

Impacto de la pandemia de Covid-19 en el diagnóstico de cáncer

PROTOCOLO
7/3/2022

COORDINADORES DEL PROYECTO

Investigador principal: David Moriña (dmorina@ub.edu)

Departament d'Econometria, Estadística i Economia Aplicada, Universitat de Barcelona
Avinguda Diagonal, 690-694 08034 Barcelona
Tel: +34 934 021 436

Equipo colaborador: Amanda Fernández Fontelo (amanda@mat.uab.cat)

Departament de Matemàtiques, Universitat Autònoma de Barcelona

1. ÍNDICE

<u>ÍNDICE</u>	3
<u>RESUMEN</u>	4
<u>TÍTULO Y RESUMEN EN INGLÉS</u>	5
<u>ANTECEDENTES, JUSTIFICACIÓN Y BIBLIOGRAFÍA</u>	6
<u>METODOLOGÍA</u>	7
<u>DATOS NECESARIOS</u>	10
<u>EXPERIENCIA DEL EQUIPO INVESTIGADOR SOBRE EL TEMA</u>	11

2. RESUMEN

La pandemia de Covid-19 sufrida a nivel global en 2020 y 2021 ha tenido un impacto notable en la mayoría de ámbitos de la actividad humana. En la literatura especializada han aparecido recientemente estudios que pretenden cuantificar el impacto específico en la infradetección de nuevos casos de cáncer como consecuencia de la suspensión o retraso en los cribados o de las restricciones de movilidad impuestas por las autoridades. Este proyecto pretende desarrollar un modelo matemático que permita determinar si, (i) por el efecto de posponer visitas y suspender programas de cribado o (ii) por las secuelas de haber sufrido el virus (Covid persistente o efectos secundarios), va a producirse un aumento de casos detectados de algunos tipos específicos de cáncer y en su caso, cuándo se producirá.

3. TÍTULO Y RESUMEN EN INGLÉS

Title: Impact of the Covid-19 pandemics on the cancer diagnosis

Abstract: The Covid-19 pandemic suffered globally in 2020 and 2021 has had a notable impact on most areas of human activity. Studies have recently appeared in the specialized literature that aim to quantify the specific impact on the underdetting of new cases of cancer as a result of the suspension or delay in screening or mobility restrictions imposed by the authorities. This project aims to develop a mathematical model to determine whether, (i) due to the effect of postponing visits and suspending screening programs or (ii) due to the sequelae of having suffered the virus (persistent Covid or side effects), there will be an increase in detected cases of some specific types of cancer and, eventually, when it will occur.

4. ANTECEDENTES, JUSTIFICACIÓN Y BIBLIOGRAFÍA

Las consecuencias derivadas de la pandemia provocada por el virus SARS-CoV-2 han afectado de manera contundente en muchos ámbitos de la actividad humana. Además de las consecuencias directas en relación con las defunciones provocadas por la enfermedad Covid-19 y la saturación de los sistemas de salud en numerosos países (incluyendo España y países de su entorno), en el año 2020 se ha detectado una disminución en el uso de los servicios del Sistema Público de Salud y de los servicios asociados a los seguros de salud privados.

Las restricciones de movilidad supusieron un declive en la utilización de los servicios sanitarios y una transformación de la interacción entre pacientes y sanitarios con una mayor utilización de la consulta telefónica. La pregunta es saber si, bien por el efecto de posponer visitas o bien por las secuelas de haber sufrido el virus (Covid persistente o efectos secundarios), va a producirse un exceso de siniestralidad en 2022 y los años sucesivos. Ya existen evidencias de una menor frecuencia de utilización de servicios de Salud en el Sistema Público en 2020, especialmente en lo relativo a cáncer (AECC 2020), y se han impulsado protocolos para revertir esta situación en 2021 (Ministerio de Sanidad 2021). Sin embargo, es difícil determinar si la mayor frecuencia de siniestralidad que se observará será igual o superior a la infra-siniestralidad que se observó durante la pandemia. Para analizarlo en el Proyecto se propone utilizar la metodología estadística de la infra-representación de casos como base de partida (Fernández-Fontelo et al. 2016; Fernández-Fontelo et al. 2019) y se extiende la misma al contexto de la sobre-representación. El planteamiento es determinar cómo es posible ver si el efecto rebote (i) se produce uniformemente o con diferentes intensidades según el tipo de cáncer o el estadiaje, (ii) se da de forma homogénea o en función de características de los pacientes o bien (iii) en qué momento del tiempo se recupera el nivel de diagnósticos que se venía observando antes del inicio de la pandemia. Es de esperar que, para monitorizar los efectos de la pandemia en los próximos años, se deban utilizar este tipo de aproximaciones, ya que no podrán compararse directamente grupos de población con características sociodemográficas diferentes, ni impactos en utilización de servicios de salud en general, ni prestaciones y coberturas diferentes. Este Proyecto pretende cuantificar el impacto de la pandemia en el diagnóstico de cáncer, y cómo evaluarlo, estimando el grado de infra-detección que se dio en 2020 principalmente y usando técnicas avanzadas de ciencia de datos desarrolladas recientemente, así como nuevos métodos e innovaciones encaminadas a valorar el potencial exceso de casos, con la finalidad de crear un sistema de seguimiento del diagnóstico que detecte el cambio en la dinámica del mismo. Es razonable pensar que los resultados y conclusiones pueden ser generalizables a otros ramos y que sirvan para valorar posibles desigualdades entre países o regiones.

