

Análisis asistido por inteligencia artificial de la eficacia, la seguridad y la eficiencia de concentrados de factor VIII pegilados en la hemofilia A basado en las fichas técnicas

Artificial intelligence-assisted analysis of efficacy, safety, and efficiency of pegylated factor VIII concentrates in hemophilia A based on product inserts

Manuel Rodríguez-López^{1*}, Michael Calviño-Suárez², Ramiro Núñez-Vázquez³
y Raquel Ocampo-Martínez¹

¹Servicio de Hematología y Hemoterapia, Hospital Universitario Álvaro Cunqueiro, Instituto de Investigación Sanitaria Galicia-Sur (IISGS), EOXI Vigo, Vigo, Pontevedra; ²Servicio de Hematología y Hemoterapia, Complejo Hospitalario Universitario de Santiago de Compostela, Santiago de Compostela, La Coruña; ³Unidad de Trombosis y Hemostasia, Hospital Universitario Virgen del Rocío, Sevilla. España

Resumen

Introducción: Existen distintos concentrados de factor VIII pegilados, con características diferentes. **Objetivo:** Evaluar el uso de la IA básica, específicamente ChatGPT-4o (OpenAI®), para comparar los distintos CFVIII pegilados ante la falta de estudios comparativos directos y las limitaciones de las comparaciones indirectas. **Material y método:** Se comparan Adynovi®, Jivi® y Esperoct®, a partir de sus fichas técnicas, con cálculos basados en un paciente promedio de 70 kg. ChatGPT4o realiza análisis estadísticos para comparar la eficacia (tasa anualizada de sangrado [TAS]) y el consumo de factor. **Resultados:** ChatGPT4o sugiere que Esperoct® (50 UI/kg, 2 veces/semana) es el régimen más costo-eficiente por consumo anual y TAS baja. Adynovi® muestra un perfil de seguridad ligeramente favorable, sin diferencias estadísticamente significativas en TAS o consumo anual de factor ($p > 0,05$) entre regímenes, sugiriendo similar eficacia entre las opciones. **Conclusiones:** Aunque las comparaciones indirectas emparejadas ajustadas son herramientas validadas para comparaciones indirectas en hemofilia, presentan limitaciones. ChatGPT4o podría ayudar a abordarlas sin sustituir la validación clínica. La inteligencia artificial, incluyendo ChatGPT4o, podría mejorar la accesibilidad y la precisión de las comparaciones indirectas emparejadas ajustadas, pero aún requiere una validación continua.

Palabras clave: Inteligencia artificial generativa. ChatGPT. CFVIII pegilados. Hemofilia A.

Abstract

Introduction: There are different pegylated factor VIII concentrates, with different characteristics. **Objective:** To evaluate the use of basic AI, specifically ChatGPT-4o (OpenAI®), to compare different pegylated CFVIII products due to the lack of direct comparative studies and the limitations of indirect comparisons. **Material and method:** Adynovi®, Jivi® and Esperoct® are compared from their data sheets, with calculations based on an average patient weighing 70 kg. ChatGPT4o performs statistical analyses to compare efficacy (annualised bleeding rate [ABR]) and factor consumption. **Results:** ChatGPT4o suggests that Esperoct® (50 IU/kg, 2 times/week) is the most cost-efficient regimen by annualised consumption and low ABR. Adynovi® shows a slightly favourable safety profile, with no statistically significant differences in ABR or annual factor consumption

*Correspondencia:

Manuel Rodríguez-López
E-mail: totsalcamp14@hotmail.com

Fecha de recepción: 05-11-2024
Fecha de aceptación: 30-01-2025
DOI: 10.24875/RHT.24000005

Disponible en línea: 08-05-2025
Hemost Trombos. 2025;2(1):24-28
www.revistahemostasiaytrombosis.com

2938-8740 / © 2025 Sociedad Española de Trombosis y Hemostasia (SETH) y Grupo Cooperativo Latinoamericano de Hemostasia y Trombosis (CLAHT). Publicado por Permanyer. Este es un artículo de acceso abierto bajo la CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

($p > 0.05$) between regimens, suggesting similar efficacy between the options. **Conclusions:** Although matching-adjusted indirect comparisons are validated tools for indirect comparisons in haemophilia, they face limitations. ChatGPT4o could help address them without replacing clinical validation. Artificial intelligence, including ChatGPT4o, could improve the accessibility and accuracy of matching-adjusted indirect comparisons, but still requires ongoing validation.

