Hemostasia y Trombosis
REVISTA IBEROAMERICANA DE TROMBOSIS Y HEMOSTASIA

PERMANYER

Check for updates

ORIGINAL RESEARCH ARTICLE

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

# 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

*Manuel Rodríguez-López[1]*, Michael Calviño-Suárez[2], Ramiro Núñez-Vázquez[3], and Raquel Ocampo-Martínez[1]*

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

## 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 methods:** *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 ($p > 0.05$) 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 FVIIIC. Hemophilia A.*

## 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 métodos:** *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*

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

Date of reception: 05-11-2024
Date of acceptance: 30-01-2025
DOI: 10.24875/RHT.M25000036

Available online: 21-05-2025
*Hemost Trombos* (Eng). 2025;2(1):23-26
www.revistahemostasiaytrombosis.com

2938-8740 / © 2025 Sociedad Española de Trombosis y Hemostasia (SETH) y Grupo Cooperativo Latinoamericano de Hemostasia y Trombosis (CLAHT). Published by Permanyer. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

*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.*

## Introduction

In Spain, three extended half-life factor VIII concentrates (FVIII-C) are currently available, where the technique used to prolong plasma half-life is pegylation. These pegylated FVIII-C (rurioctocog alfa pegol, damoctocog alfa pegol, and turoctocog alfa pegol) exhibit structural and pharmacokinetic characteristics that allow them to be categorized as extended half-life products[1]. However, this only translates into a marginal improvement in extending the dosing interval (Table 1) for most patients. Pharmacokinetic data, as well as efficacy results, are not necessarily comparable across different clinical trials conducted with these molecules due to methodological differences and the patient profiles studied. These products have demonstrated safety and efficacy for the treatment and prevention of bleeding in patients with hemophilia A in various settings, both in clinical trials and real-world scenarios. In the absence of head-to-head studies comparing different concentrates, adjusted indirect comparison methods have proven useful for aiding clinical decision-making in hemophilia A. However, their utility is tied to methodological rigor and maintaining an appropriate clinical context. Lack of consistency and contextual measures may limit the applicability of results, and these should never replace clinical judgment[2].

The objective of this study was to evaluate the potential use of basic artificial intelligence (AI), in this case ChatGPT4o (OpenAI®), as a tool to compare different pegylated FVIII-C, given the absence of direct comparative studies and the limitations of adjusted indirect comparisons.

## Method

Data reported in the approved product summaries by the European Medicines Agency[3-5] were used as a reference for comparison. Statistical analyses proposed by ChatGPT4o were performed to compare the safety, efficacy and efficiency profile, considering an average patient weighing 70 kg. Based on the results, ChatGPT4o was asked to recommend the treatment with the best cost-efficiency ratio, assuming equal pricing (Table 1).

## Results

First, based on the results in figure 1, ChatGPT4o performed a descriptive analysis and concluded that, to maximize efficacy and minimize dosing frequency, Esperoct® at a dose of 50 IU/kg every 4 days would be preferable in terms of bleeding control. If annual cost is a concern and slightly lower efficacy is acceptable, Jivi® weekly could be a viable option due to its lower IU/kg consumption. According to the differences observed in annual consumption and annualized bleeding rate (ABR), the least recommended regimen would be Esperoct® 75 IU/kg weekly, due to its high ABR (3.57). Although annual consumption is relatively low, its effectiveness in reducing bleeding is lower, which could imply insufficient control of bleeding episodes in patients (Fig. 1).

ChatGPT4o was, then, asked to perform a deeper statistical analysis and justify the selected techniques. It chose ANOVA and the Kruskal-Wallis test. The use of ANOVA was justified because it is particularly useful if the data meet assumptions of normality within each analysis group and homogeneity of variances across groups, being a robust parametric analysis for efficiently comparing multiple groups. The Kruskal-Wallis test was considered as a non-parametric alternative to ANOVA, ideal when data do not meet normality or homoscedasticity assumptions. The t-test was not considered because multiple products were being compared, nor was the chi-square test, as continuous rather than categorical data were being analyzed.

Regarding ABR, the ANOVA analysis showed a p value of 0.80 (> 0.05), indicating no statistically significant differences in mean ABR between the products and regimens analyzed. The Kruskal-Wallis test yielded a p value of 0.64, confirming that the observed differences in ABR between products and regimens are not statistically significant ($p > 0.05$).

For annual factor consumption, the ANOVA analysis showed a p value of 0.22, indicating no statistically significant differences in annual IU/kg consumption

**Table 1.** Pegylated factor VIII concentrates available in Spain

| Drug (year of authorization) | Marketing authorization holder | Active ingredient | Origin/cell culture | Half-Life F8 (h) | Dosage, prophylaxis |
|---|---|---|---|---|---|
| Adynovi® (2018) | Baxalta Innovations GmbH | Rurioctocog alfa pegol | CHO | 14.7 | 3 to 4 days |
| Jivi® (2019) | Bayer AG | Damoctocog alfa pegol | BHK | 17 | 3 to 7 days |
| Esperoct® (2019) | NovoNordisk A/S | Turoctocog alfa pegol | CHO | 14-20 | 4 to 7 days |

