

# Model-based impact evaluation of new tuberculosis vaccines in aging populations under different modeling scenarios: the case of China.

Mario Tovar,<sup>1,2</sup> Joaquín Sanz,<sup>1,2</sup> and Yamir Moreno<sup>1,2,3</sup>

<sup>1</sup>*Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Zaragoza 50009, Spain*

<sup>2</sup>*Department of Theoretical Physics, University of Zaragoza, Zaragoza 50009, Spain*

<sup>3</sup>*Centai Institute S.p.A, 10138 Torino, Italy*

(Dated: September 27, 2023)

The slow descent in TB burden, the COVID-19 pandemic, along with the rise of multidrug-resistant strains of *Mycobacterium tuberculosis*, seriously threaten TB control and the goals of the End TB strategy. To fight back, several vaccine candidates are under development, with some of them undergoing the phases 2B and 3 of the development pipeline. The impact of these vaccines on the general population needs to be addressed using disease-transmission models, and, in a country like China, which last year ranked third in number of cases worldwide, and where the population is undergoing a fast process of demographic aging, the impact of TB vaccination campaigns may depend heavily upon the age of targeted populations and with the mechanistic descriptions of the TB vaccines. For these reasons, transmission models need to capture the coupling between TB dynamics and demographic evolution, as well as to be able to accommodate different mechanistic descriptions of TB vaccine protection. In this work, we studied the potential impact of a new TB vaccine in China targeting adolescents (15-19 y.o.) or elderly people (60-64 y.o.), according to varying vaccine descriptions that represent reasonable mechanisms of action leading to prevention of disease (PoD), or prevention of recurrence (PoR), each of them targeting specific routes to TB disease. To measure the influence of the description of the coupling between transmission dynamics and aging in TB transmission models, we explored two different approaches to compute the evolution of the contact matrices, which relate to the spreading among different age strata. Our results show that the magnitude of model-based impact estimates substantially depends upon the vaccine profile, and it is also strongly related to the modeling approach chosen to describe the time evolution of contact matrices. In spite of these sources of uncertainty, our results also show, in line with previous modeling works, that elder vaccination is a suitable option in China to reduce the incidence of TB.

## I. INTRODUCTION

Tuberculosis (TB) is an infectious disease caused by the bacterium *Mycobacterium tuberculosis* (*M.tb.*) that usually affects the lungs. It is a complex but preventable disease with a high global burden that requires early detection and long treatments, and that is still among the leading causes of death for a single pathogen worldwide. In 2014, the World Health Assembly introduced the World Health Organization (WHO) "Global Strategy and Targets for tuberculosis prevention, care, and Control after 2015", labeled as the End TB Strategy, which consists of completing a reduction of TB incidence and mortality rates by 90% and 95%, between 2015 and 2035 [1].

Arguably, the measures and interventions currently under use for TB control are effective, but also insufficient to meet the goal of the End TB strategy. Thus, although a decay in TB incidence and mortality has been achieved worldwide since 1990 [2], its yearly rate of reduction is still too slow. Furthermore, during 2020 and 2021, and for the first time in decades, the world witnessed a surge in global TB burden levels with respect to previous years due to the COVID-19 emergency that led to underdiagnosis and under-treatment of TB, along with the saturation of most healthcare systems [3–6]. During those years alone, the WHO estimated that TB was the cause of death of more than 1.5 and 1.6 million people respectively, worldwide, combining HIV-negative and positive cases [3, 7].

The recent increase in TB burden observed in 2020-2022 due to the irruption of the COVID-19 pandemic, threatens, in high-burden countries like India or Indonesia, to raise the TB-related mortality back to even higher levels than before in the next few years [8, 9]. Moreover, the ever-increasing rates of emergence of drug resistance [10] evidences the necessity of new tools against the disease, including new and better drugs as well as improved diagnosis methods. Among these new resources, the development of a new vaccine that either boosts or replaces the current bacillus Calmette-Guerin (BCG) is commonly referred to as the potentially most impactful single intervention to halt TB transmission, given the limited and variable efficacy levels observed for BCG against the more transmissible respiratory forms of the disease in young adults [11]. Consequently, the TB vaccine development pipeline is populated by a number of novel candidates of different types, based on a variety of immunological principles and vaccine platforms [12]. For estimating and comparing the potential impact of each of these candidates on halting the TB transmission chain, the development of epidemiological models arises as a powerful tool. Refining these models and addressing the main sources of uncertainty and bias in their architecture constitutes an important step toward the development of new TB vaccines.

In this work, we forecast the impact of the introduction of a new TB vaccine in China, which, as of 2021, represented 7.4% of the total number of TB cases worldwide [3]. To produce robust estimates for vaccine impact in this country, it is important to consider aging as a key demographic determinant where China differs from the majority of high-burden countries in TB. According to the UN population division estimates, among the top-8 countries with the highest number of incident TB cases in 2021 (China, India, Indonesia, Philippines, Pakistan, Nigeria, Bangladesh, and the Democratic Republic of Congo), China is the one where population aging in the years to come will be more pronounced, going from a median age of 39.0 in 2023 to 50.7 years in 2050. Considering this, recent modeling studies have suggested that targeting elder population groups in a vaccination campaign may produce a greater impact than targeting children or young adults [13].

Estimating the impact of a vaccine on an aging population must be done considering several technical aspects. From a modeling perspective, demographic aging couples with TB transmission dynamics critically in the age-mixing contact matrices [14]. These objects capture the relative frequency of social and/or physical contact between individuals of different age strata and constitute a great tool for representing contact patterns within epidemic models. When working with TB, whose time scales are comparable to those of demographic evolution, models need to incorporate sensible heuristics to describe the evolution of those matrices over time, which appears as a consequence of demographic evolution itself. In previous works [15], we identified different methods that can be used to adapt the contact matrices that were measured in a given population, as the population ages over time. Two of these methods are often found in modeling studies of TB [13, 14, 16]. First, contact matrices can be adapted to ensure that the symmetry of the encoded information is preserved, namely, the number of contacts per unit of time between two age groups,  $i$  and  $j$ , should be the same, when calculated from the number of contacts per capita, from  $i$  to  $j$  and from  $j$  to  $i$ . This method is commonly referred to as the pairwise correction method and ensures symmetry in reported numbers of contacts but does not produce contact networks whose mean connectivity is stable over time. To solve this issue, these matrices can be further corrected to ensure that not only the symmetry but also the average connectivity of the networks is preserved across time while underlying populations are aging. This second method, firstly proposed in [15], is referred to in this study as the intrinsic connectivity method, and produces contact structures that feature controlled contact densities on average, which are stable over time. The adoption of each of these methods to describe the evolution of contacts within TB transmission models may affect model outcomes regarding the impact of a new vaccine.

Another relevant aspect through which populations' aging and vaccine impact forecasts may couple in ways that are hard to predict a priori is the vaccine mechanisms of action, combined with its protection profiles. A successful TB vaccine may confer either POD or POR through a variety of mechanisms, including halting fast progression towards primary TB upon a recent, first infection event, diminishing the endogenous reactivation rates upon latent infection, diminishing the risk of reinfection and/or preventing recurrence after recovery of a previous disease episode. Furthermore, the ability of a vaccine to confer protection through each of these mechanisms may depend upon the previous exposure of vaccine recipients to *Mycobacterium tuberculosis*, ascertained by an interferon-gamma release assay (IGRA). Since the fraction of individuals who have been previously exposed to the pathogen varies across age, and, in an aging population, that changes across time, exploring the effects of vaccine protection profiles and mechanisms of action at once is crucial to compare different vaccination strategies targeting different age groups in China.