Keywords: Generative artificial intelligence. ChatGPT. Pegylated FVIII. Hemophilia A.

Introducción

En España, actualmente están disponibles tres concentrados de FVIII (CFVIII) de vida media extendida, en los cuales la técnica empleada para la prolongación de la semivida plasmática ha sido la pegilación. Estos CFVIII pegilados (rurioctocog alfa pegol, damoctocog alfa pegol y turoctocog alfa pegol) presentan características estructurales y farmacocinéticas que permitirían definirlos como productos de vida media extendida¹, aunque esto solo se traduce en una mejora marginal en cuanto a la ampliación del intervalo de administración (Tabla 1) en la mayoría de los pacientes. Los datos farmacocinéticos, así como los resultados de eficacia, no son necesariamente comparables entre distintos ensayos clínicos realizados con estas moléculas por las diferencias metodológicas y del perfil de pacientes estudiados. Estos productos han demostrado eficacia y seguridad para el tratamiento y la prevención del sangrado en pacientes con hemofilia A en distintos escenarios, tanto en ensayos clínicos como en la vida real. Ante la falta de estudios *head to head* que comparen los diferentes concentrados, las comparaciones indirectas emparejadas ajustadas han demostrado su utilidad para ayudar en la toma de decisiones clínicas en la hemofilia A. Con todo, la utilidad está ligada al rigor metodológico y al mantenimiento de un contexto clínico adecuado. La falta de consistencia y de medidas de contexto puede limitar la aplicabilidad de los resultados, y estos nunca deben sustituir al juicio clínico².

El objetivo del presente estudio fue evaluar el potencial de uso de la inteligencia artificial (IA) básica, en este caso ChatGPT4o (OpenAI®), como herramienta para comparar los distintos CFVIII pegilados, ante la ausencia de estudios comparativos directos y por las limitaciones de las comparaciones indirectas emparejadas ajustadas.

Método

Se toman como referencia para la comparación los datos comunicados en las fichas técnicas aprobadas por la Agencia Europea de Medicamentos³⁻⁵ y se procede a efectuar los análisis estadísticos propuestos por

ChatGPT4o, con el fin de comparar la eficacia, la seguridad y la eficiencia, considerando para los cálculos un paciente promedio de 70 kg de peso. De acuerdo con los resultados, se le solicita, a igualdad de precio, que recomiende cuál es el tratamiento con mejor ratio coste-eficiencia (Tabla 1).

