

Indirect comparison of two real-life studies in immune thrombocytopenia using artificial intelligence

Comparación indirecta de dos estudios en la vida real en trombocitopenia inmune con uso de inteligencia artificial generativa

Manuel Rodríguez-López^{1*}, Raquel Ocampo-Martínez¹ y Ramiro J. Núñez-Vázquez²

¹Servicio de Hematología y Hemoterapia, Hospital Universitario Álvaro Cunqueiro, Vigo; ²Servicio de Hematología y Hemoterapia, Hospital Universitario Virgen del Rocío, Seville. Spain

Abstract

Artificial intelligence (AI) is here to stay in medicine. This paper explores the use of generative AI tools (ChatGPT4o and Perplexity AI) to compare the efficacy and safety of two new drugs for chronic primary immune thrombocytopenia (ITP), fostamatinib and avatrombopag, using data from real-life studies. The comparison aims to aid in treatment selection for adult patients with ITP who have not responded to previous treatments. Both treatments are effective and safe, but avatrombopag shows a higher response rate and lower incidence of common adverse events. The choice of treatment may depend on additional clinical factors, patient preferences and economic considerations. ChatGPT4o suggests an initial economic advantage for fostamatinib, while Perplexity AI favours avatrombopag in the long term in the Spanish setting. It is concluded that AI can be a valuable tool for treatment comparison, but validation and human oversight are essential to ensure the quality of the information generated. The analysis suggests that both therapeutic options are valid for the management of patients with chronic ITP, highlighting the importance of a collaborative approach between AI and haematologists to maximise outcomes.

Keywords: Immune thrombocytopenia. Artificial intelligence. Real life.

Resumen

La inteligencia artificial (IA) ha llegado a la medicina para quedarse. Este artículo explora el uso de herramientas de IA generativa (ChatGPT4o y Perplexity AI) para comparar la eficacia y la seguridad de dos nuevos fármacos para la trombocitopenia inmunitaria primaria (también conocida como púrpura trombocitopénica idiopática [PTI]) crónica, el fostamatinib y el avatrombopag, utilizando datos de estudios en la vida real. La comparación pretende ayudar en la selección del tratamiento para pacientes adultos con PTI que no han respondido a tratamientos previos. Ambos fármacos son eficaces y seguros, pero el avatrombopag muestra una mayor tasa de respuesta y menor incidencia de eventos adversos comunes. La elección puede depender de factores clínicos adicionales, las preferencias del paciente y consideraciones económicas. ChatGPT4o sugiere una ventaja económica inicial para el fostamatinib, mientras que Perplexity AI favorece al avatrombopag a largo plazo en el contexto español. Se concluye que la IA puede ser una herramienta valiosa para la comparación de tratamientos, pero la validación y la supervisión humana son esenciales para garantizar la calidad de la información generada. El análisis realizado sugiere que ambas opciones terapéuticas son válidas para el manejo de pacientes con PTI crónica, destacando la importancia de un enfoque colaborativo entre la IA y los hematólogos para maximizar los resultados.

Palabras clave: Trombocitopenia inmunitaria. Inteligencia artificial. Vida real.

*Correspondence:

Manuel Rodríguez-López
E-mail: manuel.rodriguez.lopez@sergas.es

Date of reception: 19-07-2024

Date of acceptance: 26-07-2024

DOI: 10.24875/RHT.M24000034

Available online: 11-12-2024

Hemost Trombos (Eng). 2024;1(1):23-31

www.revistahemostasiaytrombosis.com

2938-8740 / © 2024 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/>).

Introduction

The concept of intelligent machines capable of assisting humans is over 70 years old. In 1950, Alan Turing wrote his seminal article “Computing Machinery and Intelligence,” and in 1956, the term “artificial intelligence” (AI) was coined by Marvin Minsky and John McCarthy. In 2023, with the launch of the ChatGPT (Generative Pre-trained Transformer) application by OpenAI, the concept has become popular. AI is a general concept that encompasses several advanced technologies, such as machine learning, natural language processing, and deep learning, methods that facilitate the extraction of patterns and knowledge from vast amounts of data.

Within AI, ChatGPT is a type of linguistic model based on transformers, capable, among other things, of generating texts similar to those produced by humans, although the model still requires improvement in terms of generating scientific texts¹ and never shows its sources of information. There are publications on the use of ChatGPT in the area of thrombosis and hemostasis, for example, to generate recommendations on thromboprophylaxis in spinal surgery² or develop an accurate population pharmacokinetic model for standard half-life factor VIII from literature data³.

