

## Nota metodológica

## Germina: herramienta para la detección automática de agrupaciones de casos de COVID-19



Lluís Forcadell-Díez<sup>a,b</sup>, Nacho Sánchez-Valdivia<sup>a</sup>, Maria Piñero Casas<sup>a</sup>, Marc Pila<sup>a</sup>, Alexia Reyes Segura<sup>a,b</sup>, Raúl Salobral<sup>a</sup>, Marc Olivella Cirici<sup>a,b</sup>, Sílvia Fernández-Barrés<sup>a</sup> y Glòria Pérez<sup>a,b,c,d,\*</sup>

<sup>a</sup> Agència de Salut Pública de Barcelona, Barcelona, España

<sup>b</sup> Universitat Pompeu Fabra, Barcelona, España

<sup>c</sup> CIBER de Epidemiología y Salud Pública (CIBERESP), España

<sup>d</sup> Institut d'Investigació Biomèdica Sant Pau, Barcelona, España

## INFORMACIÓN DEL ARTÍCULO

## Historia del artículo:

Recibido el 21 de febrero de 2023

Aceptado el 20 de junio de 2023

On-line el 9 de septiembre de 2023

## Palabras clave:

COVID-19

Informática en salud pública

Manejo de datos

Métodos epidemiológicos

Brotos de enfermedades

## RESUMEN

La pandemia de COVID-19 evidenció que la vigilancia epidemiológica no disponía de recursos para responder a los aumentos de casos ni a los brotes. La alta transmisibilidad comunitaria entre la población escolar en la ciudad de Barcelona al inicio de la sexta ola tensionó la unidad de vigilancia de COVID-19 de la ciudad. Mediante metodología SCRUM se desarrolló e implementó Germina, una herramienta informática capaz de capturar, armonizar, integrar, almacenar, analizar y visualizar diariamente datos de múltiples fuentes de información. Germina permite identificar agrupaciones de tres o más casos escolares y calcular indicadores epidemiológicos. La implementación de Germina facilitó la respuesta epidemiológica a la sexta ola de COVID-19 en el ámbito escolar en Barcelona. Esta herramienta es aplicable a otros ámbitos de exposición y a otras enfermedades transmisibles. El uso de herramientas informáticas automatizadas, como Germina, mejora los sistemas de vigilancia epidemiológica y apoya la toma de decisiones basada en la evidencia.

© 2023 SESPAS. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Germina: tool for automatic detection of clusters of COVID-19 cases

## ABSTRACT

The COVID-19 pandemic showed that epidemiological surveillance was under-resourced to respond to increases in cases and outbreaks. The high community transmissibility among the school population in the city of Barcelona at the beginning of the sixth wave strained the local COVID-19 surveillance unit. Using SCRUM methodology, Germina was developed and implemented as a software tool capable of capturing, harmonizing, integrating, storing, analysing and visualizing data from multiple information sources on a daily basis. Germina identifies clusters of three or more school cases and calculates epidemiological indicators. The implementation of Germina facilitated the epidemiological response to the sixth wave of COVID-19 in the school setting in the city of Barcelona. This tool is transferable to other exposure settings and communicable diseases. The use of automated informatics tools such as Germina, improves epidemiological surveillance systems and supports evidence-based decision making.

© 2023 SESPAS. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Keywords:

COVID-19

Public health informatics

Data management

Epidemiological methods

Disease outbreaks

## Introducción

La pandemia de COVID-19 se ha caracterizado por oleadas cíclicas de gran volumen de casos, contactos estrechos y brotes. Esto ha evidenciado las limitaciones de la vigilancia epidemiológica, que en España estaba infradimensionada<sup>1</sup>, en particular sus sistemas de información. Durante la pandemia se desarrollaron herramientas

para manejar datos y obtener indicadores de vigilancia epidemiológica, como los *dashboards*<sup>2</sup>. Sin embargo, no ha habido un desarrollo homólogo en epidemiología de campo. Las herramientas disponibles han sido insuficientes para actuar ágilmente en el manejo y la gestión de casos, contactos estrechos y brotes en situaciones de alta transmisibilidad comunitaria<sup>1</sup>.

En la ciudad de Barcelona, este déficit fue más evidente durante la sexta ola de la COVID-19 (2/11/2021 a 28/3/2022), que afectó particularmente al ámbito escolar. La incidencia acumulada en 14 días en Barcelona en personas de 0 a 14 años se mantuvo por encima de los 1000 casos por 100.000 personas del 18/12/2021 al 20/02/2022<sup>7</sup>. El pico fue de 11.681 casos el 26/01/2022. Durante la sexta ola, el 12% de las personas de 0-14 años escolarizadas en

\* Autora para correspondencia.

