

INSTITUTO COLOMBIANO DE ENDODONCIA



Conferencia de actualidad
Julio 20 - 2026

www.incoendo.com

*Historia natural de las enfermedades humanas:
Pasado y Futuro.*

Natural History of Human Diseases: Past and Future.

«Que te ignoren, que te ataquen, que te cuestionen, pero que nunca te silencien.»

Bertrand Arthur William Russell (1872-1970). Filósofo, matemático, lógico y escritor británico. Premio Nobel de literatura (1950)

Article

Learning the natural history of human disease with generative transformers

<https://doi.org/10.1038/s41586-025-09529-3>

Received: 18 May 2024

Accepted: 13 August 2025

Published online: 17 September 2025

Artem Shmatko^{1,2,3,13}, Alexander Wolfgang Jung^{2,4,5,6,13}, Kumar Gaurav^{2,13}, Søren Brunak^{4,7},
Laust Hvas Mortensen^{5,7,8}, Ewan Birney^{2,9}, Tom Fitzgerald^{2,9} & Moritz Gerstung^{1,2,9,10,11,12}

Shmatko A, Jung AW, Gaurav K, Brunak S, Mortensen LH, Birney E, Fitzgerald T, Gerstung M. Learning the natural history of human disease with generative transformers. **Nature**. 2025 Nov;**647**(8088):248-256. doi: 10.1038/s41586-025-09529-3. Epub 2025 Sep 17. Erratum in: Nature. 2025 Nov;**647**(8091):E8. doi: 10.1038/s41586-025-09879-y. PMID: 40963019; PMCID: PMC12589094.

Affiliations

1Division of AI in Oncology, German Cancer Research Centre DKFZ, Heidelberg, Germany.

2European Molecular Biology Laboratory, European Bioinformatics Institute EMBL-EBI, Hinxton, UK.

3Faculty of Biosciences, Heidelberg University, Heidelberg, Germany.

4Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.

5Statistics Denmark, Copenhagen, Denmark.

6Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland.

7Department of Public Health, University of Copenhagen, Copenhagen, Denmark.

8ROCKWOOL Foundation, Copenhagen, Denmark.

9European Molecular Biology Laboratory, European Bioinformatics Institute EMBL-EBI, Hinxton, UK. birney@ebi.ac.uk.

10European Molecular Biology Laboratory, European Bioinformatics Institute EMBL-EBI, Hinxton, UK. tomas@ebi.ac.uk.

11Division of AI in Oncology, German Cancer Research Centre DKFZ, Heidelberg, Germany. moritz.gerstung@dkfz.de.

12European Molecular Biology Laboratory, European Bioinformatics Institute EMBL-EBI, Hinxton, UK. moritz.gerstung@dkfz.de.

13Faculty of Mathematics and Computer Science, Heidelberg University, Heidelberg, Germany. moritz.gerstung@dkfz.de.

14Robert Bosch Center for Tumor Diseases, Stuttgart, Germany. moritz.gerstung@dkfz.de.

15Medical Faculty, Eberhard-Karls-University, Tübingen, Germany. moritz.gerstung@dkfz.de.

16University Hospital Tübingen, Tübingen, Germany. moritz.gerstung@dkfz.de.

#Contributed equally.

PMID: 40963019

Keywords.

Autor-Presentación:

Jorge Orlando Cortés Torres
Profesor Asociado de la Facultad de Odontología.
Departamento de medicina oral, ciencias básicas y cirugía.
Universidad Nacional de Colombia, Bogotá D.C., Colombia – Sur América
Correspondencia: jocortest@unal.edu.co
incoendo@gmail.com

Odontólogo (1983); Especialista en Endodoncia (1988); Magister Evaluación en Educación (1999); Diplomado en Urgencias Odontológicas Ambulatorias (2012); Diplomado de entrenamiento de Implantología Básica y Avanzada (2013); Diplomado en Implantes. Par académico del Ministerio de Educación Nacional; Profesor fundador de cinco postgrados de Endodoncia de los cuales, dos como creador y director.
Autor de 13 libros de Endodoncia.
Práctica clínica Institucional (Fuerza Aérea Colombiana durante 20 años) y práctica clínica desde la graduación en 1983.
Profesor universitario en varias universidades de forma ininterrumpida desde 1986.

Formato y metodología:

Shmatko A, Jung AW, Gaurav K, Brunak S, Mortensen LH, Birney E, Fitzgerald T, Gerstung M. Learning the natural history of human disease with generative transformers. **Nature**. 2025 Nov;647(8088):248-256. doi: 10.1038/s41586-025-09529-3. Epub 2025 Sep 17. Erratum in: Nature. 2025 Nov;647(8091):E8. doi: 10.1038/s41586-025-09879-y.

Se han mantenido los conceptos literalmente iguales a su publicación en el idioma original. Para evitar confusión, los análisis e intervenciones (argumentadas con referencia científica verificable) del profesor JO. Cortés-Torres van precedidos y finalizados como se explica a continuación indicando el inicio y fin del comentario y con fondo de color en el texto.

