ARAŞTIRMA MAKALESİ / *RESEARCH ARTICLE*

Comparing Predictive Accuracy of Bleeding in Total Abdominal Hysterectomy Among Anesthesiologists, Gynecologists and AI: A Clinical Observational Study

Total Abdominal Histerektomide Kanama Tahmini: Anesteziyologlar, Jinekologlar ve Yapay Zekâ Üzerine Klinik Gözlemsel Çalışma

Ali Selcuk Yeniocak[1], Can Tercan[1], Emrah Dagdeviren[1], Emrullah Akay[1], Osman Murat Guler[1], Alperen Ince[1], Necmiye Ay[2], Havva Betul Bacak[3], Enes Serhat Coskun[4]

[1]Department of Obstetrics and Gynecology; [2]Department of Anesthesiology, Basaksehir Cam and Sakura City Hospital; [3]Department of Obstetrics and Gynecology, Gaziosmanpasa Taksim Education and Research Hospital, Istanbul; [4]Department of Obstetrics and Gynecology, Simav Doc. Dr. Ismail Karakuyu City Hospital, Kutahya, Türkiye

ABSTRACT

Aim: *Accurate prediction of perioperative blood loss is critical for optimizing outcomes in total abdominal hysterectomy (TAH). Traditional estimation methods by clinicians are subjective and prone to variability, while artificial intelligence (AI) offers a potential data-driven alternative. This study compares the predictive accuracy of anesthesiologists, gynecologists, and the AI algorithm ChatGPT4.0 for blood loss during TAH.*

Material and Methods: *This single-center, prospective observational study evaluated 50 patients who underwent TAH for benign conditions in 2023. Clinical data, including uterine size, surgical duration, and surgeon experience, were retrospectively collected. Participating gynecologists and anesthesiologists predicted intraoperative blood loss based on anonymized patient data. Predictions were compared to ChatGPT4.0's outputs and actual recorded blood loss, categorized into mild, moderate, and severe bleeding levels. Sensitivity, positive predictive value, and overall accuracy were analyzed using statistical tests appropriate for data distribution.*

Results: *Anesthesiologists achieved the highest overall accuracy (40%), excelling in moderate bleeding predictions. Gynecologists demonstrated moderate performance across all categories, with 38% accuracy. ChatGPT4.0 showed the lowest overall accuracy (34%) but outperformed clinicians in predicting severe bleeding (37.5% positive predictive value). Variability in clinician predictions highlighted the challenges of subjective estimation, while AI predictions demonstrated consistency but limited precision.*

Conclusions: *AI offers promise in enhancing objective blood loss prediction, particularly for severe cases. However, its performance remains inferior to clinician estimates in most scenarios, underscoring the need for further algorithm refinement and integration into clinical workflows. Future research should focus on long-term validation and addressing ethical challenges in AI adoption.*

Key words: *AI; anesthesiology; artificial intelligence; blood loss prediction; ChatGPT; gynecology; total abdominal hysterectomy*

ÖZET

Amaç: *Perioperatif kan kaybının doğru tahmini, total abdominal histerektomi (TAH) sonuçlarını optimize etmek için kritik öneme sahiptir. Klinik uzmanlar tarafından yapılan geleneksel tahmin yöntemleri öznel olup değişkenliğe yatkınken, yapay zekâ (YZ) veri odaklı bir alternatif sunma potansiyeline sahiptir. Bu çalışma, TAH sırasında anestezistlerin, jinekologların ve ChatGPT4.0 adlı YZ algoritmasının kan kaybı tahmin doğruluğunu karşılaştırmayı amaçlamaktadır.*

Gereç ve Yöntem: *Bu tek merkezli, prospektif gözlemsel çalışmada, 2023 yılında benign durumlar için TAH geçiren 50 hasta değerlendirildi. Uterus boyutu, cerrahi süresi ve cerrahın deneyimi gibi klinik veriler retrospektif olarak toplandı. Katılımcı jinekologlar ve anestezistler, anonimleştirilmiş hasta verilerine dayanarak intraoperatif kan kaybını tahmin etti. Tahminler, ChatGPT4.0'ın sonuçları ve gerçek kaydedilmiş kan kaybı ile karşılaştırılarak hafif, orta ve şiddetli kanama seviyelerine göre kategorize edildi. Verilerin dağılımına uygun istatistiksel testler kullanılarak duyarlılık, pozitif prediktif değer ve genel doğruluk analiz edildi.*

