

Modelling the impact of novel tuberculosis vaccines: from clinical trials design to spreading modeling and back

Trabajo Fin de Master Física

Junio de 2018

Trabajo realizado bajo la dirección de Joaquín Sanz y Yamir Moreno (Instituto de Biocomputación y Física de Sistemas Complejos de Zaragoza)

Agradecimientos

Tras un largo período de trabajo intenso, no sólo en virtud de mi TFM sino del desarrollo mismo del máster, escribo estas líneas poniendo punto y final a otra etapa de mi aprendizaje y es por ello que me gustaría dedicar unas palabras de agradecimiento a todos aquellos con los que he compartido este tiempo.

En primer lugar, quiero agradecer a mis compañeros el apoyo que ha supuesto pasar juntos estos meses de trabajo, cada uno en un tema diferente, pero siempre dispuestos a apoyarnos entre todos. También los buenos ratos en los descansos y fuera de cualquier actividad académica. Me gustaría agradecer especialmente a aquellos con los que he compartido asignaturas, pues hemos tenido que pasar muchas horas trabajando juntos, lo cual lejos de suponer un problema, ha resultado tremendamente beneficioso.

Agradezco también a mi familia por haberme motivado siempre a seguir un camino que es un tanto oscuro hoy en día, como es el de la investigación. Sin su apoyo, y en particular el de mi hermana, no hubiera sido posible presentar hoy este trabajo.

Finalmente, me gustaría dar las gracias a mi director Yamir Moreno y a mi co-director Joaquín Sanz por el buen trato recibido y por su disposición y dedicación para solventar dudas y problemas siempre que fue necesario. Gracias a ellos la experiencia de trabajo próxima al desarrollo de la investigación ha resultado una experiencia enriquecedora y muy formativa.

Lista de acrónimos

- Tuberculosis: TB
- Randomized controlled trial: RCT
- Eficacia en infección: VE_{inf}
- Eficacia en desarrollo de la Tuberculosis: VE_{dis}
- Prevención ante infección (prevention of infection): POI
- Prevención ante desarrollo de enfermedad (prevention of disease): POD

Índice

I	Introducción y objetivos	1
1.	El problema del ciclo de transmisión de la TB: progresión rápida a enfermedad vs TB latente	2
2.	Eficacia de una vacuna: Ensayos de prevención de infección y enfermedad.	3
II	Modelo analítico para la descripción de la dinámica de progresión a infección y enfermedad en un ensayo clínico de eficacia vacunal en TB.	5
3.	Resolución analítica: obtención de $f(\varepsilon_p, \varepsilon_r)$	7
4.	El problema de la degeneración: curva de parámetros compatibles con una determinada VE_{dis}	10
5.	Rompiendo la degeneración: estimación independiente de ε_r	11
6.	Método de estimación de ε_p : Parámetros relevantes.	13
III	Verificación de la validez del método.	14
7.	Hipótesis de transiciones rápidas.	15
7.1.	Flujos de individuos en el modelo	15
8.	Validación para vacunas 100-0 y 0-100	16
8.1.	Análisis de fallos del método	18
8.2.	Caso con protección frente a infección	20
9.	Robustez en escenarios alternativos	22
9.1.	Barrido en el tiempo de seguimiento	23
9.2.	Barrido en la fuerza de infección	24
9.3.	Barrido en p	25
9.4.	Barrido en la protección frente a reinfección	26
10.	Estimación de impacto	27

Parte I

Introducción y objetivos

Existen enfermedades en las que no se puede realizar un ensayo inmunológico sencillo que permita determinar si un individuo estará protegido o no tras aplicarse una vacuna. En ellas, el diseño de un ensayo de campo para probar la eficacia de una vacuna es largo y costoso, tanto en tiempo como económicamente. Este es el caso de la Tuberculosis (TB), una de las enfermedades en las que no existe el denominado “correlate of protection” (signos medibles del grado de inmunidad de un individuo a una determinada enfermedad tras una vacunación exitosa).

Por esto es necesario un buen diseño de los ensayos, con una buena planificación y elección del país o países en los que se va a desarrollar. Por ejemplo, no tiene ningún sentido plantear un ensayo para probar una vacuna contra la malaria en el norte de Europa, pero sí en países asiáticos en los que la incidencia todavía es elevada.

Usualmente se recurre a lo que se conoce como ensayos controlados aleatorios. Estos ensayos consisten en dos cohortes con el mismo número de individuos asignados aleatoriamente. Las cohortes se denominan de control o vacunada según se le administre dosis de vacuna o placebo a sus integrantes. De estas cohortes se obtiene el ritmo al que se infectan los individuos durante un tiempo de seguimiento para tratar de extraer conclusiones sobre la eficacia de la vacuna.

En este contexto, una de las enfermedades con mayor casuística y que más problemas dan a la hora de realizar los ensayos es, como comentamos, la Tuberculosis (TB). Esta enfermedad suele presentar períodos de incubación y latencia muy largos. En este tipo de enfermedades, los ensayos deben ser lo suficientemente largos como para disponer de un número estadísticamente significativo de sujetos enfermos con los que poder aplicar las técnicas de análisis estadístico.

La falta de garantías de éxito al no existir “correlate of protection” supone un gran riesgo financiero, dado que los fondos en el campo son muy limitados, por lo que garantizar que las metodologías de diseño y análisis de ensayos funcionen es una prioridad absoluta.

No obstante, particularizando al caso de esta enfermedad, por ser el que nos interesa, las técnicas clásicas de análisis de supervivencia no siempre son las más adecuadas para estimar la eficacia de la vacuna probada en el ensayo. De hecho, en el caso de la TB, el mecanismo por el cual una vacuna puede mostrar una determinada eficacia ante enfermedad, i.e, casos de TB activa prevenidos al final del ensayo, está degenerado, pudiendo ocurrir por dos vías.

Esta degeneración es responsable del fallo en la correcta caracterización de los mecanismos sobre los que actúa la vacuna, por lo que es necesario introducir nuevas formas de analizar los resultados del ensayo que den cuenta de la degeneración y sean capaces de romperla para lograr una caracterización completa.

Esto implica una necesidad de desarrollo de nuevos métodos de prevención y tratamientos que permitan luchar de un modo más efectivo contra la TB. Se suma el hecho de que la efectividad de la vacuna que se emplea actualmente, la BCG, lleva tiempo siendo cuestionada por parte de la comunidad científica, ya que todo parece indicar que presenta niveles de protección inconsistentes contra las formas más infecciosas de TB en individuos adultos, que llegan a oscilar entre 0 y 80% de eficacia[6].

Por todo lo anterior, los objetivos del este trabajo incluyen introducir la una metodología novedosa que permita tratar con el mecanismo degenerado en los ensayos de vacunas de la TB. Con ese objetivo se desarrolla un modelo analítico adecuado a un caso real de ensayo clínico y se lleva a cabo una implementación numérica para validar el método que se propone.

1. El problema del ciclo de transmisión de la TB: progresión rápida a enfermedad vs TB latente

La TB es una enfermedad causada por la *Mycobacterium tuberculosis*, una bacteria que, en su forma infecciosa, ataca los pulmones de la víctima. La TB se transmite de persona a persona a través del aire. Cuando un enfermo de tuberculosis pulmonar tose, estornuda o escupe, expulsa bacilos al aire. Estos bacilos pueden infectar fácilmente a otro individuo en caso de que éste los inhale.

Por otro lado, pese a la importante reducción en el impacto de la TB durante los últimos 20 años a lo largo del mundo, la TB sigue siendo una de las mayores amenazas para la salud pública. Según el Informe mundial sobre la tuberculosis de la Organización Mundial de la Salud (OMS) [1], se estima que 10.4 millones de personas desarrollaron la enfermedad durante 2016 y que la enfermedad es responsable de la muerte de cerca de 1.7 millones de personas. Además, las perspectivas no son muy optimistas cuando tenemos en cuenta los efectos de la interacción de la TB con el VIH en pacientes coinfectados [2][3], con más de un millón de casos en 2016 [1], y la aparición de cepas de la TB resistentes a múltiples fármacos (MDR - TB) [4][5], con casi 0.5 millones de casos.

Esta amenaza radica en varios factores, entre los que destaca la compleja historia natural de la enfermedad, que la hace que modelarla sea altamente complicado. Esta complejidad implica que las previsiones que la epidemiología puede dar con modelos sencillos acerca de su propagación y sus previsiones de impacto no sean muy fiables. Recientemente se han producido varios avances en el campo que han dado lugar a un modelo más rico y que da cuenta de fenómenos interactuantes con la TB que no se tomaban en cuenta normalmente, como la evolución de la demografía[9].

La estructura subyacente de los estados de la TB que se adopta en este modelo, presente en la figura 1, nos permite visualizar una parte de esta complejidad. En este modelo en particular se contemplan las dos vías de latencia que presenta la TB, lenta y rápida (según el periodo de incubación) a las que accede un individuo sano inicialmente que se infecte. Las dos vías de latencia dejan paso a 3 posibles tipos de enfermedad, dos pulmonares y una no pulmonar. La enfermedad, si tratada, puede resultar en distintos estados según el fracaso o el éxito del tratamiento, e incluso la posibilidad de muerte está incluida. Esto supone un marco muy complejo que justifica la necesidad de modelos específicos para tratar con TB.

Por otro lado, se calcula que una tercera parte de la población mundial tiene tuberculosis latente; es decir, dichas personas están infectadas por el bacilo pero todavía no han enfermado ni pueden transmitir la infección.

La existencia de los estados de infección latente, asociados a periodos de incubación rápida, o estados de latencia lenta, con periodos de incubación largos, es una de las características más notables del ciclo de transmisión de la TB. La importancia dinámica de estos estados ha llevado

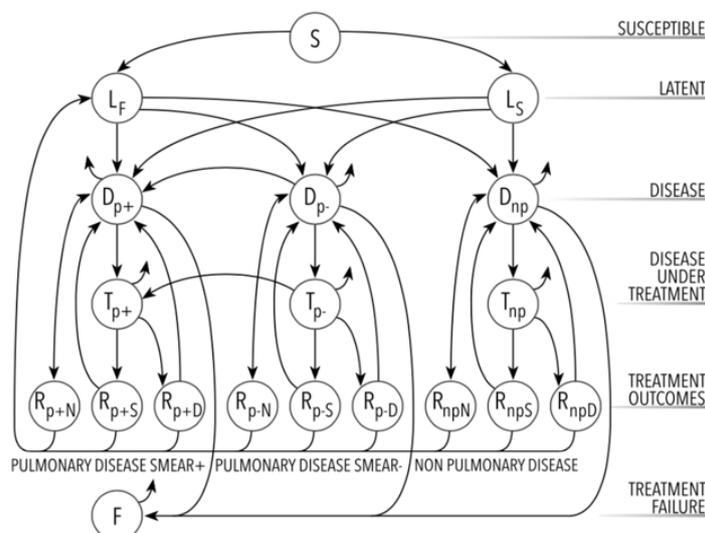


Figura 1: Esquema de la historia natural de la TB en modelos de propagación. Se distinguen tres ramas según la infección sea pulmonar ($p+$ y $p-$) o no pulmonar. Las flechas entre estados contemplan las posibles transiciones existentes en el modelo.

a la comunidad a plantearse la cuestión de determinar el objetivo óptimo de una nueva vacuna: la prevención de la infección en sí (transición susceptible a latencia) o la prevención de la enfermedad (transición de latencia a enfermedad).

2. Eficacia de una vacuna: Ensayos de prevención de infección y enfermedad.

Teniendo a disposición una nueva vacuna para una determinada enfermedad, apuntando a estimar su eficacia es necesario plantear un escenario controlado en el que poder obtenerla. Una estrategia típica que se ha escogido para realizar estudios de campo que permitan obtener datos que procesar mediante técnicas estadísticas corresponde a reclutar participantes de características similares que posteriormente se dividen en dos grupos (cohortes)[7] con aproximadamente el mismo número de individuos.

A esta forma de proceder se le denomina ensayo controlado aleatorio (RCT) dado que los participantes se distribuyen entre las cohortes de modo aleatorio. En ellos se asigna a cada cohorte un tratamiento, la de control recibe placebo mientras que la vacunada recibe dosis de la vacuna, con vistas a poder obtener información del efecto de la vacuna comparando con la de control.

La necesidad de que los participantes compartan características (lo que en principio queda garantizado por la asignación aleatoria de los participantes), viene justificada por el hecho de que las enfermedades no tienen por qué comportarse igual ante neonatos, adolescentes o adultos, o ante fumadores y no fumadores, entre otros.

Con esta configuración de RTC, mediante técnicas clásicas de análisis de supervivencia se puede estimar la protección frente a infección que dicha vacuna otorga. Para ello se realizan controles periódicos a los participantes para conocer su estado y situar, cuando la haya, una fecha de infección. Con estos tiempos y mediante el uso de métodos clásicos se obtiene la denominada

eficacia en infección, a la que llamaremos VE_{inf} . Pero para la TB esto no es suficiente.

En el caso de la Tuberculosis, existe otra eficacia que es necesario introducir, dado que tenemos 3 estados posibles en los que un participante de un ensayo clínico puede acabar al final del tiempo de seguimiento: sano, infectado o enfermo. Estas posibles fases de la enfermedad se denominan “end-points” en notación de ensayos. Al “end-point” enfermo (i.e, individuos que desarrollan fase activa de la TB en el tiempo de seguimiento) le corresponde una segunda eficacia de la vacuna a la llamaremos VE_{dis} . Esta eficacia tiene su origen en la protección que puede ofrecer una vacuna al ralentizar el paso de un infectado a enfermedad, lo que significa que la transición a fase activa ocurre más lentamente.

En el contexto de los ensayos de vacunas puede que estemos tratando con vacunas destinadas a evitar la infección (asociadas a VE_{inf}), en cuyo caso los ensayos se denominan “POI” (prevention of infection) o puede que el efecto de la vacuna sea aumentar enormemente el periodo de incubación (asociadas a VE_{dis}). En este segundo caso a los ensayos destinados a estimar la eficacia se les denomina “POD” (prevention of disease).