Los **casos clínicos presentados** (todos del profesor JO Cortés-Torres) han sido con el consentimiento informado del paciente que permiten recrear y comprender un poco mejor la presentación.

Comentarios JO Cortés-Torres

Fin de los comentarios

Tipo de publicación: Conferencia – Un análisis crítico

Introducción

Fuente de la información:

<https://aunova.net/en/blog/delphi-2m-ai-revolutionizing-disease-prediction/>

En el panorama de la tecnología sanitaria, en constante evolución, el lanzamiento de **Delphi-2M** marca un hito para la medicina predictiva.

Presentado este septiembre, **Delphi-2M** es un modelo de inteligencia artificial de código abierto capaz de predecir el riesgo de una persona de padecer más de 1000 enfermedades, extendiendo sus predicciones años, e incluso décadas, hacia el futuro.

¿Qué es Delphi-2M?

Desarrollado por un consorcio internacional que incluye al Centro Alemán de Investigación del Cáncer, **Delphi-2M** utiliza una arquitectura de IA transformadora —similar a la que impulsa los chatbots avanzados— pero está diseñado específicamente para registros médicos.



En lugar de analizar el lenguaje, Delphi-2M trata los diagnósticos y eventos médicos como **"tokens"** en el historial de salud de cada paciente.

Al aprender cómo se desarrollan estos eventos a lo largo del tiempo en cientos de miles de registros anonimizados, el modelo puede detectar patrones sutiles que escapan a las herramientas de riesgo tradicionales.

Prediciendo el futuro de su salud.

Con base en los diagnósticos, tratamientos y factores de estilo de vida previos de un paciente, Delphi-2M puede generar un pronóstico estadístico sobre el momento probable y el riesgo de su próximo evento de salud importante.

¿El resultado? Una **cronología** de posibles afecciones (**neumonía, insuficiencia cardíaca, diabetes e incluso la muerte**) con probabilidades predichas.

Para enfermedades con progreso bien definidas, Delphi-2M **iguala o supera** a las calculadoras especializadas, pero su mayor ventaja radica en su amplitud: ofrece un pronóstico general para miles de enfermedades a la vez.

Seguro, Confiable y Abierto

La apertura es un valor fundamental de Delphi-2M.

Su código, arquitectura de modelo y conjuntos de datos sintéticos —estadísticamente similares a los de pacientes reales, pero que respetan la privacidad— están disponibles gratuitamente en GitHub para que cualquiera pueda usarlos, adaptarlos o desarrollarlos.

El acceso directo a datos médicos reales, como los del Biobanco del Reino Unido, requiere solicitudes formales y estrictas medidas de protección de la privacidad.

Los **datos sintéticos** de Delphi-2M permiten realizar experimentos sólidos sin comprometer la confidencialidad.

El futuro: Privacidad criptográfica para la IA médica

Si bien Delphi-2M actualmente se basa en datos sintéticos y estrictos controles de acceso, las tecnologías criptográficas emergentes podrían revolucionar la forma en que los modelos de IA médica se entrenan y operan con datos confidenciales.

El **cifrado totalmente homomórfico** (FHE: [Fully Homomorphic Encryption](#)) permite, en teoría, que los modelos de IA se entrenen y realicen predicciones directamente sobre datos de pacientes cifrados, sin exponer nunca los historiales médicos reales, ni siquiera durante el cálculo.

Combinados con algoritmos de firma digital (DSA: [Digital Signature Algorithms](#)), en particular variantes resistentes a la computación cuántica como ML-DSA, los sistemas de IA médica podrían proporcionar garantías criptográficas de privacidad y autenticidad, asegurando que las predicciones provengan de modelos verificados e inalterados, a la vez que se mantiene la **absoluta confidencialidad del paciente**.

Este **enfoque criptográfico** podría expandir drásticamente el desarrollo de IA médica de código abierto.

En lugar de requerir aprobaciones institucionales complejas y la generación de datos sintéticos, investigadores de todo el mundo podrían mejorar de forma colaborativa modelos como **Delphi-2M** utilizando datos médicos reales cifrados de múltiples instituciones.

Sin embargo, las implementaciones actuales de cifrado de historial completo (FHE: [Fully Homomorphic Encryption](#)) presentan importantes limitaciones de rendimiento —son entre 1000 y 10.000 veces más lentas que los cálculos estándar—, lo que las hace poco prácticas para arquitecturas transformadoras a gran escala que procesan más de 400.000 registros de pacientes.

A medida que la aceleración del hardware y las mejoras algorítmicas avancen durante la próxima década, esta base criptográfica podría permitir una investigación de IA médica verdaderamente global que preserve la privacidad, al tiempo que mantiene las rigurosas exigencias de seguridad del sector sanitario.

¿Por qué es importante?

Delphi-2M transforma la atención médica, **pasando de** un enfoque reactivo a uno proactivo.

Al proporcionar a médicos, hospitales y responsables políticos información temprana sobre riesgos para la salud, el modelo podría ayudar a orientar la atención preventiva, asignar recursos de manera eficiente e impulsar nuevas formas de medicina personalizada.