Perplexity AI is another generative AI tool that allows information to be obtained from various sources, and unlike ChatGPT, it shows the sources used to obtain the generated information, enabling their evaluation.

Both tools are very interesting, but they are not stranger to limitations and require a critical evaluation of the information provided. It is not recommended to use these models to generate complete content or replace a proper research process. However, they can be useful for gaining initial knowledge, generating ideas, and conducting complementary research. There are key aspects in any process involving AI, such as the quality of the data with which the system will work or the need for validation and supervision of the results, tasks in which human intervention is crucial⁴.

Primary immune thrombocytopenia (also known as idiopathic thrombocytopenic purpura [ITP]) is a disease with an estimated prevalence of 9.5 cases per 100,000 adults and an incidence of 3.3 cases per 100,000 adult-years, which increases with age⁵, with no differences between sexes except between the ages of 30 and 60, when it is more prevalent among women⁶. Some authors believe that, given the rapidly accelerating aging of the world's population (from 461 million people older than 65 years in 2004 to about 2 billion in 2050),

it is likely that ITP will increasingly become a disease of the elderly⁷. It is often diagnosed in older adults, with a chronic course (60% up to 80%), an insidious onset, or different patterns of clinical expression, and has proven resistant to various treatments (80%)^{5,6,8,9}. In patients older than 70, it predominantly affects men and has a higher mortality rates, with a prevalence which is 3-fold greater than in younger adults.

The treatment of ITP has changed considerably in the last 15 years with the arrival, first, of thrombopoietin receptor agonists (the latest of which is avatrombopag)¹⁰, and finally, of a spleen tyrosine kinase inhibitor, fostamatinib¹¹, capable of reducing the antiplatelet activity of phagocytes. Both are considered first-line therapies in the second line of ITP¹², along with romiplostim and eltrombopag. According to the recommendations of the Spanish ITP Working Group⁶ (GEPTI), fostamatinib is especially indicated as the first-line therapy in the second line in patients with high thromboembolic risk, although this is debated even among professionals who developed the guidelines¹³. For other authors^{14,15}, avatrombopag would not only be the second-line agonist of choice in ITP, but they consider it the optimal second-line therapy in adults, potentially associated with greater efficacy due to its more favorable balance of benefits and acceptability.

Real-world results of the experience in Spain with these 2 therapies have been published: the work by González-López et al.¹⁶ on fostamatinib and that of Pascual et al.¹⁷ on avatrombopag. Both show very good safety and efficacy results, making it difficult to choose one drug over the other. This is where AI, even the most basic generative AI, could be a valuable tool in helping the treatment selection process.

Objective

The objective of this work is to make an indirect comparison in terms of safety and efficacy between fostamatinib and avatrombopag to help select treatment in adult patients with ITP who require a new line of therapy due to the failure of previous ones. We proceed to analyze each study individually and then compare them using generative AI (ChatGPT⁴ and Perplexity AI) to try to establish, with the available data, recommendations on which treatment to select.

Results

Fostamatinib, a spleen tyrosine kinase inhibitor, has been approved by the Spanish Agency for Medicines

and Health Products (AEMPS) for the treatment of chronic ITP in adult patients who are refractory to other treatments¹¹. The reported response rates are 80% in the second-line therapy and 40% in multi-refractory patients¹², generally early. Real-world results of fostamatinib use in Spanish patients have been reported¹⁶. This study evaluated its safety and efficacy profile in patients with chronic ITP who had received, at least, 1 previous line of treatment. It included a total of 146 patients from 42 Spanish centers, with a median age of 66 years (interquartile range [IQR], 56-80), contrasting with the 53 years (IQR, 20-88) reported in the phase III clinical trial¹⁸. The median of previous treatments received was 4 (IQR, 2-5), compared to 3 (IQR, 1-13) in the clinical trial¹⁸. The most common previous treatments were eltrombopag (76.1%), romiplostim (57.2%), IV immunoglobulins (44.2%), and rituximab (29.0%), and 13.8% of patients had previously undergone splenectomy (Table 1).

