in Portugal, their transmission patterns are still as uncontrolled as that of sensitive strains, stressing the need to reinforce surveillance and containment strategies.

Keywords: Multidrug-resistant tuberculosis, Epidemiology, Transmission, Risk factor

* Correspondence: teresarito@med.uminho.pt
[1]Population Health Research Domain, Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Gualtar Campus, 4710-057 Braga, Portugal
[2]ICVS/3B, PT Government Associate Laboratory, 4710-057 Braga, 4805-017 Guimarães, Portugal
Full list of author information is available at the end of the article

Background

Tuberculosis (TB) remains a high burden disease worldwide with persistent areas where elimination is still a distant goal. Despite declining incidence of TB in the last decades, multidrug-resistant TB (MDR-TB) poses a major threat for WHO's 2035 goal of TB elimination.[180]

MDR-TB, defined by resistance of Mycobacterium tuberculosis (Mtb) to isoniazid and rifampicin (RR), emerges as consequence of ineffective treatment or incompletion or inappropriate following up of the cases, translating into the evolution of resistant strains with consequent increments in patient morbidity and mortality, and further transmission.[181] WHO estimated a worldwide incidence of around 560,000 cases of MDR/RR-TB per year and a rate of 7.4 cases per 100,000 individuals.[182] In Europe, MDR/RR-TB incidence rate was 12.0 per 100,000 individuals, the highest among the regions considered by WHO,[182] but results could be biased towards regions with more detailed monitoring programmes on drug-susceptibility testing (DST).

MDR-TB notification rate in Portugal has been increasing since 2012 at a rate of 0.8 and 3.7% among new cases and previously treated patients, respectively.[183]

TB incidence has been decreasing in Portugal. In 2017, TB notification-rate was 17.8 per 100,000 individuals,[102] however 24.8 and 26.0 in the large urban centres of Lisbon and Porto respectively. The proportion of MDR-TB remains at 1% of all TB cases. 80% of MDR-TB patients had no previous TB treatment, suggesting mostly primary transmission of MDR strains but without supporting epidemiological studies confirming this scenario.

Previous genetic studies, evaluating clinical isolates collected in Lisbon Health Region, reported high prevalence of MDR-TB with most cases concentrating in two monophyletic clades (Lisbon 3 and Q1).[184,185]

A database developed by Perdigão and colleagues[186] constitutes the largest collaborative effort to catalogue Mtb diversity in Portuguese-speaking countries. It includes 423 MDR-TB isolates (129 from Portugal) within a larger dataset of 1447 clinical samples, validating Latin-American-Mediterranean (LAM: lineage 4) as the most common sub-lineage in Portugal. These studies are mostly from Lisbon hospitals,[184–186] but information on epidemiological characteristics of the cases is lacking, undermining the design of control strategies for preventing and reducing MDR-TB incidence.[187] Although MIRU-VNTR is becoming outdated in the study of TB transmission, mostly due to the overestimation of recent

transmission,[172,188–192] it is nevertheless the most used genotyping method implemented across Europe by Public Health authorities in a recent survey.[193] Although conclusions on transmission should be performed with extreme caution, the large body of work accumulated for MIRU-VNTR across Europe and worldwide allows statistics to be methodically compared for regions and scenarios still lacking whole genome data.

In this study we combined an in-depth genetic analysis of strain genotyping already published from Portugal,[188] with epidemiological data from MDR-TB cases diagnosed between 2014 and 2017, collected as part of the routine functions of Public Health services. We aim to assess the dynamics of MDR-TB emergence and transmission, including the identification of associated risk factors, and last to establish the rate of probable recent transmissions against newly developed resistant strains.

These statistics will be compared with results for sensitive strains in Portugal and other MDR-TB transmission scenarios in other developed countries using similar approaches in order to assess the relative efficiency of public health measures for containing MDR-TB in Portugal.

Methods

Data collection

We identified and extracted data from all culture confirmed MDR-TB cases, diagnosed between 2014 and 2017, from the national TB Surveillance System (SVIG-TB), containing epidemiological, clinical and laboratory data. Epidemiological information collected includes gender, age at time of TB diagnosis, country of origin, place of residence (parish), presence or absence of alcohol or drug misuse, HIV status, description of previous TB treatment and clinical characteristics of TB presentation (site of disease). Information about previous contact of patients with other TB cases was provided by Public Health services, using as linking variables date of birth, sex and place of residence, in order to identify possible epidemiological links between MDR-TB cases.

Mycobacterial interspersed repetitive unit-variable number tandem repeat (MIRU-VNTR) genotypes for sensitive and MDR strains were collected from genotypic works that characterised strains isolated from Portugal, namely [3,185,194]. Linked epidemiological and MIRU-VNTR data was only possible to MDR-TB cases described between 2014 and 2017.

Ethical approval

This work was carried out in accordance with the recommendations by the Ethics Sub-Commission of Life and Health Sciences from the University of Minho (SECVS 135/2015), by the Ethics Committee for Health of Lisbon and Tagus Valley Region Health Administration (9854/CES/2018) and the Ethics Committee for Health of the Northern Region Health Administration (ARSN 127/2018). All procedures were in accordance with the ethical standards of the responsible committees and with the Helsinki Declaration, as revised in 2008.

Cluster analysis

Clustering analysis was performed as described before in [195]: a cluster was defined as two or more cases with the same MIRU-VNTR profiles. The proportion of recent transmission was calculated by the "n minus one" method,[196] using number of cases clustered – number of clusters/number of cases with a strain type.

An epidemiological link was defined between two cases when cases had identified others as contacts, or when cases shared a family or household connection.

"Possible" epidemiological links were defined as cases from the same geographical area, with common social or behavioural traits (e.g. workplace, drug use). The number of epidemiological links was used to recalculate an adjusted proportion of recent transmission taking into consideration only those clustered cases further supported by an epidemiological link.

Evolutionary analysis of MDR-TB strains

The phylogenetic reconstruction of MIRU-VNTR profiles was done using median networks.[162] We used two algorithms consecutively implemented in the network software (fluxus engineering), the reduced median followed by the median joining algorithm, as previously described in [3]. We used this hybrid approach weighting the 24 loci according to their allelic diversity.[158]

Isolates that present similar MIRU genotypes were referred as genotypically clustered, representing potential episodes of direct TB transmission. Networks were used to construct the most parsimonious trees. Cladograms were visualized in Figtree v.1.3.3 (http://tree.bio.ed.ac.uk/software/figtree/). The classification of the strains in lineages was done using miruvntrplus (http://www.miruvntrplus.org).

Statistical analysis

Data were summarized by descriptive statistics (absolute and relative frequencies or median and range) according to the nature of the variables. Chi-square or Fisher's exact tests were used to evaluate the independence between two categorical variables, while Mann–Whitney U-test compared the distributions of two independent continuous variables. We identified risk factors associated with MDR-TB transmission, comparing patient's socio-demographic and clinical characteristics associated with clustering, that were investigated using logistic regression (comparing cases with unique genotypes against those within clusters).

Statistical analyses were performed with SPSS version 18.0 (PASW Statistics 18), and p-values below 0.05 were considered statistically significant.

Results

MDR-TB epidemiology

In Portugal, from January 2014 until December 2017, 8133 TB cases were notified, of which 4175 (51.3%) cases were culture-confirmed and had results of first-line DST (Figure 11), identifying 77 MDR-TB cases (~ 1% of total cases).

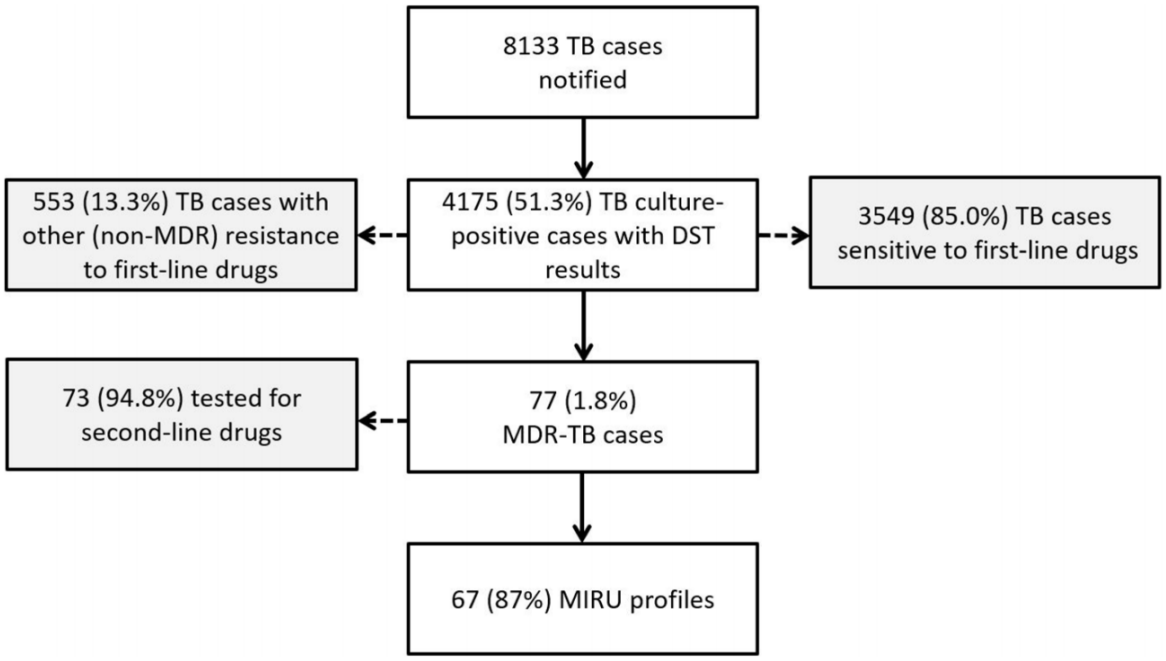


Figure 11. Tuberculosis cases notified and proportion of multidrug-resistant tuberculosis cases identified in Portugal between 2014 and 2017.

We compared patients with drug-sensitive TB and MDR-TB in the same period (Figure 12, Table 15).

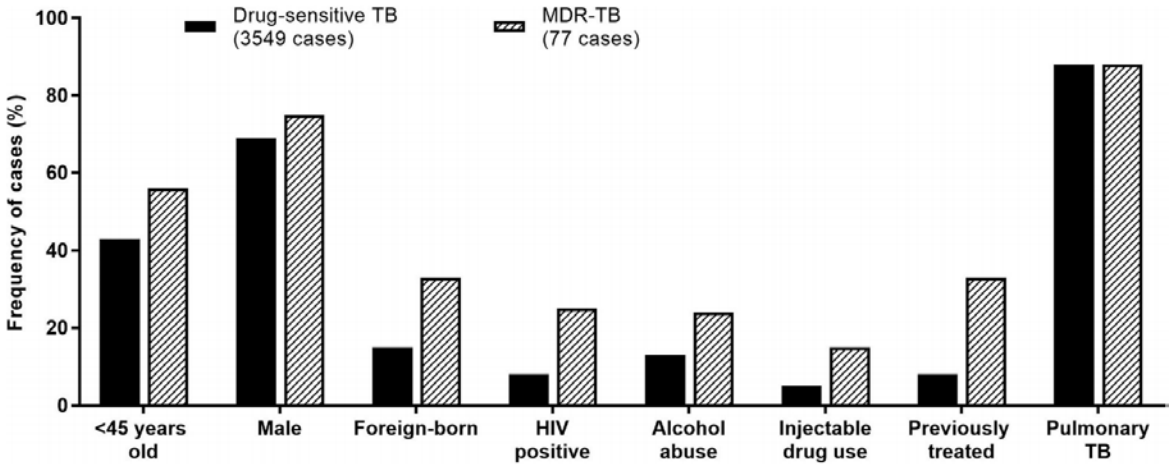


Figure 12. Differences in prevalence of socio-demographic and clinical characteristics between patients with drug-sensitive tuberculosis and multidrug-resistant tuberculosis, diagnosed in 2014-2017.

Table 15. Characteristics of drug-sensitive and MDR-TB cases diagnosed in 2014-2017.

Patients characteristic		**Unknown**	**Total**	**Drug-sensitive TB**	**MDR-TB**	**p-value**
				n (%)	n (%)	
Age, years	Median (min, max)	0	3626	48 (0-96)	43 (20-75)	0.111
Age group	<45 years old ≥45 years old	0	3626	1538 (43.3) 2011(56.7)	43(55.8) 34(44.2)	**0.038**
Gender	Female	0	3626	1098(30.9)	19(24.7)	0.292
	Male			2451(69.1)	58(75.3)	
Country of origin	Foreign-born	1	3625	543(15.3)	25(32.9)	**<0.001**
	Native			3006(84.7)	51(67.1)	
HIV status	Negative	4	3622	3265(92.0)	55(75.3)	**<0.001**
	Positive			284(8.0)	18(24.7)	
Alcohol abuse	No	218	3408	2898(86.8)	54(76.1)	**0.014**
	Yes			439(13.2)	17(23.9)	
Injectable drug use	No	223	3403	3183(95.4)	58(85.3)	**<0.001**
	Yes			152(4.6)	10(14.7)	
TB treatment hostory	Never treated	4	3622	3283(92.5)	49(67.1)	**<0.001**
	Previousley treated			266(7.5)	24(32.9)	
Site of disease	Pulmonary	4	3622	3113(87.7)	64(87.7)	1.000
	Extra-pulmonary			436(12.3)	9(12.3)	

Patients with MDR-TB were younger: the median age was 43 years (range 20–75) and 55.8% were below 45 years. MDR-TB patients presented higher frequencies of foreign-born individuals (32.9% vs. 15.3%), HIV infected (24.7% vs. 8.0%), alcohol abusers (23.9% vs. 13.2%), injectable drugs users (14.7% vs. 4.6%) and with previous TB treatment (32.9% vs. 7.5%). Among the foreign-born patients, 11 (44%) cases entered the country in the two previous years.

Cluster analysis and potential transmission links

Among 67 MDR-TB cases, 46 cases (69%) were LAM, eight (12%) cases were sublineage Haarlem, two (3%) URAL strains and one (1%) were sublineage X-type. Ten (15%) isolates

belonged to Lineage 2, sublineage Beijing. 42 cases were identified in seven MIRU-VNTR clusters ranging from 2 to 14 cases (Table 16; Figure 13).

Table 16. Characteristics of MIRU-VNTR clusters, including MIRU profile, Mycobacterium tuberculosis sub-lineage and patients' geodemographic and clinical features.

Cluster	**MIRU-24 loci profile**	***M. tuberculosis* sub-lineage**	***n***	**Country of origin**	**Residence area in Portugal**	**Risk factors**	**Disease site**	**Previous TB treatment**
1	243,244,331,234 424,153,334,332	Haarlem	4	Portugal	North	1 alcohol	3 pulmonary, 1 extra-pulmonary	No
2	244,233,352,644 425,153,353,823	Beijing	3	Guinea-Bissau	2 LVT, 1 Central	No	1 pulmonary, 1 extra-pulmonary, 1 unknown	1 no, 1 yes, 1 unknown
3	244,233,352,644 425,173,343,723	Beijing	3	Portugal	LVT	Drugs	Pulmonary	No
4	244,233,352,644 425,183,353,823	Beijing	2	Portugal	1 Madeira, 1 Central	No	Pulmonary	No
5	244,213,132,324 114,142,532,822	LAM	10	7 Portugal, 2 Cape Verde, 1 Mozambique	LVT	3 alcohol, 2 drugs, 1 shelter	7 pulmonary, 2 extra-pulmonary, 1 unknown	5 no, 4 yes, 1 unknown
6	244,213,232,324 116,143,532,822	LAM	6	5 Portugal, 1 Cape Verde	5 LVT, 1 Central	3 alcohol, 2 drugs	Pulmonary	4 no, 2 yes
7	244,213,232,424 116,143,532,822	LAM	14	12 Portugal, 2 Angola	9 LVT, 2 Central, 3 North	4 alcohol, 4 drugs	10 pulmonary, 3 extra-pulmonary, 1 unknown	6 no, 7 yes, 1 unknown

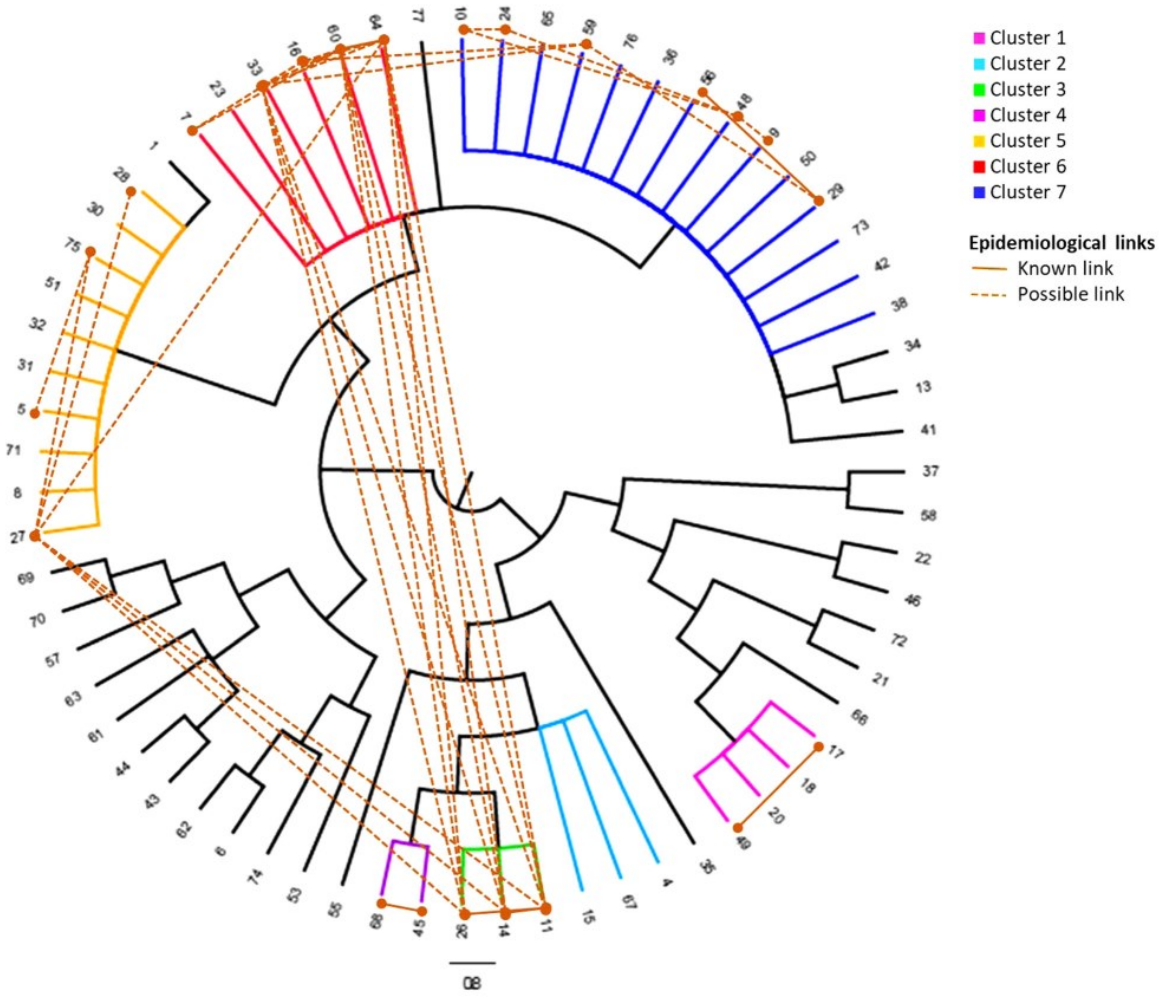


Figure 13. Cladogram representing the phylogenetic tree obtained with MIRU-VNTR profiles from 67 cases of multidrug-resistant tuberculosis and epidemiological links identified between cases.
Clusters 1 to 7 are coloured as follow: pink (cluster 1), light blue (cluster 2), green (cluster 3), purple (cluster 4), orange (cluster 5), red (cluster 6) and dark blue (cluster 7).

These MIRU-VNTR clusters showed good correspondence to previously defined clusters on the whole genome level[197] and each MIRU-defined cluster has a set of defined specific mutations in drug-resistance associated genes attesting for their significance as clades.

The three largest clusters [5–7] contained 10, 6 and 14 cases, respectively, with most of these corresponding to Portugal-born individuals, with pulmonary disease, diagnosed in the LTV region. Twenty-five cases presented unique strains, possible outside introductions or newly developed MDR-TB strains (Table 17).

Table 17. Characteristics of non-clustered cases, including MIRU profile, Mycobacterium tuberculosis sub-lineage and patients' geodemographic and clinical features.

MIRU-24 loci profile	Lineage	Sub-lineage	Number of cases	Country of origin	Residence area in Portugal	Risk factors	Site of disease	TB treatment history
244,213,132,324,114,142,532,722	4	LAM	1	Portugal	LTV	Drug use, prison	Pulmonary	Yes
132,254,322,224,125,153,322,622	4	LAM	1	Belarus	LTV	No	Pulmonary	No
214,213,232,434,116,143,532,422	4	LAM	1	Portugal	LTV	Alcohol abuse	Pulmonary	No
223,235,321,432,423,153,332,522	4	Haarlem	1	Portugal	North	Drug use, homelessness	Pulmonary	No
223,154,342,434,425,153,333,812	4	X	1	Andorra	North	No	Pulmonary	Yes
244,211,232,424,116,143,632,822	4	LAM	1	Guinea	LTV	Community residence	Pulmonary	Yes
244,213,232,424,116,143,532,622	4	LAM	1	Portugal	LTV	No	Pulmonary	No
235,237,232,244,425,113,323,632	4	URAL	1	Ukraine	LTV	No	Pulmonary	No
235,237,232,244,425,113,323,632	2	Beijing	1	Portugal	LTV	Drug use, prison	Pulmonary	Yes
214,213,232,225,126,153,332,322	4	LAM	1	Portugal	North	No	Pulmonary	Yes
214,213,232,225,126,153,332,222	4	LAM	1	Portugal	North	No	Pulmonary	No
223,235,332,534,425,153,333,632	4	Haarlem	1	Moldova	Algarve	No	Pulmonary	No
254,212,232,224,125,153,332,622	4	LAM	1	Angola	LTV	No	Pulmonary	Yes
214,213,232,225,126,153,332,422	4	LAM	1	Angola	LTV	No	Pulmonary	No
244,233,352,633,425,153,353,823	2	Beijing	1	Ukraine	LTV	Community residence	Pulmonary	Yes
134,262,332,224,126,153,322,722	4	LAM	1	Portugal	Central	No	Pulmonary	Yes
234,244,322,424,226,133,131,722	4	LAM	1	Portugal	Central	No	Pulmonary	No
243,244,332,334,425,153,334,732	4	Haarlem	1	Angola	Central	No	Pulmonary	No
254,214,232,224,125,163,333,622	4	LAM	1	Angola	LTV	No	Pulmonary	No
244,213,232,224,116,143,532,822	4	LAM	1	Portugal	LTV	No	Pulmonary	No
223,235,321,432,423,153,333,522	4	Haarlem	1	Portugal	North	No	Pulmonary	No
254,214,232,224,125,163,332,622	4	LAM	1	Angola	North	No	Pulmonary	No
235,237,132,244,425,113,323,532	4	URAL	1	Portugal	Central	No	Pulmonary	No
133,224,332,224,127,153,322,622	4	LAM	1	Angola	Central	No	Pulmonary	Yes
214,213,232,225,126,163,332,422	4	LAM	1	Unknown	LTV	Unknown	Unknown	Unknown

Genetically, the proportion of cases attributable to recent transmission was 52.2%. While this value is strongly overestimated, as MIRU clusters can correspond to clades dating to several years, it allows the identification of circulating clades of MDR strains in the Portuguese scenario.

To corroborate the established genetic connections, we analysed information provided by the Public Health services, identifying 17 links (5 known, 12 possible) between cases within five clusters (Figure 13):

- In cluster 1, containing four cases from the Northern region, two patients lived in the same parish;
- The three patients within cluster 3 were drug-users, attended the same community care facilities and two were close relatives;
- Patients of cluster 4 were friends;
- Two cases of the cluster 6 were relatives; also users of the community care facilities mentioned above for cluster 3;
- Two cases of the cluster 7 were close relatives. One was a civil construction worker with possible contact with two patients of the cluster, also construction workers in the same area.

The proportion of cases attributable to recent transmission after adjustment for detected epidemiological links was 14.9%.

Risk factors associated with clustering

To identify particular risk factors underlying active transmission of MDR-TB, we compared clustered cases (likely recent transmission cases) against unique cases. In the univariate analysis, being born in Portugal (OR 3.67; 95% CI 1.24–10.88; p = 0.019) and alcohol abuse (OR 10.15; 95% CI 1.22–84.39; p =0.032) were associated with clustering (Table 18).

Nevertheless, the table shows that alcohol abuse is a good identifier for clustered cases and its low frequency within the unique cases turns the logistic regression model impractical. After adjustment for gender and age, being Portugal-born was the only independently clustering-associated variable (adjusted OR 3.64; 95% CI 1.16–11.47; p = 0.027).

Table 18. Assessment of patient's characteristics and risk factors that could be associated with multidrug-resistant Mycobacterium tuberculosis clustering of cases, considering the cases reported between 2014 and 2017.

MDR-TB patient Characteristic		**Unknown**	**Unique (*n* = 25)**	**Clustered (*n* = 42)**	**Univariate analysis**	
			n (%)	**n (%)**	**OR (95% CI)**	***p*-value**
Age group	< 45 years old	0	16 (64.0)	21 (50.0)	Ref	0.267
	≥45 years old		9 (36.0)	21 (50.0)	1.78 (0.64–6.91)	
Gender	Female	0	7 (28.0)	9 (21.4)	Ref	0.543
	Male		18 (72.0)	33 (78.6)	1.43 (0.46–4.47)	
Country of origin	Foreign-born	1	12 (50.0)	9 (21.4)	Ref	**0.019**
	Native		12 (50.0)	33 (78.6)	**3.67 (1.24–10.88)**	
HIV status	Negative	4	21 (87.5)	27 (69.2)	Ref	0.109
	Positive		3 (12.5)	12 (30.8)	3.11 (0.78–12.46)	
Alcohol abuse	No	6	22 (95.7)	26 (68.4)	Ref	**0.032**
	Yes		1 (4.3)	12 (31.6)	**10.15 (1.22–84.39)**	
Injectable drug use	No	8	20 (87.0)	29 (80.6)	Ref	0.525
	Yes		3 (13.0)	7 (19.4)	1.61 (0.37–6.98)	
TB treatment history	Never treated	4	15 (62.5)	25 (64.1)	Ref	0.898
	Previously treated		9 (37.5)	14 (35.9)	0.93 (0.33–2.68)	
Site of disease	Pulmonary	4	24 (42.9)	32 (57.1)	n/a	n/a
	Extra-pulmonary		0 (0.0)	7 (100.0)		

Genetic contextualization of Portuguese MDR-TB strains

Aiming at a clearer picture of emergence and spread of Mtb strains, we collected and combined published Portuguese MIRU-VNTR data.[3,185,188]

The evolutionary networks are shown in Figure 14 and Figure 15. Networks represent genotypes (circles), where branch length is proportional to genetic distance. Size of the circles is proportional to frequency of cases with the same genotype, being hypothetical clusters of transmission. MDR clusters (Table 15) are highlighted in the figures. Genotypes from strains resistant to at least one second-line anti-TB drug often appear within clusters (Figure 14) but with individual strains scattered across the network, similarly to sensitive and first-line resistant strains.

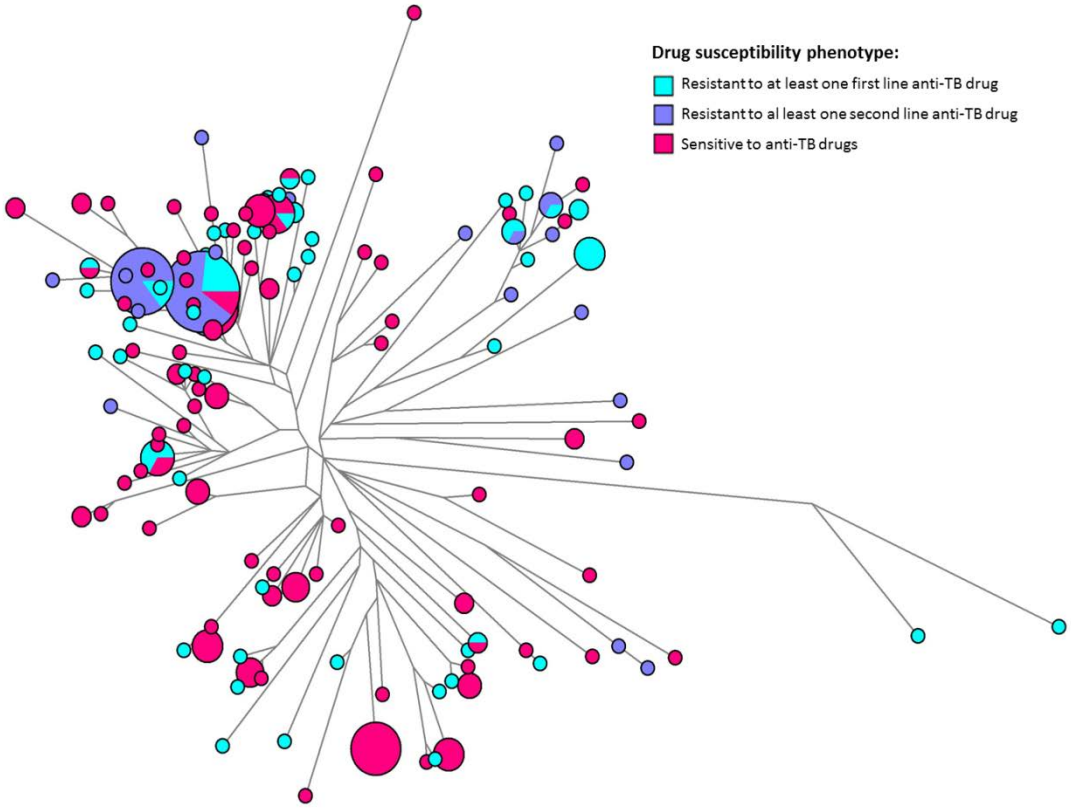


Figure 14. Reduced median network of Mycobacterium tuberculosis genotypes, representing the current scenario of multidrug-resistant tuberculosis in Portugal and how these strains are phylogenetically related with the sensitive tuberculosis strains.