Bulgular: *Anestezistler, orta şiddetli kanama tahminlerinde üstünlük sağlayarak en yüksek genel doğruluğu (%40) elde etti. Jinekologlar tüm kategorilerde orta düzey performans sergileyerek %38 doğruluk sağladı. ChatGPT4.0, genel doğruluk açısından en düşük performansı (%34) gösterdi ancak şiddetli kanama tahminlerinde (%37.5 pozitif prediktif değer) klinisyenlerden daha iyi sonuç verdi. Klinisyen tahminlerindeki değişkenlik, öznel tahminin zorluklarını ortaya koyarken, YZ tahminleri tutarlılık sergilemiş ancak sınırlı bir kesinlik göstermiştir.*

Sonuç: *Yapay zekâ, özellikle şiddetli vakalar için objektif kan kaybı tahmini sağlamada vaat sunmaktadır. Bununla birlikte, mevcut performansı çoğu senaryoda klinisyen tahminlerinin gerisinde kalmakta olup, algoritmanın daha fazla iyileştirilmesi ve klinik iş akışlarına entegrasyonu gerekmektedir. Gelecek araştırmalar, uzun vadeli doğrulama ve YZ'nin benimsenmesindeki etik zorlukların ele alınmasına odaklanmalıdır.*

Anahtar Kelimeler: *yapay zekâ; anesteziyoloji; kan kaybı tahmini; ChatGPT; jinekoloji; total abdominal histerektomi*

İletişim/Contact: *Ali Selcuk Yeniocak, Department of Obstetrics and Gynecology, Basaksehir Cam and Sakura City Hospital, Istanbul, Türkiye* • ***Tel:*** *0532 510 11 80* • ***E-mail:*** *a.s.yeniocak@hotmail.com* • ***Geliş/Received:*** *15.01.2025* • ***Kabul/Accepted:*** *15.03.2025*

ORCID: *Ali Selçuk Yeniocak : 0000-0002-8149-6348 • Can Tercan : 0000-0003-1325-6294 • Emrah Dağdeviren : 0000-0002-1730-3724 • Emrullah Akay : 0000-0003-3792-7777 • Osman Murat Güler : 0000-0001-8005-2887 • Alperen İnce : 0009-0006-0079-1457 • Necmiye Ay : 0000-0003-1787-7522 • Havva Betül Bacak : 0000-0001-8516-5493 • Enes Serhat Coşkun : 0000-0002-6166-5458*

Kafkas J Med Sci 2025; 15(1):96–103
doi: 10.5505/kjms.2025.45452

Introduction

Background

Hysterectomy is among the most frequently performed gynecologic surgeries worldwide, providing definitive treatment for a variety of benign and malignant uterine conditions. In the United States alone, hysterectomy remains the most common non-pregnancy-related major surgery for women, with over 600,000 procedures performed annually[1]. Despite its effectiveness in addressing uterine pathologies, TAH is associated with significant risks, including surgical site infections, vaginal cuff dehiscence, venous thromboembolism, injury to adjacent organs, and excessive bleeding[1,2]. Excessive intraoperative blood loss, defined as greater than 400 mL, is one of the most critical complications, with average estimated losses ranging from 300 to 400 mL and severe bleeding rates reported in up to 16.9% of cases across hospitals[2–4]. Surgery complicated by significant blood loss leads to adverse outcomes, including increased rates of transfusion, readmission, reoperation, prolonged hospital stay, and postoperative morbidity[4]. Patient and procedure-specific factors, such as uterine size, surgical duration, surgical route, and surgeon experience, contribute to variations in estimated blood loss, emphasizing the need for quality improvement initiatives[4–8]. Identifying factors and strategies that improve blood loss prediction and management remains vital for optimizing patient treatment and follow-up outcomes.