En ambos casos se suele obtener su valor aplicando regresiones de Cox, lo que constituye la metodología clásica de análisis de resultados. Sin embargo, en algunas ocasiones, la hipótesis de riesgos proporcionales puede ser violada, hipótesis que es necesaria para poder aplicar Cox. Este es el caso en TB para algunas vacunas asociadas a VE_{dis} , como se muestra en [8].

Por ello, como estimador de VE_{dis} no utilizaremos Cox, sino que la obtendremos comparando el número de enfermos de las dos cohortes al finalizar el ensayo, el cual debe durar un tiempo suficiente de acuerdo al tiempo promedio que toma infectarse^I. De este modo puede obtenerse VE_{dis} como:

$$VE_{dis} = 1 - \frac{D_v}{D_c} = 1 - \rho \quad (1)$$

siendo ρ la fracción de enfermos entre cohortes, con D_v los infectados en la cohorte vacunada y D_c los infectados en la cohorte de control. Esta no es más que otra forma de inferir la eficacia que se obtendría aplicando una regresión de Cox (estudiando las tasas de progresión de los individuos en las cohortes), como también se muestra en [8], pero asegura el uso de un método aplicable siempre.

La existencia de estos tipos diferentes de eficacias da pie al problema central que se desarrolla en este trabajo: para el caso de VE_{inf} la interpretación de los resultados es única, pues el efecto que presenta una vacuna asociada es el de reducir el riesgo de infección. Sin embargo, en el caso de una determinada VE_{dis} estimada a partir de los resultados de un RCT, existen dos posibles efectos que hayan tenido lugar. Esto es, puede que la protección provenga de que la vacuna aumenta la fracción de individuos que desarrollan un estado de latencia lenta en detrimento de la rápida o puede que el efecto sea aumentar el tiempo de incubación en los individuos que desarrollan latencia rápida.

A pesar de que una vacuna podría presentar otros efectos, como el bloqueo del ciclo de transmisión del patógeno en otros puntos, estos efectos son virtualmente imposibles de observar en el contexto de un ensayo clínico. El ejemplo más claro es que la vacuna proporcionase protección frente a la transición de latencia lenta a enfermedad, la cual se produce al ritmo de un evento por cada

^IEn el caso de enfermedades estacionales, como la gripe, puede ser suficiente con meses, mientras que para enfermedades con periodos de latencia largos, como la TB, se necesitan del orden de varios años

1300 personas al año. Registrar diferencias significativas en el ritmo de esta transición requeriría el seguimiento de varios miles de individuos por cohorte durante décadas. Esto, por supuesto, es inviable.

De un análisis convencional de VE_{dis} no es por tanto posible obtener información suficiente como para romper esta degeneración que aparece de forma natural para el caso de la TB. Esto ha sido pasado por alto en la literatura hasta ahora. La importancia de este fenómeno, como se verá, es que dos vacunas asociadas a valores equivalentes de VE_{dis} pero que actúan por los mecanismos introducidos anteriormente pueden presentar un impacto muy distinto a largo plazo al aplicarse a la población general.

Por ello, introduciremos en lo que sigue un método analítico que nos permita obtener de forma independiente la contribución a cada uno de los efectos, con el fin de romper la degeneración y caracterizar correctamente el efecto de la vacuna.

Parte II

Modelo analítico para la descripción de la dinámica de progresión a infección y enfermedad en un ensayo clínico de eficacia vacunal en TB.

Para tratar de romper la degeneración que encontramos en el análisis de los resultados de un ensayo clínico para nuevas vacunas de la TB debemos plantear una forma novedosa de análisis de datos que nos permita estimar de forma independiente la protección de la vacuna sobre cada una de las dos posibles vías a enfermedad.

La idea radica en que si somos capaces de encontrar una relación analítica entre VE_{dis} y los parámetros de protección de la vacuna en cada vía, sólo será necesario estimar uno de ellos de forma independiente para encontrar el otro aplicando dicha relación. Con objeto de calcularla planteamos la situación del ensayo de la MVA85A [7], básicamente un RCT, y proponemos un modelo analítico simplificado de la dinámica de progresión de la TB, basándonos en el esquema de la figura 2.

En este modelo consideramos 4 posibles estados para los individuos del ensayo: “S”, correspondiente a “Sano”, “F”, correspondiente a latencia rápida, “L”, correspondiente a latencia lenta y finalmente “D”, que corresponde a desarrollar la enfermedad.

Ahora, dentro del modelo las transiciones permitidas entre estados reproducen las fases iniciales de la historia natural de la enfermedad a las que se enfrentan los individuos reclutados en un ensayo, siendo condición *sine qua non* que no hayan estado expuestos a la bacteria previamente (esto es un requisito para el reclutamiento de individuos en el ensayo, como se propone en [7]). Los individuos de las dos cohortes comienzan así el ensayo en estado sano y conforme este avanza, se infectan con probabilidad β . Una vez infectados, con probabilidad p , pasan a estado de latencia

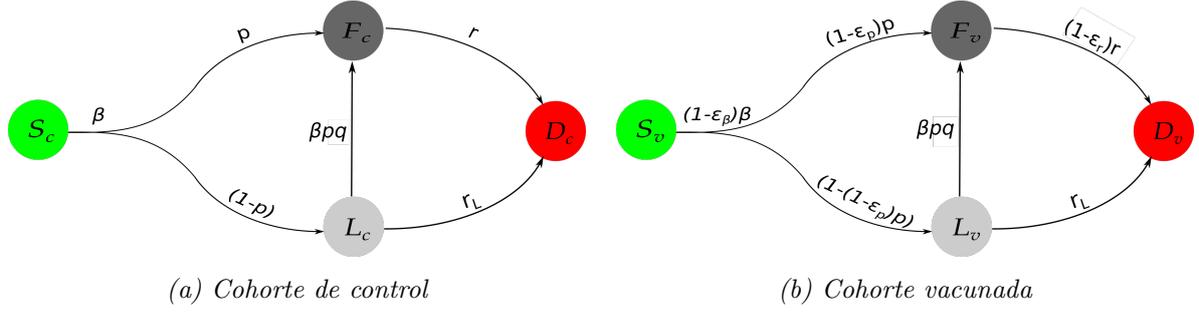


Figura 2: Esquema del modelo analítico planteado para modelar el ensayo. Los individuos que contraen la TB desarrollan uno de los posibles estados de latencia, desde los cuales desarrollan la enfermedad a ratios r y r_L . La posibilidad de reinfecciones viene contenida en la transición βpq .

rápida, o a latencia lenta con probabilidad $(1-p)$. Desde cualquiera de estos dos estados se puede alcanzar el estado final D en el que el individuo desarrolla la fase activa de la TB, mostrando sus síntomas. La diferencia radica en la probabilidad con la que se pasa a este estado desde cada estado de latencia, siendo mucho mayor para el caso de latencia rápida.

Finalmente, destacamos la posibilidad de que un individuo en estado de latencia lenta sufra una segunda infección, lo que le lleva irremediamente a latencia rápida. Este efecto, que ocurre con probabilidad βpq , siendo q la protección que ofrece la propia enfermedad frente a reinfecciones.

Para romper la degeneración, en la cohorte vacunada se introducen 3 posibles efectos de protección suministrados por la vacuna. En el caso en que ésta ofrezca protección frente a infección, lo hará a través de ε_β . Adicionalmente, su efecto sobre el transito a fase activa de la TB (estado D) queda contenido en las dos posibles vías que existen. Si la vacuna trata de aumentar la fracción de individuos que desarrollan latencia lenta, actuará vía ε_p , mientras que si aumenta el tiempo de incubación para latencia rápida, actuará a través de ε_r .

Mediante este modelo establecemos unas pautas con las que seremos capaces de romper la degeneración en VE_{dis} :

1. Primero debemos obtener una relación analítica entre VE_{dis} y ε_p , ε_r , a la que llamaremos $f(\varepsilon_p, \varepsilon_r)$.
2. Conocida dicha relación, debemos encontrar un método que nos permita conocer de forma independiente ε_p ó ε_r . En este trabajo se introduce un método para estimar ε_r a partir de las tasas de transición rápida. Adicionalmente será necesario estimar otros parámetros, como ε_β , los cuales trataremos más adelante.
3. Dados los pasos anteriores, conocida $f(\varepsilon_p, \varepsilon_r)$ y ε_r podremos obtener ε_p , encontrando así el parámetro asociado a las vacunas con acción en p .

Estas pautas quedan recogidas en el esquema presentado en la figura 3, en donde se muestra cómo a partir de la información del ensayo puede estimarse ε_r y así poder obtener ε_p mediante la ligadura que los une.

Los métodos presentes en esta figura serán introducidos en las subsiguientes secciones, sirviendo aquí únicamente con proposito de ilustrar las pautas propuestas en el trabajo.

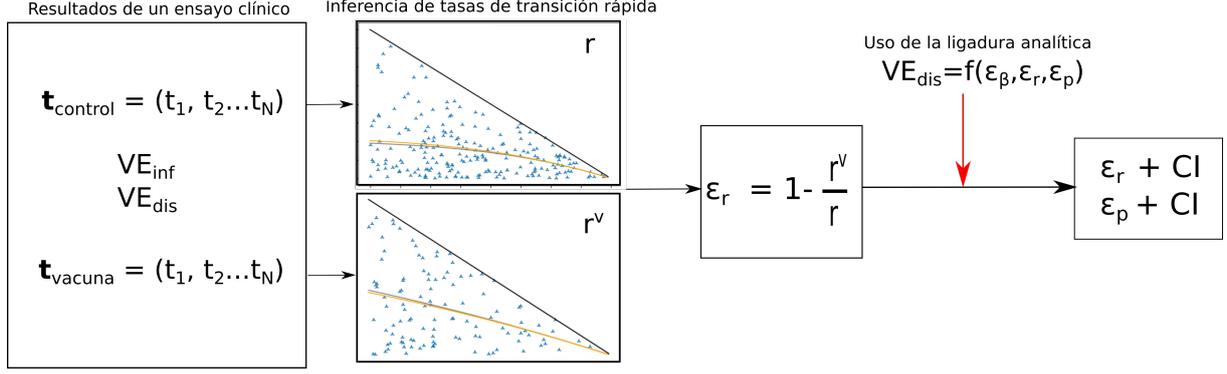


Figura 3: Esquema de la ruta de trabajo planteada. Se registran los tiempos de transición a enfermedad de los individuos que contraen la TB desde estados de latencia en las dos cohortes. Con ellos se estiman las tasas de transición r y r^v con las cuales se estima ϵ_r . Finalmente se utiliza su ligadura $f(\epsilon_p, \epsilon_r)$ para obtener ϵ_p .

3. Resolución analítica: obtención de $f(\epsilon_p, \epsilon_r)$

El primer paso necesario para encontrar una relación analítica que ligue VE_{dis} con los distintos parámetros de protección ($f(\epsilon_p, \epsilon_r)$) es resolver la dinámica del ensayo, encontrando la evolución temporal de cada grupo. En el marco planteado, basándonos en el modelo esquematizado de la figura 2, una cohorte recibe la dosis mientras que otra, la de control, queda sin vacunar.

Dado que este modelo es compartimental (figura 2), podemos plantear una serie de ecuaciones diferenciales para las cohortes con las cuales podamos estudiar la evolución de cada una de las poblaciones involucradas. Para el caso de la cohorte de control se tendrá:

$$\begin{aligned}
 \frac{dS_c(t)}{dt} &= -\beta S_c(t) \\
 \frac{dL_c(t)}{dt} &= \beta(1-p)S_c(t) - (r_L + \beta pq)L_c(t) \\
 \frac{dF_c(t)}{dt} &= \beta p S_c(t) - r F_c(t) + \beta pq L_c(t) \\
 \frac{dD_c(t)}{dt} &= r F_c(t) + r_L L_c(t)
 \end{aligned} \tag{2}$$

Podemos resolver las ecuaciones diferenciales mediante las ecuaciones presentes en 24 y en 25, puesto que se trata de ecuaciones diferenciales inhomogéneas de primer orden.

Para el grupo $S_c(t)$ la evolución queda:

$$S_c(t) = N_0 e^{-\beta t} \tag{3}$$

Grupo $L_c(t)$:

$$L_c(t) = \frac{\beta(1-p)N_0}{r_L + \beta pq - \beta} \left[e^{-\beta t} - e^{-(r_L + \beta pq)t} \right] \tag{4}$$

Grupo $F_c(t)$:

$$F_c(t) = K e^{-rt} + \frac{\beta p N_0}{r - \beta} e^{-\beta t} + \frac{\beta^2 pq (1-p) N_0}{r_L + \beta pq - \beta} \left[\frac{\exp(-\beta t)}{r - \beta} - \frac{e^{-(r_L + \beta pq)t}}{r - r_L - \beta pq} \right] \tag{5}$$

Con:

$$K = -\beta p N_0 \left[\frac{r - r_L - \beta p q - q \beta (1 - p)}{(r - \beta)(r - r_L - \beta p q)} \right] = -\beta p N_0 \left[\frac{r - r_L - q \beta}{(r - \beta)(r - r_L - \beta p q)} \right] \quad (6)$$

Finalmente, la evolución de $D_c(t)$ puede obtenerse sin necesidad de resolver la ecuación diferencial, puesto que:

$$D_c(t) = N_0 - S_c(t) - L_c(t) - F_c(t) \quad (7)$$

Ya que la población total de la cohorte “ N_0 ” se debe mantener constante a lo largo del ensayo a no haber posibilidad de muerte o de abandono del seguimiento considerada en el mismo.