In terms of efficacy, fostamatinib showed an overall response rate of 79.0%, and a complete response in 53.6% of patients. The median time to response was 11 days (IQR, 7-21). The duration of the response was significant: 83.3% of the time in response during the study period. At 3 months, 80% of patients maintained the response, and at 6 months, the rate remained at 65.7%. These results contrast positively with those reported by the phase III clinical trial¹⁸, which reports a stable response of only 18% and an overall response of 44%, with higher responses in patients whose treatment with thrombopoietin agonists had previously failed. These differences in response rates could be due to less stringent inclusion criteria and the different characteristics of the patients included in both studies. Of note that 83 patients (60.1%) received fostamatinib as monotherapy, achieving a response rate of 85.4%.

Regarding safety, 48.5% of patients experienced adverse effects, generally mild and easily manageable. The most common were diarrhea (20%) and hypertension (15%), which is similar to data from the clinical trial¹⁸. Serious adverse events, such as deep vein thrombosis and acute myocardial infarction, were reported with an incidence of 0.91 per 100 patients per year. In the clinical trial¹⁸, serious adverse events were reported in 23% of patients, including a transient ischemic attack that resolved spontaneously. A total of 13.8% of patients required bailout therapy during the observation period.

The authors conclude that fostamatinib is an effective treatment for patients with chronic ITP, achieving a rapid response and long-term maintained efficacy.

The data show that the real-world study shows a higher overall response rate (79% vs 43%) and complete response rate (53.6% vs 17%) vs the phase III trial. The time to response is also shorter in the real-world study. Although generally well tolerated, adverse effects were observed that require monitoring, but still, fostamatinib is a viable option for patients with ITP who have received various previous treatments.

Avatrombopag has been approved by AEMPS¹⁰ for the treatment of primary chronic ITP in adult patients who do not respond to other treatments. The AVESPA study¹⁷ evaluated the safety and efficacy profile of avatrombopag in the treatment of chronic ITP, including patients who had already received multiple lines of treatment. The sample included a total of 268 patients from 28 Spanish centers, treated from January 2022 through November 2023, with a median age of 59 years (IQR, 42-73). The median number of previous treatments was 3 (IQR, 2-4): 73% had received 2 or more previous lines of treatments and 40% 3 or more. The most common previous treatments were eltrombopag (72%), romiplostim (58%), intravenous immunoglobulins (50%), rituximab (30%), and splenectomy (15%).

In terms of efficacy, avatrombopag achieved an overall response rate of 92%, and a complete response in 82% of the registered patients. Additionally, 65% of patients maintained the response for at least 6 months. The mean time to achieve the initial response is not clearly specified in the study, but more than 90% achieved the response within the first 3 weeks, which is consistent with what was reported in the clinical trial¹⁹, in which 65% of patients showed a partial response by day 8 and up to 85% by day 28. In 65% of patients, no adverse events were reported; the most common were headache (10%), fatigue (8%), and nasopharyngitis (6%), generally mild to moderate. Nine deep vein thromboses were reported, 3 of them associated with thrombocytosis, which constitutes an incidence rate of 2.42 cases per 100 patient-years. This reinforces the idea of adjusting the treatment to the minimum effective dose while maintaining a platelet count between 50 and 150 × 10³/mL. At the end of the study, a total of 69% of patients were still on avatrombopag, and only 20% required the maximum weekly dose (280 mg); 67% were on < 140 mg per week.

In the phase III clinical trial¹⁹, the most common adverse events reported were headache and fatigue. Although a total of 4 thromboembolic events were reported, 3 of these patients had multiple risk factors for developing them, yet the episodes were independent of platelet count (39 to 271 × 10³/mL) and the dose

Table 1. Data used by ChatGPT and Perplexity AI for the analyses^{16,17}

	Fostamatinib	Avatrombopag
Number of patients/median age, years (IQR)	138/66 (56-80)	268/59 (42-73)
Female sex	55.8%	57.8%
Primary ITP	88.4%	83.8%
Global response rate*	79%	92%
Complete response rate [†]	53.6%	82%
Time to response, days (IQR)	11 (7-21)	> 90% (max. 21)
Duration of response At 3 months At 6 months	83.3% 80% 65.7%	65%
Previous treatments, median (IQR)	4 (2-5)	2 (1-3)
Previous treatments	Eltrombopag 76.1% Romiplostim 57.2% IVIG 44.2% Rituximab 29.0% Splenectomy 13.8% 95.1%	CCS 95.1% One or more TPO agonists 59.3% IVIG 49.2% Rituximab 15% Fostamatinib 14.6% Splenectomy 13.1%
Response rate in patients with > 3 previous treatment lines	78.4%	84%
Common adverse effects	48.5% (20% diarrhea, 15% HTN)	35%
Incidence of thrombotic or ischemic events per 100 patient-years	0.91	2.42
Treatment discontinuation	24.8%	15%
Minimum daily dose (package insert)	100 mg/12 h	20 mg/day
Response rate in patients with > 3 previous treatment lines	78.4%	84%

*Avatrombopag > 50,000/mL and fostamatinib > 30,000/mL and, at least, double the initial count with concurrent resolution of bleeding symptoms and absence of any rescue intervention during the previous 8 weeks.