Correo electrónico: [gperez@aspb.cat](mailto:gperez@aspb.cat) (G. Pérez).

@salutpublicaben

**Tabla 1**

Fuentes de datos y variables integradas en la herramienta Germina para la detección de agrupaciones de casos de COVID-19 en el entorno escolar de la ciudad de Barcelona (2021-2022)

Fuente de información	Descripción	Principales variables
Registro de casos de COVID-19 del Departament de Salut	Información relativa a diagnóstico, estado inmunitario, información clínica y encuesta epidemiológica, y censo de contactos estrechos	Identificador individual anonimizado (Id) Sexo Fecha de nacimiento Presencia o ausencia de sintomatología Fecha de inicio de síntomas Fecha de consulta médica Fecha de diagnóstico Prueba diagnóstica realizada Resultado de la prueba diagnóstica Estado vacunal Fecha de administración de vacuna COVID-19 (desglosado por número de dosis) Vacuna administrada Contacto estrecho de un caso confirmado previo Variables de la encuesta epidemiológica Variables del censo de contactos estrechos
Registro Central de Salud del Departament de Salut	Información personal y sociodemográfica	Identificador individual anonimizado (Id) Sexo Fecha de nacimiento Lugar de residencia (dirección)
Censo escolar del Consorci d'Educació de Barcelona	Información de individuos vinculados a centros educativos	Identificador individual anonimizado (Id) Centro educativo (nombre e identificador) Grupo de convivencia estable (nombre e identificador) Perfil (alumnado, personal docente, personal externo) Nivel (educación infantil, educación primaria, educación secundaria, educación posobligatoria, otras enseñanzas) Localización del centro (dirección, código postal, barrio, distrito)
Padrón de la ciudad de Barcelona Oficina Municipal de Datos de Barcelona	Información del lugar de residencia Indicadores ecológicos del área de residencia	Localización de la residencia (dirección, código postal, barrio, distrito) Índice de renta familiar disponible del barrio de residencia

Barcelona fueron positivas en COVID-19<sup>3</sup>. Siguiendo la estrategia testeo-trazado-aislamiento-cuarentena (TTAC)<sup>4</sup>, se actuó rápidamente para identificar y gestionar brotes escolares de COVID-19. Para detectar agrupaciones de casos, se evidenció la necesidad de manejar de manera eficaz datos masivos diariamente<sup>5</sup>.

Esta nota metodológica expone el proceso de diseño, desarrollo e implementación en 2021 de la herramienta informática Germina, acrónimo de *Giny Epidemiològic per a la Revisió Metòdica i Notificació d'Agrupacions de casos de Covid-19* (Herramienta epidemiológica para la revisión metódica y notificación de agrupaciones de casos de COVID-19). Germina fue concebida en la Agència de Salut Pública de Barcelona (ASPB) para automatizar la detección y el seguimiento de agrupaciones de casos escolares de COVID-19 en la ciudad de Barcelona.