De forma más compacta podemos escribir $D_c(t)$ como:

$$D_c(t) = N_0 [1 - \alpha(\beta, r, r_L, p, q) e^{-\beta t} + \Omega(\beta, r, r_L, p, q) e^{-r t} + \gamma(\beta, r, r_L, p, q) e^{-(r_L + \beta p q) t}] \quad (8)$$

Siendo respectivamente:

$$\begin{aligned} \alpha(\beta, r, r_L, p, q) &= r r_L + \beta p q r - \beta r_L - \beta p r + \beta p r_L \\ \Omega(\beta, r, r_L, p, q) &= \beta p \left[\frac{r - r_L - q \beta}{(r - \beta)(r - r_L - \beta p q)} \right] \\ \gamma(\beta, r, r_L, p, q) &= \beta r - \beta r_L - \beta p r + \beta p r_L \end{aligned} \quad (9)$$

Para el caso de la cohorte vacunada las ecuaciones se resuelven haciendo uso de las mismas herramientas, pero hay que cambiar los coeficientes de las soluciones ya que el efecto de la vacuna es reducir algunos de ellos de modo que:

$$\begin{aligned} \beta &\longrightarrow (1 - \varepsilon_\beta) \beta \\ p &\longrightarrow (1 - \varepsilon_p) p \\ r &\longrightarrow (1 - \varepsilon_r) r \end{aligned} \quad (10)$$

Las ecuaciones diferenciales para la cohorte vacunada son:

$$\begin{aligned} \frac{dS_v(t)}{dt} &= -(1 - \varepsilon_\beta) \beta S_v(t) \\ \frac{dL_v(t)}{dt} &= (1 - \varepsilon_\beta) \beta (1 - (1 - \varepsilon_p) p) S_v(t) - (r_L + (1 - \varepsilon_\beta) \beta (1 - \varepsilon_p) p q) L_v(t) \\ \frac{dF_v(t)}{dt} &= (1 - \varepsilon_\beta) \beta (1 - \varepsilon_p) p S_v(t) - (1 - \varepsilon_r) r F_v(t) + (1 - \varepsilon_\beta) \beta (1 - \varepsilon_p) p q L_v(t) \\ \frac{dD_v(t)}{dt} &= (1 - \varepsilon_r) r F_v(t) + r_L L_v(t) \end{aligned} \quad (11)$$

Cuyas soluciones son las mismas que las de la cohorte de control pero llevando a cabo las sustituciones que aparecen en 10:

Grupo $S_v(t)$:

$$S_v(t) = N_0 e^{-(1 - \varepsilon_\beta) \beta t} \quad (12)$$

Grupo $L_v(t)$:

$$L_v(t) = \frac{(1 - \varepsilon_\beta) \beta (1 - (1 - \varepsilon_p) p) N_0}{r_L + (1 - \varepsilon_\beta) \beta (1 - \varepsilon_p) p q - (1 - \varepsilon_\beta) \beta} \left[e^{-(1 - \varepsilon_\beta) \beta t} - e^{-(r_L + (1 - \varepsilon_\beta) \beta (1 - \varepsilon_p) p q) t} \right] \quad (13)$$

Grupo $F_v(t)$:

$$\begin{aligned}
F_v(t) = & K e^{-(1-\varepsilon_r)rt} + \frac{(1-\varepsilon_\beta)\beta(1-\varepsilon_p)pN_0}{(1-\varepsilon_r)r - (1-\varepsilon_\beta)\beta} \exp[-(1-\varepsilon_\beta)\beta t] \\
& + \frac{((1-\varepsilon_\beta)\beta)^2(1-\varepsilon_p)pq(1-(1-\varepsilon_p)p)N_0}{r_L + (1-\varepsilon_\beta)\beta(1-\varepsilon_p)pq - (1-\varepsilon_\beta)\beta} \left[\frac{\exp[-(1-\varepsilon_\beta)\beta t]}{(1-\varepsilon_r)r - (1-\varepsilon_\beta)\beta} \right. \\
& \left. - \frac{\exp[-(r_L + (1-\varepsilon_\beta)\beta(1-\varepsilon_p)pq)t]}{(1-\varepsilon_r)r - r_L - (1-\varepsilon_\beta)\beta(1-\varepsilon_p)pq} \right] \quad (14)
\end{aligned}$$

Siendo K igual a:

$$K = -(1-\varepsilon_\beta)\beta(1-\varepsilon_p)pN_0 \left[\frac{(1-\varepsilon_r)r - r_L - (1-\varepsilon_\beta)\beta(1-\varepsilon_p)pq - q(1-\varepsilon_\beta)\beta(1-(1-\varepsilon_p)p)}{((1-\varepsilon_r)r - (1-\varepsilon_\beta)\beta)((1-\varepsilon_r)r - r_L - (1-\varepsilon_\beta)\beta(1-\varepsilon_p)pq)} \right] \quad (15)$$

Finalmente, para el grupo $D_v(t)$:

$$D_v(t) = N_0 - S_v(t) - L_v(t) - F_v(t) \quad (16)$$

De forma compacta podemos escribir $D_v(t)$ como:

$$D_v(t) = N_0[1 - \alpha' e^{-(1-\varepsilon_\beta)\beta t} + \Omega' e^{-(1-\varepsilon_r)rt} + \gamma' e^{-(r_L + (1-\varepsilon_\beta)\beta(1-\varepsilon_p)pq)t}] \quad (17)$$

Siendo respectivamente:

$$\begin{aligned}
\alpha'(\beta, r, r_L, p, q, \varepsilon_\beta, \varepsilon_p, \varepsilon_r) = & (1-\varepsilon_r)rr_L + (1-\varepsilon_\beta)\beta[(1-\varepsilon_p)pq(1-\varepsilon_r)r - r_L - \\
& (1-\varepsilon_p)p(1-\varepsilon_r)r + (1-\varepsilon_p)pr_L] \\
\Omega'(\beta, r, r_L, p, q, \varepsilon_\beta, \varepsilon_p, \varepsilon_r) = & (1-\varepsilon_\beta)\beta p \left[\frac{(1-\varepsilon_r)r - r_L - q(1-\varepsilon_\beta)\beta}{((1-\varepsilon_r)r - (1-\varepsilon_\beta)\beta)((1-\varepsilon_r)r - r_L - (1-\varepsilon_\beta)\beta(1-\varepsilon_p)pq)} \right] \\
\gamma'(\beta, r, r_L, p, q, \varepsilon_\beta, \varepsilon_p, \varepsilon_r) = & \beta[(1-\varepsilon_r)r - r_L - (1-\varepsilon_p)p(1-\varepsilon_r)r + (1-\varepsilon_p)pr_L] \quad (18)
\end{aligned}$$

Como hemos introducido anteriormente, la gran ventaja de disponer de una solución analítica a nuestro modelo radica en que podremos plantear una ecuación que nos permita relacionar ε_p y ε_r con la fracción de enfermos entre cohortes al fin del ensayo ρ y la protección frente a infección ε_β . Disponiendo de dicha ecuación, como ρ siempre se puede obtener por simple inspección de las cohortes, si somos capaces de estimar de forma independiente dos de los otros 3 parámetros, esta ecuación nos permitirá obtener el restante de forma unívoca.

Planteamos la fracción de enfermos del mismo modo que en el estudio original, pero empleando las ecuaciones 7 y 16 evaluadas en $t = T$, tiempo en el que concluye el período de seguimiento de los pacientes en el ensayo.

$$\rho(T) = \frac{D_v(T)}{D_c(T)} = \frac{[1 - \alpha' e^{-(1-\varepsilon_\beta)\beta T} + \Omega' e^{-(1-\varepsilon_r)rT} + \gamma' e^{-(r_L + (1-\varepsilon_\beta)\beta(1-\varepsilon_p)pq)T}]}{[1 - \alpha e^{-\beta T} + \Omega e^{-rT} + \gamma e^{-(r_L + \beta pq)T}]} \quad (19)$$

Como el valor de $\rho(T)$ lo podemos obtener de nuestras simulaciones y el de ε_β mediante análisis de supervivencia, será siempre posible obtener una solución de forma gráfica, no así despejar un expresión analítica para ε_p .

En nuestro caso decidimos implementar métodos de resolución numérica de ecuaciones para hallar las raíces. Para asegurar la convergencia del método y que la función sea suave la definimos como:

$$f(\rho, \varepsilon_\beta, \varepsilon_p, \varepsilon_r) = \rho \cdot D_c(T) - D_v(T, \varepsilon_\beta, \varepsilon_p, \varepsilon_r) \quad (20)$$

Esta función así definida cambia de signo en torno a cero por lo que el método de la Bisección es aplicable para estimar el valor de ε_p que nos da la raíz 0. Destacamos, finalmente, que todas las ecuaciones anteriores han sido verificadas mediante una implementación numérica del método de Runge-Kutta, el cual puede encontrarse en el anexo I de este trabajo.

4. El problema de la degeneración: curva de parámetros compatibles con una determinada VE_{dis}

La degeneración que aparece en el análisis clásico mediante tiempos de supervivencia (que sólo nos proporciona información sobre VE_{dis} pero no sobre ε_r y ε_p) queda patente en nuestro modelo al existir dos posibles vías específicas de bloqueo de la fase activa de la TB, estado D en nuestras ecuaciones. Estas dos vías existen al margen de la prevención de la infección que dan las vacunas que se basan en β , siendo respectivamente:

- Vacunas basadas en r : En esta ruta la acción de la vacuna consiste en ralentizar la transición rápida (i.e. aumentar su periodo de incubación). Queda recogida en el parámetro ε_r .
- Vacunas basadas en p : La acción de la vacuna consiste en provocar que los individuos que se infectan de la TB desarrollen latencia lenta. La idea que subyace a este tipo de vacunas es que el estado de latencia lenta es un estado con mucha inercia, en el que el desarrollo de la enfermedad puede llegar tras tiempos del orden de décadas, si es que llega a producirse, siendo además proporcionada una protección adicional frente a segundas infecciones. El efecto de la vacuna por esta vía queda recogido en el parámetro ε_p .

Por supuesto, no existe ninguna restricción a que una vacuna proporcione protección en las dos vías al mismo tiempo, lo que permite diseñar vacunas mixtas y analizar el impacto de estas a posteriori para localizar aquella que es más efectiva a largo plazo.

A efectos de detallar mejor este efecto, introducimos para las diversas vacunas la notación $\alpha - \gamma$, donde α y γ representan el porcentaje de efecto que una vacuna tiene sobre cada vía. En esta notación α está asociado a la vía de latencia lenta (basada en ε_p). Por su parte, γ está asociado a la vía de latencia rápida (vacuna basada en ε_r). En este contexto resaltamos que una vacuna que actúa únicamente sobre la latencia rápida será denominada 0-100, mientras que una que base su efecto en la latencia lenta será denominada 100-0.

Ahora, en relación a los ensayos de nueva vacunas, la información de la que disponemos proviene de los controles periódicos que se realizan a los individuos de las cohortes. Por ejemplo, en [7], los test de infección (el llamado Quantiferon, que es el test comercial para diagnóstico de infección por M.tuberculosis más aceptado hoy en día) se realizaron cada 3 meses. En este trabajo adoptamos la misma periodicidad. De los controles podemos obtener información sobre el tiempo que tarda cada individuo en infectarse de la TB (si lo hace) y el tiempo que tarde en desarrollar la enfermedad (i.e llegar a estado D en el modelo). Al final del ensayo también disponemos del valor

de la ratio entre enfermos en cada cohorte ρ , que nos permite construir VE_{dis} . Ya introdujimos la distinción entre VE_{inf} y VE_{dis} : el primero da la lectura de la eficacia de la vacuna frente a infección (i.e, transiciones entre sano y cualquier estado de latencia) mientras que el segundo da información sobre la eficacia frente a desarrollo de TB (en fase activa). Este último se construye a partir de ρ como $VE_{dis} = 1 - \rho$.

Sabemos no obstante que VE_{dis} no da ninguna información acerca de la vía sobre la cual está actuando la vacuna y de hecho, existe toda una curva de posibles pares de ε_p - ε_r que son compatibles con una lectura de VE_{dis} . Esta relación nos la da la curva dada por la ecuación 19 para unos ρ y ε_β fijados. Si tomamos como referencia $VE_{dis} = 0.5$ ($\rho = 0.5$) y $\varepsilon_\beta = 0.0$ y resolvemos numéricamente la ligadura encontramos la curva de parámetros compatibles de la figura 4. En

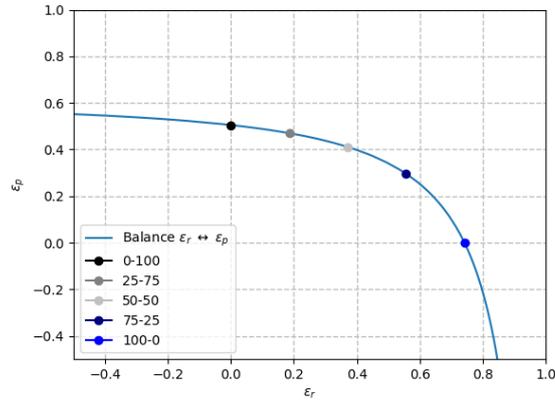


Figura 4: Curva de parámetros ε_p - ε_r compatibles con una lectura de $VE_{dis} = 0.5$ obtenida tras la realización de un ensayo para el caso de una vacuna cuya protección es exclusivamente contra enfermedad ($\varepsilon_\beta = 0$). Sobre la curva se marcan 5 vacunas de referencia correspondientes a vacunas mixtas o basadas únicamente en una vía (100-0 para r y 0-100 para p).

ella hemos marcado 5 vacunas de referencia en las que el efecto viene determinado por una u otra vía en ciertas proporciones. Dichos parámetros de las vacunas nos servirán de referencia para lo que sigue, principalmente para calibrar el modelo en la cohorte vacunada y poder realizar simulaciones in silico con las que obtener las lecturas que obtendríamos en el mundo real al realizar un ensayo de una nueva vacuna.