Para los investigadores, los **datos sintéticos** de **Delphi-2M** superan el obstáculo de la privacidad de los datos, democratizando los estudios sobre el riesgo de enfermedades y la ética de la IA.

¿Qué sigue?

- Si bien **Delphi-2M** establece un nuevo estándar, **aún no está listo** para su implementación clínica.
- Los investigadores trabajan para **integrar** la genómica, imágenes y biomarcadores **en tiempo real** en futuras versiones, validándolos al mismo tiempo en cuanto a seguridad y confiabilidad.
- A medida que las herramientas de IA como **Delphi-2M** maduran, la atención médica predictiva, preventiva y universalmente accesible se acerca cada vez más.

Delphi-2M ofrece un vistazo a la medicina del futuro: abierta, innovadora y centrada en mantener a las personas sanas, no solo en tratarlas cuando enferman.

El futuro, al parecer, se presenta un poco más predecible.

Fin de los comentarios - Introducción

PARTE 1: Historia Natural de la Enfermedad



Franklin White

Department of Community Health and Epidemiology, Dalhousie University, Halifax, NS, Canada; Pacific Health and Development Sciences, Victoria, BC, Canada



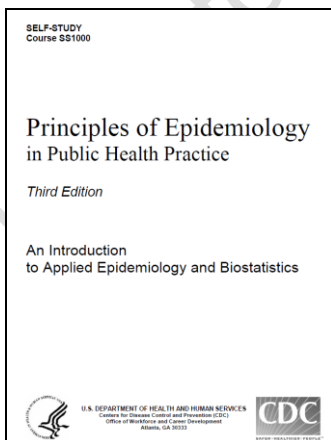
White F. Application of Disease Etiology and Natural History to Prevention in Primary Health Care: A Discourse. *Medical Principles and Practice*. 2020;29(6):501-513. doi: 10.1159/000508718. Epub 2020 May 18. PMID: 32422632; PMCID: PMC7768156.

Aspectos destacados del estudio

- La **etiología y la historia natural** son claves para identificar oportunidades de intervención.
- La prevención se define en tres niveles: primario, secundario y terciario.
- Las **estrategias** incluyen la promoción de la salud, la reducción de factores de riesgo, la detección de casos, el cribado, la atención a la salud funcional, la prevención de errores y los modelos integrados.
- La **alfabetización en salud y el lenguaje** son consideraciones fundamentales para la **comunicación** pública y la atención centrada en el paciente.

Historia natural de la enfermedad

La historia natural de la enfermedad se refiere al progreso de un proceso patológico en un individuo a lo largo del tiempo, en ausencia de tratamiento [DCD, 2012].



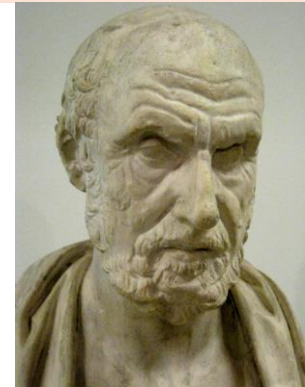
Natural History and Spectrum of Disease

Natural history of disease refers to the progression of a disease process in an individual over time, in the absence of treatment. For example, untreated infection with HIV causes a spectrum of clinical problems beginning at the time of seroconversion (primary HIV) and terminating with AIDS and usually death. It is now recognized that it may take 10 years or more for AIDS to develop after seroconversion.⁴³ Many, if not most, diseases have a characteristic natural history, although the time frame and specific manifestations of disease may vary from individual to individual and are influenced by preventive and therapeutic measures.

- Centers for Disease Control and Prevention. Principles of epidemiology in public health practice, third edition. 9. Natural history and spectrum of disease. Atlanta: CDC; 2012.

«Uno de los primeros en considerar la enfermedad como un fenómeno natural y no sobrenatural».

Pappas G, Kiriaze IJ, Falagas ME. Insights into infectious disease in the era of Hippocrates. Int J Infect Dis. 2008 Jul;12(4):347-50. doi: 10.1016/j.ijid.2007.11.003. Epub 2008 Feb 21. PMID: 18178502.



Hipócrates de Cos
(460 a.C. - 370 a.C.)

En la actualidad, el conocimiento sobre la **historia natural** de cualquier enfermedad se basa principalmente en observaciones de personas afectadas a las que se les ha dado seguimiento a lo largo del tiempo.

Idealmente, esto requiere estudios de cohortes definidas que comiencen con el inicio de la enfermedad; sin embargo, aunque se han realizado estudios de este tipo para numerosas enfermedades, el rigor puede ser difícil de lograr, especialmente para las afecciones crónicas de inicio insidioso, que representan una carga cada vez mayor a nivel mundial. **Por ejemplo**, es prácticamente imposible reclutar cohortes de inicio para la **esclerosis múltiple**, ya que el inicio clínico suele ser discordante con el inicio biológico y la duración de la enfermedad puede exceder la trayectoria profesional del investigador, incluso su propia esperanza de vida.

- Ebers GC. Natural history of multiple sclerosis. J Neurol Neurosurg Psychiatry. 2001 Dec;71 Suppl 2(Suppl 2):ii16-9. doi: 10.1136/jnnp.71.suppl_2.ii16. PMID: 11701779; PMCID: PMC1765574.

