The Official Journal of the Italian Society for the Study of Hemostasis and Thrombosis
SISET
Società Italiana per lo Studio dell'Emostasi e della Trombosi
Corso Nazionale SISET
Bergamo | 23-25 ottobre 2025

29TH NATIONAL CONFERENCE OF THE ITALIAN SOCIETY FOR THE STUDY OF HEMOSTASIS AND THROMBOSIS, 2025

ABSTRACT N. CO39

MALATTIE EMORRAGICHE CONGENITE E ACQUISITE

ARTIFICIAL INTELLIGENCE IN HEMOPHILIA MANAGEMENT: CHALLENGING THE NEED FOR EXPERTS.

G. Sottilotta[1], D. Megalizzi[1], F. Luise[2].

1UOSD Microcitemie - Emostasi e Trombosi. Grande Ospedale Metropolitano "Bianchi-Melacrino-Morelli" di Reggio Calabria; [2]Laboratorio Analisi. Grande Ospedale Metropolitano "Bianchi-Melacrino-Morelli" di Reggio Calabria.

Background and Aims: Hemophilia is a rare genetic disorder characterized by deficiencies in Factor VIII or Factor IX, which can lead to spontaneous and prolonged bleeding episodes, as well as bleeding caused by minor trauma. This occurs most frequently in patients with severe hemophilia who have not initiated early prophylactic treatment or who do not receive prompt administration of clotting factor concentrates in response to bleeding events. Traditionally, specialized hematologists and dedicated Hemophilia Treatment Centers have been essential for accurate diagnosis and effective management of the disease. However, the recent advancement of artificial intelligence (AI) in healthcare has raised important questions regarding its potential to replace specialist physicians in disease management. This study aimed to evaluate the diagnostic and therapeutic accuracy of five AI models (ChatGPT (OpenAI), Gemini (Google), Copilot (Microsoft), DeepSeek AI and Perplexity AI) in responding to fundamental knowledge-based questionnaires on hemophilia and complex clinical scenarios.

Methods: Two sets of multiple-choice questionnaires were presented to five AI models. The first set of ten questions, generated by ChatGPT, covered essential aspects of hemophilia A and B, including diagnosis, traditional and novel treatments, disease incidence, and possible complications. The second set, independently created by an experienced hematologist, contained another ten questions on the same topics. Each question had four possible answers, with only one being correct. Following this, two complex clinical scenarios were proposed to assess the AI models' ability to provide an accurate medical opinion. The first scenario involved the sudden loss of therapeutic response to prophylactic treatment in a severe hemophilia patient. The second scenario sought guidance on the optimal prophylactic regimen for an adolescent hemophiliac engaged in semi-professional sports activities.

Results: All five AI models demonstrated perfect accuracy, achieving a response rate of 100% (20/20 correct answers) across both questionnaire sets. To validate the consistency of the results, a chi-square statistical analysis was conducted comparing AI performance across the two sets of questions confirming no variance between responses ($p>0.05$). In response to the two complex clinical queries, each AI provided detailed and structured analyses, considering multiple possibilities, exploring differential diagnoses, and ultimately arriving at accurate and guideline-compliant conclusions. The AI-generated responses exhibited a level of depth and reasoning comparable to that of an experienced hematologist.

Conclusions: The findings of this study, resumed in table 1, highlight the impressive capabilities of AI in medical knowledge retrieval and clinical reasoning within hemophilia. Given the rarity of the disease and the historical necessity of specialized experts, this raises a profound question: has AI already reached a level where it could fully replace hematologists in hemophilia management? While AI demonstrates remarkable competency, further research and ethical considerations are required before determining whether it can autonomously guide patient care without human oversight. The integration of AI into clinical practice should be carefully assessed to ensure optimal patient outcomes while preserving the irreplaceable experience and judgment of medical specialists.

Email: *gianluca.sottilotta@virgilio.it*

Summary of AI Performance in Hemophilia Knowledge and Clinical Decision-Making

Aspect	Findings
Number of Questions	20 multiple-choice questions (10 AI-generated, 10 hematologist-created)
AI Models Tested	ChatGPT (OpenAI), Gemini (Google), Copilot (Microsoft), DeepSeek (DeepSeek AI), Perplexity (Perplexity AI)
Accuracy in Questionnaires	100% (100/100 correct responses)
Statistical Validation	No variance between AI responses across question sets ($p > 0.05$)
Complex Clinical Cases	AI models provided detailed, structured, and guideline-compliant responses
Comparability to Human Expertise	Responses matched the depth and reasoning of experienced hematologists
Implication for Hemophilia Care	AI demonstrates potential in assisting or even replacing human specialists if properly utilized