Por otro lado, este ejercicio puede repetirse para recalibrar los parámetros de las vacunas dada una cierta eficacia en infección (i.e, $\varepsilon_\beta \neq 0$), caso que corresponde a haber recibido previamente algún otro tratamiento o alguna otra vacuna. La curva de parámetros compatibles en esta situación, dado $VE_{dis} = 0.5$ ($\rho = 0.5$) y $\varepsilon_\beta = 0.25$ viene dada en la figura 5.

Los parámetros de las distintas vacunas en ambos casos quedan recogidos en la tabla 1.

5. Rompiendo la degeneración: estimación independiente de ε_r

En la metodología propuesta anteriormente disponemos de una ecuación de ligadura entre ρ , ε_r , ε_β y ε_p , por lo que conociendo 3 de los 4 parámetros, mediante resolución numérica, deberemos ser capaces de obtener el restante. El problema viene cuando nos damos cuenta de que es necesario estimar ε_r de forma totalmente independiente.

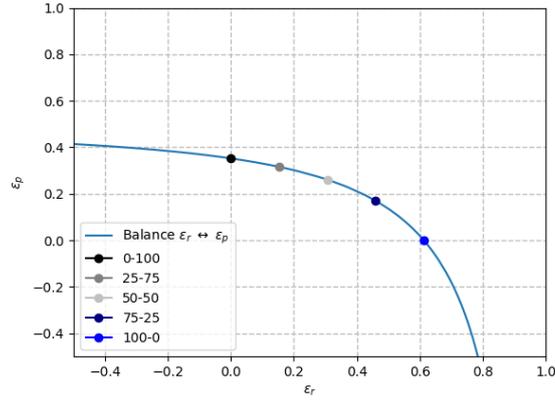


Figura 5: Curva de parámetros ε_p - ε_r compatibles con una lectura de $VE_{dis} = 0.5$ obtenida tras la realización de un ensayo en el que la vacuna proporciona además protección frente a infección ($\varepsilon_\beta = 0.25$). Sobre la curva se marcan 5 vacunas de referencia correspondientes a vacunas mixtas o basadas únicamente en una vía (100-0 para r y 0-100 para p).

$\varepsilon_\beta = 0.0$	ε_r	ε_p
100-0	0.0	0.495
75- 25	0.19	0.47
50-50	0.37	0.41
25-75	0.56	0.30
0-100	0.742	0.0

$\varepsilon_\beta = 0.25$	ε_r	ε_p
100-0	0.0	0.352
75- 25	0.153	0.316
50-50	0.306	0.261
25-75	0.459	0.169
0-100	0.613	0.0

Tabla 1: Parámetros de vacunas basadas en vías determinadas o en combinaciones de ambas para situaciones sin protección ante infección $\varepsilon_\beta = 0$ y con protección ante infección $\varepsilon_\beta = 0.25$. En ambos casos de parte de una lectura de $VE_{dis} = 0.5$

Para lograr nuestro objetivo vamos a utilizar la duración de las transiciones a enfermedad como estimador estadístico. Esto implica que debemos registrar en el ensayo el tiempo en el cual un individuo se infecta y el tiempo que le cuesta desarrollar la enfermedad en el caso en que lo haga. La justificación viene cuando nos damos cuenta de que las escalas temporales entre los procesos analizados en el estudio son muy diferentes, como queda recogido en la relación $r \gg r_L$, siendo las transiciones rápidas del orden de un año y las reinfecciones y transiciones lentas del orden de décadas. Esto nos permite asumir, a la hora de analizar el tiempo que han tardado los individuos en desarrollar la TB, que prácticamente todos los individuos lo han hecho por vía rápida.

Asumiendo que los tiempos de estos individuos siguen una distribución de Poisson, la PDF de la variable $t = t_{dis} - t_{inf}$ es básicamente una exponencial dada por $f(t|r) = re^{-rt}$, siendo el valor esperado $\langle t \rangle = 1/r$. En los ensayos de campo, no obstante, el tiempo de seguimiento no puede aumentarse *ad infinitum*, por lo que existe una cota máxima con la cual debemos truncar la distribución para dar cuenta de que $t_{max} = T - t_{inf}$

$$\langle t \rangle (t_{inf}) = \frac{\int_0^{t_{max}} tf(t|r)dt}{\int_0^{t_{max}} f(t|r)dt} = \frac{1}{r} - \frac{(T - t_{inf}) \cdot e^{-r(T-t_{inf})}}{1 - e^{-r(T-t_{inf})}} \quad (21)$$

Las ecuación anterior describe cómo se espera que los tiempos de transición observados de la infección a la enfermedad estén sesgados hacia valores más bajos a medida que las infecciones

ocurren más tarde durante el ensayo. Esto se debe simplemente a que cuanto más tarde se produce la infección, queda menos tiempo disponible para observar una transición a enfermedad. Estas expresiones nos permiten aislar el efecto de ese sesgo e inferir, usando solo datos de individuos que desarrollan TB durante el ensayo, la tasa de transición r dentro de la cohorte de control, usando MLE (paquete `bbmle` de R) junto con sus intervalos de confianza (al 95 %).

Repetiendo este ejercicio para la cohorte vacunada podremos obtener la tasa de transición r^v correspondiente, la cual, si el efecto de la vacuna se da sobre esta vía, tendrá que ser inferior a la reportada en la cohorte de control. Tras este proceso y mediante una sencilla relación obtenida por simple inspección de las figuras en 2, podemos obtener ε_r :

$$r^v = (1 - \varepsilon_r) \cdot r \longrightarrow \varepsilon_r = 1 - \frac{r^v}{r} \quad (22)$$

Cuyo intervalo de confianza puede obtenerse sin más que propagar los obtenidos tras realizar el MLE para las tasas de transición al ser independientes. Un ejemplo de los ajustes mencionados puede verse en el panel de la figura 6. En ella mostramos los ajustes que se realizan para la cohorte de control y para la cohorte vacunada para las vacunas 0-100 y 100-0 compatibles con una lectura de VE_{dis} en un ensayo.

Para demostrar la viabilidad de estos ajustes, hemos aplicado nuestra metodología de inferencia de ε_r sobre datos de ensayos clínicos simulados *in silico*. El procedimiento numérico que seguimos para simular dichos ensayos se explica en detalle en la sección III de este TFM. Como se puede ver en la figura 6, en un ensayo clínico de las dimensiones y características del completado para la MVA85A en Worcester, Sudáfrica, nuestro método produciría estimaciones satisfactorias de los parámetros r y r^v , y por tanto de la eficacia vacunal con respecto al ritmo de transición, ε_r .

Esto implica finalmente que ahora disponemos de una herramienta capaz de estimar de forma independiente la protección en una de las dos posibles vías a enfermedad, siendo ésta clave para romper la degeneración en los resultados de los ensayos.

6. Método de estimación de ε_p : Parámetros relevantes.

Obtuvimos antes una ligadura para ε_p y ε_r dada por la ecuación 19. No obstante, esta ecuación presenta todavía dos dependencias que no conocemos: ρ y ε_β , al margen de las propias ε_p y ε_r .

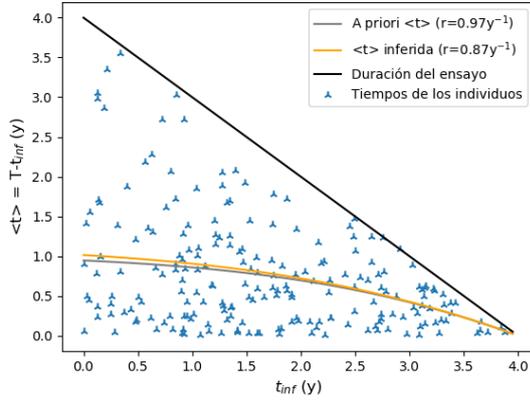
La obtención de ρ es directa del resultado de un ensayo clínico, pues se define como $\rho = \frac{D_v}{D_c}$, es decir, el cociente entre los individuos que desarrollan fase activa de la TB en el tiempo de seguimiento.

Por su parte, ε_β también puede obtenerse de forma sencilla a través de VE_{inf} , ya que es simplemente $\varepsilon_\beta = VE_{\text{inf}}$, que se calcula aplicando una regresión de Cox a los tiempos de infección.

En otras palabras, ε_β corresponde trivialmente con la protección frente a infección, la cual puede ser obtenida mediante el paquete “`survival`” de R, que incluye funciones para realizar este tipo de ajustes, llamados regresión de Cox. De este ajuste obtendremos ε_β así como su intervalo de confianza.

Esto implica que ya hemos obtenido todos los ingredientes necesarios para poder resolver la ecuación 19: VE_{inf} , VE_{dis} y ε_r .

(a) Cohorte de control



(b) Cohorte vacunada

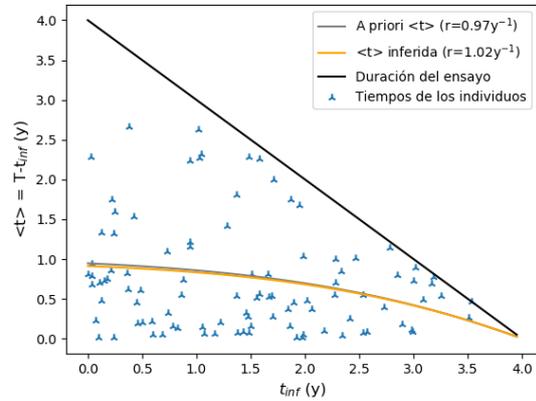
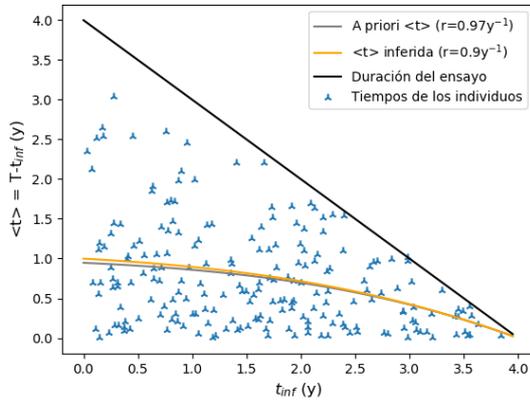
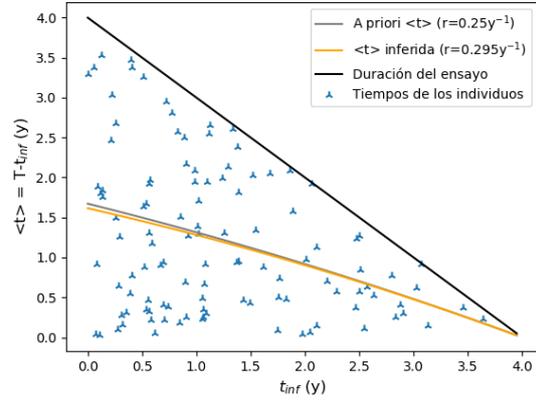


Figura 6: Ajustes por MLE a la PDF de los tiempos de transición rápida para las dos cohortes. Las dos figuras superiores muestran el caso de la vacuna basada en r (0-100) y las figuras inferiores muestran el caso de la vacuna basada en p (100-0). Los valores de r y r^v teóricos responden a la calibración del modelo estocástico que se describe en la sección III.

Parte III

Verificación de la validez del método.

Una implementación numérica para generar los resultados que daría un ensayo de campo como el planteado en Worcester para la evaluación de la eficacia de la vacuna MVA85A [7], introducida en el Anexo II, nos permite disponer de una herramienta de verificación de la metodología propuesta en este trabajo. Con ella podremos simular los ensayos y disponer de la información que obtendríamos de estos, información a la que aplicaremos nuestro método de ruptura de la degeneración para obtener la protección de la vacuna en cada vía.

Para implementar dicha plataforma primero han de calibrarse los parámetros que describan la situación epidemiológica en el lugar del ensayo que se pretende simular y para las poblaciones

representativas de las cohortes del ensayo. Todo este proceso se detalla en el anexo II, en el que se introduce el modelo numérico así como su calibración en Worcester.

La validación del modelo esta dividida en varias secciones, puesto que primero vamos a verificar la validez de la hipótesis de que todas las transiciones a enfermedad vienen dadas por la latencia rápida durante el tiempo de seguimiento (duración del ensayo).

7. Hipótesis de transiciones rápidas.

Hemos realizado anteriormente, al describir el método de obtención de ε_r , la hipótesis de que la contribución principal al número de individuos que desarrollan la TB al finalizar el ensayo viene por la vía de latencia rápida, siendo las contribuciones de reinfecciones y latencia lenta prácticamente marginales.

El efecto que cabe esperar de ellas es el de añadir eventos en la parte superior izquierda de las gráficas de eventos de la figura 6. Esto es, las reinfecciones corresponden a transiciones de tiempo medio muy elevado, dado que un individuo debe pasar por latencia lenta, lo que implica infectarse, y además sufrir una segunda reinfección, la cual será menos probable. Esta reducción en la probabilidad de infección viene marcada por el factor q y tiene origen en la activación del sistema inmune frente a la infección al haber sufrido una primera.

En el caso que nos ocupa, al estar trabajando con ensayos truncados a 4 años, y dado que el factor βpq correspondiente es absurdamente pequeño ($\beta pq = 0.00543 \text{ y}^{-1}$, un orden de magnitud por debajo de las infecciones primarias), no esperamos que dichas transiciones representen un factor relevante. Para verificar esta hipótesis podemos seguir la evolución de los individuos, rastreando las transiciones que realizan para analizar el peso que cada tipo tiene en el total de enfermos al finalizar el ensayo.

7.1. Flujos de individuos en el modelo

Podemos verificar esta hipótesis mediante la implementación del modelo estocástico presente en el anexo II, con el que simulamos *in silico* el desarrollo de un ensayo para las dos cohortes. En él, para cada individuo que enferma podemos rastrear la ruta por la que lo hizo sin más que registrar los distintos estados por los que éste pasa. Las posibles rutas que existen se muestran en la figura 7.