Resultados

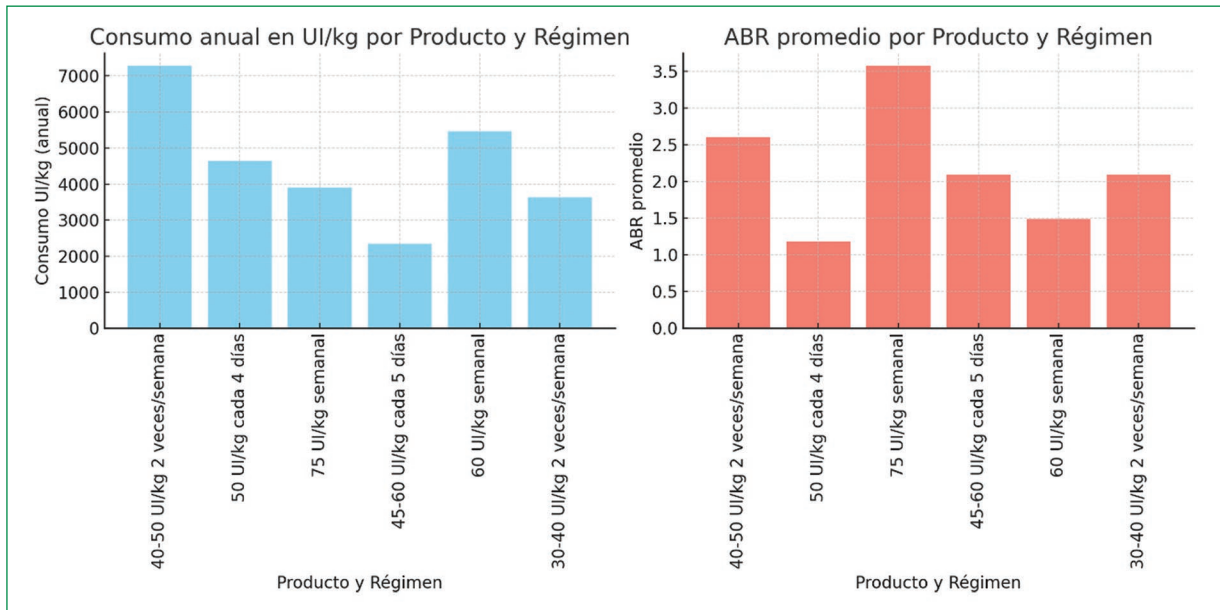
En primer lugar, basándose en los resultados de la figura 1, ChatGPT4o efectúa un análisis descriptivo y concluye que, para maximizar la eficacia y minimizar la frecuencia de administración, Esperoct® en dosis de 50 UI/kg cada 4 días sería preferible en términos de control del sangrado. Si el costo anual es una preocupación, y una eficacia levemente menor es aceptable, Jivi® semanal podría ser una opción viable por su menor consumo de UI/kg. Según las diferencias observadas en el consumo anual y la tasa anualizada de sangrado (TAS), el régimen menos recomendable sería el de Esperoct® 75 UI/kg semanal, debido a su TAS elevada (3,57). Aunque el consumo anual es relativamente bajo, la efectividad en la reducción de sangrados es menor, lo cual podría implicar un control insuficiente de los episodios hemorrágicos en los pacientes (Fig. 1).

Se le pide a ChatGPT4o que efectué un análisis estadístico más profundo y justifique las técnicas seleccionadas para llevarlo a cabo. Selecciona un test de ANOVA y la prueba de Kruskal-Wallis. ¿Por qué? Justifica el uso de ANOVA porque es especialmente útil si los datos cumplen supuestos de normalidad en cada grupo de análisis y homogeneidad de varianzas entre los grupos, siendo un análisis paramétrico robusto para comparar múltiples grupos de manera eficiente. Considera el uso de la prueba de Kruskal-Wallis por ser una alternativa no paramétrica al ANOVA, ideal cuando los datos no cumplen los supuestos de normalidad o homocedasticidad. No considera la prueba t porque está comparando varios productos, ni la χ^2 al estar considerando datos continuos, no categóricos.

En relación con la TAS, el análisis con ANOVA muestra un valor de p de 0,80 ($> 0,05$), razón por la que

Tabla 1. Concentrados de factor VIII pegilados disponibles en España

Medicamento (año de autorización)	Laboratorio titular	Principio activo	Origen/células cultivo	Vida media F8 (h)	Posología, profilaxis
Adynovi® (2018)	Baxalta Innovations GmbH	Rurioctocog alfa pegol	CHO	14,7	3 a 4 días
Jivi® (2019)	Bayer AG	Damoctocog alfa pegol	BHK	17	3 a 7 días
Esperoct® (2019)	NovoNordisk A/S	Turoctocog alfa pegol	CHO	14-20	4 días

**Figura 1.** Grafica de consumos y tasas anualizadas de sangrado por producto y régimen.

indica que no existen diferencias estadísticamente significativas en las TAS promedio entre los productos y regímenes analizados. La prueba de Kruskal-Wallis arroja un valor de p de 0,64, con lo que ambos análisis indican que las diferencias observadas en las TAS entre los productos y regímenes no son estadísticamente significativas ($p > 0,05$).