The goal of this study is to investigate the role of these aspects on the vaccine impact foreseen from two vaccination campaigns targeting adolescents (15-19 y.o) and elder individuals (60-64 y.o.) in China, starting in 2025. Capitalizing on previous TB transmission models developed by the authors [14, 16], we produced in-silico impact evaluations of a series of vaccines with different protection profiles, acting through different mechanisms, evaluated in different simulations where contact patterns evolve according to different methods. In what concerns the distribution of vaccine-mediated protection across vaccinated individuals, we modeled vaccines as all-or-nothing (AON), as schematized in Figure 1, as it is widely used in the modeling literature [17–19]. AON vaccines confer perfect protection to a fraction of the vaccinated individuals but are ineffective for the remaining fraction of vaccinated individuals. The share between those fractions is related to the overall efficacy of the vaccine. To estimate the impact, we make use of two runs of the mathematical model (control and vaccine), and we analyze the results in terms of the targeted population, the protection profile of the vaccine, and the method used for updating the contact matrix.

## II. MATERIALS AND METHODS

### A. Modeling the effect of the vaccine

The efficacy estimates obtained for a given vaccine in a clinical trial can be mapped onto mechanistic descriptions within a transmission model in a number of different ways. Elucidating what are the specific mechanisms at place that are most likely compatible with trials' results for a given vaccine is not a trivial task, and the architecture of the

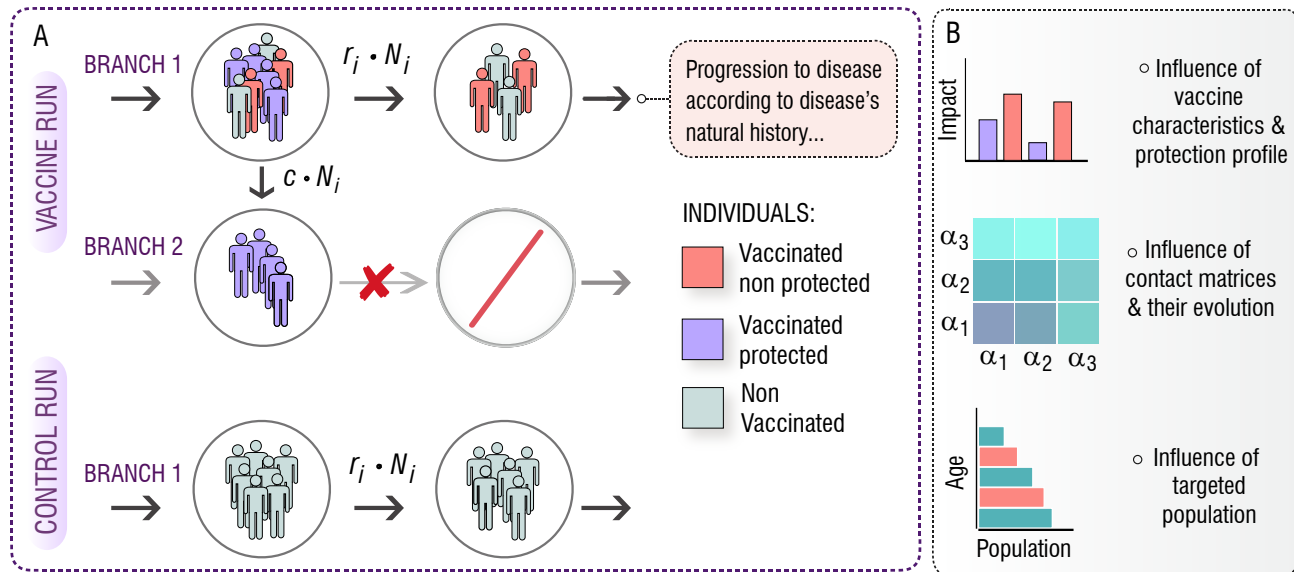


FIG. 1. **A.** In-silico simulations for the introduction of an AON vaccine in a mathematical model of TB transmission. The model is run twice, one as a control run and the second time with the introduction of the vaccine. In the **control run** (bottom) the natural history of the disease remains unaltered, and every transition between states of the model is present. In the **vaccine run**, there are two different branches that evolve in parallel. On **branch 1**, the natural history of the disease remains unaltered, but a  $c$  fraction of the individuals that receive the vaccine is moved towards **branch 2**, in which the natural history is modified according to the effect of the vaccine, thus, neglecting some transitions and conferring protection. This way, a fraction,  $c$ , of the vaccinated individuals becomes protected. **B.** Implementing this general approach for different vaccine characteristics and protection profiles, under different descriptions of contact matrices evolution, we aim to study the effect of these aspects on the impact foreseen by our computational model for a vaccine applied either in adolescents (15-19 years old), or in older individuals (60-64 years old).

110 model, as well as the characteristics of the population enrolled and detailed data analysis of the results are needed  
 111 for extrapolating the efficacy levels observed in a trial into transmission models describing TB dynamics in entire  
 112 populations[16, 20]. In a disease such as TB, a vaccine conferring POD or POR may base its protective effects on  
 113 interfering with different processes throughout the natural history of the disease, preventing individuals' progression to  
 114 TB by halting specific routes to disease. In the lack of direct evidence concerning the specific dynamic mechanisms in  
 115 place in a given vaccine, modelers often implement vaccine descriptions where all the main routes to disease putatively  
 116 affected by the vaccine are equally impacted. Instead, in this work, we aim to compare the impact of different vaccines  
 117 whose protection acts through different dynamical mechanisms. Formally, we assume that a vaccine can reduce the  
 118 risk of progressing further from a given state toward disease, thus conferring protection to vaccinated individuals at  
 119 different stages in the natural history of the disease. Capitalizing on our model, we identify four different basic vaccine  
 120 mechanisms that can act either alone or combined:

- 121 •  $E_p$ : Protection against primary TB: The vaccine confers protection against fast progression to disease upon  
 122 a recent first infection event. This mechanism is present in a POD vaccine that prevents fast latency toward  
 123 active disease (see supplementary appendix).
- 124 •  $E_{r_l}$ : Protection against endogenous reactivation of LTBI: Vaccine confers protection against endogenous reactivation  
 125 of bacilli in individuals with latent TB infection (LTBI). This mechanism is present in a POD vaccine  
 126 that prevents slow latency towards active disease.
- 127 •  $E_q$ : Protection against TB upon reinfection: Vaccine confers protection against exogenous reactivation caused  
 128 by a secondary infection event in subjects who had been previously infected. This mechanism is present in a  
 129 POD vaccine that prevents progression towards active disease upon reinfection, for individuals who had already  
 130 been exposed to the pathogen before (either LTBI or recovered individuals).
- 131 •  $E_{relapse}$ : Protection against TB relapse: This mechanism is present in a POR vaccine that prevents endogenous  
 132 reactivation in individuals who had a past episode of active TB.

- All: Every vaccine's mechanism acts at the same time.

The interaction between vaccines conferring protection at any of the previous mechanisms, and the natural history of the disease is sketched in Supplementary Figure 3. In short, the fraction of the vaccinated individuals who get protection from the vaccine will face a modified version of the natural history according to the effect of the vaccine, where some key transitions are halted.

## B. Model-based impact evaluations of TB vaccines

We estimate the impacts of vaccines using an adapted version of the model in [14], which is a deterministic, age-structured model based on ordinary differential equations, where individuals belonging to different age strata are considered to experiment different levels of epidemiological risk. This translates, in general, into age-specific parameter values, as described in the supplementary appendix. The architecture that defines the disease dynamics within each age group represents the natural history of TB (see Supplementary Figure 1). The model also includes aging dynamics, which is key in countries undergoing fast demographic changes (see Supplementary Appendix, and Supplementary Figure 2).