Para ello lanzamos 500 simulaciones de cada cohorte con un modelo de transmisión simple (ver Anexo II) y analizamos la contribución a cada flujo utilizando como parámetros de la vacuna para la cohorte vacunada los correspondientes a una vacuna $\alpha = 50\% - \gamma = 50\%$ ($\varepsilon_\beta = 0$, $\varepsilon_r = 0.41$, $\varepsilon_p = 0.37$) obtenidos de la curva de compatibilidad con $\text{VE}_{\text{dis}} = 0.5$ al finalizar el ensayo. En este caso la vacuna utilizada no es relevante ya que sólo nos proponemos estudiar la influencia de las transiciones lentas en el resultado final. Los valores de los parámetros asociados a las transiciones, por su parte corresponden a los mostrados en la tabla 2 y los resultados de realizar 500 simulaciones se muestran en la figura 8.

Se observa muy claramente que la contribución que suponen las transiciones lentas es prácticamente nula, pues para las dos cohortes se encuentran por debajo del 1% de los eventos

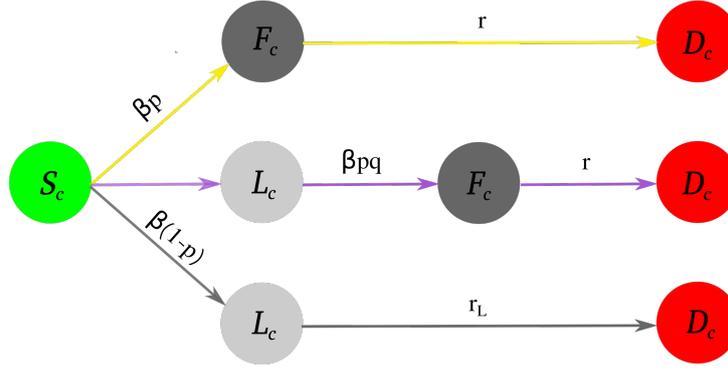


Figura 7: Posibles rutas hasta desarrollo de la enfermedad para la cohorte de control en el modelo del ensayo completo. Distinguimos por color: Latencia rápida (amarillo), Reinfeción (morado) y Latencia lenta (gris).

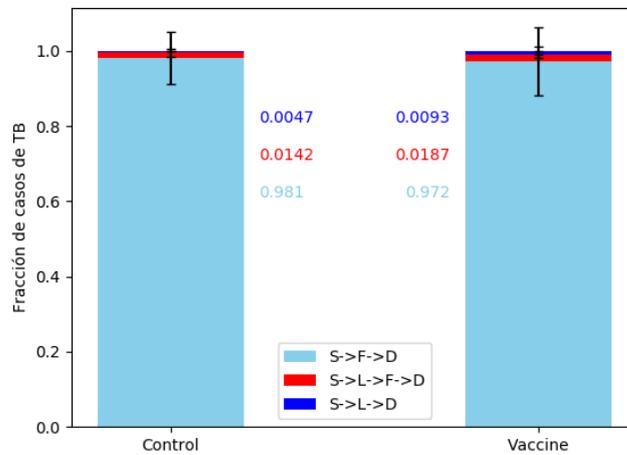


Figura 8: Contribución al número de individuos que desarrollan la TB al final del ensayo ($T = 4.0$ y). Distinguimos por color: Latencia rápida (azul claro), Reinfeción (rojo) y Latencia lenta (azul oscuro).

registrados.

Este resultado es muy relevante dado que legitima nuestro método de inferencia de ε_r . Sin importar lo compleja que sea la dinámica real de transmisión de la TB, en el contexto de un ensayo, prácticamente la totalidad de las transiciones a enfermedad corresponden a progresión desde latencia rápida.

8. Validación para vacunas 100-0 y 0-100

Vamos a emplear toda la potencia computacional de la que disponemos para aplicar el método que hemos descrito en la parte anterior destinado a romper la degeneración en la estimación de los parámetros que caracterizan la vacuna. Emplearemos para ello el modelo estocástico descrito en el Anexo II con el cual simulamos un ensayo completo, tanto de la cohorte de control como de la vacunada, de forma independiente. De esta simulación obtenemos los tiempos de transición de cada individuo. Con ellos podemos obtener las gráficas presentes en la figura 6, utilizando el paquete bbmle de R para realizar el ajuste a la ecuación 21 y así obtener los valores de las tasas

r y r^v . A continuación obtenemos la protección en latencia rápida ε_r mediante la ecuación 22 y su error propagando el de las tasas que hemos obtenido mediante R.

También obtenemos de forma independiente ε_β mediante análisis clásico de supervivencia (regresión de Cox), con lo que disponemos de suficiente información para resolver la ligadura entre ρ , ε_p , ε_β y ε_R , i.e, obtenemos ε_p cuyo error hemos de propagar numéricamente ya que no hemos podido despejar ε_p explícitamente de la condición de ligadura dada por la ecuación 19.

En este punto hemos caracterizado completamente la vacuna para esta ejecución del modelo, siendo el proceso resumido en la figura 9.

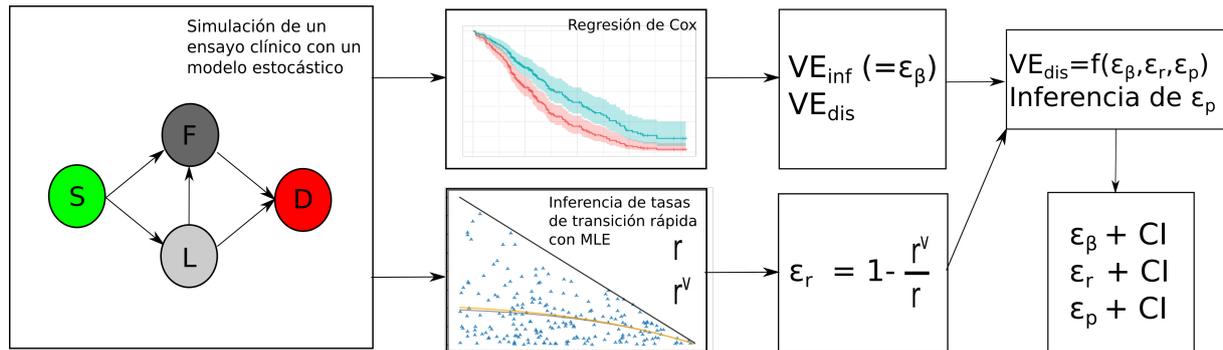


Figura 9: Esquema de la ruta de trabajo planteada. Se registran los tiempos de transición a enfermedad de los individuos que contraen la TB desde estados de latencia. Con ellos se estiman las tasas de transición entre cohortes r y r^v , con las cuales se estima ε_r . Finalmente se utiliza su ligadura $f(\varepsilon_p, \varepsilon_r)$ junto al análisis por Cox para obtener ε_p y su CI se obtiene propagando numéricamente las incertidumbres en los otros parámetros.

Como estamos trabajando con un modelo estocástico serán necesarias muchas iteraciones del proceso completo para estimar los parámetros inferidos a partir de suficiente estadística. Tomamos como ruta a seguir la realización de 500 iteraciones completas de los que obtendremos 500 estimaciones independientes de los parámetros, los cuales podremos combinar para dar lugar al resultado final. La figura 10 ilustra cómo planteamos la utilización de nuestro método para inferir globalmente los parámetros de una vacuna.

Notamos que cada una de las iteraciones con las que se obtiene una caracterización completa de la vacuna corresponde a un ciclo como el mostrado en la figura 9.

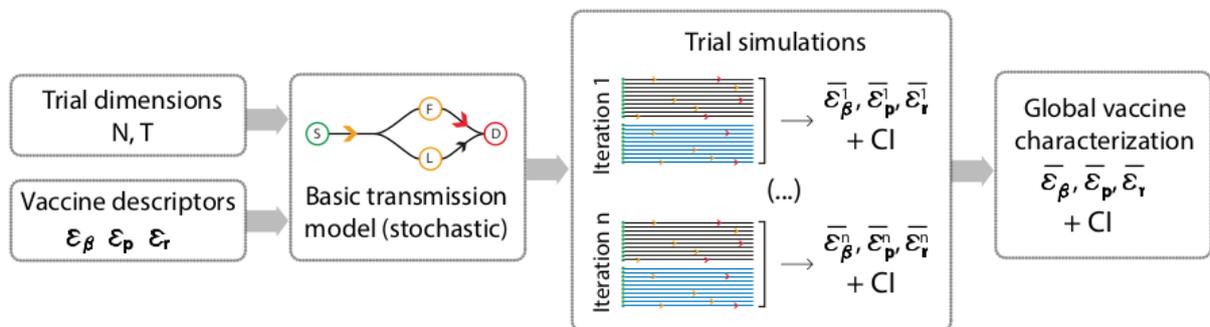


Figura 10: Esquema de trabajo orientado a la validación de la metodología propuesta. Empleando el modelo estocástico como herramienta para simular ensayos obtenemos los tiempos de transición rápida y VE_{dis} . Se aplica nuestro método para inferir la protección de la vacuna y este proceso se reitera por n iteraciones, cuyos resultados se combinan para dar una caracterización global de la vacuna.

Respecto a la obtención de los valores finales, será necesario combinar las incertidumbres de todos parámetros de protección obtenidos. Para ello, asumiremos que estos parámetros provienen de distribuciones normales, con las que se obtienen los σ 's asociados. Estos se promedian para dar lugar a una distribución promedio de los parámetros, con la que se obtienen los intervalos de confianza asociados a las medianas de las distribuciones, las cuales se usan como estimador.

Para los parámetros calibrados en la tabla 2, i.e, la situación en el ensayo propuesto en por D.Tameris *et al* [7] nuestro modelo da lugar a los resultados obtenidos en la figura 11 para el caso de una vacuna que actúe totalmente sobre latencia rápida compatible con una observación de $VE_{dis} = 0.5$ al finalizar el ensayo ($\varepsilon_\beta = 0$, $\varepsilon_r = 0.74$, $\varepsilon_p = 0$).

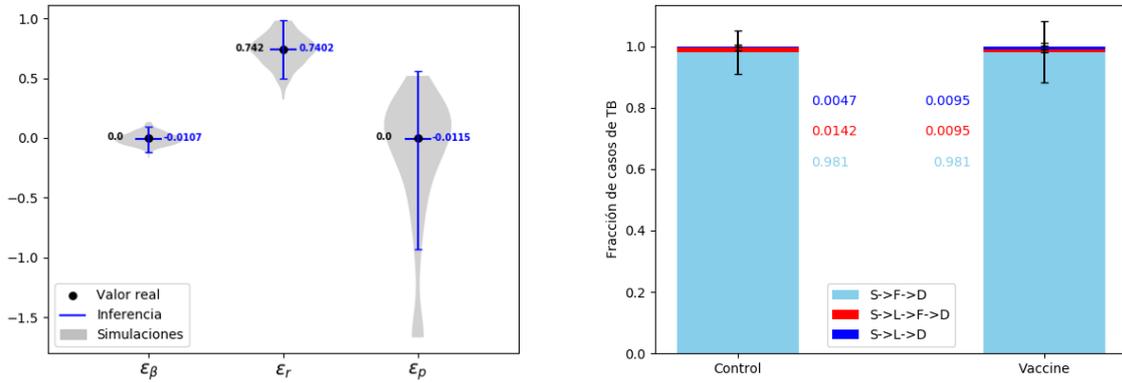


Figura 11: Resultados de la ejecución del modelo con 500 iteraciones para una vacuna compatible con $VE_{dis} = 0.5$. El panel izquierdo muestra la inferencia de los parámetros de la vacuna basada en ε_r ($\varepsilon_\beta = 0$, $\varepsilon_p = 0$, $\varepsilon_r = 0.74$) combinando los resultados de las 500 iteraciones. El panel derecho muestra el promedio a las 500 iteraciones de los flujos de individuos para verificar que la contribución principal corresponde a latencia rápida.

Por otro lado, para el caso de una vacuna que actúe bloqueando el desarrollo de latencia rápida, compatible con una observación de $VE_{dis} = 0.5$ al finalizar el ensayo ($\varepsilon_\beta = 0$, $\varepsilon_r = 0$, $\varepsilon_p = 0.5$) obtenemos los resultados de la figura 12.

Observamos a raíz de los resultados en 11 y 12 que somos capaces de inferir de forma no sesgada el valor real de la protección que proporciona la vacuna sobre cada una de las posibles vías para vacunas basadas al 100 % en alguna de las vías.

No obstante, la dispersion de los parámetros inferidos en distintas realizaciones de la simulación del ensayo es grande. Esto es en gran medida esperado, puesto que, dado lo limitado de la estadística, las variaciones entre ensayos completados en el mismo lugar y población pueden ser muy importantes.

8.1. Análisis de fallos del método

Existen en el modelo algunos criterios que se espera que cumplan los valores de eficacia de la vacuna. Se podría esperar que éstas estuviesen comprendidas entre 0 y 1, donde 1 significaría una eficacia perfecta y 0 ningún efecto. Sin embargo, es posible que una vacuna tenga un efecto negativo. En el caso de las eficacias que afectan las transiciones (es decir, ε_β y ε_r) no hay un límite

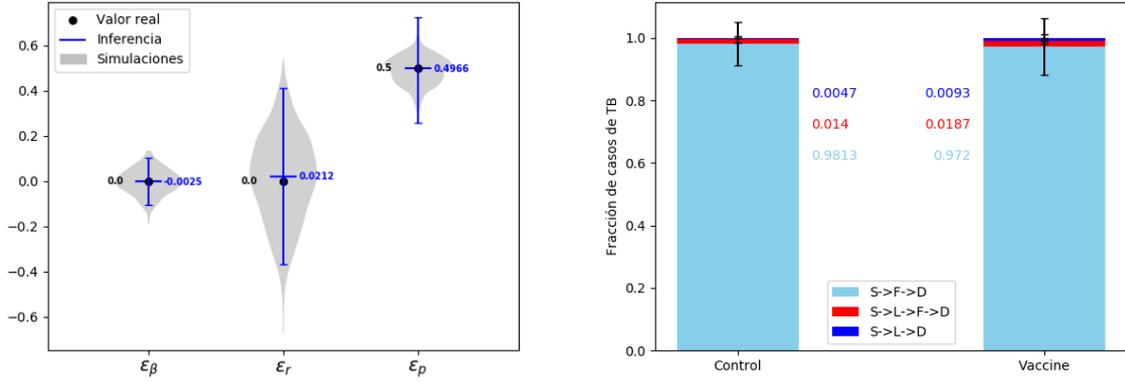


Figura 12: Resultados de la ejecución del modelo con 500 iteraciones para una vacuna compatible con $VE_{dis} = 0.5$. El panel izquierdo muestra la inferencia de los parámetros de la vacuna basada en ϵ_p ($\epsilon_\beta = 0$, $\epsilon_p = 0.5$, $\epsilon_r = 0$) combinando los resultados de las 500 iteraciones. El panel derecho muestra el promedio a las 500 iteraciones de los flujos de individuos para verificar que la contribución principal corresponde a latencia rápida.

inferior formal y una tasa modificada $((1 - \epsilon_r)r)$ igual a ∞ (asociada a $\epsilon = -\infty$) significaría una transición instantánea.