The network displays 67 profiles of multidrug-resistant strains [194], 144 from the Northern region [3] and 56 strains with different levels of resistance [185]. Clusters detected in previous analyses and reported throughout the paper are highlighted.

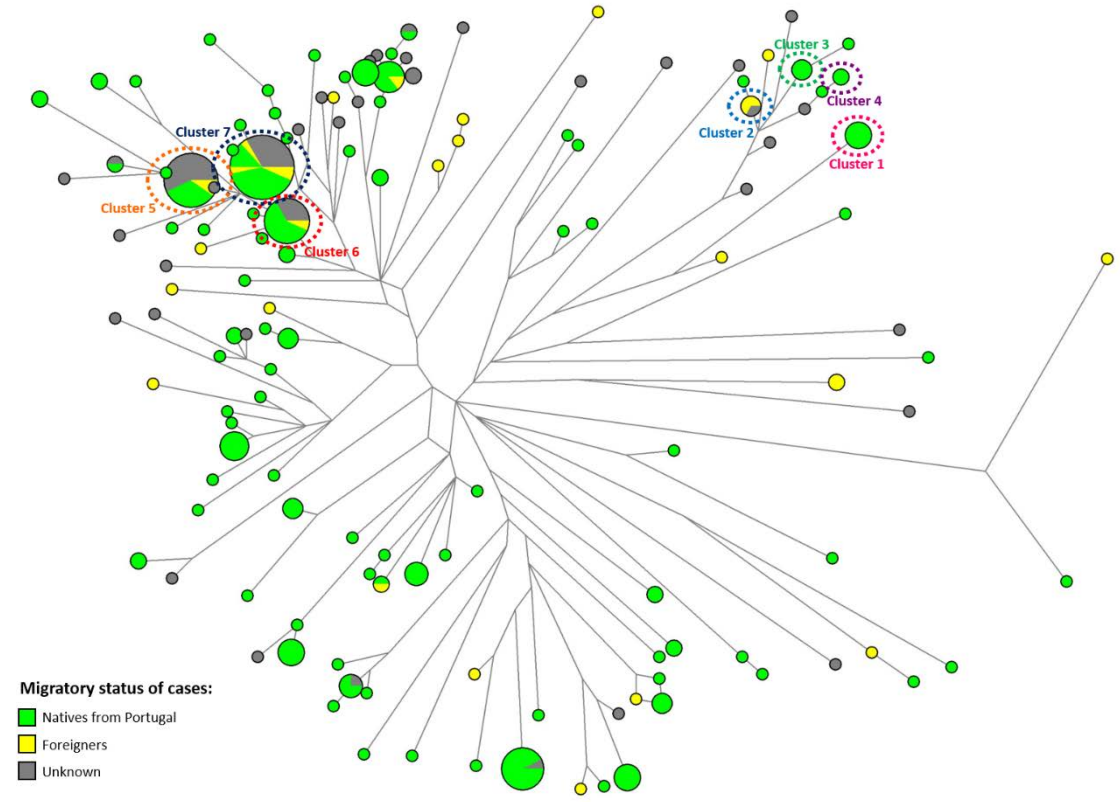


Figure 15. Reduced median network, representing the current scenario of multidrug-resistant TB in Portugal and how these strains are grouped together with the sensitive TB strains.

The network displays 67 profiles of multidrug-resistant strains,[194] 144 from the Northern region[3] and 56 strains with different levels of resistance.[185] Clusters detected in previous analyses and reported throughout the paper are highlighted. The migratory status of the cases is highlighted: green cases are natives from Portugal, yellow are foreigners and grey are individuals that on diagnosis (or publication) did not disclosure migratory status.

One aspect to explore with caution, given the low resolution of MIRU-VNTR markers, is that deeper clusters often include sensitive, first-line resistant and MDR strains within the Portuguese population (being cluster 1 the exception). A common ancestry of MDR and sensitive strains in the same area suggests that MDR strains are likely the result of evolution of strains occurring within local chains of transmissions in Portugal and not brought from abroad, often without time for differentiation between both in MIRU markers.

From the network analysis, MDR-TB cluster distribution is similar to the pattern observed with drug sensitive strains, as described, for example, for Porto Urban area,[3] with the existence of several prominent clusters in both suggesting active transmission, or at least the maintenance of several circulating MDR-TB clades within the population. Moreover, general clustering statistics are very similar to that dataset, corresponding to 63% of the MDR-TB cases being clustered, against 59.7% in the drug-sensitive TB cases within Porto ($p = 0.883$, not significant). Considering both datasets, the percentage of cases attributed

to recent transmission based on MIRU-VNTR data is 52.2% against 43.8%, higher in MDR cases. While the statistics are overestimated using MIRU-VNTR they are directly comparable and both suggest the circulation of specific Mtb clades within the population in terms of sensitive and MDR strains.

For refining the geography of MDR-TB, we inspected the origin of the drug-resistance cases. The main clusters described in Lisbon,[185] are also present in the Centre and Northern regions of the country. Also, while many of the single genotypes are present in immigrants (Figure 15) possible recent introductions, clustered cases, often clustering with sensitive (Figure 14), are mostly present in autochthonous individuals, reinforcing a possible evolution and transmission of MDR strains within indigenous transmission chains.[3]

History of previous TB treatment

Twenty-three (34%) of the 67 MDR-TB cases with MIRU-VNTR profiles reported previous TB treatment, lacking DST and genotypic data on these past infections.

One individual reported three previous TB treatments. This individual was Portuguese, inmate and injectable drug user. He perished during the fourth treatment with a strain resistant to all first-line drugs and 6 s-line drugs tested, appearing isolated in the genetic networks, pinpointing this case expectedly as one of emergence of resistance. Six other cases reported two previous treatment courses but display no specific evolutionary pattern, being either clustered or isolated.

Within the 23 MDR-TB cases that reported a previous TB treatment, nine had an unique genotype, appearing isolated in the network, possibly representing independent evolution of MDR strains. In opposite, 14 genotypes were in clusters, likely transmission events. The frequency of individuals with previous TB treatment in clustered genotypes is basically the same as the frequency in non-clustered individuals, 36 and 38%, respectively (Table 18). This might reflect the fact that risk of reinfection and inadequate treatment are likely associated with the same risk groups, making the individuals likely vessels for either acquiring new strains or development of MDR.

Discussion

There are various initiatives worldwide aiming to contain and minimise the impact of MDR-TB. Nevertheless, an integrative scenario of MDR-TB transmission dynamics in Portugal is virtually inexistent. During a 4-year period (2014–2017) MDR-TB patients were younger, with higher prevalence of foreign-born individuals, HIV infected, alcohol abusers, injectable drug users and with previous TB treatment, when compared with drug-sensitive patients. Previous TB treatment is known as a strong risk factor for MDR-TB,[198–201] and young age, HIV infection, being foreign-born and frequent consumption of alcohol have also been reported in several countries associated with development of MDR-TB.[198,199,201,202] Several of these risk factors are underlying ineffective completion of previous treatments. We identified seven MIRU-VNTR clusters, where the three largest clusters belonged to sublineage LAM, mostly corresponding to Portugal-born individuals, with pulmonary disease, diagnosed in the LTV region. Being Portugal-born was the independent risk factor statistically associated with MDR-TB recent transmission, suggesting that contrarily to the scenario in other developed countries MDR-TB is not strongly associated with foreign individuals. This trend is also supported by the frequency of the main genotyped lineages.

According to the European Centre for Disease Control (ECDC) surveillance report, Beijing was the most common genotypic lineage among MDR-TB strains isolated in 2015 in Europe (60.0%), followed by LAM (17.6%),[203] despite LAM being more prevalent in sensitive strains (as in Ireland[204] and Belgium[205], for example). By contrast, here, the most common genotypic lineage in MDR-TB was LAM (69%), followed by Beijing (15%) and Haarlem (12%), similarly to a study on sensitive strain in Northern Portugal (LAM represented 61.1%).[3] The similar prevalence of LAM in sensitive and MDR-strains strengthens the hypothesis that MDR-TB in Portugal is evolving within autochthonous chains of transmission and being transmitted similarly to drug-sensitive TB.

Clusters 5, 6 and 7 correspond to previously identified MDR-TB clusters Q1, Lisboa3-A and Lisboa3-B in the LTV region,[184,185] reflecting continuous circulation of these strains in the region extending to other regions of Portugal and abroad (Q1 and Lisboa3-B were included in cross-border clusters reported by ECDC, being present in the UK and France[203]). While we suggested an origin of these clusters within the autochthonous transmission chains, determining exact source and direction of transmission between countries would require higher discriminatory power and deeper epidemiological investigations.[206]

The proportion of clustered cases in our study (63%) was generally higher than most studies using similar methodologies, for example in the UK, Switzerland and USA[195,207,208] but

similar to a Portuguese sensitive sample (59.7%) following the same criteria.[3] We estimated recent transmission to 52.2% using MIRU-VNTR data and 14.9 after adjustment for epidemiological data.

While MIRU-VNTR data largely overestimates recent transmission estimates [193,209], the epidemiological data likely underestimates direct transmission given the difficulty in assessing all relevant epidemiological information using only conventional contact tracing data in complex transmission chains.[3] Nevertheless, these 14.9% were still twice as high as what was reported in UK[195] and Switzerland[208] using similar approaches.

The higher prevalence of MDR-TB (above 50%) displaying no previous treatment provides evidence for high rate of primary transmission of MDR-TB in Portugal, also supported by a higher estimated rate of transmission in MDR than sensitive strains in Portugal, against the expected European trend. Independently of the clustering being overestimated, in the sense that MIRU clusters could date up to decades in some instances, it reflects the existence of circulating genetic clades being maintained in the population.

Conclusion

This study has limitations in terms of analysed period and genotyping, but it offers nevertheless a strong effort to assess MDR-TB scenario in Portugal using a more detailed combined MIRU-VNTR and epidemiological analysis. While the recent transmission rates are widely overestimated by the application of MIRU-VNTR data, there is nevertheless a striking scenario that relates to a high percentage of circulating strains from the same clades (when compared with sensitive strains and MDR-TB in other countries) than what could be expected given the public health efforts to contain MDR-TB. Taking into account the possible scenarios of emergence of MDR-TB, it is necessary to readjust measures to decrease transmission, including for example, improvement of earlier diagnosis and better adherence to treatment in specific regions of the country, including the support of community institutions focusing on specific population groups, and a faster and integrative genotyping protocol for early identification of clustered cases.

3.3. FAILURE TO COMPLETE TREATMENT FOR LATENT TUBERCULOSIS INFECTION IN PORTUGAL, 2013–2017: GEOGRAPHIC-, SOCIODEMOGRAPHIC-, AND MEDICAL-ASSOCIATED FACTORS

European Journal of Clinical Microbiology & Infectious Diseases (2020) 39:647–656
https://doi.org/10.1007/s10096-019-03765-y

ORIGINAL ARTICLE

Check for updates

Failure to complete treatment for latent tuberculosis infection in Portugal, 2013–2017: geographic-, sociodemographic-, and medical-associated factors

Alexis Sentís[1,2,3,4,5] · Paula Vasconcelos[1,2] · Rita Sá Machado[1,2] · Joan A. Caylà[6] · Mònica Guxens[3,4,5,7] · Vasco Peixoto[1,2] · Raquel Duarte[1,8,9,10] · Isabel Carvalho[1] · Carlos Carvalho[1,11,12]

Received: 11 September 2019 / Accepted: 7 November 2019 / Published online: 3 December 2019

Abstract

There is conflicting evidence about factors associated with failure to complete treatment (FCT) for latent tuberculosis infection (LTBI). We aim to identify the geographic, sociodemographic, and medical factors associated with FCT in Portugal, highlighting the two main metropolitan areas of Porto and Lisbon. We performed a retrospective cohort study including LTBI patients that started treatment in Portugal between 2013 and 2017. We calculated adjusted odds ratios (aOR) and 95% confidence intervals (95% CI) using multivariable logistic regression to identify geographic, sociodemographic, and medical factors associated with FCT. Data on completion of treatment were available for 15,478 of 17,144 patients (90.3%). Of those, 2132 (13.8%) failed to complete treatment. Factors associated with FCT were being older than 15 years (aOR, 1.65 (95% CI = 1.34–2.05) for those aged 16 to 29), being born abroad (aOR, 2.04 (95% CI = 1.19–3.50) for Asia; aOR, 1.57 (95% CI = 1.24–1.98) for Africa), having a chronic disease (aOR, 1.29 (95% CI = 1.04–1.60)), alcohol abuse (aOR, 2.24 (95% CI = 1.73–2.90)), and being intravenous drug user (aOR, 1.68 (95% CI = 1.05–2.68)). Three-month course treatment with isoniazid plus rifampicin was associated with decreased FCT when compared with 6- or 9-month courses of isoniazid-only (aOR, 0.59 (95% CI = 0.45–0.77)). In Lisbon metropolitan area, being born in Africa, and in Porto metropolitan area, alcohol abusing and being intravenous drug user were distinctive factors associated with FCT. Sociodemographic and medical factors associated with FCT may vary by geographical area and should be taken into account when planning interventions to improve LTBI treatment outcomes. This study reinforces that shorter course treatment for LTBI might reduce FCT.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10096-019-03765-y) contains supplementary material, which is available to authorized users.

✉ Carlos Carvalho
ccarvalho@arsnorte.min-saude.pt

[1] Directorate-General of Health, Lisbon, Portugal

[2] European Programme for Intervention Epidemiology Training (EPIET), European Centre for Disease Prevention and Control, (ECDC), Stockholm, Sweden

[3] ISGlobal, Barcelona, Spain

[4] Pompeu Fabra University (UPF), Barcelona, Spain

[5] Spanish Consortium for Research on Epidemiology and Public Health (CIBERESP), Instituto de Salud Carlos III, Madrid, Spain

[6] Foundation of Tuberculosis Research Unit of Barcelona, Barcelona, Spain

[7] Department of Child and Adolescent Psychiatry/Psychology, Erasmus University Medical Centre–Sophia Children's Hospital, Rotterdam, The Netherlands

[8] Public Health Science and Medical Education Department, Faculty of Medicine, University of Porto, Porto, Portugal

[9] EPIUnit, Institute of Public Health, University of Porto, Porto, Portugal

[10] Pulmonology Department, Hospital Centre of Vila Nova de Gaia/Espinho EPE, Vila Nova de Gaia, Portugal

[11] Department of Public Health, Northern Regional Health Administration, Porto, Portugal

[12] Multidisciplinary Unit for Biomedical Research (UMIB), Institute of Biomedical Sciences Abel Salazar, University of Porto, Porto, Portugal

Background

Approximately one-fourth of the worldwide population is affected by latent tuberculosis infection (LTBI), 5–10% of whom will develop active tuberculosis (TB) during their lifetime mainly in the first 2 years after infection.[210,211]

LTBI cases are asymptomatic and potential future new cases of TB; therefore, if they are not diagnosed and treated, TB elimination will be not attained.[151,210] In addition to being a close contact of individuals with active TB, those living in communities with high TB incidence rates and in poorly ventilated and crowded environments have an increased risk of acquiring LTBI.[210,212]

Portugal has one of the highest TB notification rates in the European Union (EU) (17.8 TB cases/100,000 inhabitants in 2016, vs. an average of 11.4 in the EU).[213] From 2012 to 2016, TB cases have decreased an average of 20% in EU countries and 30% in Portugal.[213] All cases of LTBI that started treatment should be notified to the Portuguese TB surveillance system, but those who did not start treatment will not be reported.

The number of LTBI cases reported increased from 600 in 2000 to 3246 in 2013,[149] reflecting national efforts to better screen, notify, and treat LTBI.[102]

It has been estimated that the treatment of LTBI could avoid approximately 90% of the cases that would result from progression to active TB.[214] In the EU, in spite of the high treatment initiation rates, the proportion of individuals failing to complete (FCT) LTBI treatment broadly differs within all the risk groups to develop TB.[215]

Although with conflicting results, different factors such as age, gender, country of birth, substance abuse, HIV coinfection, treatment regimen, treatment adverse effects, missing early clinical visits during treatment, and receiving information about disease and treatment have been associated with FCT.[216–219] Geographical differences in the TB risk factors have also been described, between regions, metropolitan areas, rural and urban spaces, or even in smaller geographic areas.[220,221] In Portugal, the metropolitan areas of the two main cities (Lisbon and Porto, which are respectively part of the Lisbon and Tagus Valley (LTV) and Northern Health Regions) present the highest TB notification rates in the country,[153,222] but risk factors for TB seem to be different.[149,153]

In this study, we aim to identify the most relevant factors associated with FCT in Portugal between 2013 and 2017, and to identify possible differences between the two main

metropolitan areas, Porto and Lisbon, in order to inform specific preventive measures to reduce LTBI and, in this way, prevent the TB development.

Methods

Study design and participants

We performed a retrospective cohort study using nationwide anonymized data from the national TB clinical notification and follow-up surveillance system (SVIG-TB), between the years 2013 and 2017. These data are routinely collected from paper notification forms, used mainly in TB outpatient centres, where all patients are followed up until the end of treatment. The Lisbon metropolitan area includes 18 municipalities and had 2,833,679 inhabitants in 2017. The Porto metropolitan area includes 17 municipalities and had 1,719,702 inhabitants.

LTBI can be defined as an asymptomatic infection characterized by an immune response to Mycobacterium tuberculosis antigens.[223] In Portugal, all patients diagnosed LTBI are offered treatment, but the screening of LTBI is opportunistic (mainly after contact with an infectious case of TB) and might vary between cities, regions, and healthcare centres.

All the patients that were diagnosed with LTBI and started treatment in Portugal between the years 2013 and 2017 were included in the study. Treatment outcome was dichotomized as follows: (i) success to complete treatment (SCT): LTBI patients that started and finished the treatment during the study period, and (ii) failure to complete treatment (FCT): LTBI patients that started but did not complete the treatment, with a recorded treatment outcome of “therapeutic failure” or “interruption or withdrawal”. LTBI patients that started treatment and did not have a treatment outcome recorded were classified as “unknown treatment outcome” and excluded from the analysis, including 51 patients with unknown status that died after the treatment started (cause of death not known). All LTBI records in the database had at least one confirmatory test with a positive result, an IGRA or TST. Treatment regimen was categorized as isoniazid-only (6-month course if duration of treatment was more than 4 months and less than 8 months, 9-month course if treatment was longer than 8 months, other courses if treatment was shorter than 4 months), rifampicin only (4-month course), and combination of isoniazid and rifampicin (3-month course).

Variable definition and potentially associated factors

Sociodemographic potentially associated factors included sex, age, continent/region/country of birth, incarceration at start of treatment, and living in a community health centre. Medical factors included LTBI treatment regimen, previous treatment for TB or LTBI, and previous chronic disease or other factors potentially associated with treatment, such as chronic liver disease (LCD), chronic kidney disease (CKD), chronic obstructive pulmonary disease (COPD), diabetes mellitus (DM), alcohol abuse, intravenous or other drugs used (IDU), and directly observed treatment (DOT). We computed the variables "region" and "metropolitan areas" of Lisbon and Porto (according to the European Nomenclature of Territorial Units for Statistics—NUTS 2 and NUTS 3, respectively) from the variable "municipality of residence". The variable "region" was used in the descriptive analysis in order to obtain a regional picture and "metropolitan area" to highlight the main metropolitan areas (Porto and Lisbon), where different risk factors for TB have been described,[149,153] and performed separated models. Failure to complete treatment (FCT) was defined as having a treatment outcome of "therapeutic failure" (if the patient developed active TB during LTBI treatment) or "interruption or withdrawal" (when for adverse effects or patient personal reasons the treatment was stopped) recorded.

Statistical analyses

We described the LTBI patients in terms of geographic, sociodemographic, and medical factors using proportions and measures of central tendency. We characterized the cases with unknown treatment outcome (missing information) in order to identify possible biases. We tested associations between independent variables and FCT using logistic regression. All those variables that we found to be significantly associated with FCT at the univariable analysis ($p < 0.05$) were included in a logistic regression model for the multivariable analysis. We calculated odds ratios and their 95% confidence intervals for FCT, adjusted for the co-variables included in the model. We repeated the analysis restricting the population from Lisbon and Porto metropolitan areas. In Lisbon, the best fitting regression model included age, country of birth, having a chronic disease, and treatment regimen; in Porto, the model included age, alcohol abuse, IDU, and users of other drugs, living in a community health centre and treatment regimen. Analyses were performed using STATA (version 14; Stata Corporation, College Station, TX, USA).

Results

Descriptive analysis

From 2013 to 2017, overall, 17,144 LTBI cases were recorded in the national TB clinical notification and follow-up surveillance system (SVIG-TB), from whom 15,478 had available data on treatment outcome (90.3%). From those, 13,346 (86.2%) succeeded to complete treatment (SCT) and 2132 (13.8%) failed to complete treatment (FCT).

Women and men were similarly represented. 79.8% of the LTBI cases were between 30 and 69 years old. The Northern and LTV regions were the most represented, with 9003 (52.5%) and 4265 (24.9%) of the recorded LTBI cases and 20.8% and 25.3% of cases failing to complete treatment, respectively. The country of birth was Portugal approximately in 90% of the individuals in all groups. Within those people not born in Portugal, the highest percentages were people born in Africa and in Asia. Among patients that completed treatment, the most common LTBI treatment regimen was isoniazid-only for 6 or 9 months (54.9 and 32.2%, respectively), followed by the combination of isoniazid and rifampicin for 3 months (7.3%), rifampicin-only for 4 months (2.6%), and finally isoniazid-only courses for 4 months or less (2.0%). The proportion of LTBI patients treated under DOT was 8.1% (Table 19 and Table 20).

Table 19. Geographic and sociodemographic characteristics of patients initiating treatment for latent tuberculosis infection in Portugal between 2013 and 2017, according to treatment outcome (N = 15,478)

	SCT, N (%)[a]	FCT, N (%)[a]
	13,346	**2132**
Sex		
Women	6921 (51.9)	1069 (50.1)
Men	6425 (48.1)	1063 (49.9)
Age, Median (IQR)	43 (30-55)	44 (30-55)
Missing[b]	9 (0.1)	3 (0.1)
Age, categorized		
0-15 years	1366 (10.2)	158 (7.4)
16-29 years	1908 (14.3)	358 (16.8)
30-49 years	5166 (38.7)	820 (38.5)
50-69 years	4184 (31.4)	658 (30.9)
≥70 years	713 (5.3)	135 (6.3)
Missing[b]	9 (0.1)	3 (0.1)
Region of residence		
North	7134 (53.5)	1151 (54.0)
Lisbon and Tagus valley	3184 (23.9)	569 (26.7)
Central	1684 (12.6)	185 (8.7)
Algarve	722 (5.4)	143 (6.7)
Alentejo	345 (2.6)	52 (2.4)
Açores	190 (1.4)	15 (0.7)
Madeira	87 (0.6)	17 (0.8)
Metropolitan area		
Lisbon	2628 (19.7)	498 (23.3)
Porto	3290 (24.6)	590 (27.7)
Rest of the country	7428 (55.7)	1044 (48.0)
Region/Country of birth		
Portugal	12,491 (93. 6)	1940 (91.0)
Africa	473 (3.5)	118 (5.5)
Eastern Europe	101 (0.8)	22 (1.0)
South and Central America	102 (0.8)	14 (0.7)
Asia	70 (0.5)	22 (1.0)
Western/central Europe,USA	84 (0.6)	13 (0.6)
Missing[b]	25 (0.2)	3 (0.2)
Incarceration		
No	12,960 (97.1)	2087 (97.9)
Yes	181 (1.4)	12 (0.6)
Missing[b]	205 (1.5)	33 (1.5)
Living in a community health centre		
Yes	425 (3.2)	61 (2.9)
Missing[b]	228 (1.7)	33 (1.5)

[a]N (%)= total number (percentage), [b]those variables with no missing category=no missing values were found, IQR= Interquartile range, [c]COPD=Chronic obstructive pulmonary disease, [d]DOT=Directly Observed Therapy. Those variables with no missing category=no missing values were found. NA=Not available (no date of treatment completion) SCT: success to complete treatment; FCT: failure to complete treatment.

Table 20. Medical characteristics of patients initiating treatment for latent tuberculosis infection in Portugal between 2013 and 2017, according to treatment outcome (N = 15,478)

	SCT, N (%)[a]	FCT, N (%)[a]
	13,346	**2132**
Chronic liver disease		
No	13,188 (98.8)	2090 (98.0)
Yes	158 (1.2)	42 (2.0)
Chronic kidney disease		
No	13,314 (99.8)	2115 (99.2)
Yes	32 (0.2)	17 (0.8)
COPD[c]		
No	13,225 (99.1)	2101 (98.5)
Yes	121 (0.9)	31 (1.5)
Diabetes mellitus		
No	13,009 (97.5)	2070 (97.1)
Yes	337 (2.5)	62 (2.9)
Alcohol abuse		
No	12,561 (94.1)	1932 (90.6)
Yes	302 (2.3)	111 (5.2)
Missing[b]	483 (3.6)	89 (4.2)
Intravenous drug users		
No	12,857 (96.3)	2004 (94.0)
Yes	121 (0.9)	46 (2.2)
Missing[b]	368 (2.8)	82 (3.8)
Users of other drugs		
No	12,623 (94.6)	1965 (92.2)
Yes	367 (2.7)	87 (4.1)
Missing[b]	356 (2.7)	80 (3.7)
DOT[d]		
No	11,226 (84.1)	1850 (86.8)
Yes	1,119 (8.4)	137 (6.4)
Missing[b]	1,001 (7.5)	145 (6.8)
Treatment regimen		
Only isoniazid (H)	11,893 (89.1)	1988 (93.3)
≤4 months	263 (2.0)	NA
~6 months	7332 (54.9)	NA
~9 months	4298 (32.2)	NA
Only rifampicin (R)	338 (2.6)	45 (2.1)
R and H	978 (7.3)	81 (3.8)
Missing[b]	137 (1.0)	18 (0.8)
Previously treated (PT)		
No	13,218 (99.0)	2105 (98.7)
Yes	128 (1.0)	27 (1.3)
Compliance in PT		
No	46 (35.9)	15 (55.6)
Yes	52 (40.6)	9 (33.3)
Missing[b]	30 (23.5)	3 (11.1)

[a]N (%)= total number (percentage), [b]those variables with no missing category=no missing values were found, IQR= Interquartile range, [c]COPD=Chronic obstructive pulmonary disease, [d]DOT=Directly Observed Therapy. Those variables with no missing category=no missing values were found. NA=Not available (no date of treatment completion) SCT: success to complete treatment; FCT: failure to complete treatment.

The proportion of those treated with both isoniazid and rifampicin simultaneously was higher in SCT, 7.3% than in FCT, 3.8%. For those that finished the treatment (SCT), the shorter

treatment regimen was isoniazid and rifampicin simultaneously with 92 days as median (Md) (interquartile range (IQR), 91–101) followed by treatment only with rifampicin (Md, 123 days (IQR, 121–133)) and treatment only with isoniazid (Md, 191 days (IQR, 183–273)). The DOT showed higher percentage for the SCT than in the FCT, 8.38% vs. 6.43%. Cases with unknown treatment outcomes showed values between those observed in SCT and FCT (Table 21).

Table 21. Patient geographic, sociodemographic and medical characteristics in unknown outcome (UO) group in Portugal between 2013 and 2017.