## Exposición del método

### Diseño e implementación

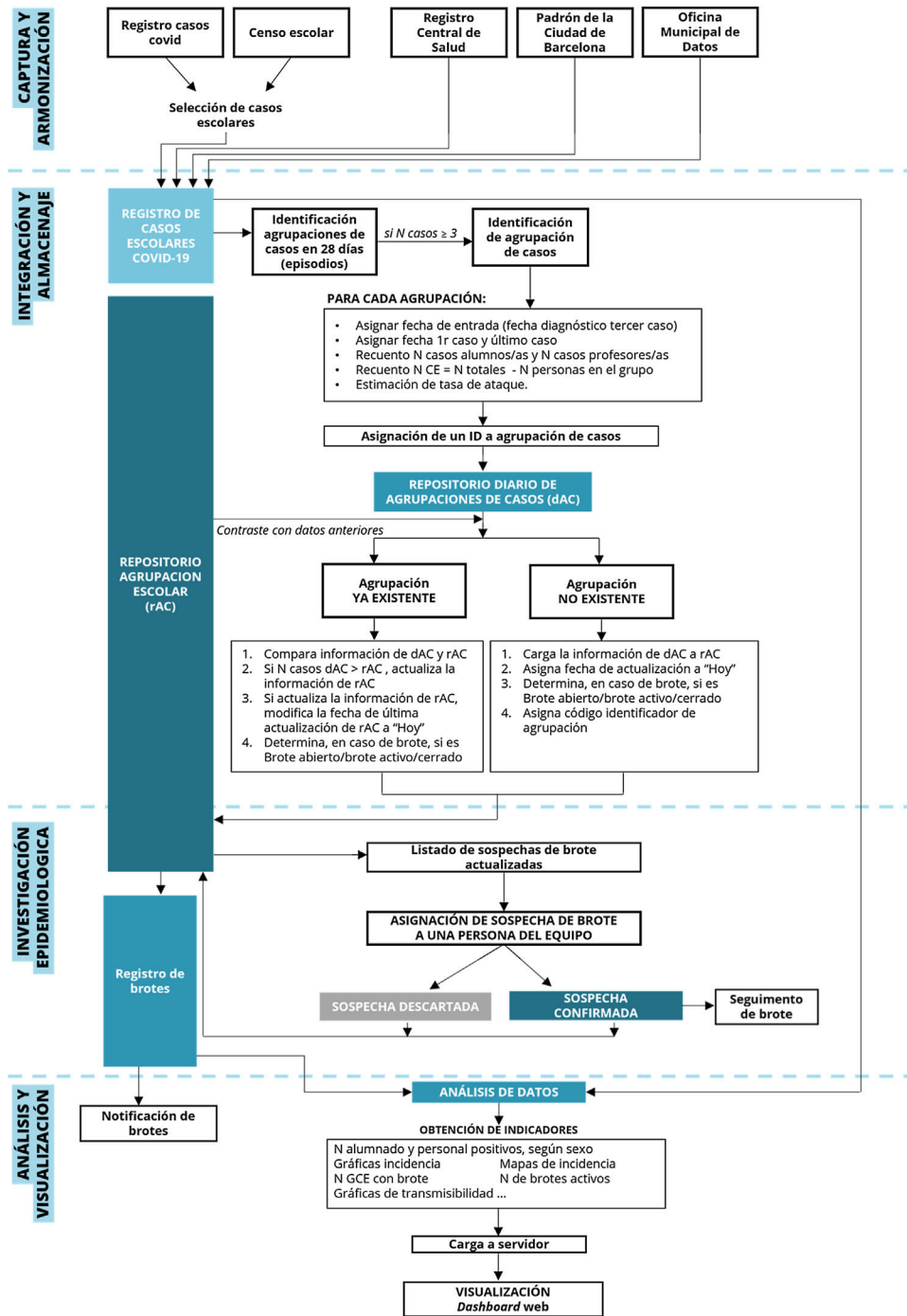
Para el desarrollo y la implementación de Germina se aplicó la metodología SCRUM, una estrategia para resolver problemas emergentes mediante el desarrollo de productos complejos. La metodología SCRUM optimiza el tiempo y los recursos de las organizaciones, y es especialmente útil para equipos multidisciplinares y flexibles. Además, garantiza la transferencia de los aprendizajes obtenidos durante el proceso<sup>6</sup>. El diseño se adaptó a las necesidades de vigilancia epidemiológica en la ciudad y se comprobaron, mediante un pilotaje inicial, su correcto funcionamiento y su aplicabilidad. Para facilitar su uso en epidemiología de campo, Germina se desarrolló como un código *do-file* estructurado en módulos para ser ejecutado diariamente con el *software* STATA-15.

Germina se concibió para la captura, la armonización, la integración, el almacenaje, el análisis y la visualización de datos procedentes de múltiples fuentes de información: el Registro

COVID-19 y el Registro Central de Salud del Departament de Salut de Catalunya, el Censo Escolar del Consorci d'Educació de Barcelona, indicadores ecológicos del área de residencia y el padrón de la ciudad de Barcelona de la Oficina Municipal de Datos del Ayuntamiento de Barcelona. Estas fuentes de información y las variables que incluían se describen en la [tabla 1](#). Germina incluyó las 327.486 personas escolarizadas o con relación profesional con un centro educativo de Barcelona en el curso 2021-2022, independientemente de su empadronamiento. La población escolar estaba distribuida en 13.504 grupos de convivencia estable (GCE) y 871 centros educativos.