Históricamente, gran parte de lo que sabemos sobre la historia natural de la mayoría de las enfermedades se ha ido construyendo de forma pragmática, a partir de observaciones clínicas perspicaces y estudios formales de fases definidas de una afección, ya sean estudios de etiología, prevención y eficacia del tratamiento, o de resultados pronósticos.

De este modo, a menudo a partir de múltiples fuentes, se construye una imagen integral de un proceso dinámico de la enfermedad (**su historia natural**).

Existen **ejemplos relevantes** para todas las categorías de enfermedades, por ejemplo:

- Lesiones, desgarros de menisco [39]
- Traumatismos infantiles [40]
- Enfermedades infecciosas, por ejemplo, tracoma [41]
- Virus del papiloma humano (VPH) [42]
- Enfermedades no transmisibles, por ejemplo, diabetes tipo 2 [43]
- Enfermedad de Crohn [44]
- Trastornos mentales y del comportamiento, por ejemplo, esquizofrenia [45]
- Trastorno bipolar [46]

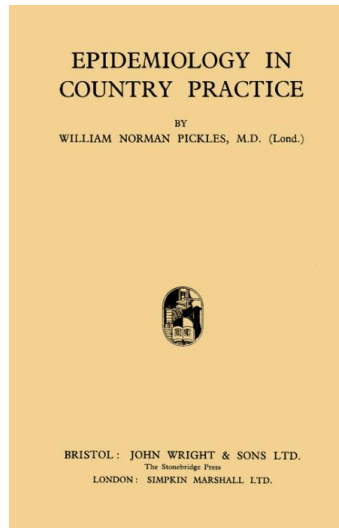
Dicho conocimiento se reorganiza continuamente a la luz de las investigaciones y la experiencia que van surgiendo, mientras que la gestión de enfermedades emergentes, como la COVID-19, se ve particularmente dificultada por el escaso conocimiento existente que permite fundamentar las intervenciones iniciales.

Existe una **venerable tradición de investigación observacional** en el campo de la medicina general, la cual se ha fortalecido con el surgimiento de la medicina familiar como especialidad.

Entre sus referentes se encuentra **William N. Pickles**, autor de la **obra clásica** titulada ***Epidemiología en la práctica rural***.

Su contribución científica consistió en documentar, a un nivel nunca antes alcanzado, el desarrollo y la evolución de las epidemias de enfermedades infecciosas.

Lo logró mediante una observación perspicaz, la elaboración de gráficos y un excelente registro de datos.



Epidemiología en la práctica rural.
Reedición de 1949



William N. Pickles
(1885-1969)

Médico general rural en Wensleydale, North Yorkshire, Inglaterra.

En **1939** escribió lo siguiente:

“Permítanme recomendar... en particular a quienes... se inician en la práctica rural, esta observación de la historia natural de las enfermedades epidémicas...

Nosotros... los médicos estamos en condiciones de aportar datos a partir de nuestras observaciones de la naturaleza, y considero que es nuestro deber aprovechar esta oportunidad única”.

Pickles WN. Epidemiology in country practice. Torquay: The Devonshire Press Ltd.; 1939.

<https://www.sciencedirect.com/book/monograph/9781483167657/epidemiology-in-country-practice>

A partir de estos orígenes tan diversos, la disciplina de la epidemiología clínica se ha formalizado progresivamente, contribuyendo significativamente a la elucidación de la etiología y la historia natural de las enfermedades, así como a nuestro conocimiento actual y creciente sobre la prevención de enfermedades.

Como se ilustra en la **Figura 1** (un modelo genérico), el paso del tiempo de una persona está representado por la línea gruesa que se mueve de izquierda a derecha. Así, una persona sana puede entrar en un proceso subclínico que, a su vez, puede cruzar un umbral para ser reconocido como una enfermedad clínica.

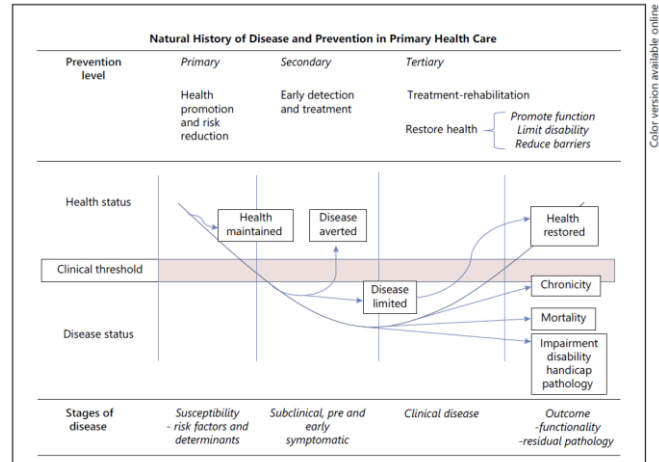
El modelo muestra cómo la persona puede salir de la situación clínica con una gama de posibles resultados, desde la muerte hasta la recuperación parcial o total.