[†]Avatrombopag > 100,000/mL and fostamatinib > 100,000/mL.

CCS: corticosteroids; HTN: hypertension; IVIG: intravenous immunoglobulin; ITP: primary immune thrombocytopenia; IQR: interquartile range.

used (10 up to 40 mg/day). This highlights that the thromboembolic risk in patients with ITP, whether or not on or thrombopoietin agonists, is multifactorial, as noted in the study conducted by Lambert et al.²⁰. A total of 18% of patients required bailout therapy during the observation period (Table 1).

In conclusion, Pascual et al.¹⁷ demonstrate that avatrombopag is an effective drug in patients with chronic ITP, achieving a significant response rate and maintaining long-term efficacy in most cases, which is consistent with what was reported in the phase III trial¹⁹. Although treatment was well tolerated—there are adverse events that require attention—it remains a viable therapeutic option for patients with ITP, including those who have received multiple previous treatments.

Discussion

As already mentioned, in any process involving AI, a key aspect is the quality of the data entered to generate equally high-quality information. When comparing the two real-life studies, reference is made to the data in table 1, which summarizes the main characteristics of both. Four key aspects are considered:

- The age of the patients: In the fostamatinib study, the median age is 66 years (IQR, 56-80), compared to 59 years (IQR, 42-73) in the AVESPA study, although patients treated with fostamatinib tend to be slightly older.
- The number and type of previous treatment lines: In the AVESPA study, the median is 3 (IQR, 2-4), while for fostamatinib it is 4 (IQR, 2-5), with previous

treatments being fairly similar and comparable, though the fostamatinib group is slightly more treated.

- The number of patients included: This item presents a greater difference, as AVESPA includes a total of 268 patients and the fostamatinib study includes 146 patients.
- The different definition of overall response: For fostamatinib, it is defined as an increase in platelets $> 30,000/\text{mL}$ for at least 2 weeks without the need for bailout therapy in the previous 8 weeks, and for avatrombopag, as an increase in platelets $> 50,000/\text{mL}$ sustained, thus making the criterion stricter and potentially implying a lower overall response rate vs a lower threshold. It is true that the fostamatinib study includes an additional criterion of not needing rescue treatment in the previous 8 weeks, which is stricter and could decrease the observed response rate. Given that avatrombopag has a higher threshold for defining the response, any direct comparison of response rates should consider these differences to avoid biases. Therefore, ChatGPT-4 adjusts the comparison of efficiency by considering the different response definitions and recalculating response rates, attempting a sensitivity analysis to evaluate how response rates change when applying the criteria of the other study, although this was not possible due to a lack of information. Nonetheless, it states that these differences can significantly affect the assessment of efficiency, because avatrombopag response definition threshold is higher and possibly more demanding. For a more precise comparison, it would be necessary to adjust the response rates considering these definitions or, preferably, conduct additional studies with standardized criteria.

With these results, ChatGPT-4 concludes that the populations in both studies are comparable in terms of age and number of previous treatments, as they have a similar median age and have undergone a similar number of treatment lines before receiving fostamatinib or avatrombopag, suggesting that any differences in the observed safety and efficacy between the 2 drugs are unlikely due to differences in the baseline characteristics of the studied populations.

Regarding sample size, it is suggested that the AVESPA trial¹⁷ is more robust due to the larger number of patients included (268 vs 146), so the results for avatrombopag may be more generalizable. On the other hand, that larger number of patients can also imply greater variability in response to avatrombopag and provide a more complete picture of its safety and efficacy in a more diverse population.

As for thrombotic risk, a common method to compare incidences between two groups is the chi-square test or a proportions test. ChatGPT-4, to determine if the difference between 0.91 and 2.42 cases per 100 patient-years is statistically significant, first opts for an initial approach with the Z-test for 2 proportions (Annex 1, supplementary data), concluding that the observed difference between the 2 studies is not statistically significant. Then it performs the comparison with the chi-square test and reaches the same conclusion: there is no statistically significant difference in the incidence of thrombotic or ischemic events between the two studies.