Regarding the impact evaluation, we use two different runs of the model that ultimately lead to an estimate of the incidence rate reduction due to the vaccine. In the first run, specific values of all the epidemiological parameters are stochastically drawn from suitable distributions. Using the specific set of parameters obtained, the model is calibrated and the spreading of the disease is forecasted in a non-intervention scenario, referred here as the control run. Then, the model is run again, using the same calibration, but introducing the vaccine in 2025. This vaccine run does not follow qualitatively the same Natural History as in the control run, as the vaccine alters it by reducing the progression risk of protected individuals in certain transitions that depend on the characteristics of the vaccine (see supplementary appendix). Finally, the impact of the vaccine is estimated by comparing those two runs through the obtention of the incidence rate reduction (IRR) at the end of 2050, as follows:

$$\text{IRR}(t = 2050)(\%) = \frac{i_{\text{control run}}(t = 2050) - i_{\text{vaccine run}}(2050)}{i_{\text{control run}}(t = 2050)} \cdot 100 \quad (1)$$

Repeating this procedure a number of  $N = 500$  times, we obtain a distribution of forecasted vaccine impacts, which allows us to build suitable expected values and confidence intervals that propagate uncertainty from model inputs to vaccine impacts.

As already mentioned, all vaccines are introduced into the model according to an all-or-nothing scheme, which means they only show efficacy in a fraction  $c$  of the vaccinated individuals, which in this context represents the vaccine efficacy under a scenario of perfect coverage. The remaining  $1 - c$  fraction of vaccinated individuals do not benefit from these effects and preserve the same dynamics as the unvaccinated individuals. Formally, this is modeled by displacing a fraction  $c$  of vaccinated individuals from a control branch to a vaccine-protection branch, where the dynamics is modified to reflect these changes.

The vaccines considered in this work feature different levels of waning. As vaccinated individuals age from their age at vaccination  $a_v$  to  $a > a_v$ , the vaccine efficacy is expected to decay, eventually becoming inefficient  $w$  years after vaccination. To implement vaccine waning in an all-or-nothing model, we introduced in the model a series of return fluxes that move individuals in the age group  $a > a_v$  back from the protected to the non-protected branch of the model. The intensity of those fluxes is given by Equation 2.

$$w_i(a, a_v) = 1 - e^{-\ln(2) \cdot \frac{5(a-a_v)}{w}} \quad (2)$$

where  $a$  is the age group that suffers the waning,  $a_v$  is the age group that is being vaccinated, and  $w$  captures the waning, in years. This formula ensures that, after  $w$  years, the vaccinated individuals will suffer a waning intensity of 50%. Moreover, for any  $a \leq a_v$ , the waning intensity is set to zero, constituting a viable approach to implement vaccination campaigns targeting one specific age stratum. Then, the waning flux is calculated as:

$$W_f(a, a_v, t) = w_i(a, a_v) \cdot X(a, t) \quad (3)$$

where  $X(a)$  is the population in the age group  $a$ , in the reservoir  $X$  at time  $t$ , where  $X$  is every reservoir that vaccinated individuals may lie into.

Besides the alterations in epidemiological risks experienced by the individuals protected by the vaccine, another important vaccine characteristic concerns the immunological status that individuals may need to fulfill before vaccination

180 for the vaccine to confer protection to them. Depending on the immunological principles of the vaccine considered,  
181 its protection may unfold contingent on the previous exposure of the vaccinated subject to the pathogen, and, as  
182 such, it is possible that new vaccines may confer protection only to susceptible, immunologically naive individuals,  
183 or to individuals who have been previously infected with *M.tuberculosis*. To illustrate the effect of the dependency  
184 between the recipient’s status and vaccine efficacy, we reproduce vaccine impact simulations where the protective  
185 effect of the vaccine, described through the displacement of a fraction  $c$  of individuals towards the protected branch,  
186 only takes place from different source reservoirs. This way, we explore four possible scenarios where only a fraction  $c$   
187 of individuals in the following reservoirs get protection: i) only susceptible individuals ii) only LTBI individuals iii)  
188 only previously exposed individuals (including LTBI and recovered subjects), and iv) all individuals.

189 Finally, the vaccination is implemented in two steps. First, a mass vaccination campaign, similar to the one proposed  
190 in [13] takes place, vaccinating annually a third of the population in the reservoirs affected by the vaccine, and for  
191 the age targeted population (15-19 or 60-64). Then, after this campaign vanishes, the vaccination continues routinely  
192 coupled to the aging.

### 193 C. Updating contact matrices with evolving demography

194 Age-mixing contact matrices play a key role in epidemic spreading [21–23], as the complete knowledge of the network  
195 of contacts is usually unreachable or impossible to implement. Thus, for modeling purposes, it is useful to study age-  
196 group interactions, where contact matrices indicate how age-strata mix between them. Usually, empirical contact  
197 matrices are obtained through statistical surveys. In these studies, participants are asked how many contacts they  
198 have during the day and with whom. This allows us to obtain the (average) number of contacts that an individual  
199 of a particular age  $i$  has with individuals of age group  $j$ . The resulting matrix is not symmetric due to the different  
200 number of individuals in each age group. However, it is precisely the demographic structure that imposes constraints  
201 in the entries of this matrix, as reciprocity of contacts should be fulfilled at any time (i.e., the total number of contacts  
202 reported by age-group  $i$  with age-group  $j$  should be equal in the opposite direction). Therefore, an empirical contact  
203 matrix, that has been measured on a specific population, should not be used directly without adapting it to the  
204 demographic structure of a different population under study.

205 This issue has important consequences in the field of disease modeling. As contact matrices play a key role in  
206 disease forecast, it is essential to assure that the matrices implemented are adapted to the demographic structure of  
207 the population considered to avoid biased estimations. For some short-cycle diseases like influenza, the time scale of  
208 the epidemic is much shorter than the typical times needed for a demographic structure to evolve [24]. The previous  
209 considerations are more troublesome for the case of persistent diseases that need long-term simulations, for which the  
210 hypothesis of constant demographic structures does not hold anymore [14]. Particularly, in the case of TB modeling,  
211 time scales are typically long, as the presence of latent individuals may lead to TB cases decades after primary  
212 infection [25]. This ultimately leads to the urge to adapt the contact matrices measured in a specific demography in  
213 such a way that they evolve accordingly to the demography of that setting. To this end, we capitalize on the methods  
214 proposed by Arregui et al. [15], which are briefly described below. We selected only methods labelled in the original  
215 article as M1 (Pairwise corrections) and M3 (Intrinsic connectivity) as the first one is typically used in the literature,  
216 also for modeling TB e.g. [13]), and is the simplest one for short-lives diseases, whereas the second one (M3) allows  
217 projecting contact matrices along with demography, which fits our needs in TB forecasting.

#### 218 1. Pairwise correction

219 The magnitude usually reported when measuring contact patterns is the mean number of contacts that an individual  
220 in age group  $i$  has with individuals in age group  $j$  during a measured period of time. Calling  $M_{i,j}$  this quantity, we  
221 observe that, in order to fulfill reciprocity,  $M_{i,j}$  should equal  $M_{j,i}$ , which is not the case with directly measured data.  
222 An immediate correction is to average those numbers, so that the excess of contacts measured in one direction is  
223 transferred to the reciprocal. Then, the matrix entry in a new demography is computed as:

$$224 M'_{i,j} = \frac{1}{N'_i} \frac{1}{2} (M_{i,j} N'_i + M_{j,i} N'_j) \quad (4)$$

225 where  $M'_{i,j}$  corresponds to the new demography under study. An example of the evolution of the contact matrix used  
226 in this study, under the pairwise correction is included in Supplementary Figure 4, panels A and B.