Geographic and sociodemographic characteristics		Medical characteristics	
	UO, N (%)[a]		UO, N (%)[a]
	1666		1666
Sex		**Chronic liver disease**	
Women	828 (49.7)	No	1,646 (98.8)
Men	838 (50.3)	Yes	20 (1.2)
Age, Median (IQR)	45 (32-57)	**Chronic kidney disease**	
Missing[b]	2 (0.1)	No	1,651 (99.1)
Age, categorized		Yes	15 (0.9)
0-15 years	145 (8.7)	**COPD[c]**	
16-29 years	206 (12.4)	No	1,651 (99.1)
30-49 years	647 (38.8)	Yes	15 (0.9)
50-69 years	523 (31.4)	**Diabetes mellitus**	
≥70 years	143 (8.6)	No	1,627 (97.7)
Missing[b]	2 (0.1)	Yes	39 (2.3)
Region of residence		**Alcohol abuse**	
North	718 (43.1)	No	1,537 (92.3)
Lisbon and Tagus valley	512 (30.7)	Yes	32 (1.9)
Central	311 (18.7)	Missing[b]	97 (5.8)
Algarve	55 (3.3)	**Intravenous drug users**	
Alentejo	55 (3.3)	No	1,566 (94.0%)
Açores	13 (0.8)	Yes	28 (1.7%)
Madeira	2 (0.1)	Missing[b]	72 (4.3%)
Metropolitan area		**Users of other drugs**	
Lisbon	419 (25.2)	No	1,559 (93.6)
Porto	287 (17.2)	Yes	32 (1.9)
Others	960 (57.6)	Missing[b]	75 (4.5)
Region/Country of birth		**DOT[d]**	
Portugal	1,538 (92.3)	No	1,433 (86.0)
Africa	79 (4.8)	Yes	81 (4.9)
Eastern Europe	8 (0.5)	Missing[b]	152 (9.1)
South and central America	15 (0.9)	**Treatment scheme**	
Asia	12 (0.7)	Only isoniazid (H)	1,522 (91.3)
Western, central Europe, USA	9 (0.5)	Only rifampicin (R)	28 (1.7)
Missing[b]	5 (0.3)	R and H	93 (5.6)
Incarceration		Missing[b]	23 (1.4)
No	1,586 (95.2)	**Previously treated (PT)**	
Yes	58 (3.5)	No	1,622 (97.4)
Missing[b]	22 (1.3)	Yes	44 (2.6)
Living in a community		**Compliance in PT**	
No	1,547 (92.9)	No	14 (31.8)
Yes	60 (3.6)	Yes	18 (40.9)
Missing[b]	59 (3.5)	Missing[b]	12(27.3)

[a]N (%)= total number (percentage), [b]those variables with no missing category=no missing values were found, IQR= Interquartile range, [c]COPD=Chronic obstructive pulmonary disease, [d]DOT=Directly Observed Therapy. Those variables with no missing category=no missing values were found.

Multivariable analysis

After adjusting for the remaining covariables included in the regression models, we found the following characteristics to be independently associated with FCT: (i) being more than

15 years old when compared with those under 15 years old, mostly when they were between 16 and 29 years or ≥ 70 years old (adjusted odds ratio (aOR), 1.65 (95% CI, 1.34–2.04)) and 1.55 (95% CI, 1.19–2.03), respectively); (ii) living in Northern, LTV, and Algarve regions when compared with living in the Central region (aOR, 1.17 (95% CI, 0.97–1.41), 1.24 (95% CI, 1.01–1.53), and 1.63 (95% CI, 1.23–2.17), respectively); (iii) being born in Africa or Asia when compared with being born in Portugal (aOR, 1.57 (95% CI, 1.24–1.98) and 2.04 (95% CI, 1.19–3.50), respectively) (Table 22).

Table 22. Adjusted associations[a] between geographic, socioeconomic and medical factors and failure to complete treatment (FCT) for LTBI (n= 15,478).

Age, categorized	**N**	**%**	**aOR[b]**	**95%CI[c]**
0-15 years	158	9.47	Ref.	
16-29 years	358	14.48	1.65	1.34-2.04
30-49 years	820	12.36	1.32	1.09-1.60
50-69 years	658	12.26	1.27	1.04-1.54
≥70 years	135	13.62	1.55	1.19-2.03
Region of residence				
Central	185	8.49	Ref.	
North	1,151	12.78	1.17	0.97-1.41
Lisbon and Tagus Valley	569	13.34	1.24	1.01-1.53
Algarve	143	15.54	1.63	1.23-2.17
Alentejo	52	11.50	1.04	0.71-1.53
Açores	15	6.88	0.77	0.42-1.42
Madeira	17	16.04	1.43	0.80-2.55
Region/Country of birth				
Portugal	1,940	12.15	Ref.	
Africa	118	17.61	1.57	1.24-1.98
Eastern Europe	22	16.79	1.52	0.92-2.51
South and central America	14	10.69	0.87	0.42-1.70
Asia	22	21.15	2.04	1.19-3.50
Western and central Europe	13	12.26	0.84	0.44-1.59
Incarceration				
No	2,087	12.55	Ref.	
Yes	12	5.58	0.37	0.17-0.82
Chronic disease				
No	1,994	12.22	Ref	
Yes	138	16.61	1.29	1.04-1.60
Alcohol abuse				
No	1,932	12.05	Ref.	
Yes	111	24.94	2.24	1.73-2.90
Intravenous drug users				
No	2,004	12.20	Ref.	
Yes	46	23.59	1.68	1.05-2.68
Users of other drugs				
No	1,965	12.17	Ref.	
Yes	87	17.90	1.27	0.93-1.73
Directly Observed Therapy				
No	1,850	12.75	Ref.	
Yes	137	10.25	0.79	0.64-0.96
Treatment regimen				
Isoniazid-only (~6 or ~9 month course)	1,988	12.91	Ref.	
Isoniazid + rifampicin (~3 months)	81	7.03	0.59	0.45-0.77
Rifampicin-only (~4 months)	45	10.95	0.75	0.52-1.07

[a]Logistic regressions models were adjusted for all the listed variables, [b]aOR=adjusted odds ratio, [c]CI=Confidence interval.

Chronic disease, alcohol abuse, and injectable drug use were also significantly associated with FCT (aOR, 1.29 (95% CI, 1.04–1.60), 2.24 (95% CI, 1.73–2.90), and 1.68 (95% CI, 1.05–2.68), respectively). Incarceration and completing treatment under direct observation by a healthcare worker were significantly associated with a lower risk of FCT (aOR, 0.37 (95% CI, 0.17–0.82)). Finally, LTBI cases treated with a 3-month course of isoniazid plus rifampicin seemed to be less likely to fail to complete treatment than those receiving a 6- or 9-month course of isoniazid (aOR, 0.59 (95% CI, 0.45–0.77)) (Table 22).

In the metropolitan areas, the factors associated with FCT were (i) in Lisbon, being born in Africa when compared with being born in Portugal (aOR, 1.41 (95% CI, 1.07–1.84)), and being treated with a 3-month course of rifampicin when compared with being treated with a 6- or 9-month course of isoniazid (aOR, 1.91 (95% CI, 1.11–3.28)), aOR and (ii) in Porto, being more than 15 years old, alcohol abuser (aOR, 2.64 (95% CI, 1.67–4.16)) and IDU user (aOR, 2.52 (95% CI, 1.10–5.79)) (Table 23 and Table 24).

Table 23. Adjusted association between potential factors and failure to complete treatment (FCT) for LTBI in Lisbon metropolitan area (n= 3,545).

	Lisbon			
Age, categorized	**FCT**		**aOR[ab]**	**95%CI[c]**
	N	**%**		
0-15 years	72	14.72	Ref.	
16-29 years	117	19.50	1.25	0.89-1.74
30-49 years	147	11.93	0.76	0.56-1.04
50-69 years	137	13.06	0.84	0.61-1.15
≥70 years	24	13.87	0.93	0.56-1.56
Region/Country of birth				
Portugal	393	13.43	Ref.	
Africa	82	18.14	1.41	1.07-1.84
Eastern Europe	6	13.33	1.00	0.41-2.41
South and central America	4	8.51	0.59	0.21-1.67
Asia	6	19.35	2.10	0.81-5.42
Western and central Europe	5	15.15	1.15	0.44-3.03
Treatment regimen				
Isoniazid-only (~6 or 9-month course)	464	13.97	Ref.	
Isoniazid + rifampicin (~3 months)	12	9.84	0.58	0.31-1.06
Rifampicin-only (~4 months)	19	27.14	1.91	1.11-3.28

[a]logistic regressions models were adjusted for all the listed variables, [b]aOR= adjusted odds ratio, [c]CI=Confidence interval.

Table 24. Adjusted association between potential factors and failure to complete treatment (FCT) for LTBI in Porto metropolitan area (n= 3,545).

	Porto			
Age, categorized	**FCT**		**aOR[ab]**	**95%CI[c]**
	N	**%**		
0-15 years	38	9.45	Ref.	
16-29 years	82	15.21	1.60	1.06-2.43
30-49 years	248	14.98	1.48	1.03-2.14
50-69 years	185	13.43	1.36	0.94-1.98
≥70 years	37	19.17	2.06	1.24-3.44
Living in a community health center				
No	572	14.03	Ref.	
Yes	9	26.47	1.84	0.66-5.11
Chronic disease				
No	553	13.90	Ref.	
Yes	37	19.47	1.40	0.94-2.10
Alcohol abuse				
No	536	13.51	Ref.	
Yes	32	29.36	2.64	1.67-4.16
Intravenous drug users				
No	558	13.82	Ref.	
Yes	14	33.33	2.52	1.10-5.79
Users of other drugs				
No	551	13.80	Ref.	
Yes	22	24.44	1.17	0.63-2.19
Treatment regimen				
Isoniazid-only (~6 or 9 month course)	578	14.43	Ref.	
Isoniazid + rifampicin (~3 months)	3	6.12	0.37	0.11-1.25
Rifampicin-only (~4 months)	9	8.33	0.48	0.23-1.01

[a]logistic regressions models were adjusted for all the listed variables, [b]aOR= adjusted odds ratio, [c]CI=Confidence interval.

Cases not living in Lisbon or Porto metropolitan areas had characteristics similar to those shown in Table 22 (Table 25).

Table 25. Adjusted association between geographic, socioeconomic and medical factors and failure to complete treatment (FCT) for LTBI in areas of Portugal other than Lisbon and Porto metropolitan areas (n= 9,432).

Age, categorized	**FCT**		**aOR[ab]**	**95%CI[c]**
	N	**%**		
0-15 years	48	6.17	1	
16-29 years	159	11.93	2.03	1.43-2.88
30-49 years	425	11.35	1.75	1.27-2.42
50-69 years	336	11.44	1.73	1.25-2.41
≥70 years	74	11.84	1.91	1.26-2.88
Region/Country of birth				
Portugal	925	10.88	1	
Africa	29	16.02	1.66	1.02-2.70
Eastern Europe	14	18.18	1.89	0.99-3.63
South and central America	5	8.47	0.68	0.21-2.24
Asia	16	21.92	2.26	1.25-4.09
Western and central Europe	4	7.02	0.35	0.09-1.48
Incarceration				
No	1,020	11.22	1	
Yes	9	4.81	0.28	0.01-0.81
Chronic disease				
No	974	10.83	1	
Yes	70	15.87	1.45	1.06-1.98
Alcohol abuse				
No	915	10.56	1	
Yes	70	24.56	2.32	1.63-3.29
Intravenous drug users				
No	959	10.70	1	
Yes	30	22.73	1.56	0.84-2.89
Users of other drugs				
No	942	10.65	1	
Yes	50	19.46	1.62	1.04-2.51
Directly Observed Therapy				
No	916	11.72	1	
Yes	46	6.34	0.56	0.40-0.79
Treatment regimen				
Isoniazid-only (~6- or ~9-month course)	946	11.72	1	
Isoniazid + rifampicin (~3 months)	66	6.73	0.61	0.45-0.81
Rifampicin-only (~4 months)	17	7.30	0.58	0.33-1.03

Discussion

We found that the FCT was associated with increasing age, region of notification (higher estimates in LTV, Northern, and Algarve regions), and the country of birth (higher for those born in Asia or Africa). Having a chronic disease, alcohol abusing, and being an IDU also were associated with higher risk of FCT. Being incarcerated and receiving LTBI treatment under direct observation by a healthcare worker were associated with lower risk of FCT. Treatment with the shorter 3-month course of isoniazid plus rifampicin was associated with a lower risk of FCT than the standard 6- or 9-month course of isoniazid-only. In the metropolitan areas of Lisbon and Porto, the factors independently associated with a higher risk of FCT were (i) in Lisbon, being born in Africa and treated with a 3-month course of rifampicin-only and (ii) in Porto, being older than 15 years old, alcohol abusing, and injecting/using other drugs.

Many studies have been published analysing factors associated with failure to complete treatment in TB but not so many regarding failure to complete treatment in LTBI. Our study is focused on failure to complete LTBI treatment, which is a crucial activity as is indicated in WHO End TB strategy.[224] This study showed a lower FCT rate (14%) than what was previously reported in other countries (varying from 20 to 60%).[218,223] We attribute this low FCT rate to the efforts performed by the different actors in Portugal in their aim to increase the number of LTBI cases diagnosed and treated as well as the treatment completion and success. Since the creation of the current National TB Programme (NTP) in 1995, LTBI screening and effective follow-up of patients in specialized TB services in primary care (TB outpatient centres) have been implemented and a robust surveillance system that included notification of not only TB but also LTBI cases was put in place. NTP current strategies and main actions to be developed can be found in the report “challenges and strategies in Tuberculosis in Portugal”.[102] On the other hand, as we only included patients that started treatment for LTBI in this study (and not all the patients that had LTBI diagnosed), it might be argued that FCT is being underestimated. Similarly, the cut-off of 4 months of treatment that we accepted to correspond to a complete 6-month course of isoniazid-only could be also underestimating FCT (although 75% of those patients received a 183-day or longer course of isoniazid-only). Uncontrolled characteristics in the 9% of patients that had unknown treatment outcome did not seem to bias our findings, as their characteristics are similar to those patients included in the analysis and had known outcome.

Lower FCT in children and teenagers is probably due to a closer follow-up by their families (frequently, other relatives are simultaneously being treated for TB or LTBI) and healthcare services increased concern when younger age groups are diagnosed LTBI. Incarcerated

patients also seem to be less prone to FCT, which could be explained by the fact that these patients are daily treated for LTBI under direct observation by a healthcare worker (more than two-thirds of the incarcerated patients were doing DOT). DOT in Portugal is recommended only to TB patients but not to LTBI patients, which could explain why other patients that usually have higher FCT rates (such is the case of IDU) showed lower FCT rates in our study (as IDU cases in methadone programs will often receive simultaneously LTBI treatment under observation).

The studies that aim to disentangle the factors associated to FCT in LTBI had shown conflicting results. Priest et al. have been describing that females had increased risk of FCT,[225] but Lobue et al. found the opposite.[226] In both cases, they described higher risk of FCT at older ages and in those non-national individuals that more recently arrived to the host community or those individuals being born abroad.[225,226]

Fiske et al. found that the FCT was not associated with race, sex, age, and place of birth but with other risk factors such as treatment with isoniazid or interview at start of treatment outside of a healthcare setting (both of these having higher FCT).[219] In our study, we did not find an association between FCT and sex, but we found it with age (higher risk for older than 15 years old) and place of birth (higher risk for those individuals born in Africa or Asia).

Ambrona de Marcos et al. suggested that patients understanding about LTBI and its treatment improved treatment completion,[218] and, in another study, Hirsch-Moverman et al. found that circumstances related to social determinants of health such as alcohol or other drug addiction, homelessness, or not being married were associated with increased FCT in LTBI treatment.[227] Actually, in a systematic review performed by Stuurman et al., it was found that shorter treatment regimens, having less adverse effects during course of treatment, not abusing alcohol, not being part of vulnerable groups, and DOT had higher completion rates.[228]

However, only shorter treatment regimens and social interventions were associated with higher completion rates, whereas DOT or incentive interventions showed contradictory results.[228] In a clinical trial, Belknap et al. found similar results in treatment completion when comparing DOT vs. weekly self-administered isoniazid and rifapentine.[229] Recently, it has been described that, in both children and adults, the 4-month regimen of rifampicin is similar in terms of efficacy to 9 months of isoniazid, and the treatment adherence is superior.[216,217] In spite that approximately 90% of our patients were being treated only with isoniazid (54.9% following the 6-month course), our results suggest the same as most of the previous studies: shorter course treatments and DOT can reduce FCT.

The facts of being dependent on alcohol or drug user showed an increased risk of FCT. Regarding place of residence, we only analysed by region/metropolitan area and not smaller geographical areas (where probably we could better analyse socioeconomic status–related variables) and we only can see how the more populated regions (LTV and Northern regions) had more LTBI patients starting treatment. These regions and Algarve showed higher FCT rates than the central region. Considering the number of cases of active TB reported in Portugal, similar number of LTBI patients would be expected in LTV and Northern regions. However, this study showed that in Northern region, the number of LTBI treatments per TB patient was twice the number found in LTV region—possibly reflecting regional policy differences that must be addressed. The FCT is slightly higher in LTV than in the Northern region, even though the proportions of IDU and alcohol-abusing patients (factors shown to be associated with FTC) are higher in the Northern region. This could be explained by the even higher FCT in those being born abroad, a subgroup of the population that tends to be more concentrated in Lisbon metropolitan area. Although not much scientific evidence is available yet, there have been studies describing the probable association between some chronic conditions or risk factors such as CKD,[230,231] DM,[232] lung cancer[233], COPD, or smoking[234] and an increased vulnerability to infection and an increased likelihood of progression from LTBI to active TB.[235] However, having a chronic disease seems to be also associated with a higher FCT rate, probably due to increased toxicity, described as one of the main causes of FCT.[236] In our study, we found that having a chronic disease seems to be associated with FCT in LTBI.

In Portugal, in 2014, the most prevalent identified risk factor for TB differed in the two biggest urban areas: in Lisbon was being migrant or living with HIV and in Porto was alcohol abuse and being IDU.[149] Our results are consistent with these data, as we found that the main determinants of FCT were in Lisbon, being born in Africa, and in Porto, alcohol abusing and intravenous drug using. These differences may help in future opportunistic LTBI screenings identifying vulnerable populations in both metropolitan areas. Therefore, we believe that there are important geographical differences that should be considered when planning specific measures to fight TB. In our study, we did not have access to HIV status. If available, these data could bring valuable additional information to our analysis as the longer 9-month isoniazid-only treatment course is recommended for these patients.

To our knowledge, few studies have analysed FCT and associated factors using all available national data in a country. The main strength of this study is the sample size, as all LTBI cases starting treatment in Portugal from 2013 to 2017 were included. The main limitation is the lack of information regarding later development of TB. This would allow us

to make further analysis to try to assess the role of FCT and associated factors in later development of TB. Further studies are needed to address this question. Factors like effectiveness of LTBI treatment (i.e. not developing active TB after completion of LTBI treatment) and drug-related adverse effects were not addressed in this study and so we cannot recommend shorter treatment regimens based solely on the decreased probability of FCT. Another limitation is that patients that are diagnosed LTBI but do not start treatment are not reported to the National Tuberculosis Program. This fact could bring a selection bias, as these patients could have different characteristics than those who effectively start treatment for LTBI.

In conclusion, we believe that our study brings new insights on failure to complete LTBI treatment, showing the relevance of some possible determinants, such as age, chronic disease, alcohol abuse, drug use, and geographical factors such are country of birth, region, or metropolitan area of residence. Our study reinforces the evidence that shorter course treatments for LTBI are more likely to be completed, particularly in individuals at an increased risk of FCT. WHO End TB strategy states that the LTBI treatment should be a priority for TB programs—our study shows some aspects that must be taken into consideration in order to plan effective public health interventions in this area.

3.4. COMPARING THE COST-EFFECTIVENESS OF TWO SCREENING STRATEGIES FOR LATENT TUBERCULOSIS INFECTION IN PORTUGAL

Pulmonology 27 (2021) 493–499

PULMONOLOGY

www.journalpulmonology.org

ORIGINAL ARTICLE

Comparing the cost-effectiveness of two screening strategies for latent tuberculosis infection in Portugal

Sofia Sousa[a,b,1,*], Diogo Rocha[c,1], Joelma C. Silva[d,1], Ana Isabel Ribeiro[e,f,1], Guilherme Gonçalves[b,1], Álvaro Almeida[g,1], Ana Maria Correia[h,1], Raquel Duarte[e,f,i,j,1], Carlos Carvalho[a,k,b,1]

[a] *Public Health Unit - South Sousa Valley, Northern Regional Health Administration, Paredes, Portugal*
[b] *Multidisciplinary Unit for Biomedical Research (UMIB), Institute of Biomedical Sciences Abel Salazar, University of Porto, Porto, Portugal*
[c] *Faculty of Medicine, University of Porto, Porto, Portugal*
[d] *Pulmonology Department of Centro Hospitalar de Trás-os-Montes e Alto Douro, Vila Real, Portugal*
[e] *Epidemiology Research Unit (EpiUnit) - Institute of Public Health, University of Porto, Porto, Portugal*
[f] *Departamento de Ciências de Saúde Pública, Ciências Forenses e Educação Médica, Faculty of Medicine, University of Porto, Porto, Portugal*
[g] *CEF.UP – Center for Economics and Finance at the University of Porto and Faculdade de Economia, Universidade do Porto*
[h] *Northern Regional Health Administration, Department of Public Health, Porto, Portugal*
[i] *Serviço de Pneumologia, Centro Hospitalar de Vila Nova de Gaia/Espinho, Vila Nova de Gaia, Portugal*
[j] *National Tuberculosis Program, Directorate-General of Health, Lisbon, Portugal*
[k] *Department of Public Health, Northern Regional Health Administration, Porto, Portugal*

Received 8 November 2020; accepted 8 April 2021
Available online 27 May 2021

KEY WORDS
Tuberculin Skin Test;
Interferon-Gamma Release Assays;
Cost-effectiveness;
Latent tuberculosis;
Screening

Abstract
Introduction and objectives: Screening for latent tuberculosis infection (LTBI) in close contacts of infectious TB cases might include Tuberculin Skin Test (TST) and Interferon-Gamma Release Assays (IGRA), in combination or as single-tests. In Portugal, the screening strategy changed from TST followed by IGRA to IGRA-only testing in 2016. Our objective was to compare the cost-effectiveness of two-step TST/IGRA with the current IGRA-only screening strategy in immunocompetent individuals exposed to individuals with respiratory TB.
Materials and Methods: We reviewed clinical records of individuals exposed to infectious TB cases diagnosed in 2015 and 2016, in two TB outpatient centers in the district of Porto. We estimated medical, non-medical and indirect costs for each screening strategy, taking into account costs of tests and health care personnel, travel distance from place of residence to screening site and employment status. We calculated the incremental cost-effectiveness ratio (ICER) as the cost difference between the two screening strategies with the difference number of LTBI diagnosis as a measure of cost-effectiveness, assuming that treating LTBI is a cost-effective intervention. We also calculated adjusted odds-ratios to test the association between diagnosis of LTBI and screening strategy and estimated the total cost for averting a potential TB case.
Results: We compared 499 contacts TST/IGRA screened with 547 IGRA-only. IGRA-only strategy yielded a higher screening effectiveness for diagnosing latent tuberculosis infection (aOR 2.12, 95%CI: 1.53 - 2.94). ICER was €106 per LTBI diagnosis, representing increased effectiveness with a slightly increased cost of IGRA-only screening strategy.
Conclusions: Our data suggests that in Portugal LTBI screening with IGRA-only is more cost-effective than the two-step TST/IGRA testing strategy, preventing a higher number of cases of TB cases.

* Corresponding author.
E-mail address: aifos.sousa@gmail.com (S. Sousa).
[1] The authors do not have a commercial or other association that might pose a conflict of interest. There was no financial support to this study.

https://doi.org/10.1016/j.pulmoe.2021.04.002

Background

Systematic diagnosis and treatment of latent tuberculosis infection (LTBI) is a key part of the TB elimination strategy in low-incidence countries.[237] Screening strategies in individuals with close contact with infectious cases of TB include the tuberculin skin test (TST) followed by the interferon-gamma-release assay (IGRA) in individuals with positive TST results (two-step strategy); single-step IGRA testing; and single-step TST testing.[238] IGRA tests have specificity greater than 95% in the diagnosis of LTBI.[239] The TST specificity is similar (97%) in populations not vaccinated with Bacillus Calmette-Guérin (BCG), but is considerably lower (60%) in those vaccinated.[239] Sensitivity of the two tests is roughly the same: 80-90% for IGRA; 80% for TST.[239,240] Some studies have shown that the progression rate (likelihood that a person with a positive test will develop active TB) is higher in IGRA-positive individuals.[241–243] A cost-effectiveness study in the United Kingdom estimated that two-step TST/IGRA screening strategy is less costly than single-step IGRA testing (£162,387 vs. £203,983 per 1000 contacts).[238] However, this study only considered medical costs. In

France, a decision analysis model, considering only direct medical costs, showed that in almost all scenarios QuantiFERON (QFT) was more effective and cost-effective than TST in detecting LTBI.[244] One study in Brazil, though, showed that the most cost-effective strategy was TST (US$ 16,021/averted case) and that the incremental cost-effectiveness ratio was US$ 227,977/averted TB case for QFT-GIT.[245] Another study of individuals entering the Dallas County Jail (Texas, United States) reported a substantially higher positivity rate of IGRA than TST: these authors suggested that sensitivity of TST screening was lower, and that IGRA was more time-efficient and associated with four-fold lower indirect costs. The overall cost per LTBI case detected was nearly three-times higher for the TST than the IGRA.[246] A recent report of the European Centre for Disease Prevention and Control used a deterministic TB transmission model to predict the impact of different LTBI screening and treatment strategies for several risk-groups, including contacts of TB cases. They concluded that from the healthcare perspective, LTBI screening is most cost-effective when done using the two-step approach (TST first and, if positive, followed by IGRA).[247,248] Given the heterogeneous results of different studies, the World Health Organization advised more research in this field.[224] The Portuguese National Health Service maintains TB outpatient centres that are responsible for TB and ITBL diagnosis, treatment, and screening across the country, under technical guidance from the National Tuberculosis Programme. In the Northern Region, TB outpatient centres switched from the two-step TST/IGRA to the single-step IGRA-only screening strategy, after shortages in tuberculin supply. Before this switch, TST was performed to exposed contacts immediately after diagnosis of TB in index case and repeated 8-10 weeks after the last exposure of risk, followed by IGRA testing (QuantiFERON Gold Plus®) whenever TST was positive. After 2016, IGRA-only (QuantiFERON Gold Plus®) was performed only once, 8-10 weeks after the last exposure to index case of TB. The IGRA assay available during the study period in the Northern Region was the QFT-Plus assay.[249,250] The objective of this study was to compare the cost-effectiveness of the two strategies described above, in terms of medical costs (for the healthcare system) and direct and indirect non-medical costs related to LTBI screening (excluding treatment costs), for LTBI screening in close contacts of confirmed cases of respiratory TB.

Material and Methods

Patient Selection

We examined clinical records of all individuals screened for TB and LTBI in two TB outpatient centres from the Northern Health Region: *Penafiel* (180,000 inhabitants, annual TB incidence rate of 47.7/10^5 in 2012-16) and *Vila Nova de Gaia* (330,000 inhabitants, annual TB incidence rate of 26.0/10^5 in 2012-16).[251] Only individuals exposed to an infectious TB patient, diagnosed from 1 January 2015 to 31 December 2016, were included. Exclusion criteria were those described in Figure 16.

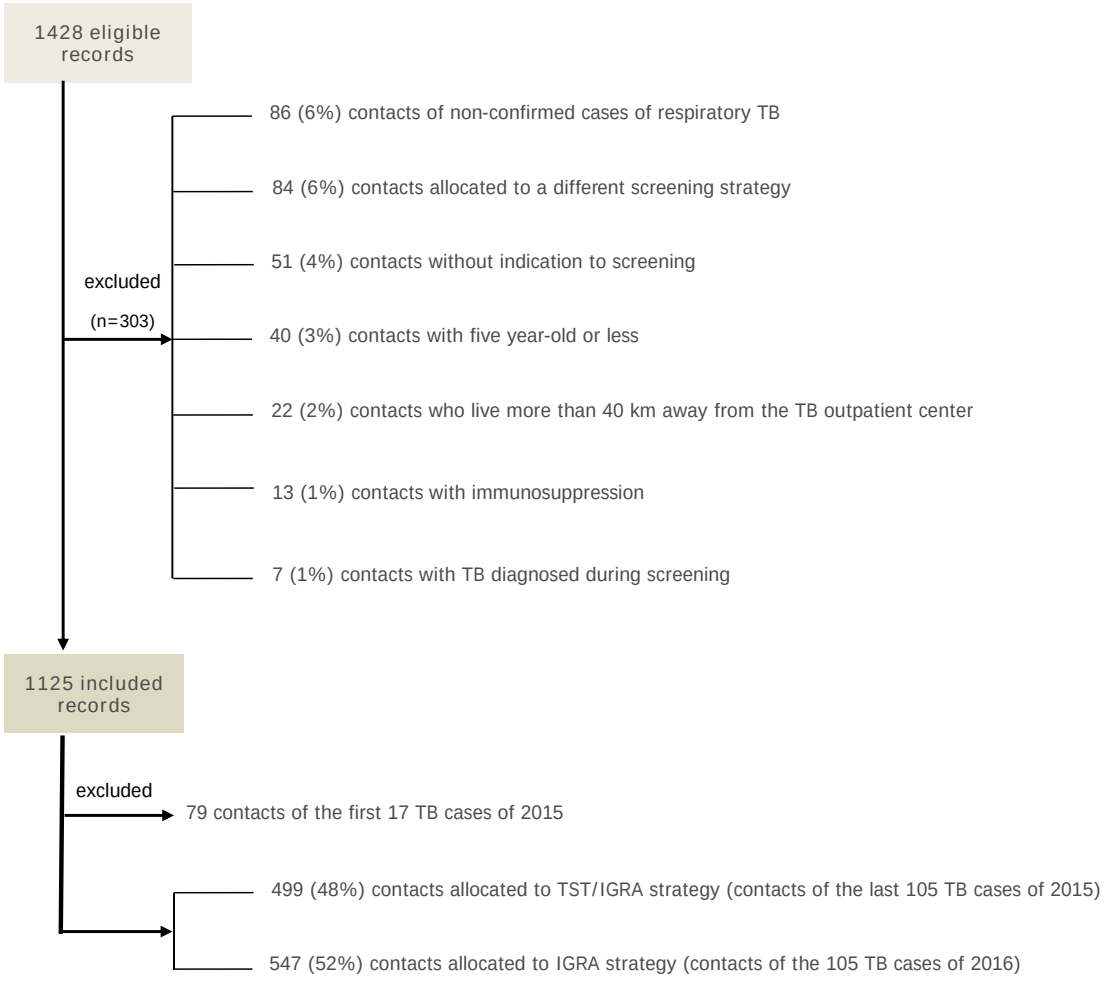


Figure 16. Eligible and included records, with reference to exclusion criteria.