Probably the most challenging part of the analysis is the assessment of efficiency, partly due to the available data from both studies, which are incomplete or absent regarding the doses used. Even so, to perform this analysis, ChatGPT-4 uses the relevant data referred to in [table 1](#). According to its efficiency analysis, if only two items (safety and efficacy) are considered, both drugs can be considered equally efficient in terms of overall and complete response. However, avatrombopag seems safer, with a lower incidence of common, mild-to-moderate adverse effects vs fostamatinib. Considering safety and efficacy, avatrombopag could be the preferred treatment due to its more favorable safety profile, although both are highly effective.

When calculating efficiency also considering financial cost, the analysis presents the same limitation, that is, absence of data or incomplete data on the doses used, especially in the AVESPA study. The prices used for the analysis correspond to the free sale price from the Ministry of Health ([Table 2](#)). The doses considered for the calculation come from the scant information reported in real-life studies.

For fostamatinib, the recommended initial dose is 100 mg twice a day; initial dose in 107 patients. After 4 weeks, it can be up titrated to 150 mg twice a day depending on platelet count and tolerability; 85 patients increased the daily dose to the maximum of 300 mg, and 31 patients started treatment directly with that maximum dose due to their extreme refractoriness to previous treatments. The median time from the start of fostamatinib to dose titration to 150 mg twice a day was 28 days (IQR, 16-39 days). The median duration of treatment for the entire cohort was 207 days (IQR, 78-449 days). However, in patients who did not respond, the duration was 48 days (IQR, 28-70 days).

For avatrombopag, the initial dose is assumed to be 20 mg/day in 100% of patients, as indicated in its technical data sheet¹⁰. At the end of the follow-up period, from January 2022 through November 2023, 67% of

Table 2. Treatment cost

	Fostamatinib	Avatrombopag
Box of 60 tablets of 100 mg	€3100 €51.66/per tablet €103.33/per day	
Box of 60 tablets of 150 mg	€4650 €77.50/per tablet €155/per day	
Box of 10 tablets of 20 mg		€925 €92.5/per tablet €92.5/per day
Box of 15 tablets of 20 mg		€1390 €92.66/per tablet €92.66/per day

Source: <https://www.farmaceuticos.com/botplus/>.

patients needed avatrombopag at doses ≤ 140 mg/week and 20% at doses of 280 mg/week. Therefore, it is unknown if 100% of patients started treatment with the recommended dose in the technical data sheet, the timing of the first adjustment, and the percentage of patients who required dose adjustment.

To perform a global efficiency evaluation of treatments, considering overall and complete response rates, adverse events, the need for bailout therapy, and financial cost, it uses a weighted approach normalizing the data to a common scale (e.g., 0 up to 1), where a higher value indicates better performance. In this way, an overall efficiency score of 0.71565 is obtained for fostamatinib and 1 for avatrombopag. Since avatrombopag has a higher overall score due to its greater efficacy and lower incidence of adverse events, it is more efficient in terms of cost per day of overall response in the minimum cost scenario. If financial cost is a critical factor, avatrombopag could be the preferred option. If clinical efficacy and reduction of adverse events are more of a priority, again avatrombopag could be the best choice, and this could have an impact in a resource-limited setting.

In an economically precarious situation, avatrombopag is more efficient in terms of cost per day of overall response than fostamatinib, at least in the minimum cost scenario. It would be preferable due to its lower cost per day of effective response (€100.54 vs €130.85 for fostamatinib). The lower treatment costs make it a more accessible option for patients and healthcare systems with limited resources (Annex 2, supplementary data) All in all, it warns that the formula used is a simplified way of combining costs and

effectiveness, and a more detailed analysis could consider more factors, such as costs of adverse events, quality of life, and other long-term effectiveness aspects.

Perplexity AI, when comparing the efficiency of both treatments using the same data (Table 1), concludes that:

- In terms of cost: Avatrombopag remains the most cost-effective treatment in terms of cost per day of overall response, despite fostamatinib having a good response rate.
- In terms of efficacy: Fostamatinib has a lower stable response rate vs avatrombopag, but its use may be more effective in refractory patients.