227

## 2. Intrinsic connectivity matrix

228 An alternative method that preserves the mean connectivity of the contact network makes use of the density matrix  
229 or intrinsic connectivity matrix. Using the original data the density matrix  $\Gamma$  is extracted as:

$$230 \quad \Gamma_{i,j} = M_{i,j} \frac{N}{N_j}. \quad (5)$$

231 The  $\Gamma$  matrix corresponds, except for a global factor, to the contact pattern in a “rectangular” demography (a  
232 population structure where all age groups have the same density). Then, introducing a new demography, the contact  
233 matrix is obtained as:

$$234 \quad M'_{i,j} = \frac{\Gamma_{i,j} N'_j N'}{\sum_{i,j} \Gamma_{i,j} N'_i N'_j}. \quad (6)$$

235 An example of the evolution of the contact matrix used in this work when using the intrinsic connectivity method is  
236 included in Supplementary Figure 4, panels C and D.

237

## III. RESULTS

238 After we implemented our computational model of TB transmission, we evaluated the impact of a TB vaccine of  
239 varying characteristics, simulated using models where contact matrices are updated using different methods, intro-  
240 duced in China in either adolescent (15-19 years old), or elder individuals (60-64 years old). In Figure 2 and Figure  
241 3 we represent the forecast impact of each vaccine using our model. As discussed in the methods section, all vaccines  
242 are modeled according to an all-or-nothing scheme, conferring different types of protection ( $E_p$ ,  $E_q$ ,  $E_{r_l}$ ,  $E_{relapse}$ ,  
243 or all at once, see Methods) to a fraction  $c$  of vaccinated individuals in certain disease reservoirs (susceptible, latent  
244 and/or recovered individuals, or all). The efficacy of the vaccine in all scenarios is set to  $c = 56\%$ , as a reference value  
245 compatible with applying a highly protective vaccine with a 70% efficacy through a high-coverage campaign reaching  
246 80% of the target population, similar to one among the most optimistic scenarios explored in previous modeling stud-  
247 ies undertaken in China[13]. The vaccination campaign in the simulation starts in 2025, and we forecast the impact  
248 of the vaccine measuring the IRR (see methods) in 2050. Individuals of the targeted age group, are vaccinated when  
249 they first enter the corresponding age group. Furthermore, selected vaccines experience waning levels of 10 years, as  
250 described in more detail in the Methods section.

251 In Figure 2, we gather the IRRs achieved by the different vaccines described above when applied to the elderly  
252 population. In this case, we found that, when vaccines are able to confer protection to immunologically naive  
253 individuals, either alone (Susceptible only) or along with the rest of individuals in the population, (whole population),  
254 vaccines featuring the largest impact are those that are able to prevent fast progression to primary TB upon recent  
255 infection ( $E_p$  vaccine featuring 24.9% IRR, 95% CI [17.8-36.1]) in panel A). Instead, if vaccine protection only unfolds  
256 on individuals who had been previously infected by the time of vaccination (protection active to either latent or  
257 latent plus recovered individuals only), the most impactful vaccines are those that are able to protect individuals  
258 from developing active TB upon endogenous reactivation of dormant bacilli ( $E_{r_l}$  vaccine featuring 6.47% IRR, 95%  
259 CI[4.88-8.15] in panel C). Furthermore, it is important to notice that, in all vaccines tested leading to an impact  
260 higher than 1% under at least one of the two methods explored for describing contact matrices evolution, the impact  
261 was systematically higher when updating the contact matrix according to the intrinsic connectivity approach.

262 In turn, in Figure 3, we present the analogous results associated with a vaccination campaign targeting the popula-  
263 tion between 15 and 19 years old. Although vaccines protecting against primary TB are still more impactful as long  
264 as susceptible individuals are protected ( $E_p$  vaccine featuring 11.11% IRR, 95% CI [8.10-13.98]) in panel A, vs. other  
265 vaccines), and vaccines halting endogenous reactivation of latent bacilli are more impactful if protection takes place  
266 after infection, in this case, the impact associated with vaccines in the latter case is comparatively lower than what is  
267 found in elders ( $E_{r_l}$  vaccine in panel C yields 2.14 IRR, 95% CI [1.65 2.76], when in elders yielded 6.47% IRR, 95%  
268 CI[4.88-8.15]). In what concerns the influence of contact matrices on forecast impacts, interestingly, we observe that  
269 the highest impacts were associated with the pairwise-corrections method, unlike what is observed in the elderly age  
270 group.

271 Admittedly, comparing the impacts from both campaigns targeting elders (Figure 2) and adolescents (Figure 3), we  
272 see that the question of what is the optimal age group to target in an immunization campaign for a new TB vaccine  
273 finds different answers depending on the combination of vaccine characteristics, protection profiles, and modeling  
274 assumptions. More specifically, in cases where previous infection is required for vaccines to elicit their protective effects,  
275 via protection against endogenous reactivation ( $E_{r_l}$ ) or against relapse ( $E_{relapse}$ ), then targeting the older age group

