

4. Calibrado con los datos reales y predicciones.

Por datos reales entendemos los registros que hace públicos el ministerio de sanidad. En la figura 15 se muestran los valores para los afectados (casos) hasta el 16 de abril



Figura 15. Evolución del número de afectados (casos) hasta el 16 de abril.

La extrapolación de los datos reales debe tener en cuenta que la curva de la epidemia es una campana y con una curvatura determinada.

Para esto, el método será comparar la “forma” de la curva de los datos reales con la forma de la curva que predice el modelo. Para comparar hay que homogeneizar, lo que requiere cambiar el **origen** y la **escala** de los resultados del modelo de forma que las desviaciones respecto a los datos reales sean las mínimas. El cambio de origen que resulte se deberá, en buena parte, a que los valores del modelo se corresponden al día de la infección y los registrados por sanidad al día en que dio la cara la enfermedad, es decir, entre ellos media el periodo de incubación. La escala da cuenta de la diferencia entre la población simulada y la real.

El procedimiento para superponer las curvas es desplazar sucesivamente los resultados del modelo en 1,2,3, ..., n días y calcular el factor de escala que minimiza las diferencias entre realidad y modelo para cada uno de estos desplazamientos, hasta encontrar la pareja (desplazamiento, escala) que produce una menor diferencia. Como es habitual la diferencia se calcula como la suma de los cuadrados de las diferencias entre todos los pares de puntos.

Los datos reales presentan un evidente cambio de tendencia. Aunque el cambio a simple vista pueda parecer pequeño, su influencia es muy grande al realizar extrapolaciones, por ejemplo, si consideramos el 24 de marzo como la fecha de corte, la figura 16 muestra los dos periodos con sus correspondientes ajustes por mínimos cuadrados. Las tendencias son claramente diferentes y la extrapolación es manifiestamente divergente.

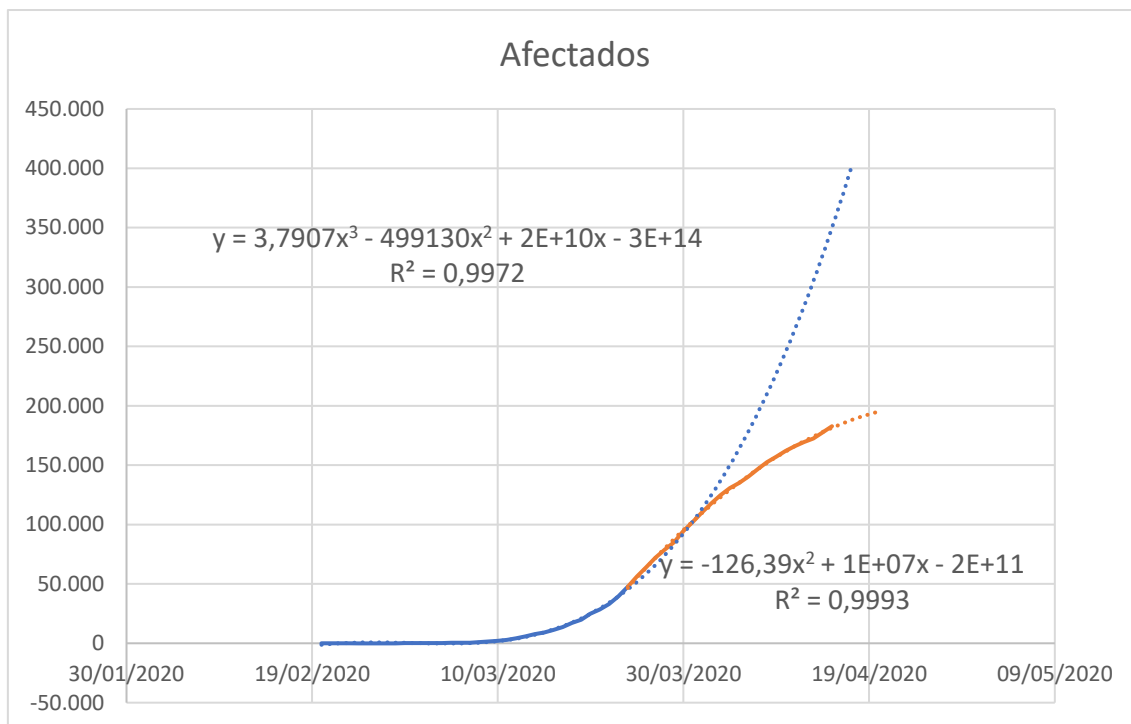


Figura 16. Cambio de tendencia en los datos reales del número de afectados a partir del día 24 de marzo.