Perplexity AI understands, like ChatGPT-4, that avatrombopag is more efficient, but it is essential to consider the clinical context and individual patient characteristics when evaluating the efficiency of any treatment. Perplexity AI agrees that avatrombopag can be seen as more efficient due to its higher response rate and safety profile. However, it warns that different analyses and assumptions can lead to different conclusions, especially if long-term factors such as treatment duration and the need for dose titration are considered. Thus, it concludes that:

- The need to increase the dose of fostamatinib to 150 mg twice a day for many patients increases costs (maximum cost per day of overall response: €196.20). The long-term effective dose assumed for avatrombopag suggests that many patients can maintain lower doses, which obviously reduces costs (Annex 2, supplementary data).
- Drug prices may vary across different autonomous communities and hospitals within the same community, and evaluations may differ depending on local costs.
- Costs associated with managing side effects and the need for rescue treatment should be considered, which will affect the overall cost assessment.

Finally, we proceeded to make a critical comparison between the conclusions made by both AI tools, from the point of view of Chat GPT-4 and from the point of view of Perplexity AI (Tables 3 and 4).

Obviously, this comparison exercise has numerous limitations: starting from observational studies, there may be discrepancies in the clinical practice followed. The concurrent use of other drugs and disease severity affect the results. The measurement of outcomes (overall response rates) and the follow-up period are different. The percentage of patients requiring bailout therapy (or which type) is not clearly defined (assumed to be 13.8% for fostamatinib and 18% for avatrombopag), and

Table 3. ChatGPT-4o: critical analysis of differences and recommendations between ChatGPT-4o and Perplexity AI

Aspect	ChatGPT-4o	Perplexity AI
Clinical efficiency	Avatrombopag is clinically more efficient due to higher global and complete response rates and fewer adverse events.	Fostamatinib, while reducing the need for rescue treatment, has a higher incidence of adverse events. Both are effective, but Avatrombopag stands out with a higher response rate. Fostamatinib might be more effective in refractory patients. Emphasizes the importance of safety in decision-making.
Treatment cost	Fostamatinib could be more costly in the short to medium term.	Avatrombopag may be more cost-effective in the long term, especially in the Spanish context, making it a more viable option compared to Fostamatinib.
General recommendations	Patients with a limited budget: recommends avatrombopag due to lower initial cost, with careful monitoring of adverse events. Prioritize clinical efficiency: recommends avatrombopag for its better response rates and lower incidence of mild to moderate adverse events. Long-term considerations: proposes a detailed cost analysis including management of adverse events and treatment duration. Individualization of treatment: emphasizes the need to consider each patient's individual profile, including prior responses and comorbidities.	Individual patient profile: critical to choose treatment based on specific patient characteristics, including severity of thrombocytopenia and comorbidities. Safety and efficacy: assess safety and efficacy profile, considering avatrombopag for higher response rates and fostamatinib for refractory patients. Cost and accessibility: avatrombopag may be more cost-effective in the long term, especially in the Spanish context. Close monitoring: recommends close monitoring of platelet counts and adverse events. Combination with other treatments: considers the possibility of combining fostamatinib or avatrombopag with other drugs to enhance efficacy and reduce dosage.
Final considerations	Although recommendations align with clinical guidelines and lack bias, it cannot replace human clinical judgment. Highlights the importance of traditional clinical studies to validate treatments and ensure personalized care.	Both are valid options, and a multidisciplinary approach involving the patient, hematologist, and primary care team is recommended to optimize therapeutic outcomes.

Table 4. Perplexity AI: critical analysis of differences and recommendations between ChatGPT-4o and Perplexity AI

	ChatGPT-4o	Perplexity AI
Clinical efficiency	Avatrombopag has higher overall and complete response rates and a lower incidence of adverse events compared to fostamatinib. Fostamatinib appears to show a lower need for bailout therapy but a higher incidence of adverse effects.	Both are effective, but emphasizes that avatrombopag might be more cost-effective in the long term. Suggests that fostamatinib might be more effective in refractory patients, introducing an important variable in the discussion of clinical efficacy.
Treatment cost	Fostamatinib is more expensive per patient; moreover, in the long term, it might be even more costly due to the management of adverse events and the need for additional treatments.	Avatrombopag may be more cost-effective in the long term, especially in our setting. Fostamatinib shows discrepancies between the efficacy reported in clinical trials and in clinical practice; this is crucial for evaluating cost-effectiveness.
Treatment recommendations	Avatrombopag is recommended for patients where efficacy and safety are priorities.	Proposes a more individualized approach; suggests that the choice should be based on specific patient characteristics, severity of thrombocytopenia, and response to previous treatments. Suggests considering combination treatments to improve efficacy.
Multidisciplinary approach	Emphasizes the importance of human clinical judgment and the need for traditional clinical studies to validate treatments.	Reinforces the need for a multidisciplinary approach involving the patient and the medical care team to optimize therapeutic outcomes.

in general, the statistical analysis performed is simplified and lacks adequate adjustments. Although the simplified analysis can provide a global view of the efficiency of both drugs, we must consider the mentioned limitations for an adequate interpretation of results. Although this analysis could be indicative for choosing between one drug or the other, considering the patient's individual characteristics, it can never replace the proper clinical judgment on which our decisions must be based.