We extracted socio-demographic information (sex, age, parish of residence, employment status), clinical information (history of TB/LTBI, immunosuppression, diagnosis of TB during

screening), LTBI screening test information (screening strategy, date[s], result[s], diagnosis, initiation of treatment), and information on the index TB cases (diagnosis date, sputum smear results).

Cost Effectiveness

Cost-effectiveness was expressed as an incremental cost-effectiveness ratio (ICER), which was calculated by dividing the cost difference between the two screening strategies with the difference number of LTBI diagnosis.[247,252] We assumed that treating LTBI is a cost-effective intervention and focused only on the differences between strategies until the moment of LTBI diagnosis.

The median costs in the TST/IGRA and TST-only groups were compared using Mann-Whitney test. The proportion of individuals diagnosed with LTBI, number of visits to the TB outpatient centre, and adherence to screening were also compared for these two groups using the Chi-squared test or Fisher's exact test, as appropriate. Adjusted odds-ratios (aORs) and 95% confidence intervals (95% CIs) for each screening strategy and diagnosis of LTBI were calculated using logistic regression. Screening strategy, sex, age, place of residence, professional status (employed/unemployed), TB outpatient centre, index case sputum smear positivity, infectious period, and site of disease in the index case were included in the initial regression models. In this procedure, a backward stepwise approach was used, and at each step, the least significant variable that was not a substantial confounder (whose removal would lead to a change of more than 20% in the OR of one or more parameters remaining in the model) was removed. Sex and index case sputum smear positivity were retained regardless of p-value.

Direct individual medical costs were calculated as the sum of estimated costs from the screening test(s), specimen transportation to the testing laboratory, and the work of healthcare professionals (collection and testing of specimens) – data provided by the Northern Regional Health Administration for 2014-2016. IGRA (QuantiFERON-TB GoldPlus) cost was €37.66 (including ELISA kit, antigen, and mitogen) and the TST cost was €1.00 (including tuberculin). Disposable material costs were estimated using online prices (including laboratory materials, blood collection materials, tuberculin needles and syringes, gloves, compresses), and were calculated as €0.31 for one TST and €0.57 for one IGRA.

Direct non-medical costs were estimated per screened individual by multiplying the number of visits to the health centre by the distance travelled (calculated taking into account

patient's address), with an estimated cost of €0.10 per km. To assess the impact of this estimate on final results, a sensitivity analysis was also performed using an estimated cost of €0.35 per km, the reference value used by the Portuguese Government. This cost was adjusted considering that contacts of the same index case could share their mean of transport if they went to the TB outpatient centre on the same day.

Indirect costs per screened individual were calculated by multiplying the number of visits to the health centre by the half-daily average income (from *Instituto Nacional de Estatística*, Statistics Portugal) for working individuals. Half-daily income was used in order to not overestimate indirect costs, considering that TB is associated with socio-economic deprivation.[253,254] A sensitivity analysis was also performed using daily average income. The number of potentially averted TB cases was estimated based on the number of individuals who started LTBI treatment, a 10% lifetime risk of developing TB, and a 70% efficacy of treatment (assuming that all patients who began LTBI treatment finished their treatment).[255]

Definitions

The following definitions were used in this study:

- Adherence to screening: proportion of individuals who showed up for screening that completed all recommended screening steps;[256]
- Confirmed respiratory TB patient: person with a positive respiratory TB diagnosis (tracheal, laryngeal, bronchial, pulmonary, and/or pleural), confirmed by a positive culture for Mycobacterium tuberculosis complex (MTC) or a positive smear plus MTC nucleic acid detection in sputum or bronchoalveolar lavage;[257]
- Close contact: person who had close contact with a patient who had respiratory TB for a cumulative time of at least 8 h if sputum smear-positive, or 40 h if sputum smear-negative (National Tuberculosis Programme recommendation);[258]
- Immunosuppressed individual: individual receiving chemotherapy, radiotherapy, an immunosuppressive drug, or infected with HIV;[256]
- Period of infectiousness: time interval during which MTC may be transferred between individuals, estimated as the number of days between onset of symptoms and TB diagnosis;[259]
- Latent tuberculosis infection: positive IGRA test in an individual who does not have active TB;[259]
- Medical direct costs: healthcare-related costs (screening tests, specimen transportation to the testing laboratory, and work of healthcare professional);[256]

- Non-medical direct costs: costs related to the transportation of an individual to a TB outpatient centre for screening;[256]
- Indirect costs: productivity loss from an individual's absence from work because of travel to the TB outpatient centre for screening.[256]

Stata®IC 15.0 (Student version) was used for statistical analyses. Ethical approval for this study was obtained from the Ethics Boards of the Institute of Public Health of the University of Porto (CE17061) and the Northern Regional Health Administration (39/2017).

Results

From 1428 eligible close contacts of infectious cases of TB, 303 (21%) met the exclusion criteria and were rejected from the analysis (Figure 16). From the remaining 1125 individuals, 578 (51%) had been screened in 2015 using the TST/IGRA strategy and 574 (49%) were screened in 2016 using the IGRA-only strategy. In order to ensure that we were comparing the costs of screening contacts of the same number of TB cases, we included all 2016 TB cases and the same number of cases diagnosed in 2015 starting from the end of the year (79 contacts screened with TST/IGRA in January/February 2015 were excluded).

Baseline Comparison of Groups

The median age of screened individuals was 38 years (interquartile range [IQR]: 25.0-51.0), 54% were women, and 58% (n = 641) were employed. Their places of residency were a median of 11 km (IQR 6.5-15.5) and 14 min (IQR 8.5-19.4) away from the visited TB outpatient centre. The two groups had no significant differences in terms of age, sex, employment status, and distance to the visited TB centre (Table 26).

However, the IGRA-only group had a significantly higher proportion of individuals who were exposed to highly infectious TB patients (positive sputum smears). The proportion of contacts diagnosed with LTBI and the adherence to screening were also greater in the IGRA-only group.

Table 26. Characteristics of screened contacts by strategy used.

	TST/IGRA strategy n=499	**IGRA strategy n=547**	**p-value**

Median age	37	39	0.108
Proportion of male individuals	49%	44%	0.118
Proportion of working individuals	59%	57%	0.689
Median distance from place of residence to TB outpatient center in kilometers	11	10	0.148
Median time needed to travel from place of residence to TB outpatient center in minutes	15	13	0.014
Average number of contacts per index case	5	6	
Proportion of contacts of a TB case with positive sputum-smear	66%	80%	<0.001
Median infectious time of TB cases	56 days	67 days	0.370
Proportion of contacts who were exposed to a pulmonary TB case (not pleural)	97%	99%	0.118
Proportion of diagnosed LTBI	15%	27%	<0.001
Median number of visits to the TB outpatient center	4	2	<0.001
Adherence to screening	81%	86%	0.038
First step screening concluded	98%		

Multivariable Analysis

After adjusting for sex, age, TB outpatient centre, index case sputum smear results, and period of infectiousness, the IGRA-only group had an increased risk for diagnosis of LTBI (aOR = 2.12, 95% CI: 1.48-2.93) (Table 27).

Table 27. Logistic regression model for diagnosis of LTBI (final model).

Variables	Reference group	Other categories	Odds Ratio	95% Confidence Interval
Screening strategy	TST/IGRA strategy	IGRA strategy	2.12	1.53 - 2.94
Sex	Female	Male	1.11	0.81 - 1.51
Age group	5-10 years	11-17 years	2.89	0.76 - 10.89
		18-29 years	2.88	0.83 - 10.01
		30-39 years	2.43	0.69 - 8.55
		40-49 years	2.11	0.60 - 7.42
		50-59 years	3.25	0.92 - 11.51
		≥60 years	3.92	1.10 - 13.94
TB Outpatient Center	Gaia TB outpatient center	Penafiel TB outpatient center	1.09	0.79 - 1.49
Infectious characteristics of index case	Negative sputum smear	Positive sputum smear	0.97	0.67 - 1.41
Index case infectious period	≤30 days	30-59 days	0.93	0.57 - 1.54
		60-89 days	1.40	0.90 – 2.21
		90-119 days	1.58	0.92 – 2.68
		>120 days	1.83	1.17 – 2.86
Constant	0.05			0.01 - 0.17

Total average costs were €42.71 per screened individual in the TST/IGRA group and €55.21 in the IGRA-only group; the corresponding median values were €43.71 and €60.23, respectively. Medical direct costs were higher in the IGRA group, but non-medical direct costs and indirect costs were higher in the TST/IGRA group (Table 28).

Table 28. Screening costs by screening strategy (median [IQR], in EUR)

Screening costs	Assumptions 1 (cost/km €0.10; half-average income)			Assumptions 2 (cost/km €0.35; average income)		
	TST/IGRA	IGRA-only	p-value	TST/IGRA	IGRA-only	p-value
Medical direct	€12.83 [12.83 – 12.83]	€49.74 [49.74 – 50.95]	<0.001	€12.83 [12.83 – 12.83]	€49.74 [49.74 – 50.95]	<0.001
Non-medical direct	€3.23 [1.33 – 4.56]	€0.91 [0.36 – 1.51]	<0.001	€11.31 [4.64 – 15.98]	€1.80 [0.50 – 3.05]	<0.001
Indirect	€17.50 [0.00 – 28.49]	€8.75 [0.00 – 14.25]	<0.001	€34.99 [0.00 – 56.99]	€17.50 [0.00 – 28.50]	<0.001
Total	**€43.71 [16.14 – 52.37]**	**€60.23 [50.64 – 66.64]**	**0.006**	**€72.20 [24.44 – 86.34]**	**€71.12 [50.64 – 84.65]**	**0.2116**

The cost per LTBI diagnosis was €280.42 in the TST/IGRA group (76 per 499 screened individuals) and €205.44 in the IGRA group (147 per 547 screened individuals). The estimated number of potentially averted cases of TB was 5 in the TST/IGRA group and 8 in the IGRA-only group. Thus, the cost per potentially averted TB case was €4412.48 in the TST/IGRA group (€21,312.29/4.83) and €3719.20 in the IGRA group (€30,199.87/8.12).

Sensitivity Analysis

After changing the previous assumptions regarding travel (cost/km of €0.35 instead of €0.10, and average daily income instead of average half-daily income), total costs were €64.48 per screened individual in the TST/IGRA group and €65.63 per individual in the IGRA-only group (median €72.20 and €71.12, respectively) (Table 28).

The cost per LTBI diagnosis was €423.36 in the TST/IGRA group (76 in 499 individuals) and €244.22 in the IGRA group (147 in 547 individuals). The cost per potentially averted TB case was €6661.60 in the TST/IGRA group (€32,175.52/4.83) and €4421.13 in the IGRA group (€35,899.61/8.12). Medical direct costs were greater in the IGRA group, but non-medical direct costs and indirect costs were greater in the TST/IGRA group.

Incremental cost-effectiveness ratio (ICER)

The calculated ICER was €106 per LTBI diagnosis, representing increased effectiveness with a slightly increased cost of IGRA-only screening strategy.

Discussion

Our comparison of two groups of close contacts of TB cases who followed different LTBI screening strategies showed that, when compared to the TST/IGRA group, the IGRA-only group had increased odds of having LTBI diagnosed (aOR = 2.12, 95% CI = 1.53-2.94). Adherence to screening was also higher in the IGRA-only group, probably because this strategy requires fewer visits to the TB outpatient centres. From a societal perspective, the IGRA-only strategy appears to be more cost-effective than TST/IGRA strategy, because it has a lower cost per diagnosed LTBI case (€205.44 vs. €280.42) and a lower cost per potentially averted case of TB (€3,719.20 vs. €4,412.48).

The odds ratio for LTBI diagnosis was greater in the IGRA-only group than in the TST/IGRA group in Penafiel (high TB-incidence) than in Vila Nova de Gaia (medium TB-incidence). There is evidence that the TST and IGRA have similar sensitivity[239,240] but the increased specificity of two-step strategies comes with a lower sensitivity. Previous studies showed that increasing age and immunosuppression are associated with false negative results, especially with TST.[260] Other preconditions, like inflammatory diseases, might be associated with IGRA false negative results.[261] Nevertheless, we used data from healthy individuals, >5 years old, without HIV infection, diabetes or pharmacological

immunosuppression. We expect very few false positive results with the IGRA-only screening strategy, because of its high specificity, but no gold-standard test is available for confirmation.

Previous cost-effectiveness studies suggested that two-step screening was less effective averting active TB cases, but more cost-effective than IGRA-only screening.[238] The present study also considered the effect of societal costs, including not only medical costs but also non-medical direct and indirect costs. Our results suggests that the IGRA-only strategy is more cost-effective, mainly because of its higher effectiveness in diagnosing LTBI (and potentially averting TB cases) and decreased indirect costs (less productivity lost by individuals and society). As expected, the IGRA-only strategy represents increased costs for health services, because of the unit cost of the IGRA test itself and associated laboratory work. The incremental cost-effectiveness ratio of €103 per LTBI diagnosis represents the amount of money spent for the outcome of interest. We considered effectiveness only for LTBI diagnosis, and not for treating LTBI (this was studied elsewhere[247]).

This may have led to an overestimation of the number of potentially averted cases of TB in both groups. A selection bias in the IGRA-only group may have occurred, because individuals in this group were exposed to more infectious TB cases (80% of index cases were sputum-smear positive in the IGRA-only group, but only 68% were sputum-smear positive in the TST/IGRA group). This might have occurred because the criteria used to identify eligible contacts were stricter for the more expensive screening strategy. However, including infectiousness in our regression model should have eliminated this bias.

Conclusion

From a societal perspective, IGRA-only screening appears to be more cost-effective than TST/IGRA screening for LTBI, with a lower cost per LTBI diagnosis and a lower cost per potentially averted TB case. These results indicate an increased effectiveness of IGRA-only screening, at an only slightly increased cost.

3.5. DECLINE OF TUBERCULOSIS NOTIFICATION RATE IN DIFFERENT POPULATIONS AND REGIONS IN PORTUGAL, 2010-2017

ARTICLE IN PRESS
JID: PULMOE [mSP6P;September 30, 2021;4:15]
Pulmonology 000 (xxxx) 1–8

PULMONOLOGY
www.journalpulmonology.org

ORIGINAL ARTICLE

Decline of tuberculosis notification rate in different populations and regions in Portugal, 2010–2017

A. Sentís[a,b], A. Prats-Uribe[c], V.R. Peixoto[d,e], J.A. Caylà[f], M.D. Gomes[g,h], S. Sousa[h,i], R. Duarte[g,j,k], I. Carvalho[h,l], C. Carvalho[i,*]

[a] *Epiconcept, Epidemiology Department, Paris, France*
[b] *Pompeu Fabra University (UPF), Barcelona, Spain*
[c] *Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences, University of Oxford, United Kingdom*
[d] *NOVA National School of Public Health, Public Health Research Centre, Universidade Nova de Lisboa, Lisbon, Portugal*
[e] *Comprehensive Health Research Centre, Universidade Nova de Lisboa, Lisbon, Portugal*
[f] *Foundation of Tuberculosis Research Unit of Barcelona, Spain*
[g] *EPIUnit, Institute of Public Health, University of Porto, Porto, Portugal*
[h] *National Tuberculosis Programme, Directorate-General of Health, Lisbon, Portugal*
[i] *Multidisciplinary Unit for Biomedical Research (UMIB), Institute of Biomedical Sciences Abel Salazar, University of Porto, Porto, Portugal*
[j] *Public Health Science and Medical Education Department, Faculty of Medicine, University of Porto, Porto, Portugal*
[k] *Pulmonology Department, Hospital Centre of Vila Nova de Gaia/Espinho EPE, Vila Nova de Gaia, Portugal*
[l] *Pediatric Department, Hospital Centre of Vila Nova de Gaia/Espinho EPE, Vila Nova de Gaia, Portugal*

Received 14 June 2021; accepted 12 August 2021
Available online xxx

KEYWORDS
Tuberculosis;
Portugal;
Notification rate;
Time trend analysis

Abstract
Background: Tuberculosis (TB) incidence declined in Portugal in recent decades, but trends differ between regions and population subgroups. We investigated these differences to inform prevention and control programmes.
Methods: We extracted TB notifications from the Portuguese National TB Surveillance System (SVIG-TB) in 2010–2017, disaggregated by region, age group, nationality and HIV status. We calculated notification rates using denominators from the Portuguese National Institute of Statistics and the Joint United Nations Programme on HIV/AIDS and performed stratified time series analysis. We estimated interannual decline percentages and 95% confidence intervals (CI) using Poisson and binomial negative regression models.
Results: The overall TB notification rate decreased from 25.7 to 17.5/100,000 population from 2010 to 2017 (5.2%/year) in Portugal. Interannual decline did not differ significantly between regions, but it was smaller amongst non-Portuguese nationals (-1.57% [CI: -4.79%, 1.75%] vs -5.85% [CI: -6.98%, -4.70%] in Portuguese nationals); children under five years of age (+1.77% [CI:

-4.61%, 8.58%] vs -5.38% [CI: -6.33%, -4.42%] in other age groups); and HIV-negative people (-6.47% [CI: -9.10%, -3.77%] vs -11.29% [CI; -17.51%, -4.60%] in HIV-positive).
Conclusions: The decline in TB notification rates in Portugal during the study period has been steady. However, the decline amongst non-Portuguese nationals, children under five years of age and non-infected-HIV patients was lower. No significant differences were observed between regions. Changes in TB epidemiology in specific risk groups and geographical areas should be closely monitored to achieve the objectives of the End TB Strategy. We recommend intensifying screening of TB in the subpopulations identified.

* Corresponding author.
E-mail address: cfcarvalho@icbas.up.pt (C. Carvalho).

https://doi.org/10.1016/j.pulmoe.2021.08.002

Please cite this article in press as: A. Sentís, A. Prats-Uribe, V.R. Peixoto et al., Decline of tuberculosis notification rate in different populations and regions in Portugal, 2010–2017, Pulmonology (2021), https://doi.org/10.1016/j.pulmoe.2021.08.002

Background

Tuberculosis (TB) incidence is steadily declining in Europe, although it remains high worldwide and in some European countries. TB is still as a major public health problem, also due to drug-resistant forms and with HIV coinfection.[129,262] In the World Health Organization (WHO) European Region and in the European Union/European Economic Area (EU/EEA) the average annual decline of the TB incidence rate was 5.1% between 2015 and 2019 (11.9 to 9.6/100,000 inhabitants from 2015 to 2019 in EU/EEA).[262] In Portugal, which has one of the highest notification rates in the EU/EEA, the average annual decline was 5.0% (from 21.2 to 17.2/100,000).[262] Out of all cases notified in 2019 in the EU/EEA, 76.9% were new cases and 9.7% had been previously treated for TB (13.4% with unknown previous history of TB).[262]

The WHO End TB strategy states that to achieve a reduction of 80% in new TB cases and 90% in deaths by 2030, collaborative TB/HIV activities, and management of comorbidities are needed.[129]

As for other infectious diseases, we can describe TB as syndemic: it is an old endemic disease, presenting a synergic interaction with and worsening the independent damages of other diseases or conditions such as HIV, diabetes, or many social determinants of health (deprivation and inequities of health amongst others).[11,263] TB affects differently specific population groups and geographical areas. Amongst others, risk groups include specific clinical conditions (such as HIV or other immunosuppression), close contact with a TB case

or migration from an area with a high incidence of TB, as well as unfavoured socioeconomic and occupational groups.[210] In some risk groups, like HIV-infected, a decline in TB incidence rates was observed in the EU/EEA and in Portugal between 2015 and 2019. The relative weight of other groups, such as children under five years of age and immigrants, increased despite the overall decline in incidence.[262] In Portugal regional differences in risk of TB have been reported, with more densely populated regions such as Lisbon and Tagus Valley (LTV) and North regions being more affected.[149,153,222]

To analyse the differences in magnitude of decline and to better understand risk from surveillance data, further analysis of the available data is needed. We aimed at identifying population groups and geographical areas in Portugal with smaller decline in TB notification rates, to inform targeted prevention and control interventions.

Methods

Study design and participants

Time trend analysis study. We performed a time series analysis of active TB cases notified to the Portuguese TB clinical notification and follow-up surveillance system (SVIG-TB), including cases diagnosed between 2010 and 2017.

Case definition and TB surveillance system (SVIG-TB)

All confirmed, probable, and possible cases of TB diagnosed in 2010-2017 were retrieved from SVIG-TB. The case definition used is in line with EU case definitions.[32] Cases are notified to SVIG-TB by any physician diagnosing and following-up TB patients, mostly at TB outpatient centres (part of public primary care). The notifications are based on paper forms that are subsequently computerised and centralised at the regional and national levels, and include clinical, epidemiological and laboratory data, from diagnosis to discharge.[33]

Definition of variables and potentially associated factors

Sociodemographic and medical characteristics of TB patients were collected from SVIG-TB. We analysed the variables Sex (Female, Male), Age (<5 and 5 or more), Nationality (Portuguese, non-nationals), Site of infection (Extrapulmonary, Pulmonary), HIV status (Positive, Negative), and region of residence (North, Centre, Lisbon and Tagus Valley [LTV], Alentejo, Algarve, Azores, Madeira).

Denominators

We obtained denominators for each analysed year by Region, Age, Sex and Nationality from the Portuguese National Institute for Statistics[264] and estimates of people living with HIV in Portugal from Joint United Nations Programme on HIV/AIDS (UNAIDS) data 2019 report.[265]

Statistical analyses

Patients were described using counts and proportions for each of the categorical independent variables included. Age was described as Median (IQR) and proportion of <5 years old. Monthly and yearly notification rates of TB were calculated as number of cases notified per 100,000 inhabitants. This was done overall and stratified by sex, age group, nationality, and region of residence. Further indicators of yearly notification rates of HIV+ and HIV- TB, Pulmonary and Extrapulmonary TB were computed overall. Interannual declines with 95% confidence intervals (CI) for each group were estimated using Poisson regression models. If overdispersion was observed the interannual decline was estimated using negative binomial regression. Analyses were performed using STATA (version 14; Stata Corporation, College Station, TX, USA).

Ethical approval

In this study, patients were not directly involved; anonymized data were extracted from the national TB clinical notification and follow-up surveillance system (SVIG-TB) and analysed by authors at Directorate-General of Health, as part of their routine functions of surveillance and control of communicable diseases. The study was performed following the indications of the Helsinki Declaration (reviewed in Tokyo, October 2004).

Informed consent

Patients' data were anonymized prior to analysis, and for that reason informed consent was not required.

Results

From 2010 to 2017, the yearly notifications of TB in Portugal decreased from 2715 to 1800, corresponding to 25.7 and 17.5 notifications per 100,000 population, respectively (global decrease of 33.7% and a yearly decrease of 5.2%).

The regions with greatest reduction in incidence were Madeira (56.0%), Algarve (39.7%) and Alentejo (35.3%) (Figure 17, Table 29).

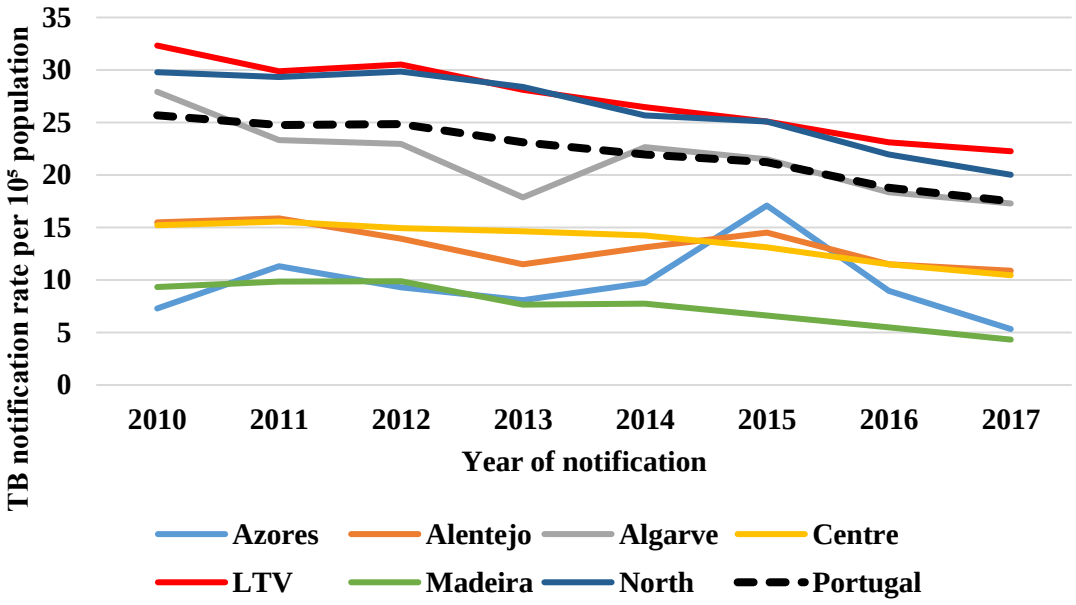


Figure 17. Tuberculosis notification rate in Portugal by region, 2010-17.

LTV: Lisbon and Tagus Valley region

Table 29. Total number of notifications and notification rate of tuberculosis in Portugal by population group and region, 2010-17.

	2010		2011		2012		2013		2014		2015		2016		2017	
Region	N	Rate	N	Rate	N	Rate	N	Rate	N	Rate	N	Rate	N	Rate	N	Rate
All regions	2,715	25.68	2,609	24.75	2,606	24.85	2,410	23.11	2,278	21.96	2,196	21.23	1,936	18.78	1,800	17.49
Azores	18	7.29	28	11.33	23	9.29	20	8.08	24	9.74	42	17.09	22	8.97	13	5.33
Alentejo	68	15.49	69	15.86	60	13.95	49	11.51	55	13.12	60	14.52	47	11.50	44	10.89
Algarve	126	27.92	104	23.31	102	22.95	79	17.86	100	22.65	95	21.50	81	18.35	76	17.29
Centre	358	15.21	364	15.56	347	14.95	337	14.64	325	14.23	298	13.10	261	11.50	235	10.45
LTV*	1,150	32.33	1,064	29.89	1,083	30.52	993	28.10	934	26.46	886	25.10	817	23.11	789	22.26
Madeira	25	9.33	26	9.84	26	9.88	20	7.65	20	7.73	17	6.63	14	5.49	11	4.32
North	970	29.79	954	29.34	965	29.85	912	28.37	820	25.66	798	25.10	694	21.94	632	20.02
Sex																
Female	897	16.25	847	15.37	911	16.59	862	15.76	829	15.21	725	13.33	669	12.33	640	11.80
Male	1,818	35.97	1,762	35.03	1,695	33.93	1,548	31.22	1,449	29.43	1,471	30.01	1,267	25.95	1,160	23.83
Age																
Under 5	18	3.60	23	4.69	32	6.65	22	4.75	17	3.78	16	3.67	21	4.90	25	5.87
>5 years	2,692	26.72	2,579	25.66	2,571	25.69	2,385	23.94	2,260	22.77	2,178	21.99	1,910	19.33	1,775	17.99
Nationality																
Portuguese	2346	23.16	2289	22.65	2228	22.12	2060	20.54	1980	19.83	1902	19.10	1617	16.31	1487	15.06
Non-nationals	362	81.71	317	72.92	354	85.38	290	72.82	221	56.65	257	66.97	310	78.89	310	74.40
HIV status																
Positive	335	1229.72	316	1121.76	301	1033.30	252	838.60	226	732.98	232	732.18	183	563.48	135	407.24
Negative	2019	19.15	1919	18.25	1976	18.89	1839	17.69	1733	16.75	1753	17.00	1514	14.73	1048	10.22
Site of infection																
Extrapulmonary	751	7.10	702	6.82	723	7.06	696	6.67	627	6.04	628	6.07	552	5.35	507	4.92
Pulmonary	1,946	18.41	1,897	17.99	1,874	17.87	1,700	16.30	1,627	15.68	1,558	15.07	1,369	13.28	1290	12.54

Madeira had the greatest yearly decrease in incidence, followed by Algarve (5.4%), North (5.4%) and Lisbon and Tagus Valley (LTV) (5.2%) (Table 30).

Table 30. Tuberculosis notification rate decline (interannual change) per 100,000 population and comparative average difference rate by region in Portugal, 2010-17.