La fecha de corte marca el punto en el que el efecto de las restricciones empieza a manifestarse y sería un error no tratar ambos periodos por separado. Además, la curva del modelo es significativamente distinta sin y con restricciones por lo que hay que usar la curva del modelo correspondiente en cada periodo.

El inconveniente es que restricciones hay que ponerle al modelo para que el efecto sea parecido a lo que ha ocurrido en la realidad, que por ahora no sabemos. En cualquier caso, es imprescindible imponer al modelo unas restricciones aunque sean aproximadas.

En primer lugar hemos procedido a ajustar las curvas real y simulada con los datos previos al 24 de marzo y obtenido los resultados de la figura 17 y tabla 4. Los ajustes se hacen sólo con la curva de afectados, para obtener la curva de enfermos del modelo calibrado se aplica el factor de **escala** a las cifras de curados y muertos que resultan del modelo y se restan de la curva de afectados con el **origen** desplazado (desplazada a la derecha).

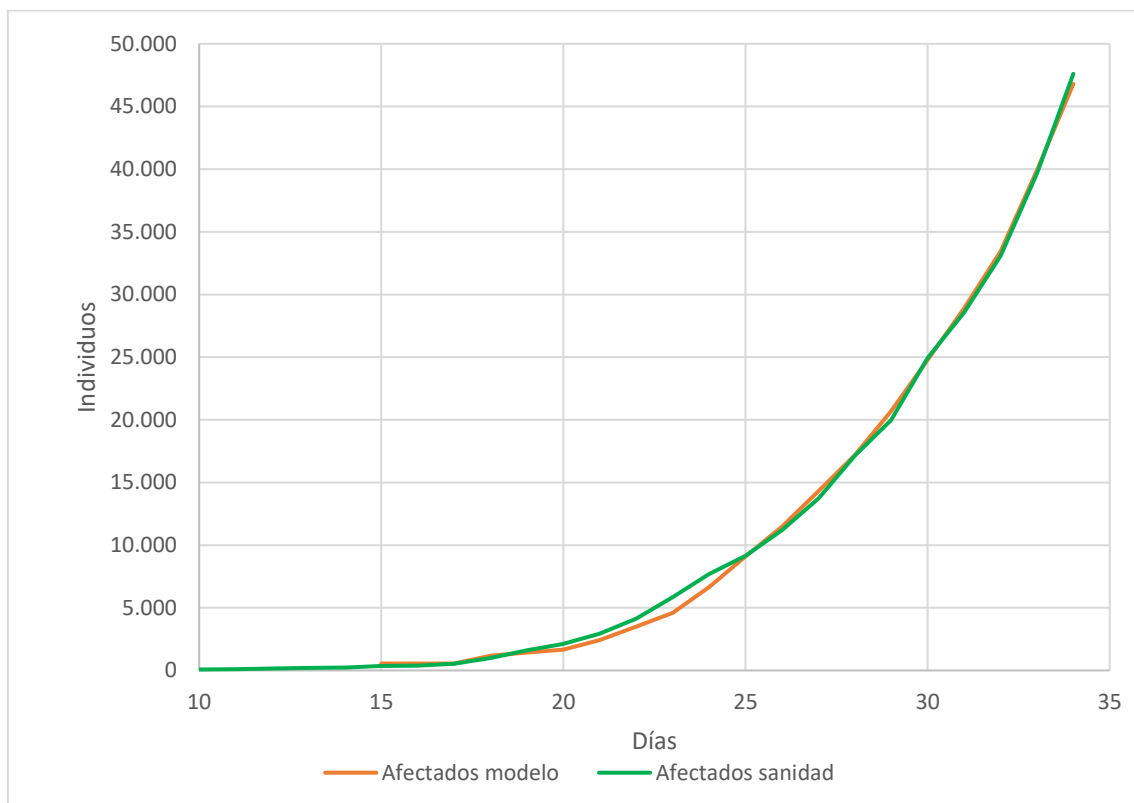
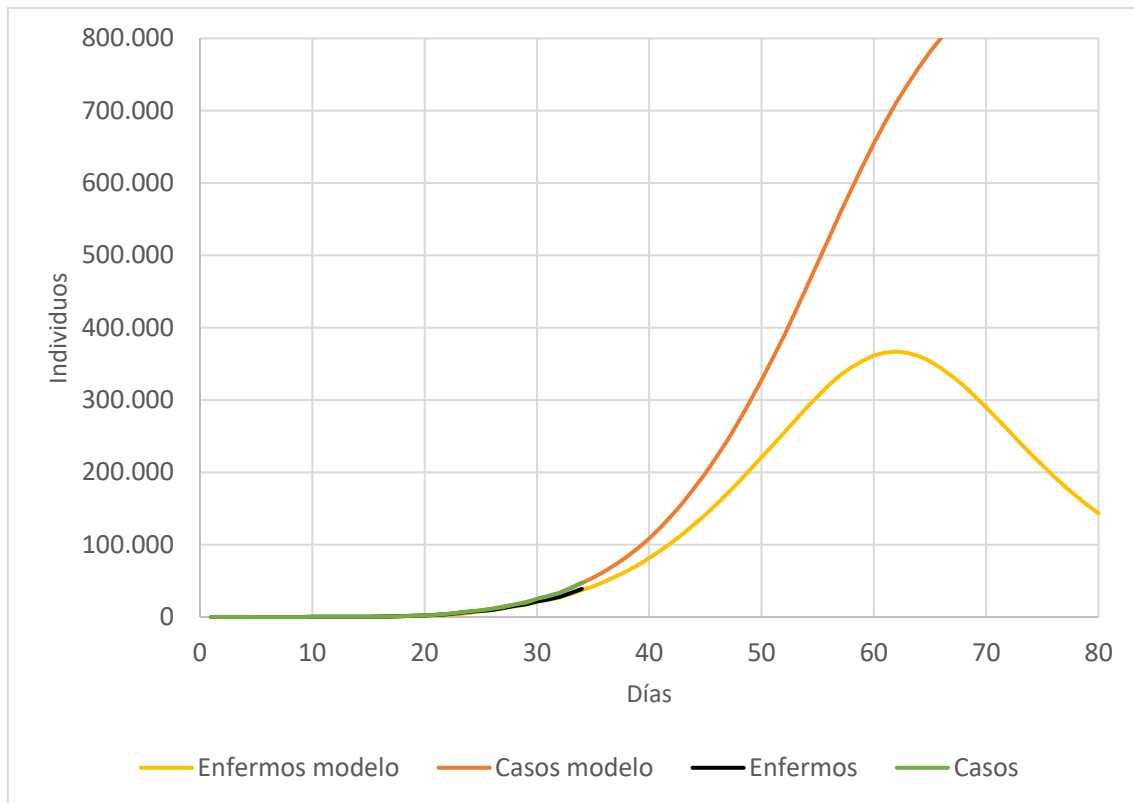