Respecto a ϵ_p , que funciona como modificador de una probabilidad, resulta que ésta debe estar comprendido entre 0 y 1, lo que introduce un límite inferior para ϵ_p :

$$\epsilon_p^{\min} = 1 - \frac{1}{p} \quad (23)$$

Notamos que para entornos estadísticos deficientes, normalmente en el caso de vacunas centradas en limitar la progresión rápida (ϵ_p), algunas realizaciones de los ensayos resultan en estimaciones que van más allá de los intervalos epidemiológicamente significativos para los parámetros. Siendo conscientes de este problema, nos proponemos cuantificar la fracción de realizaciones válidas en función de la estadística disponible (entendiendo por válidas aquellas realizaciones donde todos los parámetros están contenidos en los límites mencionados anteriormente). Para ello exploramos el efecto que tiene el tiempo de seguimiento, principal influencia de la cantidad de individuos que realizan transición a enfermedad, para las 5 vacunas de referencia exploradas en la tabla 1, en el caso $\epsilon_\beta = 0$.

Otro factor determinante en la cantidad de estadística disponible es por supuesto el tamaño de las cohortes, pues una muestra mayor también produce mayor número de transiciones. Exploramos dos entornos alternativos para el número de individuos por cohorte, con las mismas vacunas de referencia que en la figura 13.

Recuperamos un comportamiento plausible teniendo en cuenta que la mayor limitación computacional del modelo viene dada por la estadística disponible para realizar los ajustes, pues cuanto mayor sea ésta, mejor somos capaces de estimar la tasa de transición rápida con nuestro ajuste por MLE.

De este modo, cuando variamos el tamaño de la cohorte hacia valores inferiores (ver figura 14a) nuestras curvas colapsan más tarde a 1, es decir, se desplazan verticalmente hacia valores inferiores de la fracción de ensayos sintéticos válidos. Por otro lado, aumenar el tamaño de a

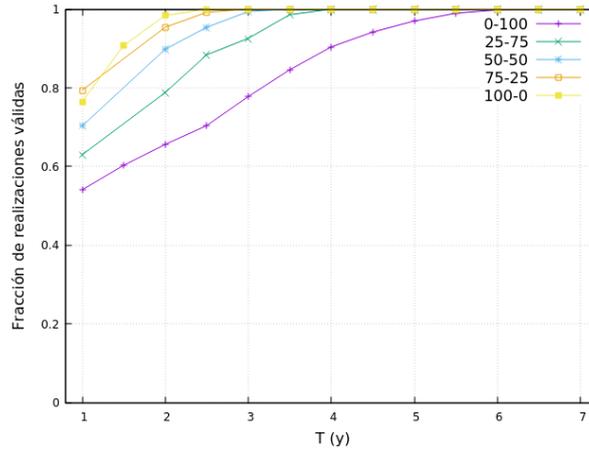
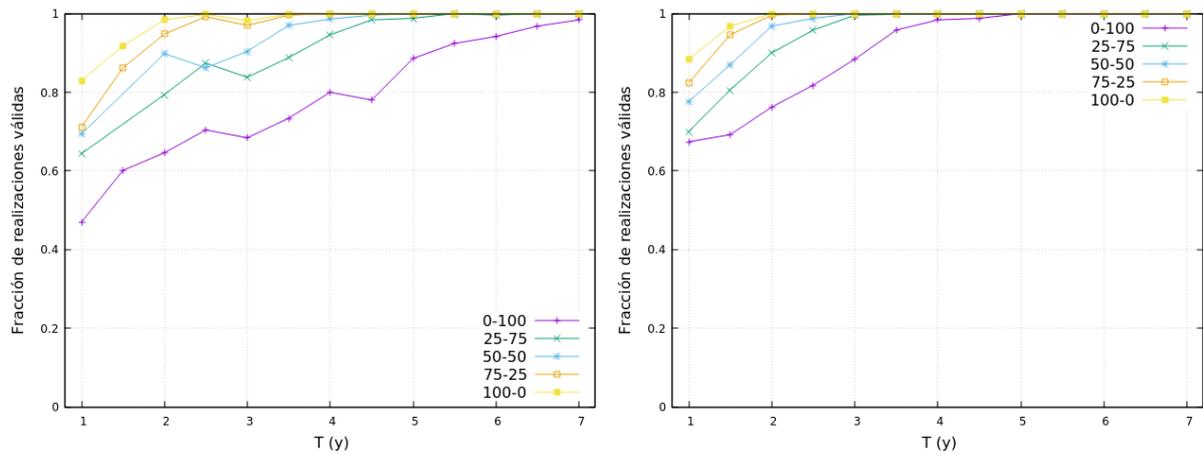


Figura 13: Fracción de realizaciones correctas de un ensayo (i.e., realizaciones a partir de las cuales obtenemos una caracterización biológicamente plausible) en función del tiempo de seguimiento, para tamaño de cohorte $N = 3000$ y para 5 vacunas diferentes



(a) Tamaño de la cohorte $N = 1000$ individuos.

(b) Tamaño de la cohorte $N = 7000$ individuos.

Figura 14: Fracción de las realizaciones correctas de un ensayo (es decir, realizaciones a partir de las cuales obtenemos una caracterización biológicamente plausible) en función del período de seguimiento, para tres tamaños de cohorte diferentes y 5 vacunas diferentes

cohorte (ver 14b) más individuos realizan la transición rápida a enfermedad, lo que aumenta los puntos con los que ajustamos y nos lleva a una mejor estimación de los parámetros de la vacuna, como vemos por el aumento significativo del número de ensayos válidos.

Este análisis nos permite también lanzar un mensaje directo a los técnicos encargados del establecimiento de un ensayo. Si plantean utilizar el método descrito en este trabajo, el análisis de fallos puede ayudarles a decidir sobre el tamaño de las cohortes así como el tiempo de seguimiento necesario para tratar de maximizar la probabilidad de que el resultado obtenido sea correcto.

8.2. Caso con protección frente a infección

Finalmente, con ánimo de demostrar la rebuetez del modelo así como su posible adaptación a escenarios alternativos, proponemos utilizarlo para calcular el efecto de las vacunas basadas en ε_p (100-0) y en ε_r (0-100) en el caso en que la vacuna ofrece protección contra infección (es decir

$\varepsilon_\beta = 0.25$).

Tal y como hemos indicado en la parte II de la memoria, estamos estimando ε_β a partir de VE_{inf} empleando una regresión de Cox (paquete “survival” de R). Esta estimación se basa en la comparación entre las transiciones de sano a cualquiera de los estados de latencia (infectado) entre las cohortes de control y vacunada. Esto implica que podemos estimar también vacunas con protección frente a infección.

Los resultados obtenidos se muestran en la figura 15 para el caso de la vacuna (0-100). Se toma un tamaño de cohorte de $N = 3000$ así como los parámetros usuales mostrados en 2.

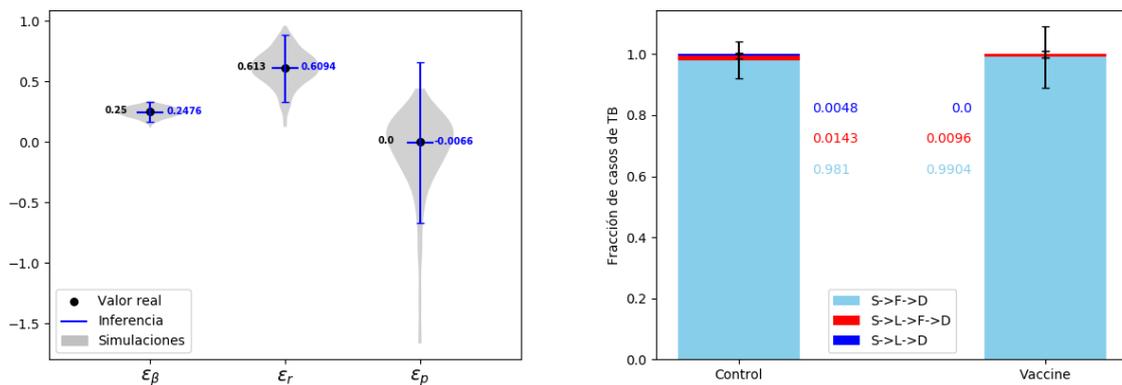


Figura 15: Resultados de la ejecución del modelo con 500 iteraciones para una vacuna compatible con $VE_{dis} = 0.5$ y con $\varepsilon_\beta = 0.25$, $\varepsilon_p = 0$, $\varepsilon_r = 0.613$. El panel izquierdo muestra la distribución de probabilidad de los parámetros de la vacuna basada en ε_r combinando los resultados de las 500 iteraciones. El panel derecho muestra el promedio a las 500 iteraciones de los flujos de individuos para verificar que la contribución principal corresponde a latencia rápida.

Por otro lado, para el caso de la vacuna basada en p (100-0) recuperamos los resultados mostrados en la figura 16.

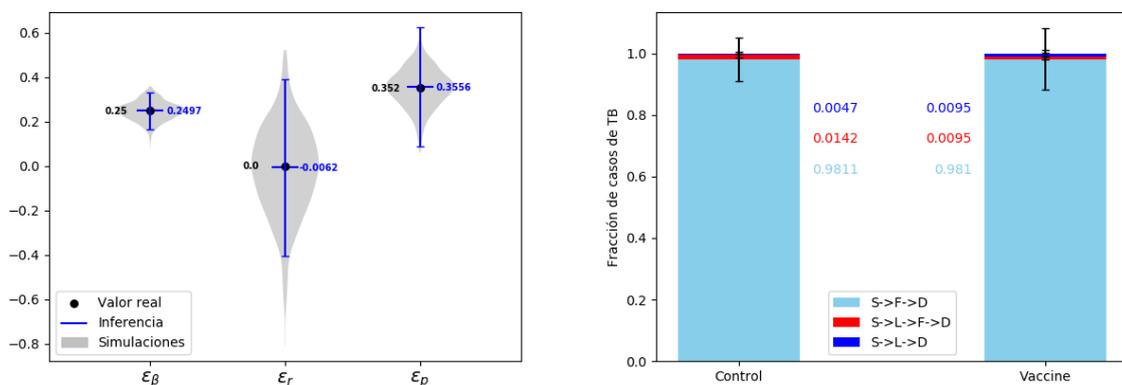


Figura 16: Resultados de la ejecución del modelo con 500 iteraciones para una vacuna compatible con $VE_{dis} = 0.5$ con $\varepsilon_\beta = 0.25$, $\varepsilon_p = 0.352$, $\varepsilon_r = 0$. El panel izquierdo muestra la distribución de probabilidad de los parámetros de la vacuna basada en ε_p combinando los resultados de las 500 iteraciones. El panel derecho muestra el promedio a las 500 iteraciones de los flujos de individuos para verificar que la contribución principal corresponde a latencia rápida.

Ambos casos muestran la versatilidad del modelo, siempre dentro de sus propias limitaciones, pero que nos permite jugar con escenarios alternativos al propuesto como hilo conductor de este trabajo. Esto hace que la metodología aquí propuesta tenga aplicación en la selección del escenario ideal para realizar un ensayo de campo.

9. Robustez en escenarios alternativos

Para verificar que el modelo que proponemos se puede calibrar y utilizar en escenarios alternativos al principal estudiado en este trabajo, nos proponemos realizar una serie de barridos en los parámetros de calibración.

Esta decisión viene motivada por el hecho de que algunos de los parámetros con los que estamos simulando los ensayos (ver tabla 2 en el anexo II) responden a un proceso de calibración respecto a un ensayo muy determinado (ver referencia [7]) y queremos demostrar que si bien el escenario principal es el de dicho ensayo, recalibraciones propias son posibles para adaptar nuestro método a otros posibles escenarios. Proponemos los siguientes parámetros sobre los que realizar los barridos:

- Tiempo de seguimiento: A la hora de realizar un ensayo de campo puede ser un factor terriblemente determinante, pues cuanto más largo sea un ensayo más recursos necesita y más compleja puede resultar su realización
- Fuerza de infección β : En nuestro caso este parámetro se ha calibrado para reproducir los resultados del ensayo MVA85A, pero calibraciones similares se pueden realizar para reproducir los resultados de otros ensayos existentes, por lo que la necesidad de variar este parámetro resulta evidente.
- Probabilidad p : Este parámetro es la estimación más reciente en nuestro conocimiento sobre la distribución de casos de latencia rápida y lenta para neonatos, pero en otros grupos de edad su valor es diferente. Resulta por ello natural plantearlo para los barridos.
- Protección frente a reinfección: El valor de este parámetro es susceptible de cambio en escenarios en los cuales puede existir interacción con otros factores externos a la propia TB y que tengan como efecto aumentar o disminuir la respuesta del sistema inmune de un individuo infectado frente a una segunda infección. Adicionalmente, este parámetro es relevante dado que no existe un total consenso en la literatura acerca de su valor, como puede verse en [9], en donde una estimación diferente es empleada.