Region	Notification rate 2010	Notification rate 2017	Interannual Change
Portugal	25.7	17.5	-5.3% (-6.2, -4.4)
Centre	15.2	10.5	-5.0% (-6.6, -3.4)
Lisbon and Tagus Valley	32.3	22.3	-5.2% (-6.1, -4.3)
North	29.8	20.0	-5.4% (-6.4, -4.4)
Alentejo	15.5	10.9	-4.4% (-8.2, -0.5)
Algarve	27.9	17.3	-5.4% (-8.3, -2.4)
Azores	7.3	5.3	-0.03% (-6.1, 6.4)
Madeira	9.3	4.3	-9.9% (-16.0, -3.5)

Amongst the 18,550 TB cases included in the analysis, 12,504 were confirmed (67.4%), 3122 probable (16.8%) and 2924 possible (15.8%). Three out of the seven Portuguese regions represented 91.6% of all the notifications (LTV 41.6%, North 36.4%, and Centre 13.6%). Proportion of male cases was 65.6% overall (highest in Algarve, 68.4%, and lowest in LTV, 63.0%). The median age was 47 years old (lowest in LTV, 44, highest in Centre region, 49); patients under 5 years of age represented 1% of all cases (highest proportions in Madeira and North region – 1.3% and 1.2%, respectively – and lowest in Centre region, 0.6%). The proportion of Portugal-born cases was 85.8% (lowest LTV, 73.3%, and highest North region and Madeira – 96.5% and 95.6%, respectively).

Regarding clinical characteristics, 10.7% of all notified cases were HIV-infected patients. HIV coinfection prevalence was highest in LTV and Algarve (16.2% and 11.8%, respectively) and lowest in Azores and Madeira (1.6% and 3.8%, respectively). Extrapulmonary TB accounted for 27.9% of all cases (proportion ranging from 9% in Azores to 30.7% in LTV) (Table 31).

Table 31. Sociodemographic and medical characteristics of tuberculosis patients notified by region in Portugal, 2010-2017 (N= 18,550).

	Portugal		**Region**													
			Centre		**LTV***		**North**		**Alentejo**		**Algarve**		**Azores**		**Madeira**	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Total	18,550		2,525	13.6	7,7	41.6	6,745	36.4	452	2.4	763	4.1	190	1.0	159	0.9
Sex																
Female	6,380	34.4	822	32.6	2,857	37	2,222	32.9	128	28.3	241	31.6	56	29.5	54	34
Male	12,170	65.6	1,703	67.4	4,859	63	4,523	67.1	324	71.7	522	68.4	134	70.5	105	66
Age																
Median (IQR)	47 (34, 61)		49 (36, 65)		44 (33, 59)		48 (36, 62)		53 (39, 71)		47 (35, 64)		46.5 (34, 56)		46 (36, 55)	
Under 5	174	1.0	15	0.6	67	0.8	81	1.2	4	0.9	5	0.7	0	0	2	1.3
>5 years	18,350	98.9	2,510	99.4	7,629	98.9	6,664	98.8	448	99.1	758	99.3	184	96.8	157	98.7
Missing	26	0.1	0	0	20	0.3	0	0	0	0	0	0	6	3.2	0	0
Nationality																
Portuguese	15,909	85.8	2,405	95.3	5,654	73.3	6,513	96.5	420	93.0	612	80.2	153	80.5	152	95.6
Non-nationals	2,421	13.1	117	4.6	1,887	24.4	229	3.4	30	6.6	151	19.8	0	0	7	4.4
Missing	220	1.1	3	0.1	175	2.3	3	0.1	2	0.4	0	0	37	19.5	0	0
HIV status																
Positive	1,980	10.7	151	6.0	1,248	16.2	460	6.8	22	4.9	90	11.8	3	1.6	6	3.8
Negative	13,801	74.4	2,040	80.8	5,209	67.5	5,353	79.4	373	82.5	635	83.2	69	36.3	122	76.7
Missing	2,769	14.9	334	13.2	1,259	16.3	932	13.8	57	12.6	38	5.0	118	62.1	31	19.5
Site of infection																
Extrapulmonary	5,186	27.9	668	26.5	2,371	30.7	1,796	26.6	101	22.3	185	24.2	17	9.0	48	30.2
Pulmonary	13,261	71.5	1,833	72.6	5,275	68.4	4,944	73.3	351	77.7	575	75.4	172	90.5	111	69.8
Missing	103	0.6	24	0.9	70	0.9	5	0.1	0	0	3	0.4	1	0.5	0	0

The interannual decline in notification rate was significantly smaller in non-Portuguese nationals (-1.57%, CI:-4.79%, +1.75%) than in Portuguese-born individuals (-5.85%, CI: -6.98%, -4.70%); in children under five years (+1.77%, CI: -4.61%, 8.58%) than in other ages (-5.38%, CI: -6.33%, -4.42%); and in HIV-negative people (-6.47%, CI: -9.10%, -3.77%) than in HIV-positive (-11.29%, CI: -17.51%, -4.60%) (Table 32).

Table 32. Tuberculosis notification rate decline (interannual change) per 100,000 inhabitants by population group in Portugal, 2010-17.

Variable	Notification rate (2010/17)	Interannual Change	95% Confidence Interval	
			lower	upper
Sex				
Female	16.3/11.8	-4.63%	-6.20%	-3.03%
Male	36.0/23.8	-5,44%	-6.18%	-4.70%
Age				
Under 5	3.6/5.9	1.77%	-4.61%	8.58%
>5 years	26.7/18.0	-5,38%	-6.33%	-4,42%
Nationality				
Portuguese	23.2/15.1	-5,85%	-6,98%	-4,70%
Non-nationals	81.7/74.4	-1,57%	-4,79%	1,75%
HIV status				
Positive	837.5/329.3	-11.29%	-17.51%	-4.6%
Negative	19.1/10.2	-6.47%	-9.10%	-3.77%
Site of infection				
Extrapulmonary	7.1/4.9	-4,80%	-5,97%	-3,62%
Pulmonar	18.4/12.5	-5,52%	-6,32%	-4,71%

Discussion

Despite the overall decreasing trend in TB notification rate in Portugal from 2010 to 2017, we found that decline in TB notification rates was smaller in non-Portuguese nationals, HIV-negative individuals, and children under five years of age (although numbers were very small for children). Most cases were notified in the three most populated Portuguese regions (Lisbon and Tagus Valley, North and Centre) but no significant differences were observed in the regional interannual declines.

Lisbon and Tagus Valley (LTV) had higher proportions of female and younger cases, non-Portuguese nationals, HIV-coinfected and extrapulmonary TB than the other six regions.

These regional differences correlate well with previous evidence that had described that the most frequent risk factors for TB, and for failure to complete latent TB (LTBI) treatment, are different in LTV (cases are more frequently migrant or living with HIV) and North region (where alcohol abuse and injecting drugs seem to be more frequently associated).[149,153,222]

In Portugal, from 2015 to 2019, the TB notification rate decreased by 18.9% (from 21.2 to 17.2 cases per 100,000 population) which is close to the 20% reduction set as interim milestone at the WHO End TB Strategy for 2015-2020.[129] According to this document, the reduction was expected to be of 50% for 2020-2025, 80% for 2025-2030 and 90% for 2030-2035. Sub-notification of TB, as it has been described in many countries,[266,267] could play a role in the observed reduction in notification rates in Portugal – although this reduction was similar to the 19% described for the WHO European region from 2015 to 2019.[268] In the U.S.A., a country with a very low incidence of TB (under 5/100,000 population) there was an average decrease of only 2.1% per year in 2012-2019.[74]

When comparing the regional declines in TB notification rates, we did not observe significant differences in Portugal. However, our results showed how those regions with higher notifications rates, mostly North and LTV, were also those with highest interannual declines. Similarly, regions with lower notifications rates (mostly Azores and Alentejo) were also the ones with lowest interannual declines. Interannual declines in TB notification rates allow a better understanding of the evolving situation than the overall decrease in the number of cases, as the latter indicator does not consider rate increases in intermediate years within the study period. This is illustrated by the regions of Madeira, Algarve, Alentejo, Azores and LTV. We believe that the interannual changes are more robust when drawing conclusions because are less affected by sporadic annual changes that can be due by a vast range of causes (including outbreaks or under-diagnosis/under-notification because of any external causes, such as the COVID-19 pandemic). Finally, the fact that some regions with lower incidence rates of TB at the start of the study period showed a smaller decline in notification rates could be partially explained by a smaller likelihood of receiving targeted interventions.

The results of this study showed that the declines in TB notification rates were smaller in children under five years old and non-Portuguese nationals, although numbers were very small for children. Factors such as improved reporting and better diagnosing tools and guidelines, might have played a role. Although with some degree of uncertainty due to the small number of cases, incidence of TB in HIV-positive patients seemed to have a greater decline than in HIV-negative, which may be due to more intensive latent TB screening and

preventive treatment, and to an increased antiretroviral coverage amongst HIV positive in the latest years. These results seem to be confirmed by the latest surveillance reports.

Despite the overall decline during the last years in the number of TB notifications, an increase of the total number or the relative weight of some specific groups has been described.[262] Number of cases in children under 15 in Portugal and EU/EEA changed from 43 (2.0%) and 2262 (3.7%) in 2015 to 63 (3.6%) and 1995 (3.9%) in 2019, respectively. In immigrants, the number of reported cases in Portugal and EU/EEA changed from 367 to 18,543 in 2015 (16.7% and 28.2% of all TB cases reported in that year, respectively) to 419 and 17,181 (23.7% and 34.5%) in 2018, respectively. In some other risk groups, such as HIV-coinfected, the number of cases in Portugal and EU/EEA declined from 232 (11.7%) and 1248 (4.6%) in 2015 to 133 (11.0%) and 502 (3.1%) in 2019, respectively. Regarding 2019 data of extrapulmonary TB, Portugal had higher proportions than the average in the EU/EEA (26.0% vs 22.5%).[262]

In our study we found steeper declines in notification rates in men than in women, although differences were not significant. Men are more frequently affected by TB, which has been attributed to the fact that they present other social, behavioural and occupational risk factors for TB more frequently than women.[73,269] However, recent evidence described a steeper decline in TB incidence rates in men and at older ages in the U.S., China and India.[270]

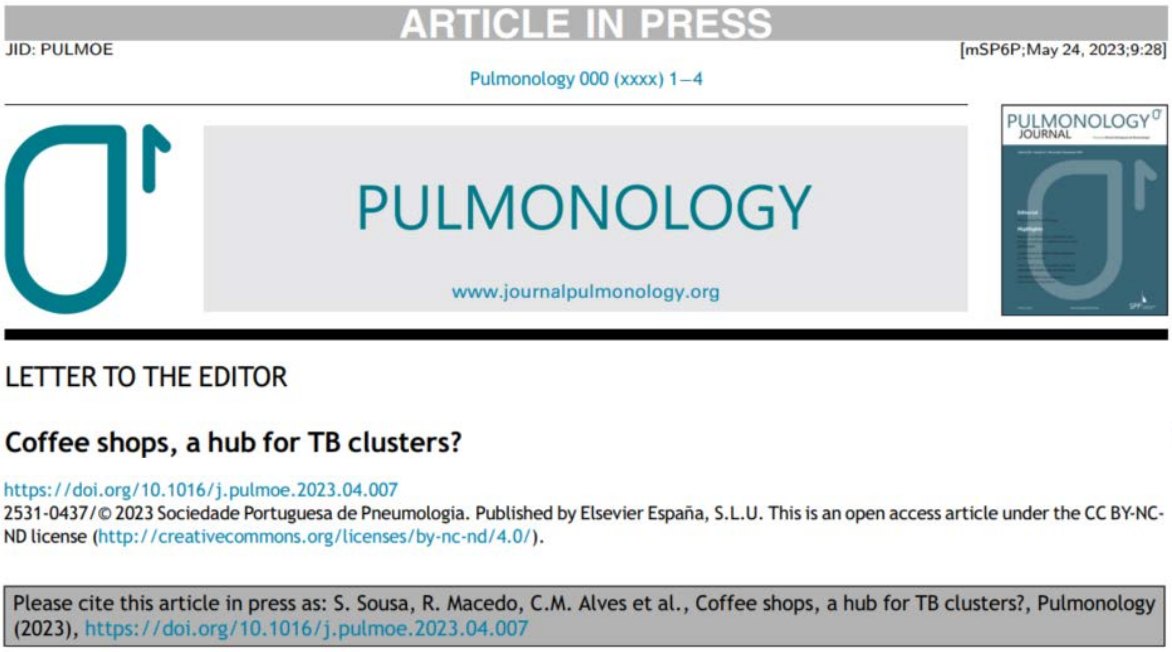
The main strength of this study is that it included high quality data at national level for some of the most important risk groups for TB. The main limitation is the lack of information about other specific risk factors such as drug-using, homelessness, or incarceration. There is a need to collect information about these conditions, amongst other determinants of health.[71,72] Active case finding in at-risk subgroups has been described as an effective way to strengthen TB prevention and control.[271] CDC and ECDC have started analysing changes and trends in different risk groups and regions,[74,262] but we still need to better monitor transmission and overlaps in these groups and obtain more precise data on the risk factors for TB as time changes.[210]

Other limitations include missing data, especially for HIV status (approximately 15% missing), which could have had biased our results. Our analysis includes only data until 2017, although at the time of submission more up-to-date data was available. In the context of the COVID-19 pandemic, it has been described that most severe COVID-19 cases also have an increased prevalence of latent tuberculosis infection (LTBI).[272] More research is needed to disentangle if COVID-19 could favour progression to TB or worsen outcomes and to learn more about the direct and indirect effects of the COVID-19 pandemic in TB

prevention and control, including those related to social determinants of health.[273,274] TB and COVID-19 have been associated with health inequities, disadvantaged socio-economic status, and poverty.[275] A recent study showed lower mortality rates in immigrants diagnosed with COVID-19 and TB, although differences were attributed to the fact that migrants tend to be younger and with fewer comorbidities.[276] Although it might be too soon to assess the impact that the COVID-19 pandemic may have had on the TB incidence, data from 33 centres in 18 countries have already showed a worldwide disruption of the TB services during the first four months of the pandemic.[277] In Barcelona, the socioeconomic and health consequences of the COVID-19 pandemic probably had a relevant impact on TB surveillance and control in the city, where a decline of almost 20% on the number of notifications and new diagnostics was observed.[278] A recent modelling study predicted long-lasting TB incidence increases in three countries with a high TB burden due to lockdown-related disruptions.[267] Either way, all resources used to screen, diagnose, and treat TB or COVID-19 must be synergic in order to tackle both conditions and optimize resources.[279] As it has been mentioned, the apparent success in Portugal achieving the WHO End TB Strategy milestone reduction from 2015 to 2020,[11] has to be monitored to end TB by 2035. Above all, the possible under-notification due to the COVID-19 pandemic in 2020 and the likely scenario of TB incidence increase "post-COVID-19" must be considered.[267]

We recommend systematic screening of TB in migrants originating from high incidence countries, as well as raising awareness of local clinicians for the most affected population groups to consider the diagnosis of TB and ensuring that new-borns with specific risk factors are timely offered BCG vaccination. We also recommend strengthening contact tracing above all in immigrants and in settings where children are more frequently involved. Community health workers should be more frequently integrated in the response to TB, both at operational and strategical levels, to facilitate and enhance contact tracing and LTBI-TB treatment in special circumstances where idiomatic or cultural barriers exist.[280] We recommend further regional-local analysis to tailor more specific interventions. We hope that the present study may also contribute to improve epidemiological surveillance, focusing on monitoring changes in risk groups and in specific geographical areas to plan prevention and control interventions more effectively, supporting the National TB Programme objective of further reducing incidence of tuberculosis in Portugal.[102]

3.6. COFFEE SHOPS, A HUB FOR TB CLUSTERS?



ARTICLE IN PRESS
JID: PULMOE [mSP6P;May 24, 2023;9:28]
Pulmonology 000 (xxxx) 1–4
PULMONOLOGY
www.journalpulmonology.org

LETTER TO THE EDITOR

Coffee shops, a hub for TB clusters?

https://doi.org/10.1016/j.pulmoe.2023.04.007

Please cite this article in press as: S. Sousa, R. Macedo, C.M. Alves et al., Coffee shops, a hub for TB clusters?, Pulmonology (2023), https://doi.org/10.1016/j.pulmoe.2023.04.007

Letter to the editor (Sousa S, Macedo R, Alves CM, **Carvalho C**, Goncalves G, Duarte R)

Stone quarry workers account for about 30% of the tuberculosis (TB) cases[281] in Penafiel and Marco de Canaveses, two municipalities in the northern region of Portugal that have the highest TB notification rates of the country.[23]

In high incidence areas, only a small number of reliable epidemiologically linked cases are identified using conventional contact investigations.[282] Whole-genome sequencing (WGS) is increasingly used to study transmission dynamics.

We conducted a retrospective study including all notified cases of TB in stone quarry workers from the municipalities of Penafiel and Marco de Canaveses from 1 January 2012 to 31 December 2019. First, we analysed classical epidemiological data from the stone quarry workers with TB diagnosed during 2012-2014. Secondly, all the available strains of M. tuberculosis isolated from 2015 to 2019 were sent for WGS in the National Reference Laboratory using a single nucleotide polymorphisms (SNP)-based approach.[197] Then we compared clustered and non-clustered cases using the Chisquared test or Fisher's exact test.

According to local public health services' records, 11 stone quarry workers with confirmed TB were notified between 2012 and 2014. Work-related exposure led to universal screening initiatives in workplaces: 135 co-workers of the same companies were screened. TB disease screening included chest-X ray and symptom questionnaire (96.3% adherence; one case found in 130 screened) and TB infection screening was performed through Tuberculin Skin Test (36.3% adherence; 30 cases found in 49 screened; two initiated preventive treatment). During 2016-2017, eight new cases were found among those previously screened individuals, i.e., 5.9% of the co-workers developed TB disease in a period of three years. The results of their screening were not screened (one), incompletely screened (four), not treated TB infection (two), previous negative screen (one). TB infection screening was not performed in more than 60% of the contacts, which may have contributed to the high proportion of screened workers that developed TB. However, we hypothesize whether this could reflect that transmission occurred in other settings besides workplaces (namely social or familiar).

A total of 76 current or former stone quarry workers diagnosed with TB in the 2015-2019 period were found, i.e., 18.8% of the notified cases of TB in 2015-2019.23 Three of those were excluded (not confirmed TB). Of the 73 included cases, 35 (47.9%) had available specimens. Genotyped and non-genotyped cases of TB had similar characteristics regarding the considered risk factors (Table 33).

Table 33. Characteristics of clustered and non-clustered TB cases

Variable	Not genotyped (n = 38)	Genotyped cases (n = 35)	*p* value	Not clustered (n = 7)	Clustered cases (n = 28)	*p* value
Age group years			0.706			0.517
21-44	8 (21.1%)	9 (25.7%)		1 (14.3%)	8 (28.6%)	
45-64	26 (68.4%)	24 (68.6%)		6 (85.7%)	18 (64.3%)	
≥65	4 (10.5%)	2 (5.7%)		0	2 (7.1%)	
Employment status			0.284			0.166
Active	25 (65.8%)	27 (77.1%)		7 (100%)	20 (71.4%)	
Inactive	13 (34.2%)	8 (22.9%)		0	8 (28.6%)	
Silicosis						1.000
Yes		18 (51.4%)		4 (57.1%)	15 (53.6%)	
No		16 (45.7%)		3 (42.9%)	13 (46.4%)	
Alcohol dependence						0.220
Yes		13 (37.1%)		1 (14.3%)	12 (42.9%)	
No		22 (62.9%)		6 (85.7%)	16 (57.1%)	
Tobacco use						0.652
Yes		24 (68.6%)		4 (57.1%)	20 (71.4%)	
No		11 (31.4%)		3 (42.9%)	8 (28.6%)	
Previous treatment of TB						0.084
Yes		10 (28.6%)		0	10 (35.7%)	
No		25 (71.4%)		7 (100.0%)	18 (64.3%)	
Sputum smear						0.155
Yes		25 (71.4%)		3 (42.9%)	22 (78.6%)	
No		10 (28.6%)		4 (57.1%)	6 (21.4%)	
Number of days between symptoms and diagnosis						0.935
3-30		6 (17.1%)		1 (14.2%)	5 (17.9%)	
61-65		13 (37.1%)		3 (42.9%)	10 (35.7%)	
≥66		16 (45.7%)		3 (42.9%)	13 (46.4%)	
Attended public places			0.683			0.401
Yes	21 (55.3%)	21 (60.0%)		3 (42.9%)	18 (64.3%)	
No	17 (44.7%)	14 (40.0%)		4 (57.1%)	10 (35.7%)	
Close contact with relatives with TB			0.468			1.000
Yes	3 (7.9%)	5 (14.3%)		1 (14.3%)	4 (14.3%)	
No	35 (92.1%)	30 (85.7%)		6 (85.7%)	24 (85.7%)	

All the cases were male, born in Portugal, with an average age of 50-years-old (median 51, interquartile range (IQR) 43-56), and had pulmonary TB. The main risk factors included tobacco use, silicosis and alcohol dependence (Table 33). No cases of HIV were found (six had not registered HIV status). Two of the genotyped TB strains were polyresistant to isoniazid and streptomycin, and 33 were susceptible to all first line drugs. The delay between onset of symptoms and diagnosis was on average of 78 days (median 63 days; IQR: 34.0-88.0).

A high molecular diversity of M. tuberculosis was found (Figure 18). Clusters included cases from 2015 to 2019, suggesting ongoing active transmission. As we did not genotype all the strains of M. tuberculosis from the community, we could be missing the remaining strains from other clusters.

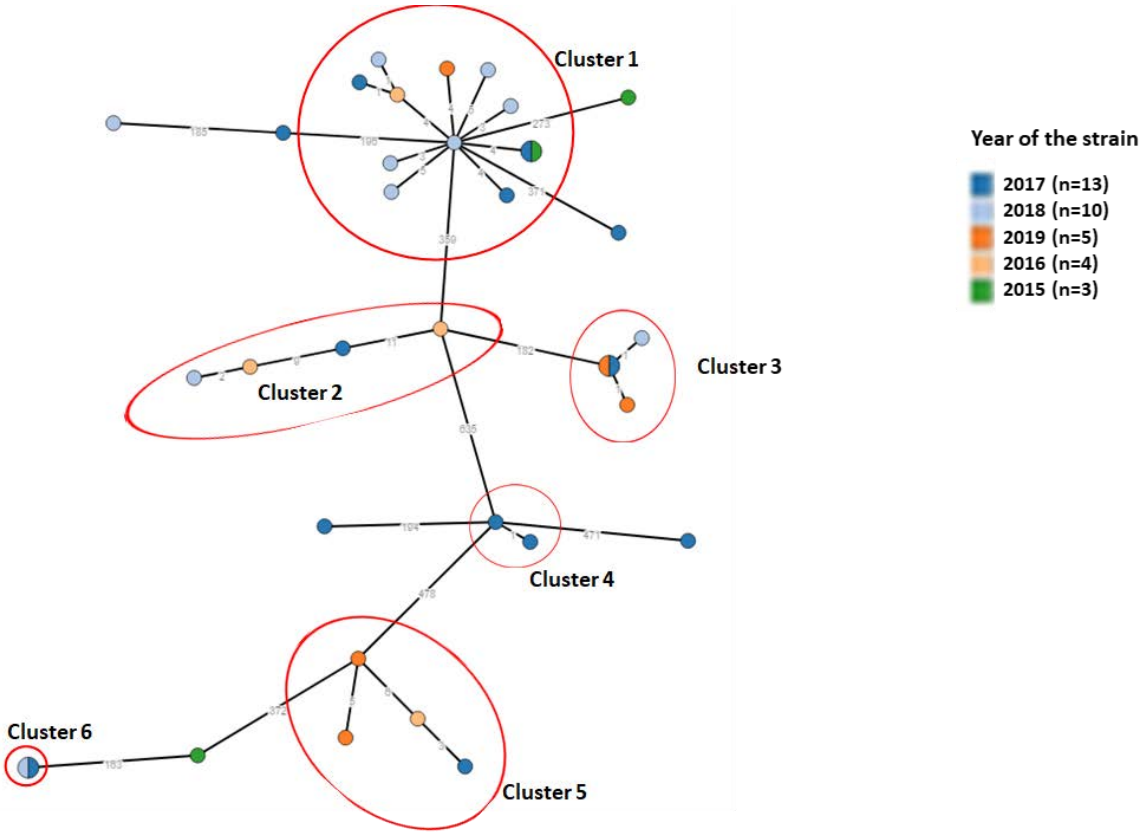


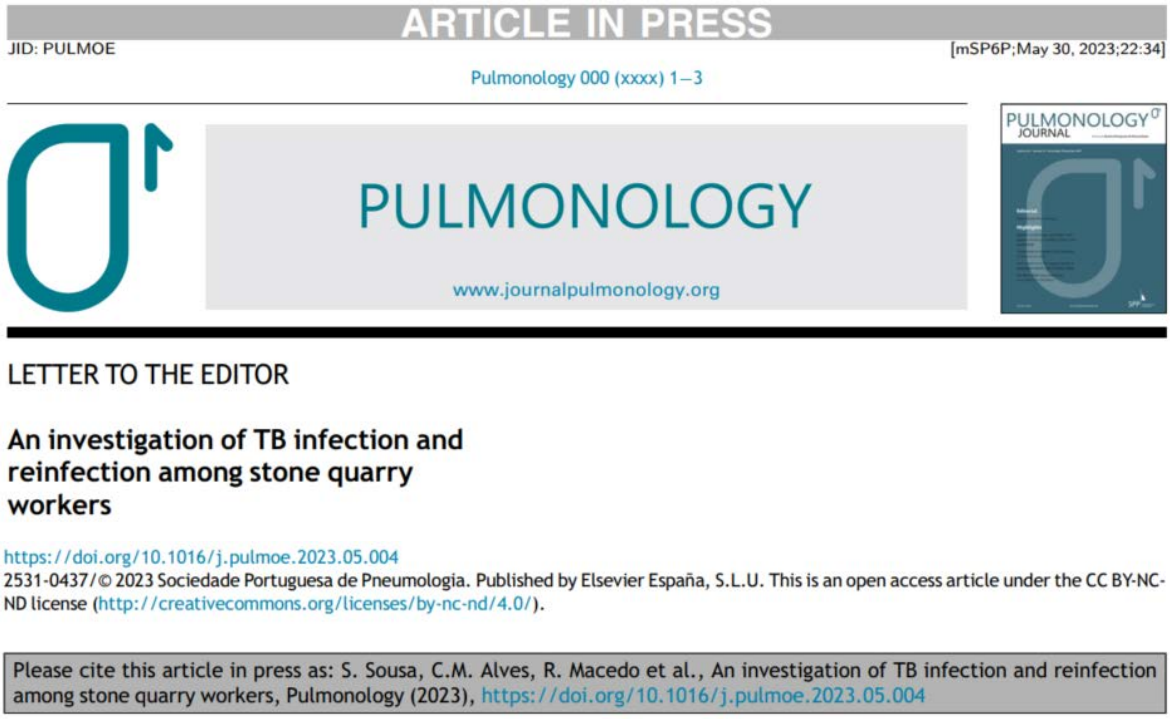
Figure 18. Network of 35 Mycobacterium tuberculosis (MTB) isolates. Phylogeny of MTB isolates was determined based on a core single nucleotide polymorphisms (SNP) approach. Each node corresponds to a single or multiple isolates (i.e., nodes with slices). Nodes are coloured by year of strain isolation and genetic clusters of closely related isolates were defined whenever the strains had less than six SNP difference and are highlighted in red circles. Clades representing determined lineages from the global M. tuberculosis tree are indicated [cluster_1: lineage4.3.4.2 (100.0%, n = 12); cluster_2: lineage4.3.2 (100.0%, n = 4); cluster_3: lineage4.3.2 (100.0%, n = 3); cluster_4: lineage4.1.2.1 (100.0%, n = 2); cluster_5: lineage4.1.1.1 (100.0%, n = 4); cluster_6: lineage4.1.1.3 (100.0%, n = 2)].

Median (IQR) time between successive cases in clustered cases was 205 (95.5-308.3) days. Clustered cases were more prone to have alcohol dependence and smoking habits, which could be associated with attending coffee shops regularly (Table 33). Likewise, being professionally inactive, having a previous episode of TB and a positive sputum smear were also more common among clustered cases. As positive sputum smear indicates higher infectiousness, that is expected. In a study that analysed MIRU-VNTR molecular clustering data from 7458 patients, cases in large molecular clusters were also more likely to have multiple social risk factos.[283]

Using a logistic regression analysis, none of the transmission settings was a significant predictor of clusters but attending public places was the better predictor (OR 1.8, 95% CI 0.254-12.449). Social contacts in community public places such as coffee shops seem to contribute to the maintenance of ongoing active transmission of TB. In other study, workers that converted from IGRA negative to positive had no co-workers with active TB and were not identified as close contacts, suggesting they could have been infected in social settings.[284]

In our opinion, the high molecular diversity found in a small sample of stone quarry workers cases suggests a complex scenario of transmission between them and the high-risk communities in which they live and work. Stone quarry workers are not only more prone to transmit TB to other people of the community; they are also more susceptible to infection and re-infection by M. tuberculosis, given their common and multiple risk factors, especially in places where epidemiological links are difficult to establish. WGS could routinely contribute to identifying public places that are hotspots of TB transmission.