Accurate prediction of perioperative blood loss is a critical aspect of surgical planning and patient safety, influencing intraoperative strategies and postoperative outcomes. Traditional methods like visual estimation have inherent limitations, including significant interobserver variability and reliance on subjective judgment. These limitations are evident in studies comparing anesthesiologists and surgeons, where differences in observation angles and clinical focus result in inconsistent and often inaccurate blood loss estimates[9]. Recent advancements in artificial intelligence (AI), specifically machine learning (ML), have introduced innovative methods to predict surgical outcomes, including perioperative blood loss. A notable study applied a random forest algorithm to predict blood loss in orthognathic surgery and demonstrated a strong correlation between predicted and actual blood loss. The model achieved an average deviation of only 7.4 mL from actual values, with a standard deviation of 172.3 ml[10]. This highlights the potential of ML to address the limitations of traditional estimation methods, offering a more objective and data-driven approach to risk stratification and clinical decision-making. While ML-based predictions have been explored in orthognathic and other surgical fields, their application to gynecologic surgery, particularly TAH, remains underexplored.

Objectives

This study aimed to evaluate the accuracy of anesthesiologists' and gynecologists' blood loss estimates for the TAH procedure according to individual risk factors and physicians' personal experiences in our retrospective study cohorts by comparing them with the ML algorithm of ChatGPT4.0 with real-world data. By bridging the gap between traditional clinical observation and AI-driven prediction, this research seeks to enhance patient management in gynecologic surgery and provide a foundation for integrating AI tools into routine clinical practice.

Material and Methods

Study design

This prospective, single-center observational cohort study was conducted to evaluate the accuracy of blood loss predictions made by anesthesiologists, gynecologists and ChatGPT4.0 for patients undergoing TAH for benign indications. The study involved a retrospective cohort of 50 patients who underwent TAH at our hospital in 2023.

Ethics approval and consent to participate

The study was conducted after receiving approval from our tertiary referral hospitals' Clinical Investigations Ethics Committee on December 10, 2024 (Ethics number: KAEK-11/30.10.2024.221), by the principles of the Declaration of Helsinki. Due to the study's retrospective nature, no additional consent was obtained from the participants in the study cohort. However, all patients provided written informed consent upon admission for using their clinical records in scientific research, with a guarantee of anonymity, as approved by the local ethics committee. Participating physicians, including gynecologists and anesthesiologists, voluntarily provided their blood loss predictions after giving electronic consent. Predictions were based on anonymized clinical data without access to the actual intraoperative outcomes.

Setting

The study was carried out at our hospital's Gynecology and Obstetrics Department. Data collection of TAH was performed retrospectively using clinical records