Conclusions

AI is here to stay, and its role in the development of medicine, at all levels, will be crucial in the future. AI tools can transform medicine by optimizing decision-making, accelerating research, and improving patient follow-up. ChatGPT-4 and Perplexity AI can help make recommendations based on scientific literature and previous study data, thus contributing to the financial sustainability of the National Health System, among other aspects. The use of this basic generative AI (ChatGPT-4o and Perplexity) is already a reality, and in the analysis of these treatments for PTI in real life, it could be a valuable instrument for comparing the safety and efficacy of different therapeutic options. Both tools agree on the importance of individualizing treatment and considering both the safety and efficacy of the drugs. They also align in their financial cost recommendations, suggesting an initial economic advantage for avatrombopag, especially in Spain. Perplexity AI offers a more nuanced analysis focused on individualizing treatment, while ChatGPT-4 focuses more on general recommendations based on costs and efficacy. This contrast suggests that, although both tools can provide valuable information, a comprehensive clinical evaluation adapted to the specific needs of each patient is necessary, requiring a collaborative approach between AI and hematologists to maximize results.

We should not forget that AI still has limitations: ensuring the quality of the provided data is crucial (in this case, insufficient or even absent for some key aspects of the analysis); proper validation of the results through human supervision of the entire process is essential to ensure the quality of the generated information. Additionally, in this specific case, as already mentioned, the formulation of hypotheses by the operator can generate biases in the responses, and even different operators could sometimes generate discordant responses.

Real-life study results, beyond the analysis performed by AI tools, show that both drugs, avatrombopag and fostamatinib, are effective in managing

chronic ITP, showing good response rates and response duration, with an adequate tolerability and safety profile. The choice of one drug or the other may depend on additional clinical factors, patient preferences (when the professional has all options at hand), the hematologist's experience, and financial aspects. Both options are valid for managing patients with chronic ITP refractory to multiple previous treatments, with very good global and complete response rates. The selection of treatment should be individualized and personalized, considering all the described factors.

Supplementary data

Supplementary data are available at DOI: 10.24875/RHT.M24000014. These data are provided by the corresponding author and published online for the benefit of the reader. The contents of supplementary data are the sole responsibility of the authors.

Funding

None declared.

Conflicts of interest

None declared.

Ethical disclosures

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

Confidentiality of data. The authors declare that no patient data appear in this article. Furthermore, they have acknowledged and followed the recommendations as per the SAGER guidelines depending on the type and nature of the study.

Right to privacy and informed consent. The authors declare that no patient data appear in this article.

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. Klang E, Levy-Mendelovich S. Evaluation of OpenAI's large language model as a new tool for writing papers in the field of thrombosis and hemostasis. *J Thromb Haemost*. 2023;21:1055-8.