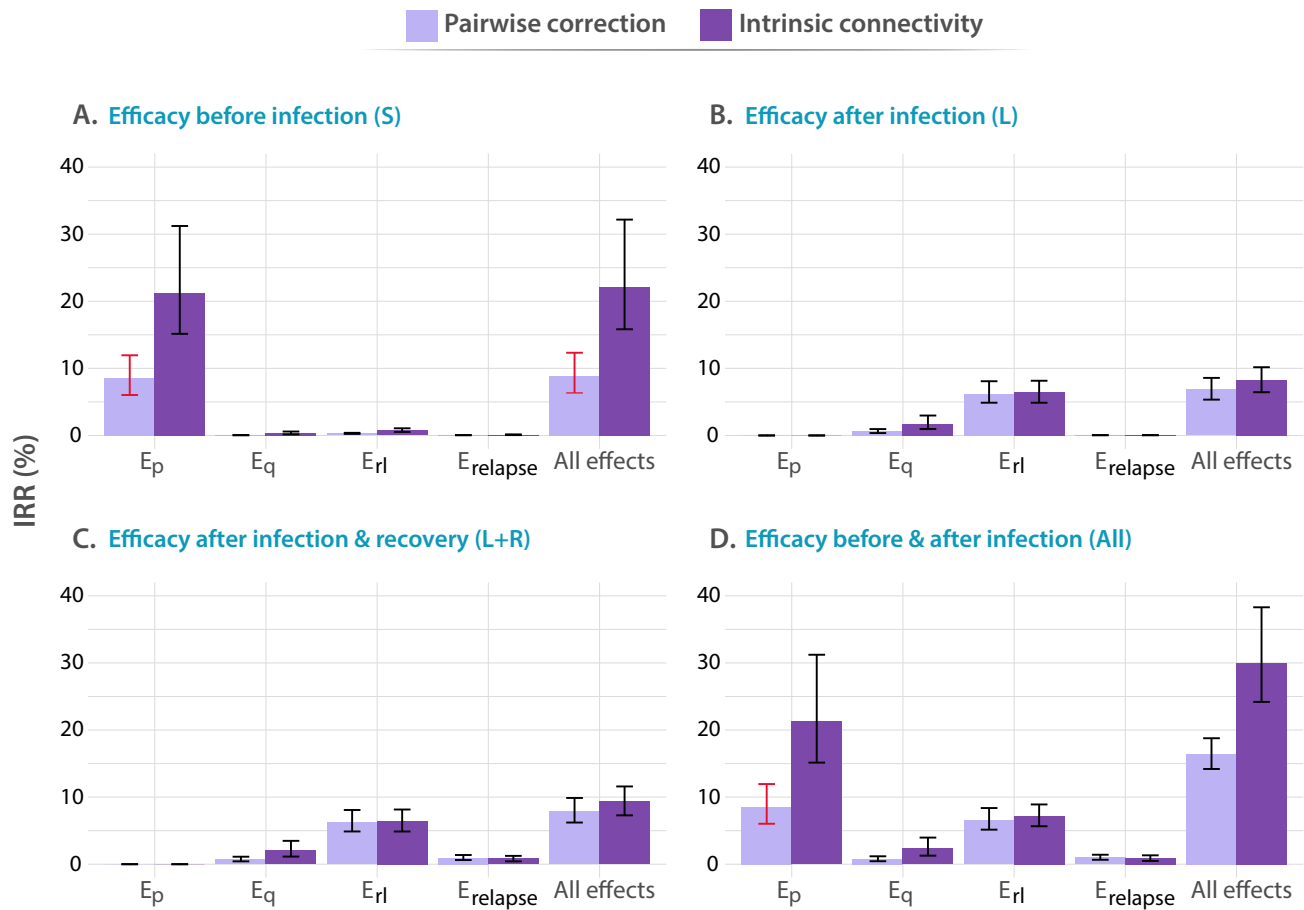


FIG. 2. Impact of different vaccines applied to individuals between 60-64 years old. In all panels, tested vaccines act in different parts of the natural history of TB, halting progression to disease in one or more of the possible routes to disease, as described in the methods:  $E_p$ : protection against primary TB,  $E_q$  protection against reinfection,  $E_{rl}$ : protection against endogenous progression to TB after LTBI,  $E_{relapse}$ : prevention of recurrence. We analyze independently the impact of vaccines whose protective effects unfold when applied to individuals belonging to different compartments of the natural history, **A**. Susceptible subjects (efficacy observed before infection). **B**. Latently infected individuals (efficacy observed after infection). **C**. Latently infected and recovered individuals. **D**. Entire population. In each case, the impact of each vaccine is evaluated for a waning level of 10 years. In all panels, bars represent median values for the IRR measured in 2050, associated with the introduction of the vaccine in 2025. Error bars capture 95% confidence intervals from a set of  $N = 500$  model outcomes in each case.

276 always appears as a superior choice according to our simulations (6.47% IRR, 95% CI[4.88-8.15] in most impactful  
 277 vaccine in elders vs. 2.14 IRR, 95% CI [1.65 2.76] in adolescents). However, as soon as protection against primary  
 278 TB is granted to susceptible individuals as one of the possible protective mechanisms of the vaccines, the quantitative  
 279 description of contagion dynamics implemented within our model becomes more crucial, and, consequently, model  
 280 forecasts are more sensitive to the adoption of either one of the two modeling approaches explored for describing  
 281 contact matrices evolution: pairwise corrections vs. intrinsic connectivity. As a result, only in some occasions when  
 282 the over-simplified pairwise correction method is adopted, the impacts foreseen for an adolescent-focused campaign  
 283 can overcome the impacts found for an elder vaccination campaign (Bars marked in red) for these vaccines. For  
 284 instance, an  $E_p$  vaccine yields 11.11% IRR (95% CI 8.10-13.98) under pairwise correction in adolescent, surpassing  
 285 the 8.54% IRR (95% CI 6.03-11.95) obtained in elders for the same method, when vaccine protection unfolds before  
 286 infection, whereas for the rest of vaccines, elders score higher impacts no matter what method is used.

287 In order to understand the influence of the vaccine mechanisms on their respective impacts, it is sufficient to analyze  
 288 the time evolution of the distribution of TB cases across the different routes to disease classically described in TB,  
 289 aggregated across age groups, as we represent in Figures 4A and 4B. These routes include fast progression to primary  
 290 TB upon a first, recent infection event; TB after endogenous reactivation from LTBI; TB upon exogenous reinfection