BIBLIOGRAFIA

1. Cáncer AEC el. El número de pacientes de cáncer nuevos bajó un 21% durante el confinamiento. Published online 2020.

2. Ministerio de Sanidad. Acuerdo del Consejo Interterritorial del Sistema Nacional de Salud de 24 de febrero de 2021 sobre la pandemia de la COVID-19 y la prevención y el control del cáncer. Published online 2021.
3. Fernández-Fontelo A, Cabaña A, Puig P, Moriña D. Under-reported data analysis with INAR-hidden Markov chains. *Stat Med*. 2016;35(26):4875-4890. doi:10.1002/sim.7026
4. Fernández-Fontelo A, Cabaña A, Joe H, Puig P, Moriña D. Untangling serially dependent underreported count data for gender-based violence. *Stat Med*. 2019;38(22):4404-4422. doi:10.1002/sim.8306

5. METODOLOGÍA

La evolución del número de diagnósticos de cáncer en las localizaciones consideradas en un periodo de tiempo considerado “normal” (pre-pandemia) X_t se estimará mediante el uso de técnicas de series temporales discretas del tipo INteger AutoRegressive (INAR), definidas cómo (Jung and Tremayne 2006):

$$X_t = \alpha_1 \circ X_{t-1} + \dots + \alpha_p \circ X_{t-p} + W_t \quad (1),$$

donde $\alpha_1, \dots, \alpha_p$ son parámetros fijados con $0 < \alpha_1, \dots, \alpha_p < 1$, y W_t sigue una distribución de Poisson de media λ . Adicionalmente, se asume que X_{t-1} y W_t son independientes para todo tiempo t . El operador \circ de la ecuación (1), llamado *thinning* binomial, se define de la siguiente manera:

$$\alpha_j \circ X_{t-j} = \sum_{i=1}^{X_{t-j}} Y_i \quad (2),$$

Donde Y_i son variables aleatorias de Bernoulli independientes e idénticamente distribuidas, con probabilidad de éxito α_j . Por tanto, si $X_{t-j} = x_{t-j}$, entonces $\alpha_j \circ X_{t-j}$ tiene distribución binomial con número de éxitos x_{t-j} . Para estimar los parámetros correspondientes a un periodo de normalidad se usarán datos correspondientes al periodo 2015-2019, segregados por localización y provincia.

La magnitud del descenso en la evolución del número de diagnósticos producida en el año 2020 a consecuencia de la pandemia de Covid-19 se estimará usando modelos basados en cadenas de Markov ocultas, considerando que el número observado en cada punto temporal Y_t es sólo una parte de lo que se observaría si el contexto fuera distinto:

$$Y_t = \begin{cases} X_t, & \text{con probabilidad } 1 - \omega \\ q \circ X_t, & \text{con probabilidad } \omega \end{cases} \quad (3)$$

Claramente, de la ecuación (3) se deriva que el número de casos diagnosticados en 2020 coincide con lo esperado en un periodo de normalidad X_t con probabilidad $1 - \omega$ y es solo una parte de éste con probabilidad ω . Los parámetros de la ecuación (3) pueden interpretarse como la frecuencia e intensidad del infra-diagnóstico producido en 2020 (ω y q respectivamente). Un valor de q cercano a 0 indica una mayor intensidad del infra-uso, mientras que si $q = 1$ se tiene que $Y_t = X_t$ para todo t , de manera que el parámetro ω sería superfluo. En contraposición, si el parámetro ω es cercano a 1 puede deducirse que el infra-uso se ha producido durante todo 2020, mientras que si $\omega = 0$ volveríamos a tener $Y_t = X_t$ para todo t , de manera que el parámetro q sería superfluo. Un modelo similar se ha usado recientemente para evaluar la subnotificación de casos de diversas enfermedades como el virus del papiloma humano, mesotelioma o botulismo (Fernández-Fontelo et al. 2016).