3.7. AN INVESTIGATION OF TB INFECTION AND REINFECTION AMONG STONE QUARRY WORKERS



ARTICLE IN PRESS
JID: PULMOE [mSP6P;May 30, 2023;22:34]
Pulmonology 000 (xxxx) 1–3
PULMONOLOGY
www.journalpulmonology.org

LETTER TO THE EDITOR

An investigation of TB infection and reinfection among stone quarry workers

https://doi.org/10.1016/j.pulmoe.2023.05.004

Please cite this article in press as: S. Sousa, C.M. Alves, R. Macedo et al., An investigation of TB infection and reinfection among stone quarry workers, Pulmonology (2023), https://doi.org/10.1016/j.pulmoe.2023.05.004

Letter to the editor (Sousa S, Alves CM, Macedo R, **Carvalho C**, Goncalves G, Duarte R)

Stone quarry workers are considered at increased risk from tuberculosis (TB) mainly because of their prolonged exposure to silica.[285] They are considered a vulnerable population group for TB in Portugal, especially in the municipalities of *Penafiel* and *Marco de Canaveses*.[23,281]

Understanding transmission patterns is essential for developing prevention strategies and prioritising resources.[286] Whole-genome sequencing (WGS) is an important tool to identify clusters.[287,288]

Data from all cases of TB in stone quarry workers from 2015 to 2019, for whom M. tuberculosis (MTB)-positive specimens were available for complementary laboratory testing (n = 35, 47.9% of the eligible sample) was analysed. The data was collected by the local public health services during their routine surveillance and public health authority duties. TB clusters were identified using WGS (single nucleotide polymorphism [SNP]-based approach[197]), and the largest cluster identified was investigated to distinguish potential exposure settings and propose strategies to reduce transmission.

Six clusters were found. The largest cluster included 12 MTB isolates from 2015 to 2019. Five of the 12 cases (41.7%) had already had at least one previous episode of TB, six (50.0%) had silicosis and alcohol dependence, and eight (66.7%) were smokers (Figure 19).

Case	Silicosis	Tobacco	Alcohol	Previous TB
C1.1	Yes	Yes	No	Yes
C1.2	No	No	Yes	No
C1.3	Yes	Yes	Yes	Yes
C1.4	No	Yes	Yes	Yes
C1.5	Yes	Yes	No	Yes
C1.6	Yes	No	No	Yes
C1.7	No	No	No	No
C1.8	No	Yes	No	No
C1.9	No	Yes	Yes	No
C1.10	Yes	No	No	No
C1.11	No	Yes	Yes	No
C1.12	Yes	Yes	Yes	No

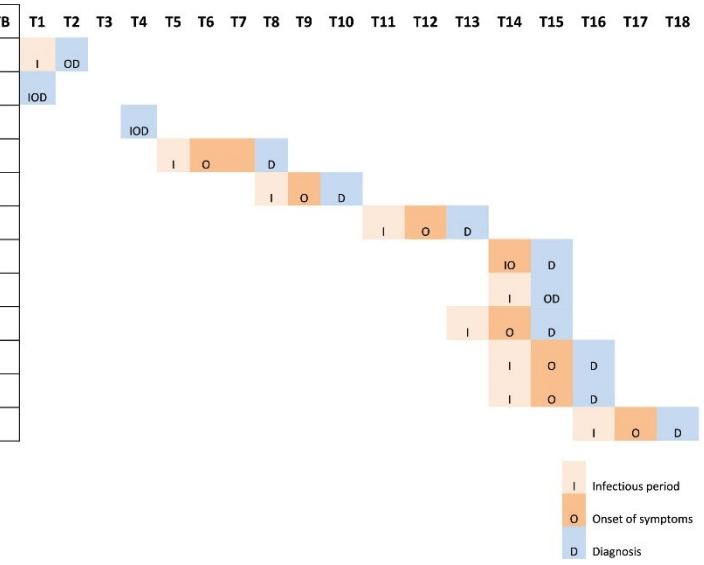


Figure 19. Detailed analysis of the 12 cases of Cluster 1 by trimester (T) regarding risk factors and time of diagnosis. Infectious period was considered for three months before symptoms onset if positive sputum smear or for one month before symptoms onset if negative sputum smear.

Out of the 12 cases, who lived in seven different parishes of the *Penafiel* and *Marco de Canaveses* municipalities, nine (75.0%) were employed in seven companies, and nine (75.0%) reported frequently attending seven different coffee shops (Figure 20).

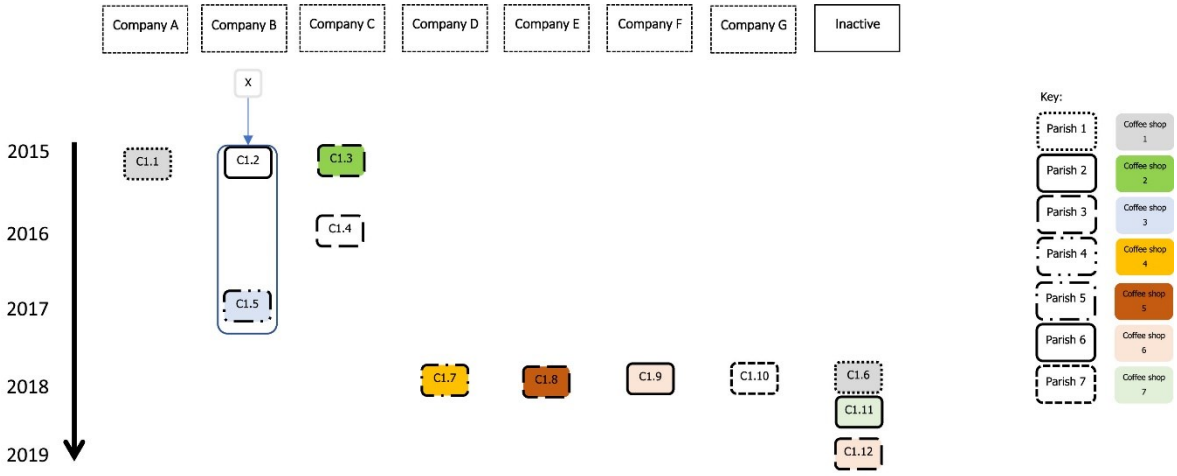


Figure 20. Detailed analysis of the 12 cases of Cluster 1.

Cases C1.2 and C1.5 were identified as close contacts of a co-worker from the same company that had had active TB in 2014 (a non-genotyped case that was considered the primary case for this cluster - X). The three worked in the central region of Portugal during the week and returned to three different parishes for the weekend. C1.2 shared the same room, and C1.5 shared transport and meals with the primary case. C1.2 had active TB in 2015, diagnosed during screening; C1.5 only had active TB in 2017 (a reinfection, previous treatment in 1999). C1.5 was identified for screening but the follow-up cannot be traced, and the screening result is not known. Failure to attend or incomplete screening was previously reported in this population,[281] with active TB cases identified late among those considered close contacts.

Case C1.5, who also had silicosis and was a smoker, had a cough for three months before the diagnosis (Figure 19); he was the suspected primary case of a community outbreak that included another six cases of TB in the first six months of 2017 in the parish where he lives. None of the other six cases were genotyped.

The primary case of this genotyped cluster (X) was considered cured in 2015 but had a second episode of active TB in 2017 at 33 years old. The new strain was genotyped and found to be associated with a different cluster suggesting that X was re-infected two years after his cure.

Case C1.9 was working abroad at the time of diagnosis and reported attending a bar also attended by C1.12 and two previous non-genotyped cases (Figure 20).

These 12 stone quarry workers had identical strains of Mycobacterium tuberculosis (with less than six differences in SNP), which suggests that they belong to the same chain of transmission. Nevertheless, it is probable that missing intermediary cases exist both in the general community and in other non-genotyped cases among stone quarry workers. The importance of household contact was not assessed in this study, but the epidemiological enquiry did not find any relatives who previously had TB.

Based on epidemiological investigation data, only close contacts in the workplace were identified. Social connections such as those occurring in coffee shops and not identified by the case, are difficult to find. However, these social contacts seem to be important in maintaining the ongoing active transmission in this high-risk population.

Active transmission of TB among stone quarry workers in this cluster was driven by multiple factors, including those related to occupation, but also those related to social habits. We highlight that this outbreak probably spread to different regions of the country and possibly

to other countries. Stone quarry workers' vulnerability, mainly due to silicosis, probably makes them more prone to TB infection and reinfection.

To better understand TB transmission dynamics in this high-risk population and the community, it would be beneficial if all the cases occurring in the country or region were genotyped. That would allow us to understand the transmission chains and different exposure settings better.

It is paramount that the local public health services explore all possible exposure settings during the epidemiological investigation whenever a new case is notified. Strategies to protect the most vulnerable should be enhanced: not only strategies addressing individuals such as promoting screening and health literacy to recognize symptoms and decrease diagnosis delay, but also environmental level strategies such as improving the ventilation conditions of the sites at which exposure occurs.

4

TIMELINESS OF THE TUBERCULOSIS SURVEILLANCE SYSTEM

The third surveillance system attribute under evaluation was timeliness, which can be defined as the interval between the occurrence of an adverse health event and one of multiple events that go from recording of diagnosis in an information system, notification to public health services, identification of trends or outbreaks, or implementation of control measures.[147]

The research protocol presented in this chapter, which received clearance from the Ethical Board of the School of Medicine and Biomedical Sciences of the Porto University (*Instituto de Ciências Biomédicas de Abel Salazar – ICBAS-UP*) and is pending assessment from the Directorate-General of Health to be implemented, will allow the study of the timeliness of the tuberculosis surveillance system, from the onset of symptoms to the date at which the notification was computerised and made available to the National Tuberculosis Programme. Previous studies have investigated the delays to start treatment, in both the component attributed to the patients and to healthcare services,[86,88,94,289–291] but this study will also provide an overview of the subsequent delay to notification.

4.1. RESEARCH PROTOCOL FOR A STUDY ON THE TIMELINESS OF TUBERCULOSIS NOTIFICATION IN PORTUGAL

Carvalho C, Gonçalves G, Sousa S, Magalhães Alves C, Correia AM, Duarte R. Research protocol for a study on the timeliness of tuberculosis notification in Portugal.

Public health surveillance plays a critical role in informed decision making through the ongoing and systematic collection, analysis, interpretation, and dissemination of health data. Therefore, the timeliness of such systems, in particular those related to infectious diseases, is key. In this study we aim at evaluating the timeliness of the National Tuberculosis Programme surveillance system (SVIG-TB) in Portugal, measured as the delays from onset of disease to notification to SVIG-TB, identifying associated factors.

Using data from SVIG-TB in 2010-2020, we will calculate the median delay (in days) between the date of onset of disease and the date at which the notification was computerised and made available to the National Tuberculosis Programme. The delay will be split in components attributable to the patient, to the diagnostic and to the notification, and stratified by year of diagnosis and other geodemographic variables. We will perform multiple regression to identify factors associated with increased delay in notification.

We expect to see a decrease in the delay of notification throughout the years from 2010 to 2020, with some degree of heterogeneity according to specific characteristics such as age and region of residence.

Delaying TB diagnosis and treatment has a great impact in public health, as the probability of having new infections increases. However, delays in notifying clinical data after conclusion of treatment also has an impact, as it hampers the ability to detect changes in epidemiology that might require prompt intervention or changes in policy. With this study we hope to formulate recommendations for targeted intervention to improve timeliness of reporting to SVIG-TB.

Background

Incidence of Tuberculosis has been declining across Europe, but it remains high and a significant public health problem in many countries.[292]

Despite the steady decline in the last decades, the TB notification rates in Portugal are still well above the average in the European Union/European Economic Area (EU/EEA). Portugal is the only country in Western Europe with a notification rate above 10 cases per 100 000 population.[292]

In order to ensure that all cases of TB are diagnosed and offered treatment, as well as outbreaks and changes in the epidemiology of the disease are quickly recognised and action is triggered to reduce the risk in the community, countries must continuously monitor and evaluate their TB surveillance systems.[11]

One of the most important components of a surveillance system is the timeliness, i.e., the interval between the occurrence of an adverse health event and the report of the event to the appropriate public health agency, the identification by that agency of trends or outbreaks, or the implementation of control measures.[147]

For most diseases, especially for those of infectious nature, other aspects of timeliness must be considered, such as the incubation/latency period (the time needed for a case to develop the disease after being exposed) but also (and even more importantly) the delay between the onset of symptoms and the decision of the patient to seeking care and the diagnosis.[293]

There is a considerable amount of literature on delay in diagnosis and treatment of TB,[86,87,289,291] but not so much about the delay to reporting to public health services and to the tuberculosis programmes. A recent meta-analysis including 198 studies from 78 countries showed a pooled mean total delay from onset of disease to starting of treatment of 88 days, to which the largest contributor was patient delay (pooled mean 81 days) followed by diagnostic and treatment delays (pooled means 30 and 8 days, respectively).[294]

Timeliness of notification (delay from diagnosis to reporting of cases) has been assessed in different countries. Still, results are difficult to compare given the differences in surveillance systems and data flows. Interestingly, a study in Ireland showed that the timeliness of notification to the National Tuberculosis Surveillance System decreased after introduction of a computerised electronic reporting system.[65]

Surveillance of tuberculosis in Portugal is supported on two national notification systems: 1) a compulsory-notification one, designed to promptly inform public health authorities and trigger public health intervention (Sistema Nacional de Vigilância Epidemiológica - SINAVE);[30] and 2) a clinical notification system, intrinsic to the National Tuberculosis Programme, designed to keep record of diagnosis and follow-up of every patient (Sistema de Vigilância da Tuberculose - SVIG-TB).[33] SVIG-TB is oriented for and operated by clinicians taking care of patients with active or latent tuberculosis and non-tuberculous mycobacterial infections in specialized TB-centres (Centros de Diagnóstico Pneumológico – CDPs), and includes clinical, epidemiological and laboratory data, from diagnosis to discharge of the patient. Unlike SINAVE, which is based on an exclusively-electronic data flow, cases are reported to SVIG-TB using specific paper forms that are only at a later stage computerised and made available for analysis.[33]

In this study we aim at describing the timeliness of national TB surveillance system in Portugal, assessing the different components of delay from the onset of disease to the moment when the case notification is digitised and made available for analysis by the National Tuberculosis Programme.

Proposed methods

Study design and participants

Descriptive study, including all active TB cases notified to the national TB surveillance system in Portugal (SVIG-TB), with date of diagnosis between 2010 and 2020.

Case definition and TB surveillance system (SVIG-TB)

All cases of active TB diagnosed in 2010-2020 will be retrieved from SVIG-TB, which includes confirmed, probable, and possible cases (following the EU case definitions and classifications for TB).[32]

Definition of variables and potentially associated factors

The variables used to compute the delays (dependent variables, in days) will be onset of symptoms, date of the first medical consultation, date of diagnosis/start of treatment, date of the end of treatment, and date of the last update submitted to SVIG-TB. Delays will be classified as attributed to the patient (date of the first consultation – date of onset of

symptoms), attributed to the healthcare services (date of diagnosis/start of treatment – date of first consultation), total delay to treatment (date of diagnosis/start of treatment – date of onset of symptoms), and delays to notification (date of the last update submitted to SVIG-TB – date of the end of treatment and date of the last update submitted to SVIG-TB – date of the end of treatment).

We will include as independent variables Year of diagnosis, Sex, Age, Region of residence, Site of infection (Pulmonary/extrapulmonary) and HIV status, investigating how these impact total and specific components of delay.

Statistical analyses

Delays will be described as Median and Inter-Quartile Ranges (IQR). This will be done overall and stratified by each of the independent variables considered. For notification delay, a multiple regression model will be implemented to adjust for co-variables. Analyses will be performed using STATA (version 15; Stata Corporation, College Station, TX, USA).

Ethical approval

In this study, patients will be not directly involved; anonymized data will be extracted from the national TB clinical notification and follow-up surveillance system (SVIG-TB). The study will be performed following the indications of the Helsinki Declaration (reviewed in Tokyo, October 2004).

All data collected will be used exclusively for the purpose of this investigation. The database used for this investigation will remain stored in an encrypted drive until conclusion of the study and destroyed afterwards.

Informed consent

Patients' data are anonymized prior to analysis and, for that reason, informed consent will not be required.

Variable list

Table 34. Variable list for the study of timeliness of tuberculosis notification in Portugal.

Variable	Type/description
Year of diagnosis	Numerical: 2010-2020
Age group	Categorical/ordinal: [0-5[[5-15[[15-25[[25-45[[45-65[65+ Unknown
Sex (assigned at birth)	Categorical: Male Female Other Unknown
Region of treatment	Categorical: Alentejo Algarve Azores Centre Lisbon Madeira North Unknown
Method of TB detection	Categorical: Post-mortem Symptomatic Contact screening Routine screening Other Unknown
Site of disease (pulmonary/extrapulmonary)	Categorical: Pulmonary TB Extrapulmonary Unknown
HIV status	Categorical: HIV positive HIV negative HIV unknown
Date of onset of symptoms	Date (dd/mm/yyyy)
Date of first medical consultation	Date (dd/mm/yyyy)
Date of start of treatment	Date (dd/mm/yyyy)
Date of end of treatment	Date (dd/mm/yyyy)
Date of first record uploaded to SVIG-TB	Date (dd/mm/yyyy)
Date of last record uploaded to SVIG-TB	Date (dd/mm/yyyy)

Expected results and outputs

We expect to see a decrease in the delay of notification throughout the years from 2010 to 2020, with some degree of heterogeneity according to specific characteristics such as age and region of residence.

The results of this study will be compiled, and a report will be provided to the National Tuberculosis Programme, including recommendations for targeted intervention to improve timeliness of reporting to SVIG-TB. A manuscript will be produced and submitted for publication in an international peer-reviewed journal.

Discussion

Timeliness of tuberculosis surveillance systems is key to enable quick intervention and minimise the risk of spread of disease. Several studies have shown the factors associated and possible impact of delay to diagnosis in Portugal, both attributable to patients and to healthcare services. Delaying TB diagnosis and treatment has a great impact in public health, as the probability of having new infections increases. However, delays in notifying clinical data after conclusion of treatment also has an impact, as it hampers the ability to detect changes in epidemiology that might require prompt intervention or changes in policy.

In this study we will focus on the delay to notification of cases diagnosed from 2010 to 2020 in Portugal, i.e. the time it takes to have clinical data recorded in the National TB surveillance system (SVIG-TB) and made available for analysis. We will also formulate recommendations for targeted intervention to improve timeliness of reporting to SVIG-TB.

Chronogram

Table 35. Chronogram for the study of timeliness of tuberculosis notification in Portugal.

Task/ Date (2022)	T0	Month 1	Month 2	Month 3
Investigation protocol approved and SVIG-TB database provided by DGS	■			
Data analysis		■		
Internal investigation report shared and discussed with NTP			■	
Article drafted and submitted to peer-reviewed journal				■

Budget and financing

No costs in addition to the investigator's time and resources are expected, no need to obtain specific funding.

5

SYNTHESIS

Tuberculosis (TB) is a serious infectious disease that affects millions of people worldwide. It is still an important public health problem, being the world's leading cause of death from a single infectious agent, the leader killer of people with HIV, and a major cause of deaths related to antimicrobial resistance.[295] In order to effectively control and prevent the spread of TB, it is essential to have robust surveillance systems in place, that are sensitive and specific enough to detect and validate new cases of disease.[296,297] This is especially true in low-incidence regions approaching elimination status, where it is crucial to monitor the incidence and prevalence of the disease, identifying trends and changes in the epidemiology, detecting outbreaks and informing public health interventions.[298]

In this chapter I summarise our findings on the evaluation of the TB surveillance system in Portugal, provide some recommendations for improvement, and close this thesis.

5.1. SENSITIVITY

In the main study of this thesis, we showed that reported notification rates of TB are probably an underestimation of the true incidence rates.[1] The availability of multiple "independent" national data sources for tuberculosis in Portugal allowed us to implement an inventory and capture-recapture methodology, using three data sources, and compare the number of TB cases meeting standard case definitions that were reported to national authorities through different notification systems, identifying duplicate and unique records through deterministic and probabilistic record-linkage. One of the main findings was that the completeness of reporting to the National Tuberculosis Programme (SVIG-TB) in 2015 could have been as low as 77%, which is in line with what was observed in similar studies in other European countries with declining TB incidence.[2,111,112] Although there are some methodological aspects that should be taken into account when interpreting the capture-recapture results (the assumptions of independence between data sources, accurate identification of cases, equal catchability, and population stability, might not hold),[104,109,114–116,118–121,123,124,137,145,299–303] the inventory methodology is very useful and produces reliable results whenever there are independent case-based data sources that can be linked.

Under notification of tuberculosis is observed in many other countries, as illustrated in another capture-recapture study covering six European countries, that I co-authored.[2] In this study we concluded that surveillance systems should be strengthened by structurally

employing electronic record linkage of high-quality TB registers, improving not only sensitivity but also data quality and completeness while avoiding redundancy.

In order to improve sensitivity of the surveillance system, we recommend modernisation through digitalisation of the notification process, working with the data providers to improve acceptability, and exploring alternative data sources that might complement or allow validation of the ones currently in use (Recommendations 1, 3, and 5, listed on Section 5.4.).

5.2. USEFULNESS

This attribute is intimately related to the objectives of the surveillance system. In addition to the periodic regional[251,304] and national[23,24] surveillance reports, SVIG-TB allows operational research and production of evidence that supports production and continuous update of clinical guidelines and public health intervention strategies, from the local to the national level. In this thesis I present seven articles that show the importance of having an integrated surveillance system that allows the combination of clinical, microbiological, and public health data sources.

In articles *3.1. A complex scenario of tuberculosis transmission is revealed through genetic and epidemiological surveys in Porto*,[3] and *3.2. A nationwide study of multidrug-resistant tuberculosis in Portugal 2014–2017 using epidemiological and molecular clustering analyses*,[4] we showed the importance of coupling molecular with epidemiological data to better understand the transmission dynamics and better inform prevention.

In articles *3.3. Failure to complete treatment for latent tuberculosis infection in Portugal, 2013–2017: geographic-, sociodemographic-, and medical-associated factors*[5] and *3.4. comparing the cost-effectiveness of two screening strategies for latent tuberculosis infection in Portugal*,[6] the focus was on screening and treatment of latent tuberculosis infection (LTBI), which are important strategies to reduce incidence in the medium/long term. The surveillance system allowed the quantification of the LTBI treatment failure, which was estimated to be 13.8%, and the identification of specific risk factors for failure to complete treatment: notably being born abroad (stronger risk factor in the Lisbon metropolitan area), alcoholism (stronger risk factor in Porto metropolitan area), or receiving longer duration treatment courses (vs. shorter 3-month course treatments).

Data from the tuberculosis surveillance system was also the basis of article *4.5. Decline of tuberculosis notification rate in different populations and regions in Portugal, 2010–2017*,[7] allowed the description of trends in different regions and population subgroups. The findings of this study were that the decline in TB incidence rates was lower amongst non-Portuguese nationals, children under five years of age and non-infected-HIV patients. Articles *4.6. Coffee shops, a hub for TB clusters?*[8] and *4.7. An investigation of TB infection and reinfection among stone quarry workers*[9] focused predominantly on investigations from local public health services in a specific setting in Northern region of Portugal (stone quarry workers), using SVIG-TB as the basis for the definition of high-risk settings, intervention priorities and identification of major TB risk factors such as silicosis and other occupational hazards.

Usefulness of the Portuguese TB surveillance should be periodically assessed, measuring not only the degree at which the surveillance objectives are achieved but also the perception of the added value it brings to the healthcare providers, from the TB outpatient centre to the coordination team of the National TB Programme (Recommendations 2, 3, and 4, listed on Section 5.4.).

5.3. TIMELINESS

Timeliness refers to the pace at which information progresses through various stages in a surveillance system, from a health-related event to its reporting. In order to assess timeliness of the TB notification system we submitted a research protocol (*5.1.b. Research protocol for a study on the timeliness of tuberculosis notification in Portugal*), that was approved by the ICBAS-UP ethical committee– unfortunately the data needed was not made available in time for analysis and publication of an article. The study will allow the measurement of delays from onset of disease to notification to SVIG-TB, split in components attributable to the patient, to the diagnostic and to the notification, and identifying factors associated to each component. This study will allow us to better understand how the timeliness has varied in recent years, and hopefully identify areas for targeted intervention to improve timeliness of reporting to the National Tuberculosis Programme. For any disease surveillance system, it is important that timeliness is balanced out with sensitivity and data quality, as well as acceptability.

The progressive digitalisation, not only of the health information systems but also of the surveillance processes, as well as the automation and integration of complementary data sources, should have a positive impact on the timeliness of the Portuguese TB surveillance system (Recommendations 1 and 3, listed on Section 5.4.).

5.4. RECOMMENDATIONS

The national tuberculosis surveillance system in Portugal can benefit from the availability of parallel case-based information systems, such as SVIG-TB and SINAVE, that have different objectives and different strengths. The questions regarding record linkage are very important not only for capture-recapture methods using different data sources, but also for the development of integrated surveillance systems that can centralise case-based data from multiple data sources (such as primary care consultations, hospital admission, laboratory results, epidemiological investigation data from the public health services, prescription data, mortality records, etc.). This does not mean that potential data protection ethical and legal issues can be overlooked. For the Portuguese tuberculosis surveillance system, where unique identifiers are usually unavailable for research purposes (which is a reality in many countries), probabilistic data linkage methods using date of birth, sex, and municipality of residence are suitable – but these methods might become obsolete if high-quality electronic TB-registers become available in the future. In any case, evaluation of disease surveillance systems is a routine public health function and therefore data should be made easily available with slightly less strict rules than for traditional research.

Evaluation is a continuous and cyclical process, that starts from an assessment and re-starts after the implementation of changes aiming at improving the system. The recommendations below should be considered for the Portuguese TB surveillance system:

Recommendation 1. Modernising the SVIG-TB notification system, replacing all paper-based elements by a fully digitalised process and ensuring integration of the existing clinical and public health information systems.

Recommendation 2. Setting up periodic evaluation of the surveillance system, including additional attributes such as data quality (i.e. internal completeness and accuracy of records) and positive predictive value (investigating cases not notified

to SVIG-TB but found in other data sources such as SINAVE, hospital records, or other).

Recommendation 3. Conducting a formal assessment of the acceptability of the SVIG-TB system by the notifiers and programme managers, ensuring adequate training and feedback that is useful for practice.

Recommendation 4. Reviewing surveillance objectives and metadata according to the changing epidemiology of tuberculosis in Portugal and the results of the system's evaluation.

Recommendation 5. Exploring alternative and/or complementary data sources for validation and evaluation of the current tuberculosis surveillance system – including for example hospital records, but also aggregated data sources (such as antibiotic consumption or other ecological indicators) which can provide valuable insights into the broader context of the disease.

5.5. CONCLUSION

Public health surveillance systems, including those for infectious diseases, have evolved greatly in recent years. Traditional notification-based systems are still very important for public health, as they tend to be flexible and sufficiently sensitive, but the notification process is burdensome and hence prone to fail. Redundancy and duplication of efforts must be avoided to prevent reporting fatigue, by ensuring that notifying physicians are aware of their responsibilities and that there is a clear and efficient process for reporting cases, as well as demonstrating the importance of the notifications and providing feedback to notifiers. Automation of the notification process can also help to reduce the risk of errors and delays, as well as minimizing the burden on healthcare providers. Digital, case-based, close-to-real-time TB surveillance systems offer several advantages over traditional paper-based reporting of aggregated data, such as the implementation of automated data quality checks, timely access to data, and the availability of individual-level data for individuals with TB infection or disease from the health facility level up to the national level. These systems also greatly simplify data analysis to inform the adaptation and targeting of response efforts both geographically and for specific population groups. The progressive digitalisation of patient health records and development of new tools that allow the automated extraction of clinical

data are progressively taking the stage, but new challenges arise in terms of system flexibility, patient confidentiality and data protection.[305,306]

This thesis was a first attempt at formally evaluating the Portuguese surveillance system for TB, considering the existing data sources and their potential integration to fulfil the national disease surveillance objectives and the End TB strategic goals. Three attributes were considered, and for each different methodologies discussed: 1) sensitivity was assessed through inventory and capture-recapture studies,[1,2] 2) usefulness through studies supporting outbreak investigation,[3,8,9] identification of disease determinants and risk factors,[4,7] and comparing screening strategies and impact of LTBI treatment,[5,6] and 3) timeliness through a research protocol that will support the study of the different components of delay from the onset of symptoms to the notification to the surveillance system. I hope this thesis might be useful to the National TB Programme, but can also support a broader discussion on performance and improvement of surveillance systems, be it in the specific field of tuberculosis, for other infectious diseases, or for other threats to public health.

R

REFERENCES

1. Carvalho, C. *et al.* Completeness of TB notification in Portugal, 2015: an inventory and capture-recapture study. *Int J Tuberc Lung Dis* **24**, 1186–1193 (2020).

2. Straetemans, M. *et al.* Completeness of tuberculosis (TB) notification: Inventory studies and capture-recapture analyses, six European Union countries, 2014 to 2016. *Eurosurveillance* **25**, 1–10 (2020).

3. Rito, T. *et al.* A complex scenario of tuberculosis transmission is revealed through genetic and epidemiological surveys in Porto. *BMC Infect Dis* **18**, 53 (2018).

4. Oliveira, O. *et al.* A nationwide study of multidrug-resistant tuberculosis in Portugal 2014-2017 using epidemiological and molecular clustering analyses. *BMC Infect Dis* **19**, (2019).