- Esta secuencia dinámica se aplica a **prácticamente todas** las afecciones:

En el caso de las enfermedades infecciosas, la transición de la enfermedad subclínica a la clínica refleja el período de incubación (intervalo de tiempo entre la invasión por un agente infeccioso y la aparición de síntomas o signos).

En el caso del cáncer (por ejemplo) es el período latente o intervalo entre la exposición y las manifestaciones (un proceso complejo de múltiples etapas en el que los cambios celulares iniciales pueden volverse irreversibles con el progreso a una neoplasia detectable).

Para las personas que superan el umbral clínico y se consideran enfermas, el tratamiento se convierte en la prioridad de intervención.



- No obstante, existen oportunidades para la intervención preventiva en cada etapa de la historia natural de la enfermedad.
- El tratamiento en sí mismo puede ser preventivo, no de la enfermedad en sí, sino de sus consecuencias en términos de cronicidad o resultado funcional;
- puede ser preventivo indirectamente, como el tratamiento de personas con una enfermedad transmisible de persona a persona, lo que a su vez puede reducir la tasa de transmisión secundaria;
- por ejemplo, las personas en tratamiento antirretroviral con una carga viral indetectable no pueden transmitir el VIH a otras personas;
- puede resultar en una decisión de cuarentena, como en la actual pandemia de COVID-19, para la cual (al momento de escribir este texto) no se ha identificado ningún fármaco o vacuna totalmente eficaz.

Toda representación de la historia natural de una enfermedad tiene limitaciones.

- **Por ejemplo**, los factores de riesgo y determinantes mostrados en la columna de prevención primaria (Fig. 1) también pueden aplicarse a las columnas siguientes, que tratan sobre la prevención secundaria y terciaria.
- **Así mismo**, si bien el diagrama se centra en la prevención en el contexto de la atención primaria de salud, podrían ser útiles diagramas más específicos para enfermedades cuya intervención se centre en los factores de su progreso.

PARTE 2. Aprender la historia natural de las enfermedades humanas con transformadores generativos.

ANTECEDENTES

La toma de decisiones en el ámbito sanitario se basa en la comprensión del estado de salud pasado y presente de los pacientes para predecir y, en última instancia, modificar su evolución futura.

- Zhu, Z. et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. Nat. Commun. 9, 224 (2018).
- Link, B. G. & Phelan, J. Social conditions as fundamental causes of disease. J. Health Soc. Behav. <https://doi.org/10.2307/2626958> (1995).
- Nyberg, S. T. et al. Association of healthy lifestyle with years lived without major chronic diseases. JAMA Intern. Med. 180, 760–768 (2020).

Article

Learning the natural history of human disease with generative transformers

<https://doi.org/10.1038/s41586-025-09529-3>

Received: 18 May 2024

Accepted: 13 August 2025

Published online: 17 September 2025

Artem Shmatko^{1,2,3,4}, Alexander Wolfgang Jung^{2,4,5,6,7}, Kumar Gaurav^{2,3}, Søren Brunak^{4,7},
Laust Hvas Mortensen^{3,8}, Ewan Birney^{9,10}, Tom Fitzgerald¹¹ & Moritz Gerstung^{1,2,3,4,5,6,7,12}

Los métodos de **inteligencia artificial (IA)** prometen facilitar esta tarea mediante el **APRENDIZAJE DE PATRONES** del progreso de enfermedades a partir de grandes conjuntos de datos de historiales clínicos

- Kraljevic, Z., Yeung, J. A., Bean, D., Teo, J. & Dobson, R. J. Large language models for medical forecasting—foresight 2. Preprint at <https://arxiv.org/abs/2412.10848> (2024).
- Yang, L. et al. Advancing multimodal medical capabilities of Gemini. Preprint at <https://arxiv.org/abs/2405.03162> (2024).

➤ Sin embargo, su potencial no se ha investigado completamente a gran escala.

En este trabajo, modificamos la arquitectura GPT6 (**Transformador Generativo Preentrenado**) para modelar el progreso y la naturaleza competitiva de las enfermedades humanas.

Entrenamos este **modelo, Delphi-2M**, con datos de 0.4 millones de participantes del Biobanco del Reino Unido y lo validamos utilizando datos externos de 1.9 millones de individuos daneses sin modificar los parámetros.

Delphi-2M predice las tasas de más de 1000 enfermedades, condicionadas al historial clínico de cada individuo, con una precisión comparable a la de los modelos existentes para enfermedades individuales.

La naturaleza generativa de **Delphi-2M** también permite el muestreo de trayectorias sintéticas de salud futuras, proporcionando estimaciones significativas de la posible carga de enfermedad hasta por 20 años y posibilitando el entrenamiento de modelos de IA que nunca han visto datos reales.

Los métodos de IA explicables...

- Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. 30, 4768–4777 (2017).

Ofrecen información sobre las predicciones de **Delphi-2M**, revelando grupos de comorbilidades dentro y entre los capítulos de enfermedades y sus consecuencias dependientes del tiempo en la salud futura, pero también ponen de manifiesto los sesgos aprendidos a partir de los datos de entrenamiento.