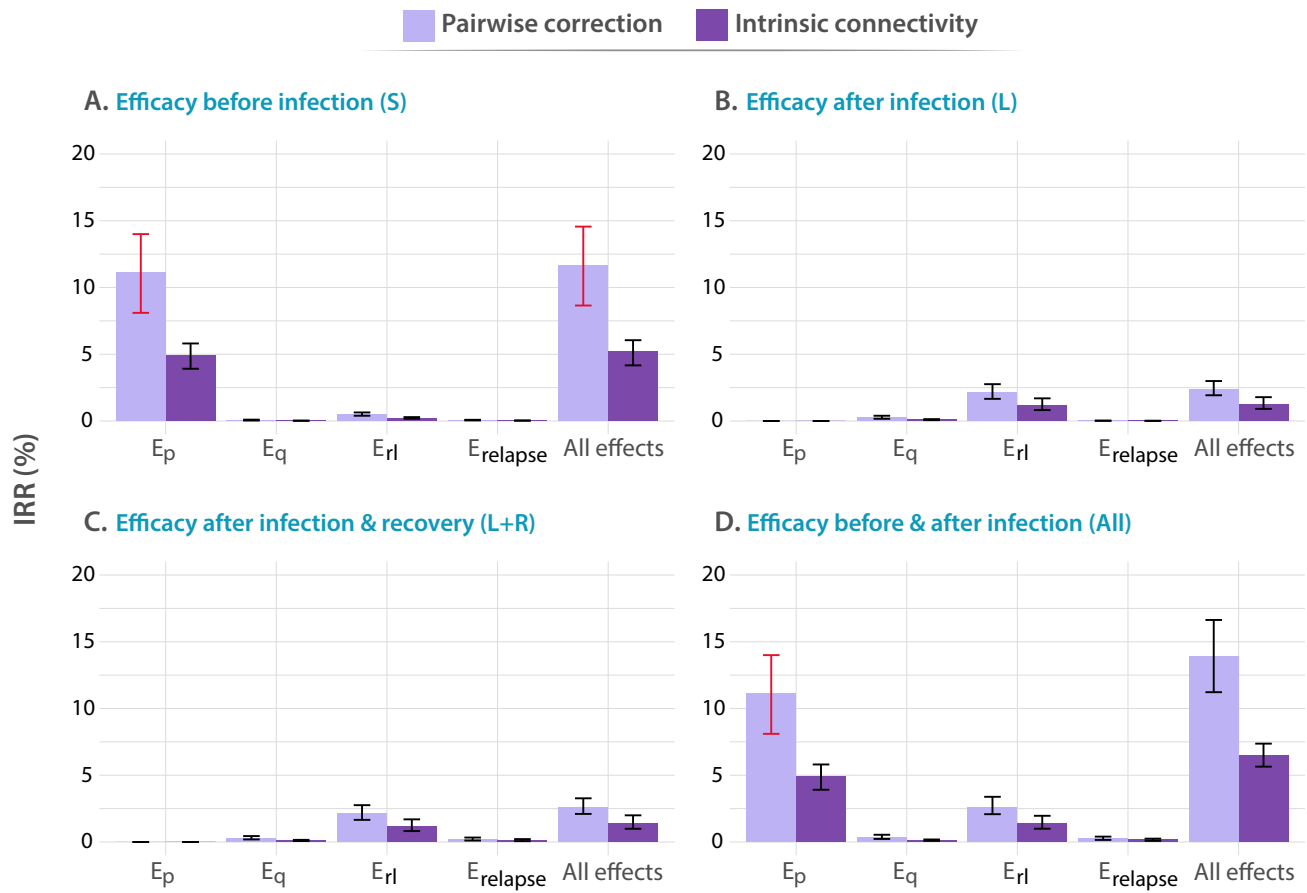


FIG. 3. Impact of different vaccines applied to adolescents, with ages between 15-19 years old. Analogous to Figure 2, in all panels, tested vaccines act in different parts of the natural history of TB, halting progression to disease in one or more of the possible routes to disease:  $E_p$ : protection against primary TB,  $E_q$  protection against reinfection,  $E_{rl}$ : protection against endogenous progression to TB after LTBI,  $E_{relapse}$ : prevention of recurrence. We analyze independently the impact of vaccines whose protective effects unfold when applied to individuals belonging to different compartments of the natural history, **A**. Susceptible subjects (efficacy observed before infection). **B**. Latently infected individuals (efficacy observed after infection). **C**. Latently infected and recovered individuals. **D**. Entire population. In each case, the impact of each vaccine is evaluated for a waning level of 10 years. In all panels, bars represent median values for the IRR measured in 2050, associated with the introduction of the vaccine in 2025. Error bars capture 95% confidence intervals from a set of  $N = 500$  model outcomes in each case.

291 and, last, TB recurrence after a previous disease event (see supplementary methods). Importantly, each of the four  
 292 vaccine mechanisms explored in this work tackles specifically each one of these routes. As seen in Figures 4A and  
 293 4B, primary TB upon recent infection is the prominent cause of TB cases during the simulated period, which makes  
 294 protection against primary TB the most impactful vaccine mechanism, at least, as long as the susceptible individuals  
 295 (who are those under a higher risk of developing primary TB upon infection[26]), could be protected by the vaccine  
 296 (see Figures 2 and 3, panels **A,D**). Furthermore, we also observe that endogenous reactivation of LTBI individuals is  
 297 the second type of event responsible for the highest share of TB cases, which in turn explains why vaccines protecting  
 298 LTBI individuals are most impactful when they protect against endogenous reactivation (see Figures 2 and 3, panels  
 299 **B,C**), and why vaccines targeting re-infection or relapse are comparatively less impactful, even when applied on older  
 300 age groups where prevalence of infected and recovered subjects is higher.

301 Furthermore, in order to understand the role of the modeling assumptions concerning contacts when comparing  
 302 impact forecasts from analogous immunization campaigns, it is important to highlight that their influence manifests  
 303 more strongly when comparing forecasts for vaccines targeting TB upon recent infection ( $E_p$  vaccines) or reinfection  
 304 ( $E_q$  vaccines), which is to say when the vaccine targets transmission. In such cases, the adoption of the most adequate  
 305 method providing intrinsic connectivity control appears systematically associated with larger impacts when we analyze



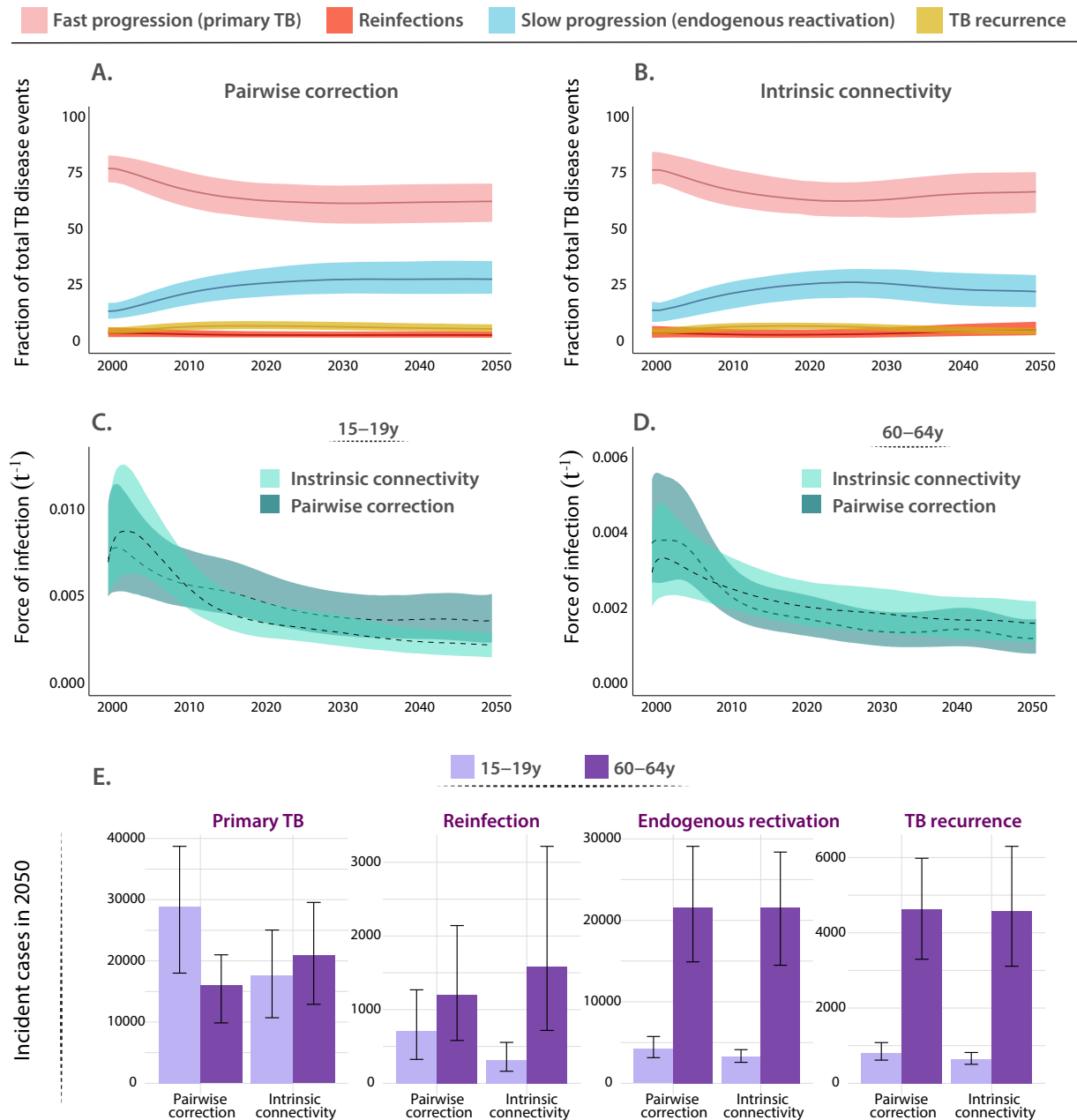


FIG. 4. **A,B.** Evolution of the Percent of TB cases associated with rapid progression upon recent infection (primary TB), endogenous reactivation of LTBI, TB upon reinfection, or TB relapse; each of them foreseen from the indicated method for describing contact matrices evolution (pairwise correction or intrinsic connectivity). Primary TB upon recent infection, followed by endogenous reactivation of LTBI individuals are the two most common types of events. Central lines are medians and the shadowed areas represent the 95% confidence intervals from  $N = 500$  model realizations. **C,D.** Evolution of the force of infection associated with individuals in age groups 15-19 and 60-64 in the period 2000-2050, as foreseen by the model using the two frameworks for describing the time evolution of contact matrices. Adopting the simpler pairwise correction method results in a lower estimate of the force of infection in the older age group, as well as an overestimation in the younger stratum. Central lines are medians and the shadowed areas represent the 95% confidence intervals from  $N = 500$  model realizations. **E.** Break down of the different contributions to the overall TB incidence pool in 2050, distributed across the routes to disease protected by the vaccines under study. The number of cases in the routes associated with already exposed individuals is systematically higher in elders than adolescents, no matter which correction for the contact matrices is at play. Bars represent median values measured in 2050. Error bars capture 95% confidence intervals from  $N = 500$  model realizations.