### Funcionamiento

En primer lugar, Germina importa los casos confirmados de COVID-19 diariamente del Registro COVID-19 ([Fig. 1](#)). Estos datos se armonizan e integran con los procedentes de las otras fuentes de información, incluyendo el centro de estudios y la información geográfica. Para su integración, para cada caso se genera en cada fuente un identificador único y anónimo a partir de las variables comunes en las diferentes fuentes: CIP, DNI o NIE, fecha de nacimiento y sexo. Esto permite que Germina pueda buscar agrupaciones de tres o más casos de COVID-19 en un mismo GCE en 14 días consecutivos. Germina asigna un identificador único a cada agrupación y obtiene el número de casos totales, desagregados por sexo, por alumnado y personal; estima la tasa de ataque, identifica las fechas del primer y del último caso, y propone fechas de inactivación y cierre de un posible brote. Germina almacena esta información en un fichero temporal diario y seguidamente la compara con la almacenada el día anterior, lo que mantiene un repositorio actualizado de agrupaciones de casos escolares. Las agrupaciones son valoradas por el Equipo de Vigilancia, Control y Prevención de la COVID-19, que las trata como sospechas de brote y confirma o descarta el vínculo epidemiológico entre los casos. Esta información se reincorpora automáticamente en Germina para actualizar el repositorio



**Figura 1.** Diagrama de flujo del procedimiento de captura, armonización, integración, almacenaje, análisis y visualización de datos realizado por Germina para la detección automática de agrupaciones de casos de COVID-19 en el entorno escolar de la ciudad de Barcelona (2021-2022) y el cálculo de indicadores epidemiológicos.

de agrupaciones y generar un listado actualizado de los brotes de COVID-19 en el ámbito escolar.

Además, Germina calcula diariamente indicadores epidemiológicos y de gestión, por casos, agregados por GCE y escuelas, por divisiones territoriales y globales. Exporta los indicadores a diferentes tablas y genera automáticamente un informe-resumen para las personas de la ciudad implicadas en la toma de decisiones. Las tablas generadas por Germina se guardan en el servidor de la ASPB y alimentan un *dashboard* público<sup>7</sup> que contiene gráficos y mapas interactivos, y que permite descargar los datos de COVID-19 en la ciudad.

### Cuestiones éticas

Los datos individuales son usados y tratados de acuerdo con las directrices del Reglamento General de Protección de Datos vigente. Siguiendo la Ley 18/2009, del 22 de octubre, de Salud Pública, las bases de datos están sometidas a requisitos especiales de protección y se mantienen en archivos propiedad de la ASPB.

### Resultados

Germina se desarrolló en 3 semanas. Conectó cinco fuentes de información diferentes, con estructuras complejas. Se ejecutó dia-

riamente e identificó 7713 agrupaciones de casos en el ámbito escolar en la ciudad de Barcelona entre enero y febrero de 2022, coincidiendo con el pico de la sexta ola. Generó cuatro repositorios de información epidemiológica de uso interno, vinculados a un cuadro de mando, y permitió la actualización automática diaria de 12 indicadores clave en un *dashbord* público<sup>7</sup>, desde el cual se pueden descargar los datos.

### Aplicación práctica

Germina se concibió para poder aplicarse en otros ámbitos de transmisión de la COVID-19 y para poder replicarse con otras enfermedades de declaración obligatoria. La estructura en módulos de Germina y su programación en STATA favorecen que esta herramienta pueda aplicarse en la vigilancia epidemiológica y el seguimiento de brotes epidemiológicos. Asimismo, la configuración de Germina permite identificar agrupaciones de casos en otros ámbitos, como domicilios, residencias geriátricas, centros de trabajo, actividades lúdico-deportivas y áreas territoriales.

La herramienta también se aplica en la generación y la visualización de datos actualizados diariamente para la comunicación de información sobre vigilancia de la salud pública<sup>8</sup>. Además, la información actualizada facilita la toma de decisiones basadas en la evidencia dirigidas a la reducción de las desigualdades sociales en la distribución de la COVID-19<sup>9,10</sup>.

### Fortalezas y limitaciones

Germina automatizó procesos manuales, optimizando así los recursos humanos. Facilitó el seguimiento de los brotes epidemiológicos y permitió obtener indicadores en tiempo real y actualizar a diario la información epidemiológica.

Sin embargo, Germina no permite valorar el vínculo epidemiológico entre casos. Además, es dependiente de la calidad, la robustez, la validez, la continuidad y la periodicidad de las fuentes utilizadas.