5. Sentís, A. *et al.* Failure to complete treatment for latent tuberculosis infection in Portugal, 2013–2017: geographic-, sociodemographic-, and medical-associated factors. *European Journal of Clinical Microbiology and Infectious Diseases* **39**, 647–656 (2020).

6. Sousa, S. *et al.* Comparing the cost-effectiveness of two screening strategies for latent tuberculosis infection in Portugal. *Pulmonology* **27**, 493–499 (2021).

7. Sentís, A. *et al.* Decline of tuberculosis notification rate in different populations and regions in Portugal, 2010–2017. *Pulmonology* (2021) doi:10.1016/j.pulmoe.2021.08.002.

8. Sousa, S. *et al.* Coffee shops, a hub for TB clusters? *Pulmonology* (2023) doi:10.1016/J.PULMOE.2023.04.007.

9. Sousa, S. *et al.* An investigation of TB infection and reinfection among stone quarry workers. *Pulmonology* (2023) doi:10.1016/J.PULMOE.2023.05.004.

10. World Health Organization (WHO). *TB a Global Emergency*. (1994).

11. World Health Organization (WHO). *The End TB Strategy*. (World Health Organization, 2015).

12. MacNeil, A., Glaziou, P., Sismanidis, C., Maloney, S. & Floyd, K. Global Epidemiology of Tuberculosis and Progress Toward Achieving Global Targets — 2017. *MMWR Morb Mortal Wkly Rep* **68**, 263–266 (2019).

13. World Health Organization (WHO). *Global tuberculosis report 2022*. (2023).

14. European Centre for Disease Prevention and Control (ECDC). *Tuberculosis surveillance and monitoring in Europe 2023 - 2021 data*. (2023). doi:10.2900/637206.

15. World Health Organization (WHO). *Global Tuberculosis Report 2021*. (2021).

16. Bagcchi, S. Dismal global tuberculosis situation due to COVID-19. *Lancet Infect Dis* **21**, 1636 (2021).

17. Kirby, T. Global tuberculosis progress reversed by COVID-19 pandemic. *The Lancet Respiratory* **9**, e118–e119 (2021).

18. Zimmer, A. J. *et al.* Tuberculosis in times of COVID-19. *J Epidemiol Community Health* **76**, 310–316 (2022).

19. Winglee, K., Hill, A. N., Langer, A. J. & Self, J. L. Decrease in Tuberculosis Cases during COVID-19 Pandemic as Reflected by Outpatient Pharmacy Data, United States, 2020. *Emerg Infect Dis* **28**, 820 (2022).

20. WHO Regional Office for Europe. *Final report on implementation of the Tuberculosis Action Plan for the WHO European Region 2016-2020*. (2020).

21. World Health Organization European Region. *Tuberculosis action plan for the WHO European Region 2023–2030*. (2022).

22. Ammon, A. & Kluge, H. Ending tuberculosis in Europe - resetting the course in the post-COVID-19 era. *Eurosurveillance* **28**, 1 (2023).

23. Programa Nacional para a Tuberculose (PNT). *Relatório de Vigilância e Monitorização da Tuberculose em Portugal 2021: dados definitivos 2020*. (Directorate-General of Health (DGS), 2022).

24. Programa Nacional para a Tuberculose (PNT). *Relatório de Vigilância e Monitorização da Tuberculose em Portugal 2022: dados definitivos 2021*. (2023).

25. Brownson, R. C., Baker, E. A., Leet, T. L., Gillespie, K. N. & True, W. R. *Evidence-Based Public Health*. (Oxford University Press, 2011).

26. World Health Organization (WHO). *WHO Tuberculosis Programme : framework for effective tuberculosis control*. https://apps.who.int/iris/handle/10665/58717 (1994).

27. World Health Organization (WHO). *Definitions and reporting framework for tuberculosis – 2013 revision: updated December 2014 and January 2020*. (2013).

28. World Health Organization (WHO). Digital, case-based, real-time surveillance for TB: status of progress. https://www.who.int/publications/digital/global-tuberculosis-report-2021/featured-topics/case-based-surveillance (2021).

29. European Union. General Data Protection Regulation (GDPR) – Official Legal Text. *Official Journal of the European Union* https://gdpr-info.eu/ (2016).

30. Directorate-General of Health Portugal. Sistema Nacional de Vigilância Epidemiológica (SINAVE). https://www.dgs.pt/servicos-on-line1/sinave-sistema-nacional-de-vigilancia-epidemiologica.aspx.

31. Direção-Geral da Saúde (DGS). *Despacho n.º 1150/2021 de 28 de janeiro: Doenças de notificação obrigatória a notificar na plataforma de apoio ao SINAVE (Sistema Nacional de Vigilância Epidemiológica) ou no SI-Vida (Sistema de informação VIH/SIDA)*. (2021).

32. European Comission. *Commission Implementing Decision (EU) 2018/945: the communicable diseases and related special health issues to be covered by epidemiological surveillance as well as relevant case definitions*. (2018).

33. Direção-Geral da Saúde (DGS). Tuberculose - Sistema SVIG-TB. https://www.dgs.pt/paginas-de-sistema/saude-de-a-a-z/tuberculose1/sistema-svig-tb.aspx.

34. Direção-Geral da Saúde (DGS). *Circular Normativa Nr. 8/DT 29/05/00: Sistema de Informação Intrínseco ao Programa de Tuberculose - registo dos casos e monitorização dos resultados*. 1–5 (2000).

35. Direção-Geral da Saúde (DGS). *Circular Normativa Nr. 6/DT 13/03/2001: Sistema de Vigilância da Tuberculose (SVIG-TB) Substituição da Aplicação Informática e Suporte do Registo Clínico dos Casos*. *Circular Normativa nº6* 1–2 (2001).

36. European Centre for Disease Prevention and Control (ECDC). *The European Surveillance System (TESSy)*. (2008).

37. Frieden, T. R. & Munsiff, S. S. The DOTS Strategy for Controlling the Global Tuberculosis Epidemic. *Clin Chest Med* **26**, 197–205 (2005).

38. European Centre for Disease Prevention and Control. *Progressing towards TB elimination*. (2010) doi:10.2900/34915.

39. German, R. R. *et al.* Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. *MMWR Recomm Rep* **50**, 1–35; quiz CE1-7 (2001).

40. European Centre for Disease Prevention and Control (ECDC). *Data quality monitoring and surveillance system evaluation - A handbook of methods and applications*. (2014).

41. Groseclose, S. L., German, R. R. & Nsubuga, P. Evaluating Public Health Surveillance. in *Principles & Practice of Public Health Surveillance* (eds. Lee, L. M., Teutsch, S. M., Thacker, S. B. & St. Louis, M. E.) 166–197 (Oxford University Press, 2010). doi:10.1093/acprof:oso/9780195372922.003.0008.

42. World Health Organization Regional Office for Europe. Tuberculosis action plan for the WHO European Region 2016–2020. 14–17 (2015).

43. World Health Organization. TB Impact Measurement - Policy and recommendations for how to assess the epidemiological burden of TB and the impact of TB control (Stop TB policy paper; No. 2). *Group (New York)* **384**, 1–58 (2009).

44. Tambo, E. *et al.* Surveillance-response systems: The key to elimination of tropical diseases. *Infect Dis Poverty* **3**, 1–11 (2014).

45. Hatherell, H. A., Simpson, H., Baggaley, R. F., Hollingsworth, T. D. & Pullan, R. L. Sustainable Surveillance of Neglected Tropical Diseases for the Post-Elimination Era. *Clinical Infectious Diseases* **72**, S210–S216 (2021).

46. Dye, C. *et al.* Measuring tuberculosis burden, trends, and the impact of control programmes. *The Lancet Infectious Diseases* vol. 8 233–243 Preprint at https://doi.org/10.1016/S1473-3099(07)70291-8 (2008).

47. Glaziou, P., Sismanidis, C., Dodd, P., Zignol, M. & Floyd, K. *Methods used by WHO to estimate the global burden of TB disease*. http://www.who.int/tb/publications/global_report/gtbr2017_online_technical_appendix_disease_burden_estimation.pdf (2017).

48. Trébucq, A. & Schwoebel, V. Numbers of tuberculosis cases: dreams and reality. *Int J Tuberc Lung Dis* **20**, 1288–1292 (2016).

49. World Health Organization. *Assessing tuberculosis under-reporting through inventory studies*. (2012) doi:ISBN 978 92 4 150494 2.

50. Pillaye, J. & Clarke, A. An evaluation of completeness of tuberculosis notification in the United Kingdom. *BMC Public Health* vol. 3 1–5 Preprint at https://doi.org/10.1186/1471-2458-3-31 (2003).

51. Keramarou, M. & Evans, M. R. Completeness of infectious disease notification in the United Kingdom: A systematic review. *Journal of Infection* **64**, 555–564 (2012).

52. Van Hest, N. A. H. *et al.* Completeness of notification of tuberculosis in The Netherlands: how reliable is record-linkage and capture–recapture analysis? *Epidemiol Infect* **135**, 1021–1029 (2007).

53. Domaszewska, T. *et al.* Completeness of tuberculosis case notifications in Germany in 2013–2017: first results of an inventory study. *BMC Infect Dis* **20**, 1–13 (2020).

54. Cojocaru, C., Van Hest, N. A., Mihaescu, T. & Davies, P. D. Completeness of notification of adult tuberculosis in Iasi county, Romania: A capture-recapture analysis. *International Journal of Tuberculosis and Lung Disease* **13**, 1094–1099 (2009).

55. Ibarz-Pavón, A. B. *et al.* Pilot study of the completeness of notification of adult tuberculosis in Athens, Greece. *Int J Tuberc Lung Dis* **20**, 920–925 (2016).

56. Jelastopulu, E. *et al.* Substantial underreporting of tuberculosis in West Greece: implications for local and national surveillance. *Euro Surveill* **14**, 19152 (2009).

57. Li, T. *et al.* Under-reporting of diagnosed tuberculosis to the national surveillance system in China: An inventory study in nine counties in 2015. *BMJ Open* (2019) doi:10.1136/bmjopen-2018-021529.

58. Kang, H. Y. *et al.* Tuberculosis Notification Completeness and Timeliness in the Republic of Korea During 2012–2014. *Osong Public Health Res Perspect* **7**, 320–326 (2016).

59. Lo, H. Y., Yang, S. L., Chou, P., Chuang, J. H. & Chiang, C. Y. Completeness and timeliness of tuberculosis notification in Taiwan. *BMC Public Health* **11**, 915 (2011).

60. Alkhalawi, M. J., McNabb, S. J. N., Assiri, A. M. & Memish, Z. A. Evaluation of tuberculosis public health surveillance, Al-Madinah province, Kingdom of Saudi Arabia, 2012. *J Epidemiol Glob Health* **6**, 37 (2016).

61. Quiroz-Ruiz, H. R., Sosa-Flores, J. L. & Hernández-Palomino, F. N. Subregistro y exhaustividad de los sistemas de vigilancia de tuberculosis en una región del Perú: un análisis de captura-recaptura. *Cad Saude Publica* **37**, e00276020 (2021).

62. Dunbar, R. *et al.* Capture-recapture to estimate completeness of tuberculosis surveillance in two communities in South Africa. *International Journal of Tuberculosis and Lung Disease* **15**, 1038–1043 (2011).

63. Da Luz, E. F. & Braga, J. U. Under-reporting of tuberculosis in Praia, Cape Verde, from 2006 to 2012. *International Journal of Tuberculosis and Lung Disease* vol. 22 258–263 Preprint at https://doi.org/10.5588/ijtld.17.0256 (2018).

64. World Health Organization (WHO). *Communicable disease surveillance and response systems: Guide to monitoring and evaluating*. (2006).

65. Chaintarli, K., Jackson, S., Cotter, S. & O'Donnell, J. Evaluation and comparison of the National Tuberculosis (TB) Surveillance System in Ireland before and after the introduction of the Computerised Electronic Reporting System (CIDR). *Epidemiol Infect* **146**, 1756–1762 (2018).

66. Saeed, K. M. I., Bano, R. & Asghar, R. J. Evaluation of the national tuberculosis surveillance system in Afghanistan تقييم النظام الوطني لرتصٌ د السل يف أفغانستان. **19**, (2013).

67. Al kalali, F. S. A., Mahyoub, E., Al-Hammadi, A., Anam, L. & Khader, Y. Evaluation of the National Tuberculosis Surveillance System in Sana'a, Yemen, 2018: Observational Study. *JMIR Public Health Surveill* **7**, (2021).

68. Garrido-Cardenas, J. A., de Lamo-Sevilla, C., Cabezas-Fernández, M. T., Manzano-Agugliaro, F. & Martínez-Lirola, M. Global tuberculosis research and its future prospects. *Tuberculosis* **121**, 101917 (2020).

69. Tetali, S. R. *et al.* Current advances in the clinical development of anti-tubercular agents. *Tuberculosis* **125**, 101989 (2020).

70. Kanipe, C. & Palmer, M. V. Mycobacterium bovis and you: A comprehensive look at the bacteria, its similarities to Mycobacterium tuberculosis, and its relationship with human disease. *Tuberculosis* **125**, 102006 (2020).

71. European Centre for Disease Prevention and Control (ECDC). *Social determinants and risk factors in tuberculosis surveillance in the EU/EEA*. (ECDC, 2018). doi:10.2900/463341.

72. de Colombani, P. & Hovhannesyan, A. Social determinants and risk factors for tuberculosis in national surveillance systems in Europe. *Public Health Action* **5**, 194–201 (2015).

73. Marçôa, R. Tuberculosis and gender – Factors influencing the risk of tuberculosis among men and women by age group. *Pulmonology* **24**, 199–202 (2018).

74. Schwartz, N. G., Price, S. F., Pratt, R. H. & Langer, A. J. Tuberculosis — United States, 2019. *MMWR Morb Mortal Wkly Rep* **69**, 286–289 (2020).

75. Jajosky, R. A. & Groseclose, S. L. Evaluation of reporting timeliness of public health surveillance systems for infectious diseases. *BMC Public Health* **4**, 29 (2004).

76. Jajosky, R. A. & Groseclose, S. L. Evaluation of reporting timeliness of public health surveillance systems for infectious diseases. *BMC Public Health* **4**, 29 (2004).

77. Najeebullah, K., Liebig, J., Darbro, J., Jurdak, R. & Paini, D. Timely surveillance and temporal calibration of disease response against human infectious diseases. *PLoS One* **16**, (2021).

78. Bastos, L. S. *et al.* A modelling approach for correcting reporting delays in disease surveillance data. *Stat Med* **38**, 4363–4377 (2019).

79. Botsis, T., Bellika, J. G. & Hartvigsen, G. New directions in electronic disease surveillance: Detection of infectious diseases during the incubation period. *Proceedings - International Conference on eHealth, Telemedicine, and Social Medicine, eTELEMED 2009* 176–183 (2009) doi:10.1109/ETELEMED.2009.9.

80. Cheng, S. *et al.* Effect of Diagnostic and Treatment Delay on the Risk of Tuberculosis Transmission in Shenzhen, China: An Observational Cohort Study, 1993–2010. *PLoS One* **8**, e67516 (2013).

81. Zhu, X. han *et al.* Association between diagnostic delay and prognosis of pulmonary tuberculosis in Shandong, China: a retrospective study. *BMC Pulm Med* **22**, (2022).

82. Bea, S. *et al.* Risk of mortality and clinical outcomes associated with healthcare delay among patients with tuberculosis. *J Infect Public Health* **16**, 1313–1321 (2023).

83. Storla, D. G., Yimer, S. & Bjune, G. A. A systematic review of delay in the diagnosis and treatment of tuberculosis. *BMC Public Health* **8**, 15 (2008).

84. Zão, I., Ribeiro, A. I., Apolinário, D. & Duarte, R. Why does it take so long? The reasons behind tuberculosis treatment delay in Portugal. *Pulmonology* **25**, 215–222 (2019).

85. Marques, J. *et al.* The effect of TB patient delay on loss to follow-up in Portugal. *Int J Tuberc Lung Dis* **27**, 537–542 (2023).

86. Soares, P., Aguiar, A., Leite, A., Duarte, R. & Nunes, C. Ecological factors associated with areas of high tuberculosis diagnosis delay. *Public Health* **208**, 32–39 (2022).

87. Santos, E., Felgueiras, Oliveira, O. & Duarte, R. Diagnosis delay of tuberculosis in the Huambo province, Angola. *Pulmonology* **24**, 294–299 (2018).

88. Seabra, B. & Duarte, R. Tuberculosis national registries and data on diagnosis delay – Is there room for improvement? *Pulmonology* (2021) doi:10.1016/j.pulmoe.2021.05.004.

89. Guimarães, M., Oliveira, O., Teixeira, C., Gaio, A. R. & Duarte, R. Delay in the diagnosis of tuberculosis. *Revista Portuguesa de Pneumologia (English Edition)* **21**, 346–348 (2015).

90. Loutet, M. G. *et al.* Delay from symptom onset to treatment start among tuberculosis patients in England, 2012–2015. *Epidemiol Infect* **146**, 1511–1518 (2018).

91. Tattevin, P. *et al.* Factors associated with patient and health care system delay in the diagnosis of tuberculosis in France. *Int J Tuberc Lung Dis* **16**, 510–515 (2012).

92. Gagliotti, C., Resi, D. & Moro, M. L. Delay in the treatment of pulmonary TB in a changing demographic scenario. *Int J Tuberc Lung Dis* **10**, 305–9 (2006).

93. Stjepanović, M. *et al.* PATIENT, HEALTHCARE SYSTEM AND TOTAL DELAY IN TUBERCULOSIS DIAGNOSIS AND TREATMENT AMONG SERBIAN POPULATION. *Acta Clin Croat* **57**, 257 (2018).

94. Getnet, F., Demissie, M., Assefa, N., Mengistie, B. & Worku, A. Delay in diagnosis of pulmonary tuberculosis in low-and middle-income settings: systematic review and meta-analysis. *BMC Pulm Med* **17**, (2017).

95. Sebastian, N. M., Haveri, S. P. & Nath, A. S. Delay in diagnosis of tuberculosis and related factors from a district in Kerala, India. *Indian Journal of Tuberculosis* **68**, 59–64 (2021).

96. Curtis, A. B., McCray, E., McKenna, M. & Onorato, I. M. Completeness and timeliness of tuberculosis case reporting: A multistate study. *Am J Prev Med* **20**, 108–112 (2001).

97. Campbell, J. R., Faust, L., Paulsen, C. & Heffernan, C. The state of tuberculosis surveillance in Canada. *Canadian Journal of Public Health* **114**, 671–675 (2023).

98. da Silva, G. D. M., Bartholomay, P., Cruz, O. G. & Garcia, L. P. Evaluation of data quality, timeliness and acceptability of the tuberculosis surveillance system in Brazil's micro-regions. *Cien Saude Colet* **22**, 3307–3319 (2017).

99. Commission Implementing Decision 2018/945. On the communicable diseases and related special health issues to be covered by epidemiological surveillance as well as relevant case definitions. *Official Journal of the European Union* 1–74 (2018).

100. Ted eTendering - Calls for tenders from the European Institutes. Assessment of tuberculosis under-reporting through inventory studies. https://etendering.ted.europa.eu/cft/cft-documents.html?cftId=1136 (2015).

101. Uplekar, M. *et al.* WHO's new end TB strategy. *Lancet* **385**, 1799–1801 (2015).

102. Directorate-General of Health Portugal. *Tuberculose em Portugal - Desafios e Estratégias - 2018*. https://www.dgs.pt/documentos-e-publicacoes/tuberculose-em-portugal-desafios-e-estrategias-2018-.aspx (2018).

103. European Centre for Disease Prevention and Control & WHO Regional Office for Europe. *Tuberculosis surveillance and monitoring in Europe 2017*. (2017) doi:10.2900/666960.

104. van Hest, R. Capture-recapture Methods in Surveillance of Tuberculosis and Other Infectious Diseases. *Erasmus MC - University Medical Center Rotterdam repository* (2007).

105. Portugal - Inspeção Geral dos Serviços Sanitários do Reino. *Regulamento Geral dos Serviços de Saúde e Beneficiência Pública*. 3–126 (1902).

106. Blasnik, M. RECLINK: Stata module to probabilistically match records. *Statistical Software Components* (2010).

107. Cormack, R. M. Log-Linear Models for Capture-Recapture. *Biometrics* (1989) doi:10.2307/2531485.

108. Fienberg, S. E. The multiple recapture census for closed populations and incomplete 2 contingency tables. *Biometrika* **59**, 591–603 (1972).

109. Hook, E. B. & Regal, R. R. Validity of methods for model selection, weighting for model uncertainty, and small sample adjustment in capture-recapture estimation. *Am J Epidemiol* **145**, 1138–1144 (1997).

110. Tocque, K., Bellis, M. A., Beeching, N. J. & Davies, P. D. Capture recapture as a method of determining the completeness of tuberculosis notifications. *Communicable disease and public health / PHLS* **4**, 141–143 (2001).

111. van hest, N. A. H. *et al.* Completeness of notification of tuberculosis in The Netherlands: How reliable is record-linkage and capture - Recapture analysis? *Epidemiol Infect* **135**, 1021–1029 (2007).

112. Van Hest, N. A. H. H. *et al.* Record-linkage and capture-recapture analysis to estimate the incidence and completeness of reporting of tuberculosis in England 1999-2002. *Epidemiol Infect* **136**, 1606–1616 (2008).

113. College Station, T. S. L. *StataCorp. 2017*. *Stata Statistical Software: Release 15* (2017).

114. Desenclos, J.-C. & Hubert, B. Limitations to the Universal use of Capture-Recapture Methods. *Int J Epidemiol* **23**, 1322–1323 (1994).

115. Dunn, J. & Andreoli, S. B. Método de captura e recaptura: nova metodologia para pesquisas epidemiológicas. *Revista de Saude Publica* vol. 28 449–453 Preprint at (1994).

116. Cormack, R. Problems with Using Capture-Recapture in Epidemiology: An Example of a Measles Epidemic. *J Clin Epidemiol* **52**, 909–914 (1999).

117. Ding, Y. & Fienberg, S. S. Multiple sample estimation of population and census undercount in the presence of matching errors. *Survey Methodol* **22**, 55–64 (1996).

118. Hook, E. B. & Regal, R. R. Capture-Recapture Methods in Epidemiology: Methods and Limitations. *Epidemiol Rev* **17**, 243–264 (1995).

119. Hook, E. B. & Regal, R. R. Recommendations for Presentation and Evaluation of Capture-Recapture Estimates in Epidemiology. *J Clin Epidemiol* **52**, 917–926 (1999).

120. Hook, E. B. & Regal, R. R. A. Effect of variation in probability of ascertainment by sources (‘variable catchability’) upon ‘capture-recapture’ estimates of prevalence. *Am J Epidemiol* **137**, (1993).

121. Tilling, K. *Capture-recapture methods-useful or misleading? International Journal of Epidemiology* vol. 30 (2001).

122. Galego, M. A., Santos, J. V., Viana, J., Freitas, A. & Duarte, R. To be or not to be hospitalised with tuberculosis in Portugal. *International Journal of Tuberculosis and Lung Disease* **23**, 1029–1034 (2019).

123. Chao, A., Tsay, P. K., Lin, S.-H. H., Shau, W.-Y. Y. & Chao, D.-Y. Y. The Applications of Capture-Recapture Models to Epidemiological Data. *Stat Med* **1**, 31–65 (2005).

124. Tilling, K. & Sterne, J. A. C. Capture-recapture models including covariate effects. *Am J Epidemiol* (1999) doi:10.1093/oxfordjournals.aje.a009825.

125. Wolter, K. M. Accounting for America’s uncounted and miscounted. *Science (1979)* **253**, 12–15 (1991).

126. Dunn, J. & Andreoli, S. B. Capture and recapture method: a new methodology for epidemiological research. *Rev Saude Publica* **28**, 449–453 (1994).

127. World Health Organization. *Towards tuberculosis elimination: an action framework for low-incidence countries*. *ISBN 978 92 4 150770 7* vol. 1 (2014).

128. WHO Regional Office for Europe. *Roadmap to implement the tuberculosis action plan for the WHO European Region 2016-2020 - Towards ending tuberculosis and multidrug-resistant tuberculosis*. (2016).

129. World Health Organization (WHO). *The End TB Strategy*. (World Health Organization, 2015).

130. Migliori, G. B. *et al.* Validation of the surveillance system for new cases of tuberculosis in a province of Northern Italy. *European Respiratory Journal* **8**, 1252–1258 (1995).

131. Mukerjee, A. K. Ascertainment of non-respiratory tuberculosis in five boroughs by comparison of multiple data sources. *Communicable disease and public health / PHLS* **2**, 143–144 (1999).

132. International Classification of Diseases Ninth revision. Basic Tabulation List with Alphabetical Index. *World Health Organization* **91**, 399–404 (1978).

133. World Health Organization. International statistical classification of diseases and related health problems, 10th revision (ICD-10), Fifth version. *World Health Organization* **1**, 332–345 (2016).

134. International Working Group for Disease Monitoring and Forecasting. Capture-recapture and multiple-record systems estimation I: History and theoretical development. International Working Group for Disease Monitoring and Forecasting. *Am J Epidemiol* (1995).

135. Bishop, Y. M., Holland, P. W. & Fienberg, S. E. *Discrete multivariate analysis theory and practice*. *Discrete Multivariate Analysis Theory and Practice* (Springer New York, 2007). doi:10.1007/978-0-387-72806-3.

136. Capture-recapture and multiple-record systems estimation II: Applications in human diseases. *Am J Epidemiol* **142**, 1059–1068 (1995).

137. Brenner, H. Use and limitations of the capture-recapture method in disease monitoring with two dependent sources. *Epidemiology* **6**, 42–48 (1995).

138. European Centre for Disease Prevention and Control. *Tuberculosis surveillance and monitoring in Europe Tuberculosis surveillance and monitoring in Europe 2018*. (2018).

139. WHO. Tuberculosis prevalence surveys: a handbook. *World Health Organization* 1–324 (2011).

140. Baussano, I. *et al.* Undetected burden of tuberculosis in a low-prevalence area. *International Journal of Tuberculosis and Lung Disease* **10**, 415–421 (2006).

141. Warris, A. & Draaisma, J. M. T. Tuberculose in Nederland. *Tijdschr Kindergeneeskd* **74**, 21–22 (2006).

142. European Parliament and Council of the European Union. *EU general data protection regulation*. *Regulation (EU) 2016/679 of the European Parliament of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data P* vol. 2014 45–62 (2018).

143. Van Loenhout-Rooyackers, J. H., Leufkens, H. G. M., Hekster, Y. A. & Kalisvaart, N. A. Pyrazinamide use as a method of estimating under-reporting of tuberculosis. *Int J Tuberc Lung Dis* **5**, 1156–60 (2001).

144. Regal, R. R. & Hook, E. B. Marginal versus conditional versus 'structural source' models: A rationale for an alternative to log-linear methods for capture-recapture estimates. *Statistics in Medicine* vol. 17 69–74 Preprint at https://doi.org/10.1002/(SICI)1097-0258(19980115)17:1<69::AID-SIM729>3.0.CO;2-Q (1998).

145. Jones, H. E. *et al.* Recapture or precapture? Fallibility of standard capture-recapture methods in the presence of referrals between sources. *Am J Epidemiol* **179**, 1383–1393 (2014).

146. Van Hest, N. A. H. *et al.* Incidence and completeness of notification of Legionnaires' disease in The Netherlands: Covariate capture-recapture analysis acknowledging regional differences. *Epidemiol Infect* **136**, 540–550 (2008).

147. Thacker, S., Trowbridge, F. & Parrish, R. A method for evaluating systems of epidemiological surveillance. *World health statistics quarterly* **41**, 11–18 (1988).

148. World Health Organization (WHO). *Global Tuberculosis Report 2017*. http://www.who.int/tb/publications/global_report/en/ (2017) doi:WHO/HTM/TB/2017.23.

149. DGS. *Portugal: Infeção por VIH, SIDA e Tuberculose em números - 2015*. (2015).

150. Paulino, J. *et al.* Tuberculosis in native-and foreign-born populations in Portugal. *International Journal of Tuberculosis and Lung Disease* **20**, 357–362 (2016).

151. Lönnroth, K. *et al.* Towards tuberculosis elimination: An action framework for low-incidence countries. *European Respiratory Journal* vol. 45 928–952 Preprint at https://doi.org/10.1183/09031936.00214014 (2015).

152. Nunes, C. & Taylor, B. M. Modelling the time to detection of urban tuberculosis in two big cities in Portugal: A spatial survival analysis. *International Journal of Tuberculosis and Lung Disease* **20**, (2016).

153. Franco, I. *et al.* Social profile of the highest tuberculosis incidence areas in Portugal. *Revista Portuguesa de Pneumologia* vol. 22 50–52 Preprint at https://doi.org/10.1016/j.rppnen.2015.08.006 (2016).

154. Bastos, H. N. *et al.* A prediction rule to stratify mortality risk of patients with pulmonary tuberculosis. *PLoS One* **11**, (2016).

155. Barnes, P. F. & Cave, M. D. Molecular Epidemiology of Tuberculosis. *New England Journal of Medicine* **349**, 1149–1156 (2003).

156. Gagneux, S. *et al.* Variable host-pathogen compatibility in Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* **103**, 2869–2873 (2006).