Figura 17. Predicciones del modelo calibrado con el período sin restricciones. La figura inferior es el detalle de la superior en el periodo que ha servido para el ajuste.

Desplazamiento		SEC	N	SEC/N	Máximo		Falta
Origen	Escala				Enfermos	Casos	Al pico
	15	55,5	5.497.051	20	274.853	366.886	711.258
							28 días

Tabla 4. Predicciones del modelo calibrado con el período sin restricciones.

Los resultados de la tabla 4 los podemos interpretar como los valores a los que se habría llegado sin restricciones. El máximo de personas que el sistema de sanidad tendría que hacer atendido simultáneamente (enfermos) habría sido algo más de 350.000 y en ese momento los afectados sumarían poco más de 700.000 personas. Estos resultados hay que tomarlos con mucha precaución porque es una extrapolación a larga distancia (28 días). Como puede verse en la figura 17, los datos reales (línea verde) corresponden a un periodo muy pequeño en relación con el necesario para alcanzar la punta.

La conclusión no puede ir más allá de suponer que el número de afectados, sin imponer restricciones, habría estado muy por encima de las capacidades del sistema sanitario. Sin pronunciarnos si las medidas realmente adoptadas han sido las más adecuadas, a este respecto, sería interesante estudiar el efecto de unas restricciones selectivas, centradas en la población más sensible (personas a partir de una determinada edad o con patologías previas) que son del orden del 15% de la población en España.

A continuación, hemos procedido a calibrar el modelo para el segundo periodo, a partir del 24 de marzo.