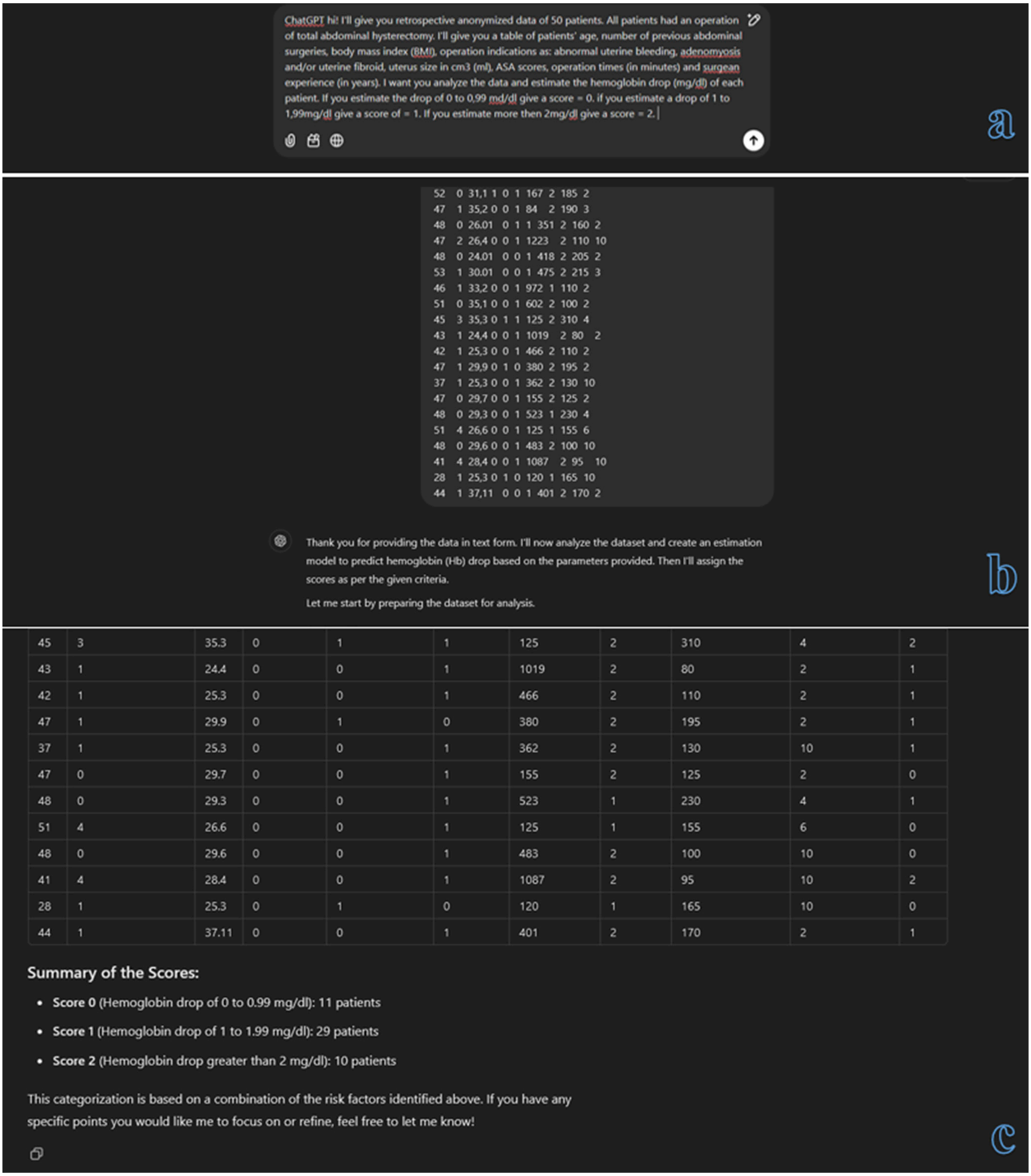

Figure 1. a–c. *Anonymized data of 50 patients was uploaded to ChatGPT4.0, and AI's predictions of blood loss for each case were recorded and analyzed. A detailed comment on ChatGPT4.0 is given for the process demanded (a). An anonymized input of 50 cases was uploaded (b). Output of AI was recorded and compared with the physician's results (c).*

from the first 50 operations performed in 2023. Predictions of intraoperative blood loss were obtained through anonymized electronic surveys distributed to participating physicians, ensuring that all data were de-identified. Also, anonymized data from 50 patients was uploaded to ChatGPT4.0, and AI was asked to predict the blood loss of each patient in the study cohort (Figure 1).

Participants

The study cohort included 50 patients who met the eligibility criteria: those who underwent TAH for benign indications in 2023, with complete clinical data available. These data included patient age, number of previous abdominal surgeries, body mass index (BMI), indications for surgery, preoperative and postoperative hemoglobin levels (mg/dl), uterine size obtained from imaging (cm^3), preoperative American Society of Anesthesiologists (ASA) score, total operative time (min) and the surgeon's years of experience are collected and analyzed. (Table S1, supporting information for full details) Patients undergoing TAH for malignant conditions were excluded from the study.

Variables

The primary outcome of interest was the accuracy of the blood loss predictions compared to the recorded intraoperative blood loss of retrospective real-world data. Predictor variables included patient age, BMI, operative indication, uterine size calculated from preoperative ultrasonographic imaging, ASA score, total operative time, and surgeon experience in years[4–8].

Data sources and measurement

Data were obtained from hospitals' digital archives and anonymized surveys. The accuracy of predictions was assessed by comparing the predicted blood loss from the physicians and ChatGPT4.0 with the recorded intraoperative blood loss documented in the patient's medical records.

Bias

Efforts were made to address potential biases by anonymizing participants and patient data and ensuring predictions were made without knowledge of actual outcomes. Reducing the study cohort to one year before the survey minimized recall bias.

Study size

The sample size was calculated using G*Power software, referencing prior research on predictive models for surgical outcomes[11.12]. With an alpha level of 0.05, an effect size of 0.495, and a statistical power of 95%, the required sample size was determined to be 46 patients. To improve reliability, 50 patients were included in the survey.