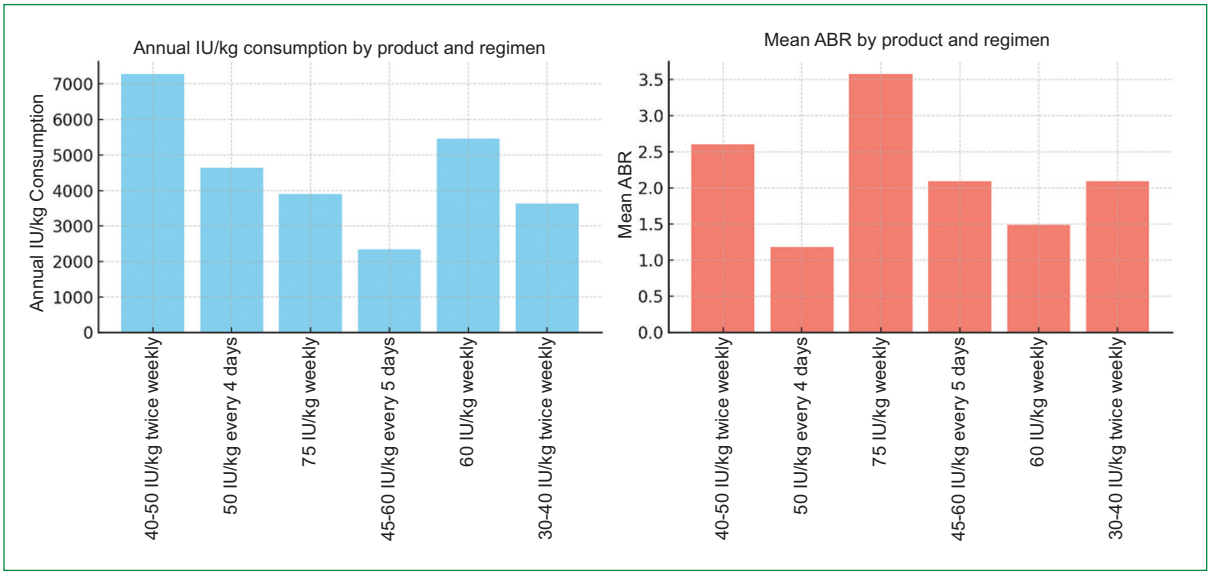


**Figure 1.** Graph of annual consumption and annualized bleeding rates by product and regimen.

between products and regimens. The Kruskal-Wallis test showed a p value of 0.30. Both analyses indicate no statistically significant differences in annual consumption between products and regimens ($p > 0.05$).

Regarding safety, Adynovi® generally has a slightly more favorable safety profile vs Jivi® and Esperoct®, particularly in terms of immune reactions and serious adverse events. Serious adverse reactions, such as anaphylaxis, are less frequent vs anti-PEG antibody-related complications observed with the other products.

When asked to perform a cost-efficiency analysis and recommend a product and regimen, assuming equal unit costs for all 3 options, ChatGPT4o concluded that the best option is Esperoct® every 5 days, followed by Jivi® every 7 days. If the goal is to maximize cost-efficiency under the same unit cost, Esperoct® at 50 IU/kg every 4 days emerges as the most cost-efficient option in terms of balancing low annual consumption (4641 IU/kg) and low ABR (1.18), followed by Jivi® at a weekly regimen of 60 IU/kg, with lower annual consumption (3276 IU/kg) than Esperoct® but a slightly higher ABR (1.49).

Analyzing whether there are statistically significant differences in efficiency, the Kruskal-Wallis test yielded a p value of 0.30 ($p > 0.05$) for annual factor consumption and 0.64 for ABR. This suggests that, in terms of efficiency, the observed differences between products and regimens are not significant enough to claim that one regimen or product is notably superior in efficiency.

## Discussion

Currently, unlike AI, adjusted indirect comparison methods are well-validated statistical tools for comparing products in hemophilia A and have been used for various FVIII-C. For example, in 2019, Batt et al.[6] concluded that, when comparing damoctocog alfa pegol with other FVIII-C such as efmoroctocog alfa, rurioctocog alfa pegol, and

octocog alfa, the ABR was similar, but factor use was 20%-40% lower. However, this adjusted indirect comparison has limitations: undocumented baseline characteristics, absence of a common reference group, and differences in treatment exposure duration, among others. ChatGPT4o could mitigate some limitations of adjusted indirect comparisons, though it could not eliminate them entirely due to the intrinsic nature of the data and the limitations of indirect studies. For example, it could not fully compensate for the lack of a common reference group or population overlap limitations, as these are structural issues. On the other hand, the quality and accuracy of the training data used are always crucial. Similar limitations have been reported in other adjusted indirect comparisons[7,8], so while results are interesting, they must be interpreted cautiously and require additional confirmation for clinical application. ChatGPT4o demonstrates the ability to consolidate large amounts of information from different sources[9] and present it accessibly, but its use still requires necessary validation for these purposes. In our opinion, the comparative analysis performed between the pegylated products is quite accurate and logical, both in structure and reported results. The 3 compared FVIII-C have proven safe and effective fulfilling their intended purpose without issues. However, due to the appropriate balance between consumption, treatment burden, and efficacy, Esperoct® at a dose of 50 IU/kg twice weekly is likely the most suitable treatment option among the compared pegylated products.

## Conclusions

ChatGPT4o and other AI platforms have the potential to improve the efficiency, accuracy, and accessibility of adjusted indirect comparisons in hemophilia A by simplifying the handling of complex data and interpretation of results. Their integration into the process can contribute to more robust comparisons and better communication of findings, benefiting researchers, clinicians, and patients alike. However, ChatGPT4o and other AI platforms require continuous validation processes to enable their implementation.

## Funding

None.

## Conflicts of interest

None.

## Ethical considerations

**Protection of human and animal subjects.** The authors declare that no experiments were performed on humans or animals for this study.

**Confidentiality, informed consent, and ethical approval**. The authors have followed their institution's confidentiality protocols, obtained informed consent from patients, and received approval from the Ethics Committee. The recommendations of the SAGER guidelines have been followed according to the nature of the study.

**Use of artificial intelligence for generating text.** The authors declare that they have not used any type of generative artificial intelligence for the writing of this manuscript, nor for the creation of images, graphics, tables, or their corresponding captions.

## References

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.