En resumen, los modelos basados en transformadores parecen ser idóneos para tareas predictivas y generativas relacionadas con la salud, son aplicables a conjuntos de datos a escala poblacional y proporcionan información sobre las dependencias temporales entre eventos de enfermedad, lo que podría mejorar la comprensión de los riesgos para la salud personalizados e informar los enfoques de medicina de precisión.

El **progreso de las enfermedades humanas** a lo largo de la vida se caracteriza por periodos de salud, episodios de enfermedad aguda y también debilitamiento crónico, que a menudo se manifiestan como agrupaciones de comorbilidades.

Los patrones de **multimorbilidad** afectan a las personas de forma **desigual** y se han asociado con:

- El estilo de vida
- Rasgos hereditarios y
- Nivel socioeconómico

- Zhu, Z. et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* 9, 224 (2018).
- Link, B. G. & Phelan, J. Social conditions as fundamental causes of disease. *J. Health Soc. Behav.* <https://doi.org/10.2307/2626958> (1995).
- Nyberg, S. T. et al. Association of healthy lifestyle with years lived without major chronic diseases. *JAMA Intern. Med.* 180, 760–768 (2020).

Comprender los **riesgos de multimorbilidad** de cada individuo es importante para:

- Personalizar las decisiones sanitarias
- Motivar cambios en el estilo de vida o
- Referir a los pacientes a programas de cribado, como en el caso del cáncer

Mukama, T. et al. Risk-adapted starting age of screening for relatives of patients with breast cancer. *JAMA Oncol.* 6, 68–74 (2020).

Tian, Y. et al. Calculating the starting age for screening in relatives of patients with colorectal cancer based on data from large nationwide data sets. *Gastroenterology* 159, 159–168 (2020).

Es **fundamental comprender** que la salud no solo puede entenderse a partir de la presentación de diagnósticos individuales, sino también en el **contexto de las comorbilidades** de cada persona y su evolución a lo largo del tiempo.

Si bien existe una amplia gama de algoritmos de predicción para enfermedades específicas, desde enfermedades cardiovasculares hasta cáncer (10-12), pocos algoritmos son capaces de predecir el espectro completo de enfermedades humanas, que reconoce más de 1000 diagnósticos en el nivel superior de la Clasificación Internacional de Enfermedades, Décima Revisión (CIE-10).

Aprender y predecir patrones de progreso de enfermedades también es importante en poblaciones que envejecen y que presentan cambios en las morbilidades demográficas subyacentes.

Por ejemplo, se ha pronosticado que, a nivel mundial:

- El número de diagnósticos de cáncer **aumentará un 77 %** para 2050 (ref. 13)
- que, en el Reino Unido, el número de personas en edad laboral con **enfermedades graves**, como depresión, asma, diabetes, enfermedades cardiovasculares, cáncer o demencia, aumentará de **3 a 3.7 millones** para 2040 (ref. 14).

Por lo tanto, modelar la carga esperada de enfermedad es fundamental para la planificación sanitaria y económica.

Además, el seguimiento continuo de la incidencia de enfermedades, junto con su probable prevalencia futura en los grupos de población, promueve un sistema sanitario mejor informado.

Los recientes avances en IA pueden ayudar a abordar algunas limitaciones metodológicas de la modelización de la multimorbilidad, que hasta ahora han resultado difíciles de superar¹⁵.

Además del elevado número de diagnósticos, estos incluyen desafíos en la modelización de las dependencias temporales entre eventos previos, la integración de datos potencialmente diversos con relevancia pronóstica y la calibración estadística de las predicciones.

Los modelos de lenguaje a gran escala (LLM: Modelos lineales Aleatorizados) (16-19) —un subcampo de la IA que permite el funcionamiento de chatbots como ChatGPT (20-21)— modelan el lenguaje como una secuencia de fragmentos de palabras (tokens).

Generado **token a token**, el nuevo texto se basa en todo el texto precedente y, con suficiente entrenamiento, las dependencias estadísticas entre estos tokens resultan suficientes para producir texto contextualizado e incluso conversacional, a menudo indistinguible del de un humano.

La analogía entre los LLM (Modelos lineales Aleatorizados) y el modelado del progreso de enfermedades, que también implica reconocer eventos pasados y explotar sus dependencias mutuas para predecir la secuencia futura de morbilidad, ha inspirado recientemente una serie de nuevos modelos de IA.

Por ejemplo, se han desarrollado modelos basados en BERT (22-25) para tareas de predicción específicas.

Los modelos Transformer, entrenados con registros electrónicos de salud, se han utilizado para predecir diagnósticos como cáncer de páncreas²⁶, autolesiones²⁵ y accidentes cerebrovasculares²⁴, así como parámetros no clínicos como la autoestima²⁷.

Sin embargo, a pesar de las prometedoras pruebas de concepto (4,28,29), el potencial para el modelado integral y generativo de la **multimorbilidad** aún no se ha evaluado completamente.