La figura 18 muestra los resultados del modelo, con restricciones, que hemos utilizado para calibrar el modelo con el segundo periodo de datos reales. Hemos supuesto que a los 30 días del inicio de la epidemia se reduce la movilidad en un 25% (a las tres cuartas partes), el número de clusters de 123 a 49 y después de 25 días se incrementan las restricciones, reduciendo la movilidad al 50% de lo normal y los clusters a 23, ambas cosas de forma indefinida.

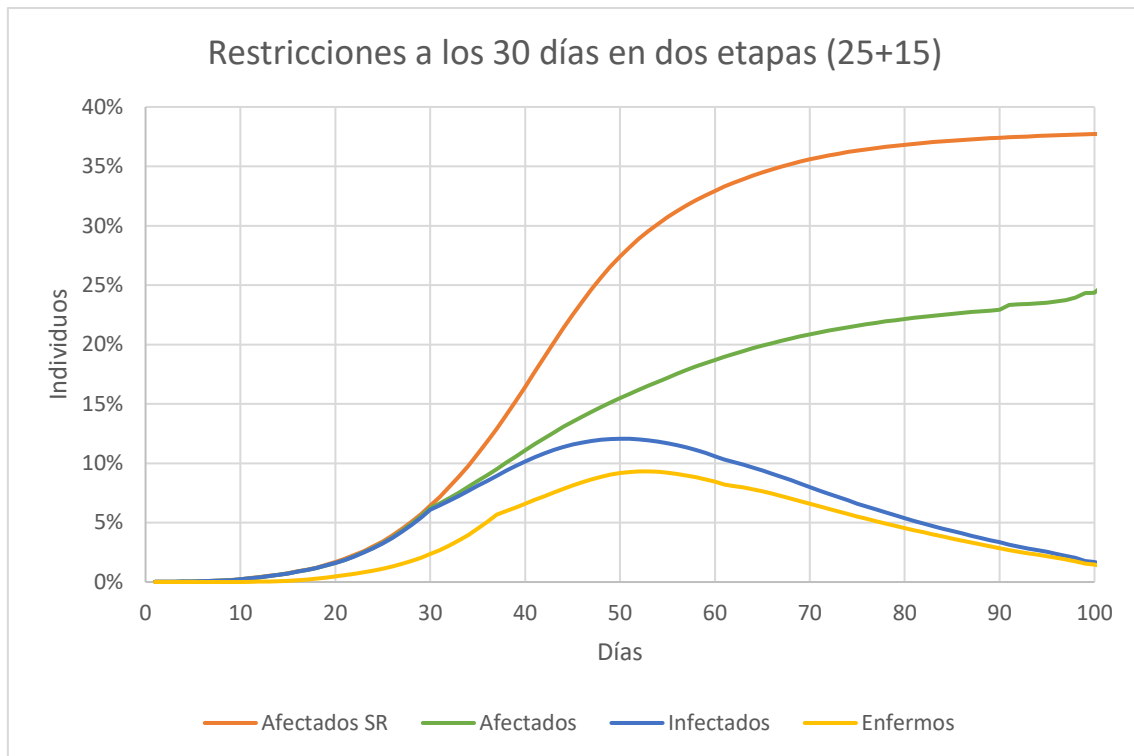


Figura 18. Resultados del modelo con restricciones en dos etapas a partir de los 30 días de inicio de la epidemia y mantenidas indefinidamente.

Los resultados del calibrado del modelo para el período a partir del 24 de marzo (incluido) se muestra en la figura 19 y la tabla 5.