306 elder-focused campaigns, as well as lower impacts when we focus on campaigns targetting adolescents. This can be  
307 further contextualized by observing the evolution of the force of infection (defined as the fraction of susceptible  
308 individuals in a given group that gets infected per year) that individuals in each of these two age strata are subject  
309 to, according to each of the two models explored to describe the evolution of contact matrices. As seen in Figures  
310 4C and 4D, using the pairwise model appears associated with an underestimation of the force of infection suffered by  
311 individuals in the older age group, and an overestimation of it among adolescents, with respect to the adoption of the  
312 more rigorous correction method based on preserving the intrinsic connectivity of contact matrices.

313 Finally, in order to contextualize the differences in impact found between elder and adolescents-focused campaigns,  
314 in Figure 4E, we present a simultaneous breakdown of TB cases predicted by 2050, in each of the age groups,  
315 associated with each of the routes to disease, according to each of the contact matrices models. In this figure, we see  
316 how both modeling approaches for contact matrices concur in assigning a higher incidence for TB cases related to  
317 LTBI reactivation, reinfection, or relapses, which can be interpreted as the main cause why, in Figures 2 and 3, we  
318 observe that  $E_{rl}$ ,  $E_q$  and  $E_{relapse}$  vaccines are systematically more impactful in elder individuals than in adolescents.  
319 However, in Figure 4E, we can also observe how the number of cases associated with primary TB is either higher, or  
320 lower in adolescents than it is in the older age group, depending on whether we adopt the pairwise correction method,  
321 or the intrinsic connectivity method, respectively.

322 Importantly, none of these observations are affected by the level of vaccine waning: while Figures 2 and 3 cap-  
323 ture impacts associated with vaccines whose protection lasts ten years, largely comparable results are obtained for  
324 longer lasting vaccines, as summarized in the supplementary appendix (see Supplementary Figure 5, corresponding  
325 to waning=20 years).

#### 326 IV. DISCUSSION

327 Mathematical disease-transmission models are a powerful tool for estimating the impact of new TB vaccines, which,  
328 if done properly, may be instrumental in comparing the potential of different vaccine candidates and immunization  
329 campaigns. This is true, especially in TB, where vaccine development must face two simultaneous hindrances. First,  
330 vaccine efficacy is harder to foresee before phases 2b/3 of the development pipeline than for other diseases, given  
331 the lack of reliable correlates of immune protection[27]. Second, the architectures of the clinical trials of vaccine  
332 efficacy that are being adopted to test novel TB vaccine candidates are highly diverse [16, 20], and the protection  
333 profiles of the tested vaccines may be equally diverse. Taken together, these issues claim the development of rigorous  
334 computational models to produce impact comparisons for different vaccines tested in trials of different characteristics  
335 and implemented through assorted immunization campaigns. These constitute extremely non-trivial tasks, which  
336 enhances the need to ensure that current TB models can handle them while minimizing bias and uncertainty.

337 In accomplishing this goal, an aspect that demands special attention is the description of the coupling between  
338 demographic aging and the evolution of TB epidemiology in a given population. This is especially true in a country  
339 such as China, where two simultaneous aspects concur, namely: an intense process of demographic aging -already  
340 ongoing, and expected to continue in the next few decades-, concomitant with a high burden of TB incidence and  
341 prevalence levels. While previous works pointed to the observation that immunization campaigns targetting older  
342 age groups (paradigmatically individuals above 60 years old) are expected to cause a stronger reduction in global  
343 TB incidence levels than campaigns targetting adolescents (16-20 years old) [13], the robustness of these results under  
344 different modeling scenarios, including different vaccine characteristics and modeling decisions concerning the evolution  
345 of contact matrices among different age-groups remained to be proven.

346 Capitalizing on a mathematical model previously developed by the team [14, 16, 20], in this work we reproduce the  
347 general observation that, in China, immunization campaigns targetting older individuals, in the age group between  
348 60 and 64 years old, are associated with promising levels of reduction in the incidence rates expected by 2050, with  
349 varying forecast impacts depending on vaccine characteristics and modeling assumptions, especially if the vaccine is  
350 able to protect already exposed individuals. This observation can be interpreted in light of the demographic shift  
351 expected in the country, where older age strata are expected to accumulate a higher fraction of total TB cases in  
352 the years to come. However, by using our model, we were able to address, for the first time in this study, how  
353 this observation may depend, in turn, on vaccine characteristics (the combination of its mechanisms of action and  
354 protection profile), as well as on modeling assumptions (the description of contact matrices over time).

355 On the one hand, when modeling TB vaccines, it is important to acknowledge the multiplicity of possible mechanisms  
356 of action a vaccine may confer protection through [16, 20]. This aspect, in turn, must be considered simultaneously with  
357 the fact that the initial immunological profile of vaccinated individuals (i.e. their IGRA status) may in turn influence  
358 the ability of the vaccine to provide its protective effects. [13]. In this work, we describe how these two aspects are  
359 coupled, generating strong interactions between the vaccine mechanisms in place and the sub-population reservoirs  
360 that may gain protective effects upon vaccination. Specifically, we observed that vaccines protecting susceptible,

immunologically naive individuals are more impactful when their mechanism of action is based upon the prevention of primary TB after infection. This result can be understood by observing that progression to primary TB upon recent infection represents not just the main epidemiological risk for susceptible individuals, but the most common route to disease in the whole population, as sketched in Figures 4A and 4B. Importantly, our simulations indicate that tackling primary TB is the most promising intervention, not only when the immunization campaign targets adolescents, but also when it targets older individuals, as long as vaccine protection unfolds for susceptible individuals at least. Furthermore, when a vaccine requires that vaccinated individuals have previously been infected, the most impactful vaccine mechanism of action is based on preventing endogenous reactivation of LTBI. This indicates that, for LTBI subjects, endogenous progression to TB represents the highest epidemiological risk, which is shown in Figures 4A and 4B, where endogenous reactivation is the second most common route to disease. These couplings between vaccine mechanisms and protection profiles should be carefully taken into account when testing and comparing vaccine candidates with different profiles and immunization strategies.

On the other hand, in a country such as China, it is key to produce model-based descriptions of TB dynamics that are robust under the scenario of fast demographic aging. Under these circumstances, the adoption of plausible description frameworks to describe the evolution of contact matrices is key. The reason for this is that these matrices capture the relative frequency of contacts that may lead to new infections among individuals of different ages, and these are bound to evolve with time in an aging population. While relatively naive descriptions of contact matrices based on symmetry preservation through pairwise corrections are enough when modeling infectious diseases during short periods of time, TB demands more sophisticated approaches that preserve not only the symmetry but also the overall connectivity of the entire contact networks [15]. The reason for this is that during the extended time windows that TB modeling requires, demographic structures are expected to vary significantly, and, with them, the frequencies of social contacts among age strata, and the entire connectivity, measured as the average number of contacts per individual, of the system.