En cuanto al consumo anual de factor, el análisis con ANOVA muestra un valor p de 0,22, lo que indica que no existen diferencias estadísticamente significativas en el consumo anual en UI/kg entre los productos y regímenes. La prueba de Kruskal-Wallis muestra un valor p de 0,30. Ambos análisis indican que no hay diferencias estadísticamente significativas en el consumo anual entre los productos y regímenes ($p > 0,05$).

Respecto a la seguridad, en general, Adynovi® tiene un perfil de seguridad ligeramente favorable en comparación con Jivi® y Esperoct®, especialmente en términos

de reacciones inmunitarias y eventos adversos graves. Las reacciones adversas graves, como la anafilaxia, son menos frecuentes en comparación con las complicaciones relacionadas con anticuerpos anti-PEG observadas con los otros productos.

Cuando se le pide que efectúe un análisis de coste-eficiencia y recomiende un producto y un régimen, asumiendo el mismo coste unitario para las tres opciones, ChatGPT4o concluye que la mejor opción es Esperoct® cada 5 días, seguido del Jivi® cada 7 días. Si el objetivo es maximizar la eficiencia en relación al costo-eficacia bajo un mismo coste unitario, Esperoct® en régimen de 50 UI/kg cada 4 días se perfila como la opción más costo-eficiente en términos de balance entre bajo consumo anual (4641 UI/kg) y baja TAS (1,18), seguido por Jivi® en régimen semanal de 60 UI/kg, con menor consumo anual (3276 UI/kg) que Esperoct®, pero con una TAS ligeramente más alta (1,49).

Analizando si existen diferencias estadísticamente significativas en términos de eficiencia, aplicando la prueba de Kruskal-Wallis se obtiene un valor de p de 0,30 ($p > 0,05$) para el consumo anual de factor y de 0,64, para la TAS. Esto sugiere que, en términos de eficiencia, las diferencias observadas entre los productos y regímenes no son lo bastante significativas como para afirmar que un régimen o producto es notablemente superior en cuanto a eficiencia.

Discusión

En el momento actual, a diferencia de la IA, las comparaciones indirectas emparejadas ajustadas son herramientas estadísticas perfectamente validadas para la comparación de productos en la hemofilia A, y se han realizado para distintos CFVIII. Por ejemplo, en 2019, Batt et al.⁶ concluyeron que, al comparar damoctocog alfa pegol con otros CFVIII, como efmo-roctocog alfa, ruriotocog alfa pegol y octocog alfa, la TAS era similar, pero el uso de factor era un 20-40% menor. Con todo, esta comparación indirecta emparejada ajustada presenta limitaciones: características basales no documentadas, ausencia de un grupo de referencia común y diferencias en la duración de la exposición al tratamiento, entre otras. ChatGPT4o podría mitigar algunas de las limitaciones de las comparaciones indirectas emparejadas ajustadas, aunque no podría eliminarlas por completo debido a la naturaleza intrínseca de los datos y las limitaciones de los estudios indirectos. Por ejemplo, no podría compensar por completo la falta de un grupo de referencia común ni las limitaciones de superposición de poblaciones, ya que estos problemas son estructurales. Por otro lado, como siempre se recuerda, son fundamentales la calidad y la exactitud de los datos de entrenamiento empleados. En otras comparaciones indirectas emparejadas ajustadas^{7,8} también se han reportado limitaciones similares, por lo que los resultados, siendo interesantes, deben interpretarse con cautela y obligan a una confirmación adicional para su aplicación clínica. ChatGPT4o muestra capacidad para consolidar grandes cantidades de información de diferentes fuentes⁹, pudiendo presentarla de manera accesible, pero su uso aún requiere la validación necesaria para estos fines. Con todo, en nuestra opinión, el análisis comparativo efectuado entre los productos pegilados es bastante exacto y entra dentro de la lógica, tanto por su estructura como por los resultados reportados. Los tres CFVIII comparados han mostrado eficacia y

seguridad, cumpliendo sin problemas la función para la que fueron diseñados. Sin embargo, por el adecuado equilibrio entre consumo, carga de tratamiento y eficacia, probablemente Esperoct® a la dosis de 50 UI/kg, dos veces por semana, resulte la opción de tratamiento más adecuada entre los productos pegilados comparados.