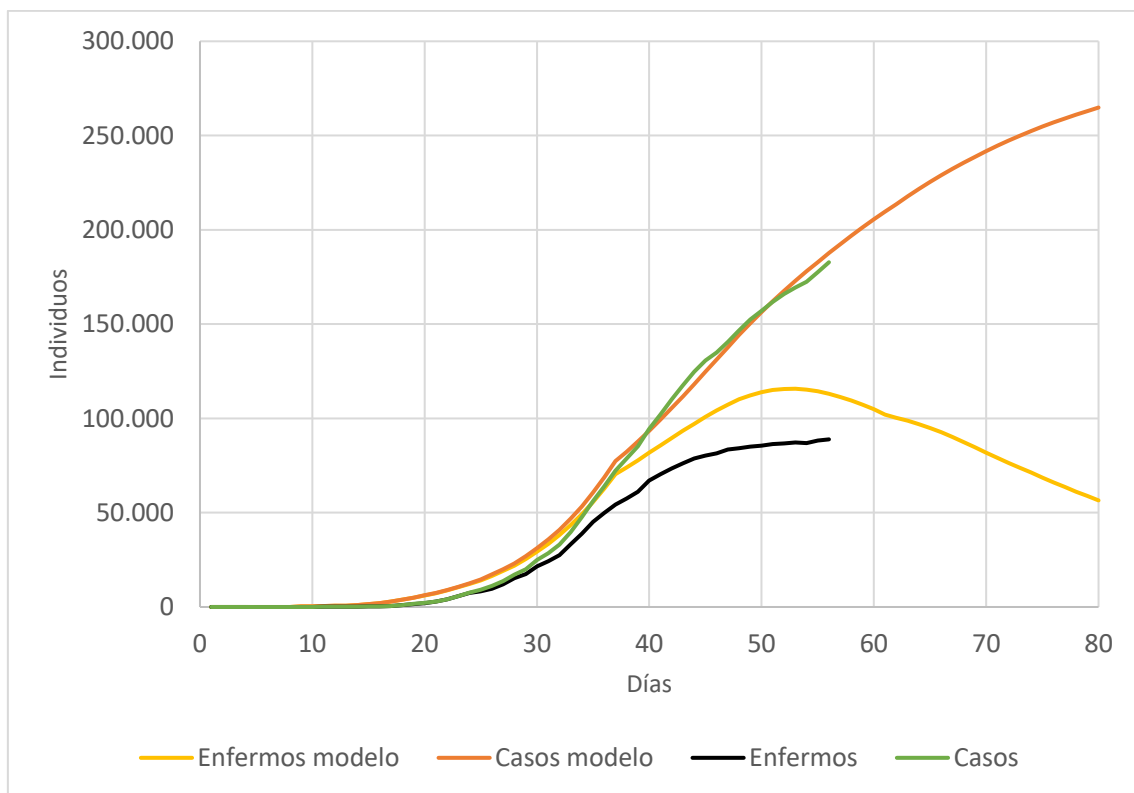


Figura 19. Predicciones del modelo calibrado con el período con restricciones.

Desplazamiento					Máximo		Falta
Origen	Escala	SEC	N	SEC/N	Enfermos	Casos	Al pico
	8	24.9	371.398.570	22	16.881.753	115.736	172.945
							-2

Tabla 5. Predicciones del modelo calibrado con el período con restricciones.

Los resultados del modelo calibrado indican que hace 2 días se alcanzó el máximo de enfermos (los que precisan de los servicios sanitarios) con 116.000 enfermos, los datos de sanidad hoy, 16 de abril, son 89.000 enfermo y muy cerca del pico, esta diferencia se debe, principalmente, a que el número de curaciones y muertes del modelo es inferior al registrado. Para unos resultados más precisos habría que incrementar la probabilidad de muerte que se usa como parámetro en el modelo. No obstante, los datos registrados de muerte y curaciones muestran una “forma” distinta a la del modelo. En el modelo, antes del pico, las curvas de curación y muertes son convexas como la de afectados, las registradas por sanidad son cóncavas.

Finalmente, hay que hacer una consideración sobre el procedimiento de superposición de curvas. Estamos comparando el inicio de la epidemia, que es muy sensible a las circunstancias que desencadenan la epidemia y que con seguridad han sido distintas de unos lugares a otros. Si el foco de la infección surge en una zona densamente poblada, con transporte colectivo masivo (aeropuertos) o con motivo de concentraciones de personas (ferias) tendrá un inicio mucho más explosivo que si no se dan estas circunstancias. Una vez que la epidemia progresa el punto inicial pierde relevancia, pero nosotros estamos haciendo la comparación en la parte inicial y necesitamos que el modelo sea representativo del inicio real en promedio de la epidemia. El modelo se comporta como la realidad, si al situar los focos iniciales al azar, son pocos o ningunos los que son encuadrados en clusters, la epidemia progresará al principio lentamente. Con una sólo simulación se depende de la suerte. Para estimar el número de afectados y contagiados a la vez en el conjunto de la epidemia esto tendrá poca importancia pero no al comparar la fase inicial.

Para evitar este sesgo es necesario un mínimo de simulaciones realizadas para cada escenario (lo cual es norma en el Método Montecarlo), el inconveniente es el tiempo de cálculo para un ordenador personal que es la herramienta de cálculo de este trabajo, creemos que las 30 simulaciones por caso que hemos realizado son suficientes.