Podremos ver que, salvo en el caso de escenarios estadísticamente deficientes (i.e $p = 0.1$ ó $T = 2$ años) somos capaces de estimar correctamente los parámetros de la vacuna en todos los escenarios para la vacuna basada en ε_r . Se ha elegido esta por ser en la que los ajustes por MLE juegan el papel determinante. Como complemento, en el anexo III se muestran los resultados para la vacuna basada en ε_p en 3 de estos escenarios.

9.1. Barrido en el tiempo de seguimiento

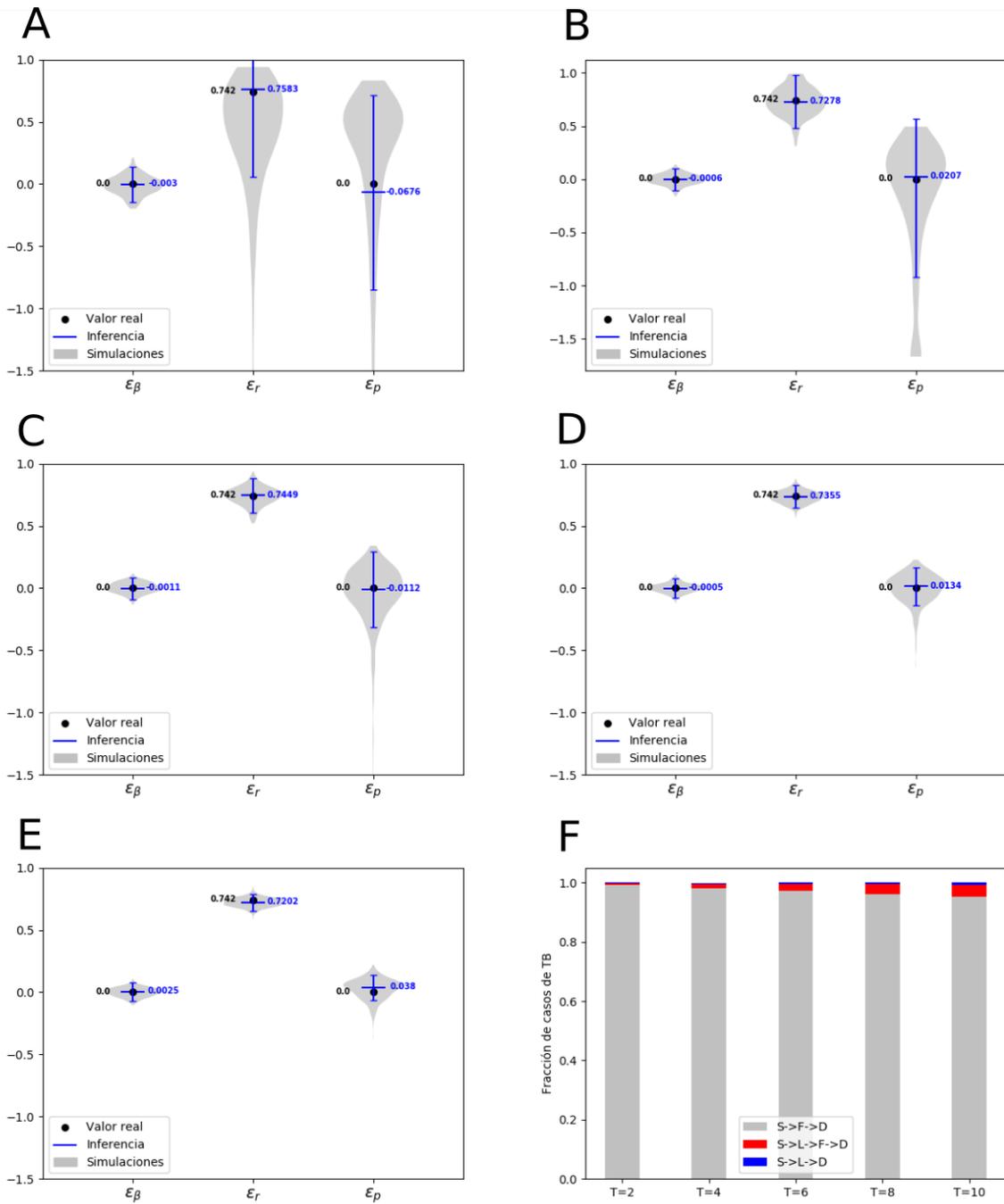


Figura 17: Resultados tras 500 iteraciones para A) $T = 2.0$ y, B) $T = 4.0$ (valor de referencia) y, C) $T = 6.0$ y, D) $T = 8.0$ y, E) $T = 10.0$ y. La figura F) muestra las contribuciones en cada caso a los individuos que desarrollan TB al final del ensayo para verificar que la hipótesis de transición rápida sigue siendo válida. En todos los casos se ha trabajado con tamaños de cohorte $N = 3000$ y con una vacuna basada en ϵ_r . ($\epsilon_\beta = 0.0$, $\epsilon_p = 0$, $\epsilon_r = 0.74$)

9.2. Barrido en la fuerza de infección

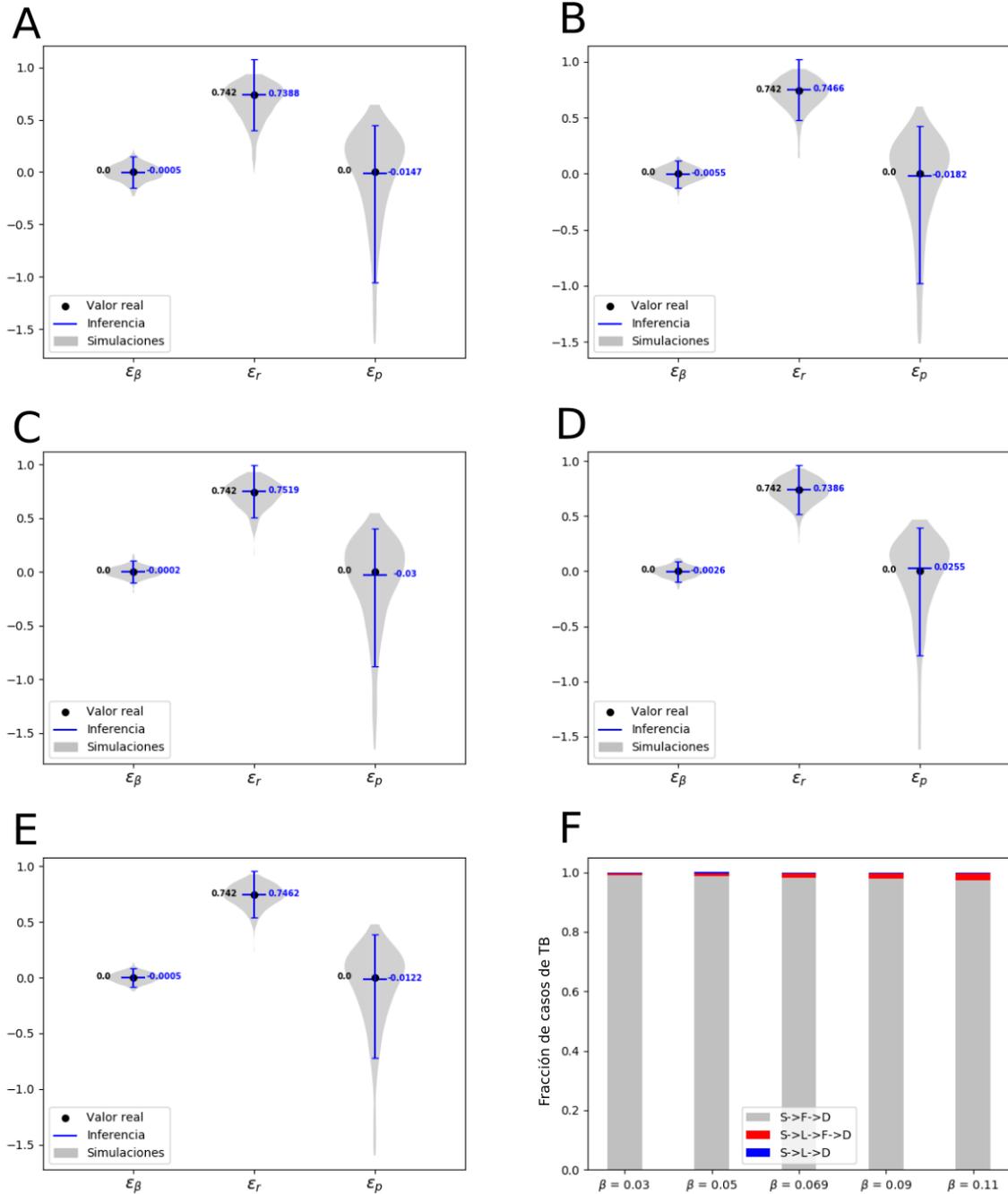


Figura 18: Resultados tras 500 iteraciones para A) $\beta = 0.03 \text{ y}^{-1}$, B) $\beta = 0.05 \text{ y}^{-1}$, C) $\beta = 0.069 \text{ y}^{-1}$ (valor de referencia), D) $\beta = 0.09 \text{ y}^{-1}$, E) $\beta = 0.11 \text{ y}^{-1}$. La figura F) muestra las contribuciones en cada caso a los individuos que desarrollan TB al final del ensayo en la cohorte de control para verificar que la hipótesis de transición rápida sigue siendo válida. En todos los casos se ha trabajado con tamaños de cohorte $N = 3000$ y con una vacuna basada en ϵ_r . ($\epsilon_\beta = 0.0$, $\epsilon_p = 0$, $\epsilon_r = 0.74$)

9.3. Barrido en p .

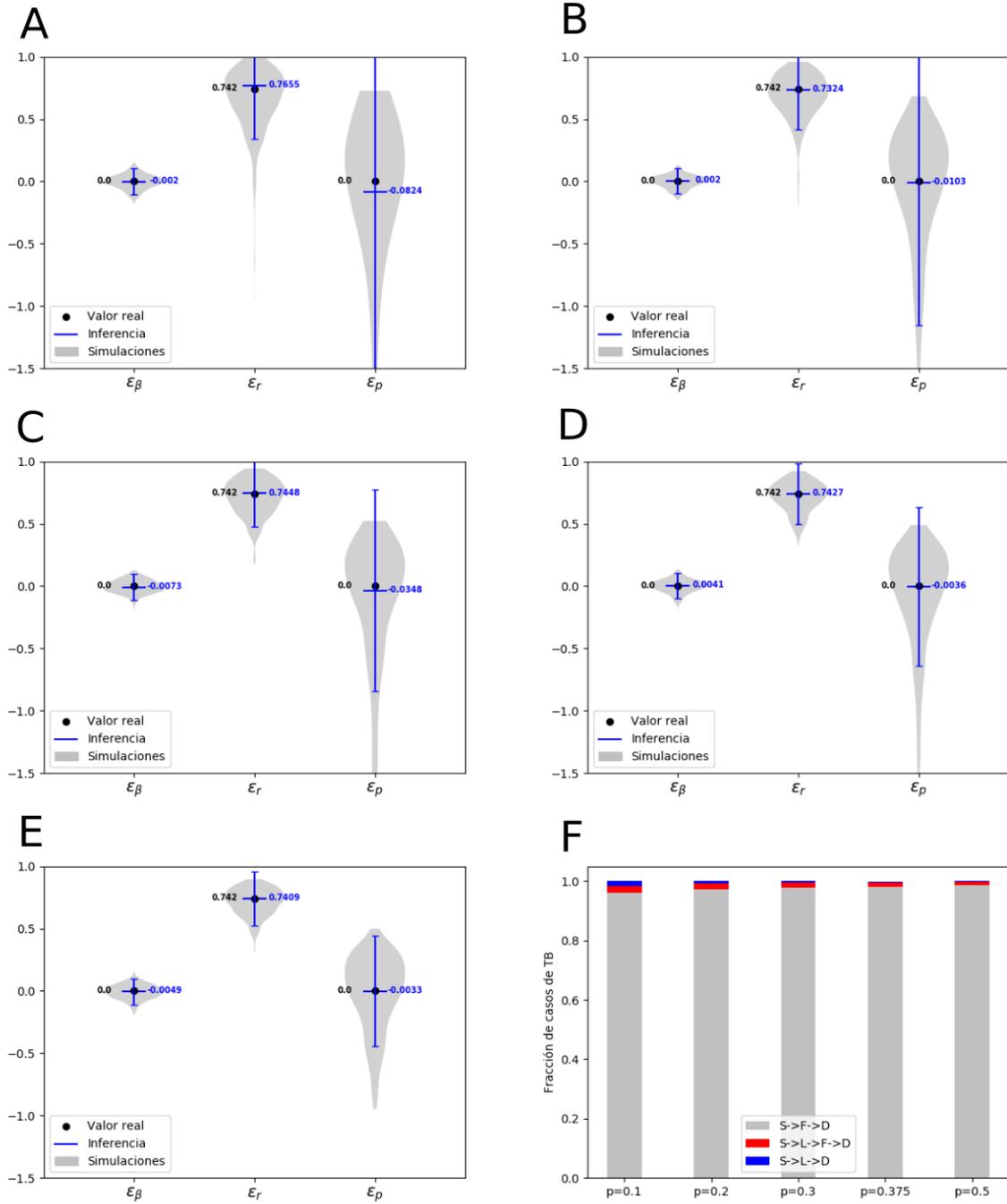


Figura 19: Resultados tras 500 iteraciones para A) $p = 0.1$, B) $p = 0.2$, C) $p = 0.3$, D) $p = 0.375$ (valor de referencia), E) $p = 0.5$. La figura F) muestra las contribuciones en cada caso a los individuos que desarrollan TB al final del ensayo en la cohorte de control para verificar que la hipótesis de transición rápida sigue siendo válida. En todos los casos se ha trabajado con tamaños de cohorte $N = 3000$, y con una vacuna basada en ϵ_r . ($\epsilon_\beta = 0.0$, $\epsilon_p = 0$, $\epsilon_r = 0.74$)