Aquí demostramos que los modelos transformadores basados en atención, similares a los modelos lineales aleatorizados (LLM), pueden extenderse para aprender trayectorias de salud a lo largo de la vida y predecir con precisión las tasas futuras de más de 1000 enfermedades simultáneamente, basándose en diagnósticos previos, factores de estilo de vida y otros datos informativos.

Nuestro modelo extendido, denominado **Delphi-2M**, se entrenó con datos del Biobanco del Reino Unido, una cohorte de investigación a escala poblacional, y se validó con registros de población daneses.

- El vocabulario del modelo incluye los códigos de diagnóstico de nivel superior de la CIE-10, así como sexo, índice de masa corporal, tabaquismo, consumo de alcohol y mortalidad.
- Delphi proporciona predicciones a nivel individual de la incidencia de múltiples enfermedades y modela trayectorias de salud futuras en cualquier momento del ciclo vital de una persona.
- Además, el modelo interno de Delphi ofrece información sobre cómo los datos pasados influyen en las tasas de enfermedades posteriores.
- Asimismo, evaluamos los sesgos y la equidad entre subgrupos demográficos y analizamos el potencial de Delphi como marco para la modelización de la atención sanitaria.

Discusión

Presentamos **Delphi-2M**, un modelo basado en GPT para el progreso de múltiples enfermedades.

- Delphi-2M extiende el modelo de lenguaje a gran escala GPT para considerar la naturaleza temporal de las trayectorias de salud.

De forma análoga a los modelos de lenguaje a gran escala **LLM (Modelos lineales Aleatorizados)**, que aprenden la gramática y la lógica contextual del lenguaje a partir de grandes volúmenes de texto, **Delphi-2M** infirió los **patrones de progreso de múltiples enfermedades** al entrenarse con datos de más de 1000 enfermedades e información de salud basal registrada en 402.799 participantes del Biobanco del Reino Unido.

Una evaluación detallada de las predicciones de **Delphi-2M** **demostró** que reproducen consistentemente los patrones de aparición de enfermedades a escala poblacional, tal como se registran en el Biobanco del Reino Unido.

Para la mayoría de las enfermedades, el modelo multienfermedad de tiempo continuo de **Delphi-2M** predijo las tasas futuras con una precisión **comparable o superior a**:

- Modelos de riesgo de enfermedades individuales establecidos
- Otros marcos de aprendizaje automático y
- Modelos basados en biomarcadores sanguíneos.

Solo se observó una ligera disminución en el rendimiento al aplicarlo a datos de registros de enfermedades daneses, lo que demuestra que los modelos, incluso sin ajustes adicionales, son ampliamente aplicables en los sistemas nacionales de salud.

Delphi-2M tiene la capacidad única de **muestrear trayectorias futuras de enfermedades**, lo que permite estimar la carga acumulada de enfermedades durante periodos de hasta 20 años, condicionada a la información de salud previa.

Cabe destacar que las predicciones de Delphi generalmente están fuertemente influenciadas por el azar estadístico y son compatibles con una variedad de resultados para un individuo determinado.

La capacidad de generar datos sintéticos también puede ayudar a crear conjuntos de datos que preserven los patrones de coocurrencia estadística sin revelar ningún dato específico, lo que podría facilitar el desarrollo de modelos de IA adicionales con un menor riesgo de revelar información personal.

- **Delphi-2M** ofrece información sobre los modos de progreso de las enfermedades.

Su capacidad para agrupar los riesgos de enfermedad puede ser útil para estudios de asociación genómica centrados en comorbilidades o estratificados según los riesgos derivados de trayectorias de salud.

*La capacidad de **Delphi-2M** para cuantificar la influencia temporal de datos de salud previos reveló que los cánceres aumentan la mortalidad de forma sostenida, **mientras que** los efectos del infarto de miocardio o la septicemia remiten en un plazo de 5 años.*

Análisis similares también revelaron grupos de comorbilidades persistentes, como **trastornos de salud mental**.

Si bien **Delphi-2M** parece ser capaz de modelar dependencias dirigidas temporalmente, advertimos sobre la interpretación de estas como relaciones causales que podrían utilizarse para modificar la evolución futura de la salud.

También existen varias **LIMITACIONES** que deben tenerse en cuenta.

Un análisis detallado de los datos de primera aparición del Biobanco del Reino Unido reveló varios sesgos, reflejados en **Delphi-2M**.

Además del sesgo de selección de voluntarios sanos y el sesgo de selección de participantes antes del reclutamiento, la diversidad de las fuentes de datos de salud afectó la predicción de Delphi, ya que se aprovecharon los patrones de datos faltantes específicos del Biobanco del Reino Unido para inferir las tasas de enfermedad.

Asi mismo, **Delphi-2M** predice diferentes tasas de enfermedad en subgrupos según el origen étnico y los índices de privación, pero no se observa ninguna tendencia entre las medidas de estilo de vida y el año de nacimiento.

Estos hallazgos subrayan la necesidad de precaución al utilizar modelos de IA para la inferencia y la predicción en conjuntos de datos sanitarios heterogéneos, lo que podría indicar que son complementos útiles para los procesos de diagnóstico actuales, en lugar de sustitutos.