Conclusiones

ChatGPT4o y otras plataformas de IA tienen el potencial de mejorar la eficiencia, la precisión y la accesibilidad de las comparaciones indirectas emparejadas ajustadas en hemofilia A, al simplificar el manejo de datos complejos y la interpretación de los resultados. Su integración en el proceso puede contribuir a comparaciones más robustas y a una mejor comunicación de los hallazgos, beneficiando tanto a investigadores como a clínicos y pacientes. Con todo, ChatGPT4o y otras plataformas de IA requieren procesos de validación continua que permitan implementar su uso generalizado para análisis de estas características.

Financiamiento

Ninguno.

Conflicto de intereses

Ninguno.

Consideraciones éticas

Protección de personas y animales. Los autores declaran que para esta investigación no se han realizado experimentos en seres humanos ni en animales.

Confidencialidad, consentimiento informado y aprobación ética. Los autores han seguido los protocolos de confidencialidad de su institución, han obtenido el consentimiento informado de los pacientes, y cuentan con la aprobación del Comité de Ética. Se han seguido las recomendaciones de las guías SAGER, según la naturaleza del estudio.

Declaración sobre el uso de inteligencia artificial. Los autores declaran que no utilizaron ningún tipo de inteligencia artificial generativa para la redacción de este manuscrito.

Bibliografía

1. Mahlangu J, Young G, Hermans C, Blanchette V, Berntorp E, Santagostino E. Defining extended half-life rFVIII — a critical review of the evidence. *Haemophilia*. 2018;24:348-58.
2. Batt K, Klamroth R, Mancuso ME, Tiede A, Mantovani LG. The importance of clinical context and consistency in methodology when using matching-adjusted indirect comparisons (MAICs) to compare outcomes. *Int J Gen Med*. 2024;17:3927-32.
3. Esperoct, ficha técnica. EMA. (Consultado el 5/11/2024.) Disponible en: https://ec.europa.eu/health/documents/community-register/2023/20230915160170/anx_160170_es.pdf.
4. Adynovi, ficha técnica. EMA. (Consultado el 5/11/2024.) Disponible en: https://ec.europa.eu/health/documents/community-register/2018/20180108139593/anx_139593_es.pdf.
5. Jivi, ficha técnica. EMA. (Consultado el 5/11/2024.) Disponible en: https://ec.europa.eu/health/documents/community-register/2023/20230623159402/anx_159402_es.pdf.
6. Batt K, Gao W, Ayyagari R, Deschaseaux C, Vashi PB, Yao Z et al. Matching-adjusted indirect comparisons of annualized bleeding rate and utilization of BAY 94-9027 versus three recombinant factor VIII agents for prophylaxis in patients with severe hemophilia A. *J Blood Med*. 2019;10:147-59.
7. Kessler C, Corrales-Medina FF, Mannucci PM, Tarantino MD. Matching-adjusted indirect comparison between personalized prophylaxis with simoctocog alfa versus standard prophylaxis with emicizumab in adults with hemophilia A. *Blood*. 2023;142(Suppl 1):5482.
8. Kessler CM, Corrales-Medina FF, Mannucci PM, Jiménez Yuste V, Tarantino MD. Clinical efficacy of simoctocog alfa versus extended half-life recombinant FVIII concentrates in hemophilia A patients undergoing personalized prophylaxis using a matching-adjusted indirect comparison method. *Eur J Haematol*. 2023; 111:757-67.
9. Van de Wyngaert C, Iarossi M, Hermans C. How good does ChatGPT answer frequently asked questions about haemophilia? *Haemophilia*. 2023;29:1646-8.