157. Lopes, J. S. *et al.* SNP typing reveals similarity in Mycobacterium tuberculosis genetic diversity between Portugal and Northeast Brazil. *Infection, Genetics and Evolution* **18**, 238–246 (2013).

158. Supply, P. *et al.* Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of Mycobacterium tuberculosis. *J Clin Microbiol* **44**, 4498–4510 (2006).

159. Supply, P. *et al.* Automated high-throughput genotyping for study of global epidemiology of Mycobacterium tuberculosis based on mycobacterial interspersed repetitive units. *J Clin Microbiol* **39**, 3563–3571 (2001).

160. Vadwai, V., Shetty, A., Supply, P. & Rodrigues, C. Evaluation of 24-locus MIRU-VNTR in extrapulmonary specimens: Study from a tertiary centre in Mumbai. *Tuberculosis* **92**, 264–272 (2012).

161. Yuan, L. *et al.* Genotypic characteristics of Mycobacterium tuberculosis circulating in Xinjiang, China. *Infect Dis* **48**, 108–115 (2016).

162. Bandelt, H. J., Forster, P. & Röhl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* **16**, 37–48 (1999).

163. Salipante, S. J. & Hall, B. G. Inadequacies of minimum spanning trees in molecular epidemiology. *J Clin Microbiol* **49**, 3568–3575 (2011).

164. Bandelt, H.-J., Forster, P., Sykes, B. C. & Richards, M. B. Mitochondrial Portraits of Human Populations Using Median Networks. *Genetics* **141**, 743–53 (1995).

165. Walker, T. M. *et al.* Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: A retrospective cohort study. *Lancet Infect Dis* **15**, 1193–1202 (2015).

166. Gurjav, U. *et al.* Whole genome sequencing demonstrates limited transmission within identified Mycobacterium tuberculosis clusters in New South Wales, Australia. *PLoS One* **11**, (2016).

167. Comas, I. *et al.* Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. *Nat Genet* **45**, 1176–1182 (2013).

168. Comas, I., Homolka, S., Niemann, S. & Gagneux, S. Genotyping of genetically monomorphic bacteria: DNA sequencing in Mycobacterium tuberculosis highlights the limitations of current methodologies. *PLoS One* **4**, (2009).

169. Fitzgibbon, M. M. *et al.* A snapshot of genetic lineages of Mycobacterium tuberculosis in Ireland over a two-year period, 2010 and 2011. *Eurosurveillance* **18**, (2013).

170. Pérez Del Molino Bernal, I. C. *et al.* Genomic Diversity of Mycobacterium tuberculosis Complex Strains in Cantabria (Spain), a Moderate TB Incidence Setting. *PLoS One* **11**, (2016).

171. Oelemann, M. C. *et al.* Assessment of an optimized mycobacterial interspersed repetitive-unit- variable-number tandem-repeat typing system combined with spoligotyping for population-based molecular epidemiology studies of tuberculosis. *J Clin Microbiol* **45**, 691–697 (2007).

172. Roetzer, A. *et al.* Whole Genome Sequencing versus Traditional Genotyping for Investigation of a Mycobacterium tuberculosis Outbreak: A Longitudinal Molecular Epidemiological Study. *PLoS Med* **10**, e1001387 (2013).

173. van Deutekom, H. *et al.* Clustered Tuberculosis Cases Do They Represent Recent Transmission and Can They Be Detected Earlier? *Am J Respir Crit Care Med* 806–810 (2004) doi:10.1164/rccm.200306-856OC.

174. Fallico, L. *et al.* Four year longitudinal study of Mycobacterium tuberculosis complex isolates in a region of North-Eastern Italy. *Infection, Genetics and Evolution* **26**, 58–64 (2014).

175. Kato-Maeda, M., Metcalfe, J. Z. & Flores, L. Genotyping of Mycobacterium tuberculosis: Application in epidemiologic studies. *Future Microbiology* vol. 6 203–216 Preprint at https://doi.org/10.2217/fmb.10.165 (2011).

176. Dias, M. *et al.* Tuberculosis among the homeless: Should we change the strategy? *International Journal of Tuberculosis and Lung Disease* **21**, 327–332 (2017).

177. Moss, A. R. *et al.* Tuberculosis in the homeless: A prospective study. *Am J Respir Crit Care Med* **162**, 460–464 (2000).

178. D'Ambrosio, L. *et al.* Tuberculosis elimination: Theory and practice in Europe. *European Respiratory Journal* **43**, 1410–1420 (2014).

179. Hamblion, E. L. *et al.* Recent TB transmission, clustering and predictors of large clusters in London, 2010-2012: Results from first 3-years of universal MIRU-VNTR strain typing. *Thorax* **71**, 749–756 (2016).

180. World Health Organization. *Antimicrobial resistance: global report on surveillance*. *World Health Organization* (2014).

181. Xie, Y. L. *et al.* Evaluation of a Rapid Molecular Drug-Susceptibility Test for Tuberculosis. *New England Journal of Medicine* **377**, 1043–1054 (2017).

182. World Health Organization (WHO). *Global Tuberculosis Report 2018*. (2018).

183. ECDC. *Tuberculosis surveillance and monitoring in Europe 2019: 2017 data*. (2019) doi:10.2900/512553.

184. Perdigão, J. *et al.* Tuberculosis drug-resistance in Lisbon, Portugal: A 6-year overview. *Clinical Microbiology and Infection* **17**, 1397–1402 (2011).

185. Perdigão, J. *et al.* Unraveling genomic diversity and evolution in lisbon, portugal, a highly drug resistant setting. *BMC Genomics* **15**, 991 (2014).

186. Perdigão, J. *et al.* Clonal expansion across the seas as seen through CPLP-TB database: A joint effort in cataloguing Mycobacterium tuberculosis genetic diversity in Portuguese-speaking countries. *Infection, Genetics and Evolution* **72**, 44–58 (2019).

187. CDC. *Prioritizing Tuberculosis Genotype Clusters for Further Investigation & Public Health Action*. (2017).

188. Wyllie, D. H. *et al.* A Quantitative Evaluation of MIRU-VNTR Typing Against Whole-Genome Sequencing for Identifying Mycobacterium tuberculosis Transmission: A Prospective Observational Cohort Study. *EBioMedicine* **34**, 122–130 (2018).

189. Meehan, C. J. *et al.* The relationship between transmission time and clustering methods in Mycobacterium tuberculosis epidemiology. *EBioMedicine* **37**, 410–416 (2018).

190. Stucki, D. *et al.* Standard genotyping overestimates transmission of mycobacterium tuberculosis among immigrants in a low-incidence country. *J Clin Microbiol* **54**, 1862–1870 (2016).

191. Bjorn-Mortensen, K. *et al.* Tracing Mycobacterium tuberculosis transmission by whole genome sequencing in a high incidence setting: A retrospective population-based study in East Greenland. *Sci Rep* **6**, 33180 (2016).

192. Jajou, R. *et al.* Epidemiological links between tuberculosis cases identified twice as efficiently by whole genome sequencing than conventional molecular typing: A population-based study. *PLoS One* **13**, e0195413 (2018).

193. Andrés, M. *et al.* Molecular and genomic typing for tuberculosis surveillance: A survey study in 26 European countries. *PLoS One* **14**, e0210080 (2019).

194. Macedo, R. *et al.* Dissecting whole-genome sequencing-based online tools for predicting resistance in Mycobacterium tuberculosis: can we use them for clinical decision guidance? *Tuberculosis* **110**, 44–51 (2018).

195. Anderson, L. F. *et al.* Transmission of multidrug-resistant tuberculosis in the UK: A cross-sectional molecular and epidemiological study of clustering and contact tracing. *Lancet Infect Dis* **14**, 406–415 (2014).

196. Glynn, J. R., Vynnycky, E. & Fine, P. E. M. Influence of sampling on estimates of clustering and recent transmission of Mycobacterium tuberculosis derived from DNA fingerprinting techniques. *Am J Epidemiol* **149**, 366–371 (1999).

197. Macedo, R. *et al.* Evaluation of a gene-by-gene approach for prospective whole-genome sequencing-based surveillance of multidrug resistant Mycobacterium tuberculosis. *Tuberculosis* **115**, 81–88 (2019).

198. Lange, C. *et al.* Management of patients with multidrugresistant/ extensively drug-resistant tuberculosis in Europe: A TBNET consensus statement. *European Respiratory Journal* **44**, 23–63 (2014).

199. Günther, G. *et al.* Multidrug-resistant tuberculosis in europe, 2010-2011. *Emerg Infect Dis* **21**, 409–416 (2015).

200. Asgedom, S. W., Teweldemedhin, M. & Gebreyesus, H. Prevalence of Multidrug-Resistant Tuberculosis and Associated Factors in Ethiopia: A Systematic Review. *J Pathog* **2018**, 1–8 (2018).

201. Workicho, A., Kassahun, W. & Alemseged, F. Risk factors for multidrug-resistant tuberculosis among tuberculosis patients: A case-control study. *Infect Drug Resist* **10**, 91–96 (2017).

202. Mesfin, E. A. *et al.* Drug-resistance patterns of mycobacterium tuberculosis strains and associated risk factors among multi drug-resistant tuberculosis suspected patients from Ethiopia. *PLoS One* **13**, e0197737 (2018).

203. European Centre for Disease Prevention and Control. Molecular typing for surveillance of multidrug-resistant tuberculosis in the EU / EEA. *Stockholm* 1–6 (2017).

204. Roycroft, E. *et al.* Molecular epidemiology of multi- and extensively-drug-resistant Mycobacterium tuberculosis in Ireland, 2001–2014. *Journal of Infection* **76**, 55–67 (2018).

205. Vluggen, C. *et al.* Molecular epidemiology of Mycobacterium tuberculosis complex in Brussels, 2010-2013. *PLoS One* **12**, e0172554 (2017).

206. Fiebig, L. *et al.* A joint cross-border investigation of a cluster of multidrug-resistant tuberculosis in Austria, Romania and Germany in 2014 using classic, genotyping and whole genome sequencing methods: Lessons learnt. *Eurosurveillance* **22**, 30439 (2017).

207. Feng, J. Y. *et al.* Clinical & bacteriological characteristics associated with clustering of multidrug-resistant tuberculosis. *International Journal of Tuberculosis and Lung Disease* **21**, 766–773 (2017).

208. Somoskovi, A. *et al.* Transmission of multidrug-resistant tuberculosis in a low-incidence setting, Switzerland, 2006 to 2012. *Eurosurveillance* **19**, 20736 (2014).

209. Macedo, R. & Duarte, R. Trends of multidrug-resistant tuberculosis clustering in Portugal. *ERJ Open Res* **5**, 00151–02018 (2019).

210. European Centre for Disease Prevention and Control (ECDC). *Review of reviews and guidelines on target groups , diagnosis , treatment and programmatic issues for implementation of latent tuberculosis management*. (ECDC, 2018). doi:10.2900/318458.

211. Comstock, G. W., Livesay, V. T. & Woolpert, S. F. The prognosis of a positive tuberculin reaction in childhood and adolescence. *Am J Epidemiol* **99**, 131–138 (1974).

212. Rosales-Klintz, S. *et al.* Guidance for programmatic management of latent tuberculosis infection in the European Union/European Economic Area. *The European respiratory journal* vol. 53 Preprint at https://doi.org/10.1183/13993003.02077-2018 (2019).

213. European Centre for Disease Prevention and Control (ECDC). *Tuberculosis surveillance and monitoring in Europe 2018 - 2016 data*. (2018) doi:10.2900/1940.

214. Fox, G. J., Dobler, C. C., Marais, B. J. & Denholm, J. T. Preventive therapy for latent tuberculosis infection—the promise and the challenges. *International Journal of Infectious Diseases* vol. 56 68–76 Preprint at https://doi.org/10.1016/j.ijid.2016.11.006 (2017).

215. Sandgren, A. *et al.* Initiation and completion rates for latent tuberculosis infection treatment: A systematic review. *BMC Infect Dis* **16**, (2016).

216. Menzies, D. *et al.* Four Months of Rifampin or Nine Months of Isoniazid for Latent Tuberculosis in Adults. *New England Journal of Medicine* **379**, 440–453 (2018).

217. Diallo, T. *et al.* Safety and Side Effects of Rifampin versus Isoniazid in Children. *New England Journal of Medicine* **379**, 454–463 (2018).

218. Ambrona de Marcos, V. *et al.* Cumplimiento del tratamiento de la infección tuberculosa latente en una cohorte de contactos de enfermos de tuberculosis. *Rev Esp Salud Publica* (2018).

219. Fiske, C. T. *et al.* Risk factors for treatment default in close contacts with latent tuberculous infection. *International Journal of Tuberculosis and Lung Disease* **18**, 421–427 (2014).

220. Goswami, N. D. *et al.* Geographic Information System-based Screening for TB, HIV, and Syphilis (GIS-THIS): A Cross-Sectional Study. *PLoS One* **7**, (2012).

221. Sun, W. *et al.* A spatial, social and environmental study of tuberculosis in China using statistical and GIS technology. *Int J Environ Res Public Health* **12**, 1425–1448 (2015).

222. Felgueiras, M., Cerqueira, S., Gaio, R., Felgueiras, Ó. & Duarte, R. Comparative Study Between Tuberculosis Incidence Rates in the Two Largest Metropolitan Areas of Portugal. *Archivos de Bronconeumologia* vol. 54 595–596 Preprint at https://doi.org/10.1016/j.arbres.2018.05.003 (2018).

223. Hirsch-Moverman, Y., Daftary, A., Franks, J. & Colson, P. W. Adherence to treatment for latent tuberculosis infection: Systematic review of studies in the US and Canada. *International Journal of Tuberculosis and Lung Disease* **12**, 1235–1254 (2008).

224. World Health Organization (WHO). *Latent tuberculosis infection : updated and consolidated guidelines for programmatic management*. *World Health Organization* vol. 25 (2018).

225. Priest, D. H., Vossel, L. F., Sherfy, E. A., Hoy, D. P. & Haley, C. A. Use of intermittent rifampin and pyrazinamide therapy for latent tuberculosis infection in a targeted tuberculin testing program. *Clinical Infectious Diseases* **39**, 1764–1771 (2004).

226. LoBue, P. A. & Moser, K. S. Use of isoniazid for latent tuberculosis infection in a public health clinic. *Am J Respir Crit Care Med* **168**, 443–447 (2003).

227. Hirsch-Moverman, Y., Bethel, J., Colson, P. W., Franks, J. & El-Sadr, W. Predictors of latent tuberculosis infection treatment completion in the United States: An inner city experience. *International Journal of Tuberculosis and Lung Disease* **14**, 1104–1111 (2010).

228. Stuurman, A. L. *et al.* Interventions for improving adherence to treatment for latent tuberculosis infection: A systematic review. *BMC Infect Dis* **16**, (2016).

229. Belknap, R. *et al.* Self-administered versus directly observed once-weekly isoniazid and rifapentine treatment of latent tuberculosis infection. *Ann Intern Med* **167**, 689–697 (2017).

230. Shu, C. C. *et al.* Predictors and prevalence of latent tuberculosis infection in patients receiving long-term hemodialysis and peritoneal dialysis. *PLoS One* **7**, (2012).

231. Shu, C. C. *et al.* Comparison of the prevalence of latent tuberculosis infection among non-dialysis patients with severe chronic kidney disease, patients receiving dialysis, and the dialysis-unit staff a cross-sectional study. *PLoS One* **10**, (2015).

232. Lee, M. R. *et al.* Diabetes Mellitus and Latent Tuberculosis Infection: A Systemic Review and Metaanalysis. *Clin Infect Dis* **64**, 719–727 (2017).

233. Fan, W. C. *et al.* Latent TB infection in newly diagnosed lung cancer patients - A multicenter prospective observational study. *Lung Cancer* **85**, 472–478 (2014).

234. Van Zyl Smit, R. N. *et al.* Global lung health: The colliding epidemics of tuberculosis, tobacco smoking, HIV and COPD. *European Respiratory Journal* vol. 35 27–33 Preprint at https://doi.org/10.1183/09031936.00072909 (2010).

235. European Centre for Disease Prevention and Control. *Programmatic management of latent tuberculosis infection in the European Union*. (2018).

236. Hirsch-Moverman, Y. *et al.* Latent tuberculous infection in the United States and Canada: Who completes treatment and why? *International Journal of Tuberculosis and Lung Disease* **19**, 31–38 (2015).

237. World Health Organisation. TB Elimination in Low-Incidence Countries. (2014).

238. Pooran, A. *et al.* Different screening strategies (single or dual) for the diagnosis of suspected latent tuberculosis: A cost effectiveness analysis. *BMC Pulm Med* **10**, 1–14 (2010).

239. Pai, M. & Menzies, D. Interferon-gamma release assays for diagnosis of latent tuberculosis infection. *UpToDate* (2016).

240. Rutledge, T. F. *et al.* Updated guidelines for using Interferon Gamma Release Assays to detect Mycobacterium tuberculosis infection - United States, 2010. *MMWR Recommendations and reports: Morbidity and mortality weekly report Recommendations and reports / Centers for Disease Control* **59**, 1–25 (2010).

241. Lange, C. *et al.* STATE OF THE ART Effectiveness and cost-effectiveness of screening migrants for active tuberculosis and latent tuberculous infection With the advent of the World Health Organization End. **21**, 965–976 (2017).

242. Diel, R. *et al.* Interferon-γ release assays for the diagnosis of latent Mycobacterium tuberculosis infection: A systematic review and meta-analysis. *European Respiratory Journal* **37**, 88–99 (2011).

243. Nienhaus, A., Schablon, A., Costa, J. & Diel, R. Systematic review of cost and cost-effectiveness of different TB-screening strategies. *BMC Health Serv Res* **11**, (2011).

244. Deuffic-Burban, S. *et al.* Cost-effectiveness of QuantiFERON-TB test vs. tuberculin skin test in the diagnosis of latent tuberculosis infection. *Int J Tuberc Lung Dis* **14**, 471–481 (2010).

245. Steffen, R. E. *et al.* Cost-Effectiveness of Quantiferon®-TB Gold-In-Tube Versus Tuberculin Skin Testing for Contact Screening and Treatment of Latent Tuberculosis Infection in Brazil. *PLoS One* **8**, 1–8 (2013).

246. Nijhawan, A. E., Iroh, P. A., Brown, L. S., Winetsky, D. & Porsa, E. Cost analysis of tuberculin skin test and the QuantiFERON-TB Gold In-tube test for tuberculosis screening in a correctional setting in Dallas, Texas, USA. *BMC Infect Dis* **16**, (2016).

247. Rosales-Klintz, S. *et al.* Cost-effectiveness analysis of programmatic screening strategies for latent tuberculosis infection in the EU/EEA. doi:10.2900/938014.

248. Rosales Klintz, S. *et al.* Mathematical modelling of programmatic screening strategies for latent tuberculosis infection in countries with low tuberculosis incidence. doi:10.2900/303841.

249. Won, D., Park, J. Y., Kim, H. S. & Park, Y. Comparative results of QuantiFERON-TB gold in-tube and QuantiFERON-TB gold plus assays for detection of tuberculosis infection in clinical samples. *J Clin Microbiol* **58**, (2020).

250. Pourakbari, B., Mamishi, S., Benvari, S. & Mahmoudi, S. Comparison of the QuantiFERON-TB Gold Plus and QuantiFERON-TB Gold In-Tube interferon-γ release assays: A systematic review and meta-analysis. *Adv Med Sci* **64**, 437–443 (2019).

251. Departamento de Saúde Pública do Norte. A tuberculose na região de saúde do norte 2015. *Administração Regional de Saúde do Norte; Ministério da Saúde* 29.

252. Loureiro, R. B. *et al.* Cost-effectiveness of QuantiFERON-TB Gold In-Tube versus tuberculin skin test for diagnosis and treatment of Latent Tuberculosis Infection in primary health care workers in Brazil. *PLoS One* **14**, (2019).

253. Apolinário, D. *et al.* Tuberculosis inequalities & socio-economic deprivation in Portugal. *International Journal of Tuberculosis and Lung Disease* **21**, 784–789 (2017).

254. Blas, E. & Kurup, A. S. *Equity, social determinants and public health programmes*. (World Health Organization, 2010).

255. Matteelli, A. *et al.* Tuberculosis elimination and the challenge of latent tuberculosis. *Presse Med* **46**, e13–e21 (2017).

256. Porta, M. & Last, J. M. A Dictionary of Public Health. **1**, (2018).

257. Direcção-Geral da Saúde (DGS). *Despacho n.º 12513-B/2019: Doenças de notificação, clínica e laboratorial obrigatória.* (2019).

258. Correia, A. M. & Duarte, R. *Planeamento do Rastreio de Contactos de Doentes com Tuberculose*. (Departamento de Saúde Pública da Administração Regional de Saúde do Norte, I.P., 2013).

259. Heymann, D. L. *Control of communicable diseases manual*. (American Public Health Association, 2015).

260. Mack, U. *et al.* LTBI: Latent tuberculosis infection or lasting immune responses to M. tuberculosis? A TBNET consensus statement. *European Respiratory Journal* **33**, 956–973 (2009).

261. Santos, J. A., Duarte, R. & Nunes, C. Host factors associated to false negative and indeterminate results in an interferon-γ release assay in patients with active tuberculosis. *Pulmonology* **26**, 353–362 (2020).

262. European Centre for Disease Prevention and Control (ECDC) & World Health Organization Regional Office for Europe. *Tuberculosis surveillance and monitoring in Europe 2021: 2019 data*. (WHO Regional Office for Europe, 2021).

263. Littleton, J. & Park, J. Tuberculosis and syndemics: Implications for Pacific health in New Zealand. *Soc Sci Med* **69**, 1674–1680 (2009).

264. Statistics Portugal. Statistics Portugal - Web Portal. https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_main.

265. Joint United Nations Programme on HIV/AIDS (UNAIDS). *UNAIDS data 2019*. (UNAIDS, 2019).

266. Kwak, N., Hwang, S. S. & Yima, A. J. Effect of COVID-19 on Tuberculosis Notification, South Korea. *Emerg Infect Dis* **26**, 2506–2508 (2020).

267. Cilloni, L. *et al.* The potential impact of the COVID-19 pandemic on the tuberculosis epidemic a modelling analysis. *EClinicalMedicine* **28**, (2020).

268. World Health Organization (WHO). *Global Tuberculosis Report 2020*. (2020).

269. Kyu, H. H. *et al.* The global burden of tuberculosis: results from the Global Burden of Disease Study 2015. *Lancet Infect Dis* **18**, 261–284 (2018).

270. Cui, Y. *et al.* A long-term trend study of tuberculosis incidence in china, india and united states 1992–2017: A joinpoint and age–period–cohort analysis. *Int J Environ Res Public Health* **17**, (2020).

271. Zenner, D. *et al.* Active case finding for tuberculosis among high-risk groups in low-incidence countries. *International Journal of Tuberculosis and Lung Disease* **17**, 573–582 (2013).

272. Ong, C. W. M. *et al.* Epidemic and pandemic viral infections: Impact on tuberculosis and the lung. *European Respiratory Journal* **56**, (2020).

273. Visca, D. *et al.* Tuberculosis and COVID-19 interaction: A review of biological, clinical and public health effects. *Pulmonology* **27**, 151–165 (2021).

274. Migliori, G. B., Tiberi, S., García-Basteiro, A. L. & Duarte, R. Tuberculosis and its future in the COVID-19 era: The Pulmonology series 2021. *Pulmonology* **27**, 94–96 (2021).

275. Duarte, R. *et al.* Different disease, same challenges: Social determinants of tuberculosis and COVID-19. *Pulmonology* **27**, 338–344 (2021).

276. Motta, I. *et al.* Tuberculosis, COVID-19 and migrants: Preliminary analysis of deaths occurring in 69 patients from two cohorts. *Pulmonology* **26**, 233–240 (2020).

277. Migliori, G. B. *et al.* Worldwide Effects of Coronavirus Disease Pandemic on Tuberculosis Services, January–April 2020. *Emerg Infect Dis* **26**, 2709–2712 (2020).

278. Millet, J. P. & Orcau, À. COVID y TB en Barcelona. *Enf Emerg* **20**, 27–45 (2021).

279. Comella-del-Barrio, P., De Souza-Galvão, M. L., Prat-Aymerich, C. & Domínguez, J. Impact of COVID-19 on Tuberculosis Control. *Arch Bronconeumol* **57**, 5–6 (2021).

280. Ospina, J. E. *et al.* Community health workers improve contact tracing among immigrants with tuberculosis in Barcelona. *BMC Public Health* **12**, (2012).

281. Ladeira, I. *et al.* Portugal. Clinical management of M/XDR-TB in the Northern Regional Reference Centre. *Compendium of good practices in the implementation of the Tuberculosis Action Plan for the WHO European Region 2016–2020* (2019).

282. Miyahara, R. *et al.* Risk factors associated with large clusters of tuberculosis patients determined by whole-genome sequencing in a high-tuberculosis-burden country. *Tuberculosis* **125**, (2020).

283. Smith, C. M., Maguire, H., Anderson, C., Macdonald, N. & Hayward, A. C. Multiple large clusters of tuberculosis in London: a cross-sectional analysis of molecular and spatial data. *ERJ Open Res* **3**, (2017).

284. Carvalho Sousa, S. *et al.* Tuberculosis: where and how fast are stone quarry workers infected? *Eur J Public Health* **30**, (2020).

285. WHO. WHO operational handbook on tuberculosis. Module 2: screening - systematic screening for tuberculosis disease. *Module 5: management of tuberculosis in children and adolescents* 116 (2021).

286. Guthrie, J. L. *et al.* Comparison of routine field epidemiology and whole genome sequencing to identify tuberculosis transmission in a remote setting. *Epidemiol Infect* **148**, e15 (2020).

287. Tagliani, E. *et al.* Use of a whole genome sequencing-based approach for Mycobacterium tuberculosis surveillance in Europe in 2017–2019: an ECDC pilot study. *European Respiratory Journal* **57**, (2021).

288. Mathema, B. *et al.* Drivers of Tuberculosis Transmission. *J Infect Dis* **216**, S644–S653 (2017).

289. Santos, J. A., Leite, A., Soares, P., Duarte, R. & Nunes, C. Delayed diagnosis of active pulmonary tuberculosis - potential risk factors for patient and healthcare delays in Portugal. *BMC Public Health* **21**, (2021).

290. Jiang, Y. *et al.* Duration and Determinants of Delayed Diagnosis with Tuberculosis in Shenzhen, China: A Cross-Sectional Study. *Risk Manag Healthc Policy* **15**, 1473–1481 (2022).

291. Almeida Santos, J., Soares, P., Leite, A., Duarte, R. & Nunes, C. Patient and healthcare delays in critical and non-critical pulmonary tuberculosis incidence areas in Portugal: are there differences? *Public Health* **201**, 41–47 (2021).

292. European Centre for Disease Prevention and Control (ECDC) & World Health Organization (WHO) Regional Office for Europe. *Tuberculosis surveillance and monitoring in Europe 2022: 2020 data*. (2022). doi:10.2900/488579.

293. Swaan, C. M., Wong, A., Marinović, A. B., Kretzschmar, M. E. E. & Van Steenbergen, J. E. Timeliness of infectious disease reporting, the Netherlands, 2003 to 2017: law change reduced reporting delay, disease identification delay is next. *Eurosurveillance* **24**, (2019).

294. Bello, S. *et al.* Empirical evidence of delays in diagnosis and treatment of pulmonary tuberculosis: systematic review and meta-regression analysis. *BMC Public Health* **19**, (2019).

295. World Health Organization (WHO). *World health statistics 2022: monitoring health for the SDGs, sustainable development goals*. (WHO, 2022).

296. Matteelli, A. *et al.* Tuberculosis elimination: where are we now? *European Respiratory Review* **27**, (2018).

297. Durrheim, D. N., Crowcroft, N. S. & Strebel, P. M. Measles – The epidemiology of elimination. *Vaccine* **32**, 6880–6883 (2014).

298. World Health Organization (WHO). *Setting priorities in communicable disease surveillance*. (2006).

299. Chao, A. An overview of closed capture-recapture models. *Journal of Agricultural, Biological, and Environmental Statistics* Preprint at https://doi.org/10.1198/108571101750524670 (2001).

300. Mccarty, D. J., Tull, E. S., Moy, C. S., Kwoh, C. K. & Laporte, R. E. Ascertainment corrected rates: Applications of capture-recapture methods. *Int J Epidemiol* (1993) doi:10.1093/ije/22.3.559.

301. Van Hest, R. Use of capture-recapture analysis in data sources. *Rev Saúde Pública* **42**, 377–9 (2008).

302. Hook, E. B. & Regal, R. R. Capture-recapture methods. *Lancet* **339**, 742 (1992).

303. Wittes, J. T., Colton, T. & Sidel, V. W. Capture-recapture methods for assessing the completeness of case ascertainment when using multiple information sources. *J Chronic Dis* **27**, 25–36 (1974).

304. Carvalho, C., Rodrigues, A. L., Gomes, B. & Gomes, M. *A tuberculose na região de saúde do Norte: dados definitivos 2017*. (2019).

305. European Centre for Disease Prevention and Control (ECDC). *Assessment of electronic health records for infectious disease surveillance: final mapping exercise report*. (2021).

306. Giannelos, N. EHR, the core component of eHealth | Joinup. https://joinup.ec.europa.eu/collection/ehealth/document/ehr-core-component-ehealth (2010).