5. Supresión de las restricciones.

Poco más hay que añadir a lo dicho en la primera parte del documento al analizar el escenario de FIN DE RESTRICCIONES (figuras 12 y 13).

El objetivo será que el número de afectados no aumente descontroladamente. Pretender que no aumente en absoluto no es realista, aunque parece que es la política de China, que trata a toda costa de que no haya nuevos casos, quizás como estrategia de espera para que lleguen las soluciones médicas.

Sin vacuna no hay más solución que la auto inmunización (pasar la enfermedad) hasta alcanzar un porcentaje de la población que evite la transmisión explosiva en cadena a la vez que se protege al sector de la población para los cuales la enfermedad resulta grave. Sin ese porcentaje mínimo de inmunizados el riesgo de epidemia será permanente, dicho porcentaje tiene que ser el porcentaje que el modelo indica cuando el número de contagiados se aproxima a cero, el 38% de la figura 7.

La figura 13 muestra que una vez eliminadas las restricciones, el número de afectados termina alcanzando al que se habría producido sin las mismas, esto significa que aunque el número de enfermos, de acuerdo con el modelo calibrado, es de 116.000 en el pico (gracias a las restricciones), a medio y largo plazo puede superar los 366.000 que se obtiene al calibrar el modelo sin restricciones (figura 17).

La regla para saber si la epidemia está controlada es que el incremento de afectados (nuevos casos) no supere el número de curados más muertes.

Parece inevitable un periodo de “prueba y error” que puede ser bastante largo.

6. Conclusiones.

El modelo predice que el número de afectados en una epidemia de estas características y sin restricciones llegaría a afectar el 38% de la población. A los 58 días se alcanza el máximo de enfermos (19% de la población) y en ese momento los afectados son el 32%.

Sin embargo los registros oficiales son 100 veces menores, estamos próximos al máximo de enfermos con 89.000 personas (0,2% de la población) y 183.000 afectados (0,4% de la población).

Ambas cosas no son directamente comparables porque los datos actuales son después de cuatro semanas de restricciones. Cuando se imponen restricciones al modelo, en la punta de enfermos estos son el 9% y en ese momento los afectados son el 17%.

Los datos del modelo son 50 veces los registrados oficialmente.

La explicación más sencilla es que más del 95% de los afectados lo son de forma muy leve, o incluso asintomáticos y no han requerido de los servicios sanitarios. El factor de escala del ajuste indica que el 97%.

$$50.000 \text{ (muestra simulada)} * 24,9 \text{ (escala)} / 47.000.000 \text{ (población en España)} = 3\%$$

El proceso de calibrado que hemos realizado en este documento pretende esquivar esta diferencia y se limita a extrapolar los datos oficiales suponiendo que van a seguir la misma evolución (forma de la curva) que el modelo.

El primer resultado que se aprecia es que las restricciones han frenado la epidemia de forma muy importante (a la tercera parte).

- La extrapolación de los datos oficiales antes de que las restricciones empezaran a tener efecto (24 de marzo) apuntaban a que el pico de enfermos podría llegar a las 350.000 personas y en ese momento los afectados serían 700.000.
- Los datos posteriores al 24 de marzo apuntan a 115.000 personas en el pico de contagios con 173.000 afectados.

El segundo resultado, no menos importante, es que el modelo muestra que el número de afectados, cuando se eliminan las restricciones tiende a recuperar las cifras como si no las hubiese habido, aunque más tarde (figura 13).

Si esto es cierto y la extrapolación de los datos oficiales, antes de las restricciones, apuntaba a los 350.000 enfermos y hoy, el modelo, contabiliza 115.000 enfermos, cuando se eliminen las restricciones al sistema sanitario le llegarán más de 200.000 enfermos. Para que el ritmo de afectados con necesidades sanitarias no provoque los problemas que se han vivido en el mes de marzo y primeros días de abril, la eliminación de restricciones tendrá que ser tan escalonada que puede tardar tres o seis meses. Estos plazos podrían acortarse si el número de individuos asintomáticos fuese una proporción importante de la población y se hubiesen inmunizado en gran medida antes de eliminar las restricciones (figura 14).

7. Repositorio

Todas las aplicaciones informáticas, datos y resultados obtenidos son de libre utilización y están a disposición de los interesados en el repositorio público:

<https://github.com/labotecta/covid-19>