In this work, we showed that more sophisticated modeling approaches based on imposing the preservation of the intrinsic connectivity of contact networks (instead of simpler methods based on pairwise corrections aiming only at preserving symmetry) are linked to higher vaccine impacts when immunization campaigns target transmission among elder individuals. In turn, for campaigns targeting transmission among adolescents, it is the simpler methods, based on pairwise corrections, the one yielding higher impacts. In short, our simulations indicate that vaccines whose protection mechanisms take place after infection (e.g.  $E_{rl}$ ,  $E_q$  and  $E_{relapse}$  on L, L+R, or All population), are expected to elicit higher population impacts if applied in elder individuals, as well as vaccines protecting susceptible individuals against primary TB, providing that an adequate modeling approach is used to describe the evolution of their contact matrices, ensuring intrinsic connectivity control.

We also need to mention that our approach is not exempt from limitations that affect TB transmission models. The outcomes of our model depend on a series of epidemiological parameters and initial burden estimates that are subject to strong sources of uncertainty, thus propagating this uncertainty to the results. This means that future improvements in measuring the input data are expected to impact the quantitative outcomes of our mathematical model, in the same way it would affect any other model that leans on them. Always bearing in mind the strong uncertainties that the forecasts inherit, our results highlight the importance of acknowledging the complexity of TB transmission dynamics when modeling the effects of an age-focused intervention such as the introduction of a new vaccine on a specific age group.

In closing, our results emphasize the idea that immunization campaigns for the introduction of new TB vaccines in different countries can be, and must be, tailored using mathematical models that integrate information on vaccines' profiles, population demography, and basal TB epidemiology.

#### CONFLICT OF INTEREST STATEMENT

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

#### AUTHOR CONTRIBUTIONS

M.T, J.S and Y.M designed the study. M.T performed the analysis and simulations that support the findings. M.T, J.S and Y.M analyzed and discussed the results. M.T and Y.M. wrote the manuscript. All authors revised and approved the submission of the final version of the manuscript.

411

## FUNDING

412 We thank M. Clarin for help with the figures. M.T. acknowledges the support of the Government of Aragón through  
413 a PhD contract. J.S. acknowledges support from the Spanish Ministry of Science and Innovation (MICINN) through  
414 grant PID2019-106859GA-I00 and Ramón y Cajal research grant RYC-2017-23560, as well as to the Government of  
415 Aragón, Spain, and ‘ERDF A way of making Europe’ through grant B49-23R (NeuroBioSys). Y.M was partially sup-  
416 ported by the Government of Aragón, Spain, and “ERDF A way of making Europe” through grant E36-23R (FENOL),  
417 and by Ministerio de Ciencia e Innovación, Agencia Española de Investigación (MCIN/AEI/10.13039/501100011033)  
418 Grant No. PID2020-115800GB-I00. The funders had no role in study design, data collection, and analysis, decision  
419 to publish, or preparation of the manuscript.

420

## DATA AVAILABILITY STATEMENT

421 The data that support the findings of this study are publicly available at the original sources. The code that  
422 supports the findings of this study is available from the corresponding author upon reasonable request.

- 
- 423 [1] M. Uplekar, D. Weil, K. Lonnroth, E. Jaramillo, C. Lienhardt, H. M. Dias, D. Falzon, K. Floyd, G. Gargioni, H. Getahun,  
424 *et al.*, *The Lancet* **385**, 1799 (2015).
- 425 [2] A. Kazemnejad, S. Arsang Jang, F. Amani, and A. Omid, *Journal of research in health sciences* **14**, 115 (2014).
- 426 [3] W. H. Organization *et al.*, *Global tuberculosis report 2022*. geneva: Who; 2022 (2022).
- 427 [4] U. de Investigación en Tuberculosis de Barcelona, *The covid-19 pandemic remains getting worse the tuberculosis control*  
428 (2021), [Online; posted November-2021].
- 429 [5] A. A. Malik, N. Safdar, S. Chandir, U. Khan, S. Khowaja, N. Riaz, R. Maniar, M. Jaswal, A. J. Khan, and H. Hussain,  
430 *Health policy and planning* **35**, 1130 (2020).
- 431 [6] C. F. McQuaid, A. Vassall, T. Cohen, K. Fiekert, R. White, *et al.*, *The International Journal of Tuberculosis and Lung*  
432 *Disease* **25**, 436 (2021).
- 433 [7] W. H. Organization *et al.*, *Global tuberculosis report 2021*. geneva: Who; 2021 (2021).
- 434 [8] L. Cilloni, H. Fu, J. F. Vesga, D. Dowdy, C. Pretorius, S. Ahmedov, S. A. Nair, A. Mosneaga, E. Masini, S. Sahu, *et al.*,  
435 *EClinicalMedicine* **28**, 100603 (2020).
- 436 [9] M. Tovar, A. Aleta, J. Sanz, and Y. Moreno, *Communications medicine* **2**, 1 (2022).
- 437 [10] C. Lange, D. Chesov, J. Heyckendorf, C. C. Leung, Z. Udhwadia, and K. Dheda, *Respirology* **23**, 656 (2018).
- 438 [11] P. E. Fine, *Scandinavian journal of infectious diseases* **33**, 58 (2001).
- 439 [12] C. Martin, N. Aguilo, D. Marinova, and J. Gonzalo-Asensio, *Applied Sciences* **10**, 2632 (2020).
- 440 [13] R. C. Harris, T. Sumner, G. M. Knight, T. Evans, V. Cardenas, C. Chen, and R. G. White, *The Lancet Global Health* **7**,  
441 e209 (2019).
- 442 [14] S. Arregui, M. J. Iglesias, S. Samper, D. Marinova, C. Martin, J. Sanz, and Y. Moreno, *Proceedings of the National*  
443 *Academy of Sciences* **115**, E3238 (2018), <https://www.pnas.org/content/115/14/E3238.full.pdf>.
- 444 [15] S. Arregui, A. Aleta, J. Sanz, and Y. Moreno, *PLoS computational biology* **14**, e1006638 (2018).
- 445 [16] M. Tovar, S. Arregui, D. Marinova, C. Martín, J. Sanz, and Y. Moreno, *Nature communications* **10**, 1 (2019).
- 446 [17] C. Weerasuriya, R. Clark, R. White, and R. Harris, *Journal of internal medicine* **288**, 661 (2020).
- 447 [18] P. Smith, L. Rodrigues, and P. Fine, *International journal of epidemiology* **13**, 87 (1984).
- 448 [19] M. G. M. Gomes, M. Lipsitch, A. R. Wargo, G. Kurath, C. Rebelo, G. F. Medley, and A. Coutinho, *PLoS pathogens* **10**,  
449 e1003849 (2014).
- 450 [20] M. Tovar, Y. Moreno, and J. Sanz, *Nature communications* **14**, 5312 (2023).
- 451 [21] S. Eubank, H. Guclu, V. Anil Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang, *Nature* **429**, 180 (2004).
- 452 [22] J. Wallinga, P. Teunis, and M. Kretzschmar, *American journal of epidemiology* **164**, 936 (2006).
- 453 [23] J. M. Read, K. T. Eames, and W. J. Edmunds, *Journal of The Royal Society Interface* **5**, 1001 (2008).
- 454 [24] C. J. Worby, S. S. Chaves, J. Wallinga, M. Lipsitch, L. Finelli, and E. Goldstein, *Epidemics* **13**, 10 (2015).
- 455 [25] S. Kiazzyk and T. Ball, *Canada Communicable Disease Report* **43**, 62 (2017).
- 456 [26] J. R. Andrews, F. Noubary, R. P. Walensky, R. Cerda, E. Losina, and C. R. Horsburgh, *Clinical infectious diseases* **54**,  
457 784 (2012).
- 458 [27] K. Bhatt, S. Verma, J. J. Ellner, and P. Salgame, *Clinical and Vaccine Immunology* **22**, 258 (2015).