2. Duey AH, Nietsch KS, Zaidat B, Ren R, Ndjongo LCM, Shrestha N, et al. Thromboembolic prophylaxis in spine surgery: an analysis of ChatGPT recommendations. *Spine J.* 2023;23:1684-91.
3. Cloesmeijer ME, Janssen A, Koopman SF, Cnossen MH, Mathôt RAA; SYMPHONY consortium. ¿ChatGPT in pharmacometrics? Potential opportunities and limitations. *Br J Clin Pharmacol.* 2024;90:360-5.
4. Rashidi H, Bowers K, Reyes Gil M. Machine learning in the coagulation and hemostasis arena: an overview and evaluation of methods, review of literature, and future directions. *J Thromb Haemost.* 2023;21:728-43.
5. Terrell DR, Beebe LA, Vesely SK, Neas BR, Segal JB, George JN. The incidence of immune thrombocytopenic purpura in children and adults: a critical review of published reports. *Am J Hematol.* 2010;85:174-80.
6. Schoonen WM, Kucera G, Coalson J, Li L, Rutstein M, Mowat F, et al. Epidemiology of immune thrombocytopenic purpura in the General Practice Research Database. *Br J Haematol.* 2009;145:235-44.
7. Palandri F, Santoro C, Carpenedo M, Cantoni S, Barcellini W, Carli G, et al. Management of elderly patients with immune thrombocytopenia: real-world evidence from 451 patients older than 60 years. *Thromb Res.* 2020;185:88-95.
8. Abrahamson PE, Hall SA, Feudjo-Tepie M, Mitrani-GoldFS, Logie J. The incidence of idiopathic thrombocytopenic purpura among adults: a population based study and literature review. *Eur J Haematol.* 2009;83:83-9.
9. Yong M, Schoonen WM, Li L, Kanas G, Coalson J, Mowat F, et al. Epidemiology of paediatric immune thrombocytopenia in the General Practice Research Database. *Br J Haematol.* 2010;149:855-64.
10. Avatrombopag. Ficha técnica. AEMPS. (Consultado el 06-07-2024.) Disponible en: https://cima.aemps.es/cima/dochtml/ft/1191373/FT_1191373.html.
11. Fostamatinib. Ficha técnica. AEMPS. (Consultado el 07-07-2024.) Disponible en: https://cima.aemps.es/cima/dochtml/ft/1191405002/FT_1191405002.html.
12. Mingot-Castellano ME, Canaro Hirnyk M, Sánchez-González B, Álvarez-Román MT, Báñez-García A, Bernardo-Gutiérrez A, et al. Recommendations for the clinical approach to immune thrombocytopenia: Spanish ITP Working Group (GEPTI). *J Clin Med.* 2023;12:6422.
13. Sociedad Española de Hematología y Hemoterapia. Recomendaciones PTI. Consenso Delphi. Evaluación de la práctica clínica habitual en España ante la llegada de nuevos fármacos para el tratamiento de la PTI. Disponible en: <file:///C:/Users/ccomp/Downloads/Slide-Kit-Resultados-Consenso-Delphi-PTI.pdf>.
14. Liu Y, Zhang HX, Su J, Geng QC, Lin X, Feng CX. Efficacy and incidence of treatment related adverse events of thrombopoietin receptor agonists in adults with immune thrombocytopenia: a systematic review and network meta-analysis of randomized controlled study. *Acta Haematol.* 2023;146:173-84.
15. Deng J, Hu H, Huang F, Huang C, Huang Q, Wang L, et al. Comparative efficacy and safety of thrombopoietin receptor agonists in adults with thrombocytopenia: a systematic review and network metaanalysis of randomized controlled trial. *Front Pharmacol.* 2021;12:704093.
16. González-López TJ, Bermejo N, Cardesa-Cabrera R, Martínez V, Aguilar-Monserrate G, Pérez Segura G, et al. Fostamatinib effectiveness and safety for immune thrombocytopenia in clinical practice. *Blood* 2024 Jun 6;blood.2024024250. doi: 10.1182/blood.2024024250. Online ahead of print.
17. Pascual Izquierdo C, Sánchez B, Canaro Hirnyk M, García-Donas G, Menor Gómez M, Monsalvo Saornil S, et al. Avatrombopag in ITP. A real-world study of the Spanish ITP Group (GEPTI). *EHA 2024; abstr. P2240*. Disponible en: https://library.ehaweb.org/eha/2024/eha2024congress/420327/cristina.pascual.izquierdo.avatrombopag.in.immune.thrombocytopenia.28pti29.a.html?f=menu%3D6%2Abrowseby%3D8%2Asortby%3D2%2Ace_id%3D2552%2Aot_id%3D29183%2Amarker%-3D5100%2Afeatured%3D18527.
18. Bussell JB, Arnold DM, Boxer MA, Cooper N, Mayer J, Zayed H, et al. Long-term fostamatinib treatment of adults with immune thrombocytopenia during the phase 3 clinical trial program. *Am J Hematol.* 2019;94:5 46-53.
19. Jurczak W, Chojnowski K, Mayer J, Krawczyk K, Jamieson BD, Tian W, et al. Phase 3 randomised study of avatrombopag, a novel thrombopoietin receptor agonist for the treatment of chronic immune thrombocytopenia. *Br J Haematol.* 2018;183:479-90.
20. Lambert C, Maitland H, Ghanima W. Risk-based and individualised management of bleeding and thrombotic events in adults with primary immune thrombocytopenia (ITP). *Eur J Haematol.* 2024;112:504-15.