Quantitative variables

Uterine volume was calculated using the formula: Ultrasound length (cm)=2.94 + 0.75 × Pathology length (cm), highlighting the reliability of ultrasound as a predictive tool for pathological dimensions[13]. Blood loss was estimated in mL, and corresponding hemoglobin (Hb) decreases were categorized into three groups based on clinical thresholds and prior literature. Excessive intraoperative blood loss, defined as >400 mL, typically correlates with an approximate 1 mg/dL drop in Hb. This threshold was used to delineate the first category, 0–0.99 mg/dL, representing *mild or no significant blood loss*[4]. Further, studies suggest that blood losses exceeding 800–1000 mL are associated with a Hb drop of approximately ≥2 mg/dL. Thus, the second category, 1.00–1.99 mg/dL, was chosen to reflect *moderate blood loss*, and the third category, ≥2 mg/dL, captures cases of *severe blood loss*. This categorization was selected to align with practical clinical decision-making, particularly the thresholds for interventions such as autologous blood reinfusion or ordering bank blood[14]. By basing the Hb categories on these clinical correlations, the questionnaire allows for the stratification of blood loss severity and its impact on postoperative management, providing a standardized framework for analysis.

Statistical methods

The Shapiro-Wilk test was used to assess the data's distribution. A one-way ANOVA test was used for comparisons among three independent normally distributed groups, whereas the Kruskal-Wallis test was applied for non-normally distributed groups. Variables with a normal distribution are reported as mean ± standard deviation, while non-normally distributed data are presented as median (minimum-maximum). Categorical variables were compared using the chi-square test, with results expressed as numbers and percentages.

The performance of bleeding predictions was evaluated using a 3×3 confusion matrix, and positive predictive values, sensitivities (recall) and overall accuracies were calculated to assess prediction accuracy[15]. These metrics were computed for each bleeding severity category-mild, moderate, and severe-based on the median predictions made by gynecologists, anesthesiologists, and ChatGPT4.0.

All statistical analyses were conducted with IBM Statistical Package for Social Sciences (SPSS) program version 26.0, and a P-value of less than 0.05 was considered statistically significant.

Results

Participants

The demographic and clinical characteristics of the patient cohort (n=50) are presented in Table 1. The mean age of the patients was 46.14±5.52 years. Previous abdominal surgery had a median value of 1.0. The mean BMI was 29.36±4.07 kg/m^2. The mean uterine size was 625.70±620.41 mL, and the mean operation time was 141.80±52.36 minutes. Surgeon experience was reported as a median of 4.0 years. The surgical indications were distributed as follows: 10 (20.0%) patients underwent surgery due to abnormal uterine bleeding (AUK), 7 (14.0%) for adenomyosis, and 33 (66.0%) for fibroids. American Society of Anesthesiologists (ASA) scores were distributed as 24.0% (n=12) with a score of 1, 68.0% (n=34) with a score of 2, and 8.0% (n=4) with a score of 3. The mean preoperative hemoglobin level was 10.39±0.94 mg/dL.

Descriptive data

This study compared the bleeding prediction accuracy of gynecologists, anesthesiologists, and ChatGPT4.0 during total abdominal hysterectomy. Bleeding predictions were categorized into three levels: mild bleeding (0–0.99 mg/dL), moderate bleeding (1–1.99 mg/dL), and severe bleeding (≥2 mg/dL) (Table 2).

Table 1. *Analysis of retrospective study cohort demographic data of 50 elective total abdominal hysterectomy patients in 2023 according to their postoperative hemoglobin drop levels*

Demographic data		Mild or no significant blood loss (0–0.99 mg/dL) (n=18)	Moderate blood loss (1.00–1.99 mg/dL) (n=19)	Severe blood loss (≥2 mg/dL) (n=13)	P value
Age (years)		47.27±5.08	45.05±6.18	46.15±5.14	0.48
Pr. Abd. surgery (n)		0.00 (0.00–1.00)	1.00 (0.00–1.00)	1.00 (0.00–1.00)	0.11
BMI (kg/m²)		27.76±3.54	29.62±3.69	31.17±4.70	0.06
Uterine size (ml)		685.61±629.53	520.73±464.34	696.15±811.57	0.65
Op. time (min)		135.00±37.37	146.57±67.00	144.23±48.68	0.79
Surgeon exp. (years)		3.50 (2.00–10.00)	7.00 (1.00–10.00)	3.00 (1.00–9.00)	0.26
Indication	Ab. uterine bl.	2 (20%)	4 (40%)	4 (30.8%)	0.35
	Adenomyosis	3 (42.9%)	4 (57.1%)	0 (0.00%)	
	Ut. fibroid	13 (39.4%)	11 (33.3%)	9 (69.2%)	
ASA score	1	4 (33.3%)	4 (33.3%)	4 (33.3%)	0.27
	2	13 (38.2%)	13 (38.2%)	8 (23.5%)	0.64
	3	1 (25.0%)	2 (50.0%)	1 (25.0%)	0.42
Preop. Hb (mg/dl)		10.03±0.85	10.48±0.89	10.76±1.03	0.09