Para evaluar la evolución del fenómeno en 2021 se desarrollará un modelo similar a los anteriores pero que permita determinar también si ha habido un aumento en el número de casos diagnosticados en relación con los periodos de normalidad. Para ello se desarrollará una modificación del *thinning* binomial definido en la ecuación (2) de manera que permita también que el resultado de la operación sea un valor mayor que el original. Los nuevos modelos desarrollados en el marco de este proyecto se pondrán a disposición de la comunidad científica y técnica de manera libre y gratuita en la forma de un paquete para el conocido software estadístico R (R Core Team 2021).

Además, a un nivel exploratorio, se usarán las ideas desarrolladas en (Moriña, Leyva-Moral, and Feijoo-Cid 2020) para valorar tanto si se ha producido una reducción significativa en la evolución de la utilización de servicios de salud a partir de 2020 como si se ha producido un aumento en 2021.

BIBLIOGRAFIA

1. Jung RC, Tremayne AR. Binomial thinning models for integer time series. *Stat Modelling*. 2006;6(2):81-96. doi:10.1191/1471082X06st114oa
2. Fernández-Fontelo A, Cabaña A, Puig P, Moriña D. Under-reported data analysis with INAR-hidden Markov chains. *Stat Med*. 2016;35(26):4875-4890. doi:10.1002/sim.7026
3. R Foundation for Statistical Computing. R: A Language and Environment for Statistical Computing. (2021)
4. Moriña D, Leyva-Moral JM, Feijoo-Cid M. Intervention analysis for low-count time series with applications in public health. *Stat Modelling*. 2020;20(1). doi:10.1177/1471082X18809194

6. DATOS NECESARIOS

Para alcanzar los objetivos planteados en este proyecto será necesario disponer de los datos de conteo semanal de número de casos diagnosticados de cáncer de mama, de cérvix, de colon y de cabeza y cuello, según estadiaje, sexo, grupo de edad y área básica de salud desde 2015 hasta 2022, y del conteo de personas participantes en algunos de los programas de cribado relacionados con las localizaciones mencionadas, con el mismo nivel de agregación temporal, geográfico y de grupo de edad y sexo.

7. EXPERIENCIA DEL EQUIPO INVESTIGADOR SOBRE EL TEMA

El Investigador Principal es Licenciado y Doctor en Matemáticas y Diplomado en Estadística por la Universitat Autònoma de Barcelona. Actualmente es profesor visitante del **Departamento de Econometría, Estadística y Economía Aplicada de la Universidad de Barcelona**, integrado en el grupo de investigación **Riskcenter**, de reconocida experiencia en el ámbito de la gestión del riesgo y de los seguros.

Ha participado en los congresos más prestigiosos del ámbito de la **ciencia de datos** y de la economía de la salud, y ha formado parte del Comité Científico del sexto *Workshop on Risk Management and Insurance* (RISK2015).

Es autor de más de **30 artículos publicados en revistas científicas internacionales** de reconocido prestigio (*Scientific Reports, Statistics in Medicine, Statistical Modelling, Journal of Statistical Software...*), tanto desde la perspectiva metodológica proponiendo nuevos métodos en el ámbito de la ciencia de datos como desde la perspectiva aplicada, utilizando los métodos más avanzados en diferentes contextos como la Economía de la Salud o la Epidemiología (*Plos One, Environment International, Journal of Radiological Protection...*).

A lo largo de su carrera investigadora ha participado en numerosos proyectos competitivos con financiación pública y privada, destacando su participación como **investigador principal** en dos proyectos dirigidos a la modelización de datos mal reportados en el ámbito de la Salud, el primero financiado por la beca competitiva BGSMath-Banco de Santander en el periodo 2018-2020 orientado al desarrollo de modelos con aplicaciones en Epidemiología y el segundo financiado por el Instituto de Salud Carlos III en el contexto de la convocatoria extraordinaria con cargo al fondo Covid-19, con el objetivo de cuantificar la carga real de la enfermedad incluyendo casos asintomáticos o sintomáticos leves que no requieren atención médica y que por tanto no son contabilizados por los registros de información del Sistema Nacional de Salud.