9.4. Barrido en la protección frente a reinfección

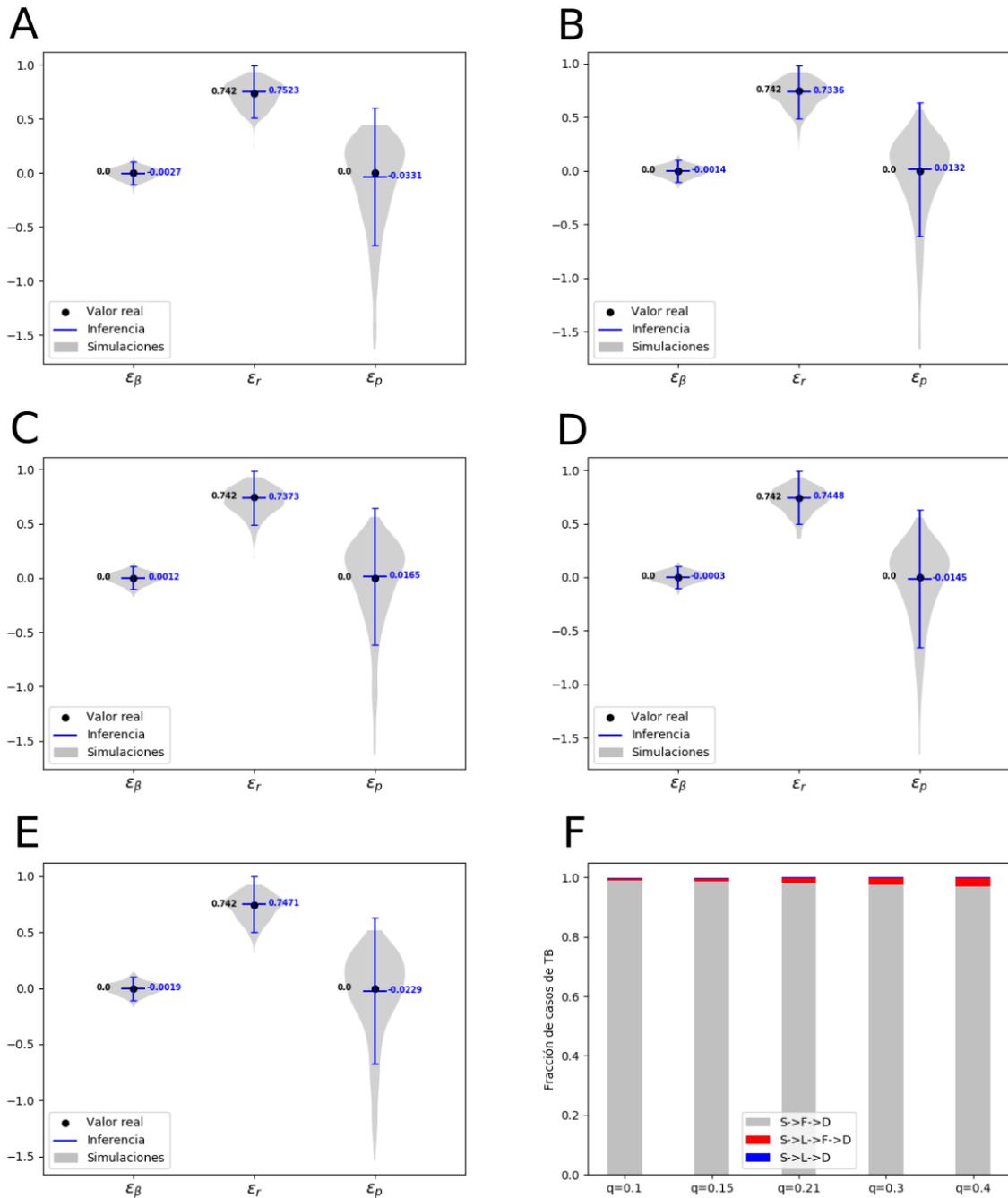


Figura 20: Resultados tras 500 iteraciones para A) $q = 0.1$, B) $q = 0.15$, C) $q = 0.21$ (valor de referencia), D) $q = 0.3$, E) $q = 0.4$. La figura F) muestra las contribuciones en cada caso a los individuos que desarrollan TB al final del ensayo en la cohorte de control para verificar que la hipótesis de transición rápida sigue siendo válida. En todos los casos se ha trabajado con tamaños de cohorte $N = 3000$ y con una vacuna basada en ϵ_r . ($\epsilon_\beta = 0.0$, $\epsilon_p = 0$, $\epsilon_r = 0.74$)

10. Estimación de impacto

Introducimos esta sección con objeto de concluir este trabajo y reforzar todavía más la idea que queremos transmitir respecto a la necesidad de lograr caracterizaciones completas y precisas de la acción de las nuevas vacunas contra la TB.

Uno de los problemas asociados a la incertidumbre sobre la vía en la cual actúa la vacuna es que el efecto de introducir en la población general una u otra vacuna puede no dar a largo plazo el mismo resultado.

Para poder ilustrar este efecto, hacemos uso de un modelo introducido por S.Arregui *et al*[9] para modelar la propagación de la TB. Con este modelo, simulamos el efecto que tiene introducir en Etiopía 3 vacunas diferentes a partir de 2025, yendo la simulación desde los años 2000 hasta 2050.^{II} Gracias al desarrollo de estos modelos de propagación podemos estudiar el efecto de vacunas diferentes en las estimaciones que se tienen para el futuro próximo.

Trabajamos con las vacunas 100-0 ($\varepsilon_\beta = 0$, $\varepsilon_r = 0$, $\varepsilon_p=0.5$) 50-50($\varepsilon_\beta = 0$, $\varepsilon_r = 0.37$, $\varepsilon_p=0.41$) y 0-100($\varepsilon_\beta = 0$, $\varepsilon_r = 0.74$, $\varepsilon_p=0$) y presentamos el impacto de cada una (es decir, los casos de TB activa prevenidos) en la figura 21.

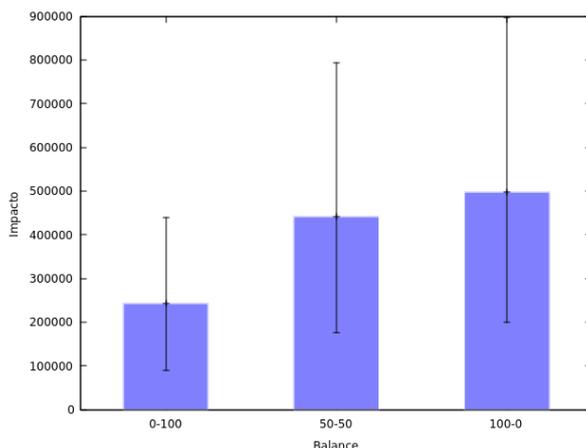


Figura 21: Resultados de introducir las vacunas 100-0 ($\varepsilon_\beta = 0$, $\varepsilon_r = 0$, $\varepsilon_p=0.5$) 50-50($\varepsilon_\beta = 0$, $\varepsilon_r = 0.37$, $\varepsilon_p=0.41$) y 0-100($\varepsilon_\beta = 0$, $\varepsilon_r = 0.74$, $\varepsilon_p=0$) en un modelo novedoso de propagación de la TB en Etiopía. La ventana de la simulación es de 50 años, empezando en el año 2000 y llegando hasta 2050. La vacuna se introduce en 2025. El eje vertical muestra el impacto de la vacuna, es decir, los casos de TB activa prevenidos respecto a una simulación en la misma ventana sin introducción de vacuna.

A la vista de los resultados en la figura 21, podemos entender la necesidad de lograr una buena caracterización. En el caso de las vacunas basadas en ε_p el impacto que se obtiene es significativamente mayor que en el caso de las vacunas basadas en ε_r .

Por todo lo anterior, consideramos necesario prestar atención a estos efectos, hasta ahora ignorados por la comunidad, con vistas a lograr las mejores vacunas posibles para una enfermedad compleja como es la TB.

^{II}La ventana entre 2000 y 2015 es la ventana de calibración del modelo de propagación, la cual ajusta los datos reales de la OMS.

Parte IV

Conclusiones

A lo largo de este trabajo se ha identificado un problema del que adolecen los ensayos clínicos de eficacia de vacunas para la TB. Esta enfermedad, al presentar dos posibles vías por las que un individuo infectado desarrolla la fase activa caracterizadas por presentar dos escalas temporales radicalmente distintas, introduce un problema de degeneración a la hora de estimar la eficacia de una vacuna frente a enfermedad.

Esto ocurre puesto que en TB existe una vía de latencia rápida y una vía de latencia lenta con una gran inercia (los individuos infectados con latencia lenta pueden llegar a vivir toda su vida sin desarrollar fase activa de la enfermedad). Las vacunas que previenen el desarrollo de TB activa tienen dos vías sobre las que actuar: pueden bien llevar a los infectados a latencia lenta, donde la propia naturaleza de la enfermedad hace que tarden décadas en desarrollar Tb, si lo hacen, o bien puede ralentizar el desarrollo de TB de los individuos que van por vía rápida.

Por supuesto, una vacuna podría también ralentizar, incluso más, la transición desde latencia lenta, pero este sería un fenómeno dinámico que en ningún caso podríamos observar en un ensayo clínico como en analizado en este trabajo, inspirado en el realizado en [7].

Entonces, mientras la vacuna prevenga la infección (POI) no hay problema en utilizar las técnicas clásicas pues nada de lo anterior entra en escena, pero cuando la vacuna se basa en prevenir el desarrollo de la fase activa (POD), dejando al individuo en un estado de latencia de vida media muy larga, encontramos estos dos posibles mecanismos por los que puede llevarse a cabo. Un análisis clásico no proporciona suficiente información como para poder romper esta degeneración en los mecanismos.

Por ello, hemos presentado un método novedoso que permite obtener el efecto que tiene una vacuna sobre la transición rápida dentro del contexto de un ensayo clínico con un tiempo de seguimiento determinado. Esta estimación de la protección en transición rápida permite, gracias a la introducción de un modelo analítico de ensayo clínico, obtener la protección en la vía restante.

Hemos podido además verificar nuestro método poniéndolo a prueba en el escenario de uno de los ensayos de campo de TB más recientes [7], además de en una serie de escenarios alternativos para dar cuenta de la versatilidad y adaptabilidad del método propuesto.

Finalmente, para concluir el presente trabajo, destacamos que el objetivo final reside en lanzar un mensaje de aviso al resto de la comunidad científica que trabaja en el desarrollo de ensayos de vacunas. En un momento en el que los ensayos de campo son pocos y costosos, es necesario disponer de los métodos de análisis de resultados más refinados, de modo que seamos capaces de obtener de la forma más precisa posible la acción de una vacuna, así como la viabilidad de la misma.

En ese aspecto, este trabajo aspira a poder introducir estos métodos novedosos en los próximos ensayos que se desarrollen, esperando obtener mejores caracterizaciones y mejores vacunas de las que se dispone actualmente.

Referencias

- [1] WHO. Global Tuberculosis Report 2017. World Health Organization, 2017.
- [2] Christopher Au-Yeung, Steve Kanters, Erin Ding, Philippe Glaziou, Aranka Anema, Curtis L Cooper, JS Montaner, Robert S Hogg, and Edward J Mills. Tuberculosis mortality in hiv-infected individuals: a cross-national systematic assessment. *Clin Epidemiol*, 3:21–29, 2011.
- [3] Cristiana J. Silva and Delfim F. M. Torres. A tb-hiv/aids coinfection model and optimal control treatment. *Discrete and Continuous Dynamical Systems*, 35(9):4639–4663, 2015.
- [4] Amita Jain and Pratima Dixit. Multidrug resistant to extensively drug resistant tuberculosis: what is next? *Journal of biosciences*, 33(4):605–616, 2008.
- [5] James M Trauer, Justin T Denholm, and Emma S McBryde. Construction of a mathematical model for tuberculosis transmission in highly endemic regions of the asia-pacific. *Journal of theoretical biology*, 358:74–84, 2014.
- [6] A Roy, M Eisenhut, RJ Harris, LC Rodrigues, S Sridhar, S Habermann, L Snell, P Mangtani, I Adetifa, A Lalvani, et al. Effect of bcg vaccination against mycobacterium tuberculosis infection in children: systematic review and meta-analysis. *Bmj*, 349:g4643, 2014.
- [7] Michele D Tameris, Mark Hatherill, Bernard S Landry, Thomas J Scriba, Margaret Ann Snowden, Stephen Lockhart, Jacqueline E Shea, J Bruce McClain, Gregory D Hussey, Willem A Hanekom, et al. Safety and efficacy of mva85a, a new tuberculosis vaccine, in infants previously vaccinated with bcg: a randomised, placebo-controlled phase 2b trial. *The Lancet*, 381(9871):1021–1028, 2013.
- [8] Sergio Arregui, Dessislava Marinova, Carlos Martín, Joaquín Sanz, Yamir Moreno. Design principles for TB vaccines’ clinical trials based on spreading dynamics. *bioRxiv*, doi: <https://doi.org/10.1101/249847>
- [9] Sergio Arregui, Joaquin Sanz, Dessislava Marinova, Maria Jose Iglesias, Sofia Samper, Carlos Martin, and Yamir Moreno. A data-driven model for the assessment of age-dependent patterns of tuberculosis burden and impact evaluation of novel vaccines. *bioRxiv*, page 112409, 2017.
- [10] Laith J Abu-Raddad, Lorenzo Sabatelli, Jerusha T Achterberg, Jonathan D Sugimoto, Ira M Longini, Christopher Dye, and M Elizabeth Halloran. Epidemiological benefits of more effective tuberculosis vaccines, drugs, and diagnostics. *Proceedings of the National Academy of Sciences*, 106(33):13980–13985, 2009.