### Conclusiones

El mecanismo de captura, armonización, integración, almacenaje, análisis y visualización de datos de Germina ha permitido la automatización y la sistematización de la identificación de agrupaciones de casos de COVID-19 en un contexto de alta demanda del sistema de vigilancia epidemiológica de la ciudad de Barcelona.

La epidemiología de campo se fortalece gracias a sistemas de vigilancia robustos, actualizados y flexibles<sup>1</sup>, que permiten una vigilancia epidemiológica eficaz<sup>5</sup>. La gestión innovadora de datos y la inteligencia epidemiológica fundamentada en herramientas informáticas, como Germina, es clave para abordar los retos presentes y futuros que puedan plantear las enfermedades transmisibles.

### Disponibilidad de bases de datos y material para réplica

Los datos están a disposición de las personas que los soliciten a la autora para correspondencia.

### Editor responsable del artículo

Carlos Álvarez-Dardet.

### Contribuciones de autoría

Todas las personas autoras han contribuido en la formulación y el diseño del artículo. L. Forcadell-Díez, N. Sánchez-Valdivia, M. Piñero Casas, M. Pila, A. Reyes Segura y G. Pérez han contribuido en la conceptualización, el diseño y la implementación de Germina. L. Forcadell-Díez y N. Sánchez-Valdivia han desarrollado, programado y validado el código de Germina. L. Forcadell-Díez, N. Sánchez-Valdivia, M. Piñero Casas, M. Pila, A. Reyes Segura y G. Pérez han realizado la escritura original del artículo. Todas las personas firmantes han contribuido a su revisión crítica con importantes contribuciones intelectuales. Todas las personas firmantes han aprobado la versión final para su publicación.

### Agradecimientos

A todas las personas que han apoyado el desarrollo de este trabajo, especialmente al equipo del Programa de Vigilancia, Control y Prevención de la COVID-19 de la Agència de Salut Pública de Barcelona.

### Financiación

Ninguna.

### Conflictos de intereses

Ninguno.

### Anexo. Material adicional

Se puede consultar material adicional a este artículo en su versión electrónica disponible en [doi:10.1016/j.gaceta.2023.102321](https://doi.org/10.1016/j.gaceta.2023.102321).

### Bibliografía

- Zurriaga-Carda R, Aginagalde Llorente AH, Álvarez-Vaca D. Epidemiología de campo en tiempos de COVID-19: retos para los servicios de salud pública. Informe SESPAS 2022. Gac Sanit. 2022;36 (Supl 1):S76-81.
- Agència de Salut Pública de Barcelona. COVID19aldiaBCN. 2020 [consultado el 6 de abril de 2023]. Disponible en: <https://webs.aspb.cat/covid19aldiabcn/metodes/>
- Departament de Salut. Decret 203/2015, de 15 de setembre, pel qual es crea la Xarxa de Vigilància Epidemiològica i es regulen els sistemes de notificació de malalties de declaració obligatòria i brots epidèmics. Barcelona: Parlament de Catalunya; 2015.
- Ashcroft P, Lehtinen S, Bonhoeffer S. Test-trace-isolate-quarantine (TTIQ) intervention strategies after symptomatic COVID-19 case identification. PLoS One. 2022;17:e0263597.
- Ibrahim NK. Epidemiologic surveillance for controlling Covid-19 pandemic: types, challenges and implications. J Infect Public Health. 2020;13:1630-8.
- Torrente G, de Souza TQ, Tonaki L, et al. Scrum framework and health solutions: management and results. Stud Health Technol Inform. 2021;284:290-4.
- Agència de Salut Pública de Barcelona. Informe de vigilància de la COVID-19 a Barcelona. 2020 [consultado el 6 de abril de 2023]. Disponible en: <https://www.aspb.cat/documents/vigilanciavid19-dades/>.
- Sierra Moros MJ, Martínez Sánchez EV, Monge Corella S, et al. Lecciones de la vigilancia de la COVID-19. Necesidad urgente de una nueva vigilancia en salud pública. Informe SESPAS 2022. Gac Sanit. 2022;36 (Supl 1):S68-75.
- Marí-Dell'Olmo M, Gotsens M, Pasarín MI, et al. Socioeconomic inequalities in COVID-19 in a European urban area: two waves, two patterns. Int J Environ Res Public Health. 2021;18:1256.
- Malmusi D, Pasarín MI, Marí-Dell'Olmo M, et al. Multi-level policy responses to tackle socioeconomic inequalities in the incidence of COVID-19 in a European urban area. Int J Equity Health. 2022;21:28.