Una característica prometedora de la implementación de **Delphi-2M** es la relativa sencillez para **incorporar capas de datos adicionales**, posible gracias a su arquitectura basada en transformadores.

Las mejoras inmediatas de **Delphi-2M** podrían incorporar datos adicionales sobre estilo de vida, estado de salud autoinformado, registros de prescripciones y análisis de sangre, información que suele estar disponible en el ámbito sanitario general.

Otras extensiones multimodales podrían incluir datos genómicos, información metabolómica más completa, datos de diagnóstico por imagen o datos de dispositivos portátiles que se pueden añadir a la capa de incrustación de **Delphi-2M**, de forma similar a como se incorporan actualmente los tokens de estilo de vida (Figura 9 de Datos Extendidos).

Además, si bien la **CIE-10** proporciona una **tokenización predefinida** de enfermedades, se ha mostrado que los modelos de lenguaje natural **LLM (Modelos lineales Aleatorizados)**, también conceptualizan el lenguaje natural, lo que hace plausible esperar que los modelos futuros puedan derivar un significado similar directamente de los registros de texto libre, permitiendo la aplicación de modelos tipo Delphi a datos no estructurados.

Por último, **Delphi-2M** podría servir como una extensión de los **LLM (Modelos lineales Aleatorizados)**.

De forma similar a los sistemas que proporcionan a los **LLM (Modelos lineales Aleatorizados)** resultados de búsqueda web relevantes para sus consultas con el fin de reducir las alucinaciones³³, **un futuro LLM** orientado a la atención médica podría recurrir a un modelo basado en Delphi para mejorar la precisión numérica de las respuestas generadas³⁴.

33. Shuster, K., Poff, S., Chen, M., Kiela, D. & Weston, J. Retrieval augmentation reduces hallucination in conversation. Findings of the Association for Computational Linguistics: EMNLP 2021 3784–3803 (2021); <https://aclanthology.org/2021.findings-emnlp.320/>.

34. Imrie, F., Rauba, P. & van der Schaar, M. Redefining digital health interfaces with large language models. Preprint at <https://arxiv.org/abs/2310.03560> (2023).

Una aplicación evidente de los modelos tipo Delphi es...

El apoyo a la toma de decisiones médicas mediante la integración racional de información procedente de diversas modalidades de datos, lo que puede suponer un reto para los profesionales sanitarios.

Entre los posibles casos de uso se incluyen la identificación de personas que se beneficiarían más de las pruebas diagnósticas o la detección de personas con un riesgo de enfermedad lo suficientemente alto como para incluirlas en programas de cribado, incluso si aún no cumplen los criterios convencionales basados en la edad.

Sin embargo, la implementación de sistemas de apoyo a la toma de decisiones clínicas requiere un marco regulatorio, que todavía se encuentra en sus primeras etapas para la IA en el sector sanitario.

Una alternativa para el uso de los modelos Delphi sería informar a los proveedores de atención médica, las aseguradoras y los responsables políticos.

Estas aplicaciones, que informan la prestación de atención médica en lugar de las decisiones de tratamiento, requieren menos regulación.

El modelado Delphi tiene una resolución a nivel individual; cuando estas predicciones se agregan, los modelos tipo Delphi pueden proporcionar beneficios sustanciales para el modelado a nivel de sistema, particularmente al proyectar la carga de enfermedad esperada a nivel local, regional y nacional, o para grupos demográficos específicos de interés.

- Esto podría revelar comunidades con necesidades de **atención médica futuras no satisfechas** en las próximas una o dos décadas y brindar la oportunidad de ajustar la prestación de atención médica.
- Estas capacidades resultan especialmente valiosas en **poblaciones que envejecen**, donde las necesidades de atención médica son cada vez más complejas y requieren más recursos.

Estas consideraciones ilustran la amplia gama de aplicaciones de los modelos generativos para la investigación biomédica y, en última instancia, también para la atención médica.

Con la capacitación y evaluación adecuadas, las futuras extensiones de modelos multimodales pueden utilizarse para la medicina preventiva, el apoyo a la toma de decisiones clínicas y la planificación de la atención médica.

Nuestro modelo y análisis representan un paso más hacia el aprovechamiento de los considerables beneficios para la atención médica de la era de la IA.

To cite this publication: Cortés-Torres JO. Natural History of Human Diseases: Past and Future. Conference at the Colombian Institute of Endodontics, Bogotá, July 20-2026

El acceso al conocimiento debe ser abierto

Esta suscripción no tiene costo.

Open access

Email: incoendo@gmail.com /

wasap:

+57 315 3517449

42 Principios Básicos: Así de simple

Jean-Baptiste Poquelin
Molière

No solo nos responsabilizamos por lo que hacemos, *sino también por lo que no hacemos.*

Poeta y dramaturgo francés, conocido en la actualidad por ser uno de los mejores escritores de la historia.
Comparado con Cervantes o William Shakespeare, algo que muy pocos otros dramaturgos de todos los tiempos pueden tan siquiera igualar.

Enero 15-1622 – Feb. 17-1673

Cortés Torres Jorge Orlando