En cuanto al impacto social y a la difusión de los resultados de investigación, cabe destacar que el Investigador Principal ha desarrollado numerosas aplicaciones web que se encuentran disponibles para la comunidad científica y el público interesado, como parte de los resultados de los proyectos en que ha participado. Por ejemplo, recientemente se ha lanzado una aplicación web que permite estimar la **evolución de la pandemia de Covid-19 en todas las Comunidades Autónomas españolas**

(<https://dmorina.shinyapps.io/UB-BSLCoV/>). Las predicciones que se muestran en esta página web forman parte de las predicciones que realiza el **European Covid-19 Forecast Hub** (<https://covid19forecasthub.eu/>), coordinado por el European Centre for Disease Prevention and Control (ECDC). Además, de forma complementaria, el Investigador Principal es autor o coautor de numerosos paquetes para el conocido programa estadístico R (acumulando más de 150.000 descargas desde enero de 2016), como forma de poner las metodologías desarrolladas al alcance de toda la comunidad científico-técnica.

El Investigador Principal también ha colaborado recientemente con el **Grupo de Trabajo de Modelos Matemáticos del Ministerio de Sanidad** en la modelización del **impacto de la vacunación contra la Covid-19 en España**, cuyos resultados más destacados se han

difundido al público general a través de la web <https://www.vacunacovid.gob.es/voces-expertas/asi-es-como-tres-modelos-matematicos-ayudan-decidir-quien-se-vacuna-primero-en-espana>

Contribuciones del Investigador Principal relacionadas con el Proyecto (2016-2021):

Artículos:

A. Fernández-Fontelo, A. Cabaña, P. Puig, **D. Moriña**. "Under-reported data analysis with INAR-hidden Markov chains". *Stat Med*. 2016;35(26):4875-4890. doi:10.1002/sim.7026

G. Hernández, **D. Moriña**, A. Navarro. "Imputing missing data in Public Health: general concepts and application to dichotomous variables". *Gac Sanit*. 2017;31-4, pp.342-345.

A. Fernández-Fontelo, A. Cabaña, H. Joe, P. Puig, **D. Moriña**. "Untangling serially dependent underreported count data for gender-based violence". *Stat Med*. 2019;38(22):4404-4422.

D. Moriña, J. M. Leyva-Moral JM, M. Feijoo-Cid. "Intervention analysis for low-count time series with applications in public health". *Stat Modelling*. 2020;20(1).

D. Moriña, A. Fernández-Fontelo, A. Cabaña, P. Puig, L. Monfil, M. Brotons, M. Diaz. "Quantifying the under-reporting of uncorrelated longitudinal data: the genital warts example". *BMC Med Res Met*. 2021;21-1.

Proyectos:

Título: Seguros de salud tras la epidemia: Medir la transformación.

Entidad financiadora: Fundación Mapfre (15,000 €)

Entidades participantes: Universitat de Barcelona, Universitat Autònoma de Barcelona

Duración: Febrero 2022 – Febrero 2023

Investigador principal: David Moriña Soler

Título: Cuantificación en tiempo real de nuevos casos no reportados de COVID-19 en España (COV20/00115)

Entidad financiadora: Instituto de Salud Carlos III (54,842.80 €)

Entidades participantes: Centre de Recerca Matemàtica, Universitat de Barcelona, Universitat Politècnica de Catalunya, Humboldt Universität zu Berlin

Duración: Mayo 2020 – Marzo 2021

Investigador principal: David Moriña Soler

Título: Modelización estadística de riesgos medioambientales, tecnológicos y de la salud (MTM2015-69493-R)

Entidad financiadora: Ministerio de Economía y Competitividad (113,135 €)

Entidades participantes: Universitat Autònoma de Barcelona

Duración: 2016-2019

Investigador principal: Pere Puig Casado / Joan del Castillo Franquet

Dirección de tesis doctorales: Propuesta metodológica para el análisis de supervivencia con eventos recurrentes en estudios epidemiológicos cuando el número de eventos previos es desconocido (Universitat Autònoma de Barcelona, 2021 Excelente Cum Laude)