The Shapiro-Wilk test was used to assess the data's distribution. A one-way ANOVA test was used for comparisons among three independent normally distributed groups, whereas the Kruskal-Wallis test was applied for non-normally distributed groups. Variables with a normal distribution are reported as mean ± standard deviation, while non-normally distributed data are presented as median (minimum-maximum). Categorical variables were compared using the chi-square test, with results expressed as numbers and percentages.
Ab. uterine bl.: Abnormal uterine bleeding, ASA: American Society of Anesthesiologists, BMI: Body mass index, kg/m²: kilograms per square meter, mg/dl: milligrams per deciliter, min.: Minutes, ml: milliliters, n: Number, Op. time: Operation time, Pr. Abd. Surgery: Previous abdominal surgery, Preop. Hb: Preoperative hemoglobin, Surgeon exp: Surgeon experience, Ut. Fibroid: Uterine fibroid.

Table 2. *Comparison of anesthesiologist's, gynecologist's and ChatGPT4.0's bleeding predictions with 3×3 confusion matrix crosstabulation*

		Mild or no significant blood loss (0–0.99 mg/dL) (n=18)	Moderate blood loss (1.00–1.99 mg/dL) (n=19)	Severe blood loss (≥2 mg/dL) (n=13)	Positive predictive value	Overall accuracy
Gynecologist's predictions (median) (n=17)	Mild or no significant blood loss	5 (27.8%)	3 (15.8%)	4 (30.8%)	41.66%	38%
	Moderate blood loss	11 (61.1%)	11 (57.9%)	6 (46.2%)	39.28%	
	Severe blood loss	2 (11.1%)	5 (26.3%)	3 (23.1%)	30.00%	
Sensitivity (recall)		27.77%	57.89%	23.07%		
Anesthesiologist's predictions (median) (n=9)	Mild or no significant blood loss	2 (11.1%)	1 (5.3%)	1 (7.7%)	50.00%	40%
	Moderate blood loss	12 (66.7%)	16 (84.2%)	10 (76.0%)	42.10%	
	Severe blood loss	4 (22.2%)	2 (10.5%)	2 (15.4%)	25.00%	
Sensitivity (recall)		11.11%	84.21%	15.38%		
ChatGPT4.0's predictions	Mild or no significant blood loss	3 (16.7%)	5 (26.3%)	3 (23.1%)	27.27%	34%
	Moderate blood loss	13 (72.3%)	11 (57.9%)	7 (53.8%)	35.48%	
	Severe blood loss	2 (11.1%)	3 (15.8%)	3 (32.1%)	37.5%	
Sensitivity (recall)		16.66%	57.89%	23.07%		

The performance of bleeding predictions was evaluated using a 3×3 confusion matrix, positive predictive values.

Gynecologists (n=17) demonstrated 27.77%, 57.89%, and 23.07% sensitivity and 41.66%, 39.28%, and 30.00% positive predictive value for mild, moderate and severe bleeding, respectively, with a 38% overall accuracy. Anesthesiologists (n=9) performed 11.11%, 84.21%, and 15.38% sensitivity and 50.00%, 42.10%, and 25.00% positive predictive value for mild, moderate and severe bleeding, respectively, with an overall accuracy of 40%. ChatGPT4.0 demonstrated a lower overall accuracy of 34% compared to the other groups but performed better in predicting severe bleeding with a sensitivity of 23.07% and 37.5% positive predictive value. ChatGPT4.0 achieved a sensitivity of 16.66% and 23.07% for mild and moderate bleeding and a positive predictive value of 27.27% and 35.48%, respectively.

Discussion

Integrating artificial intelligence (AI) into clinical practice has shown significant promise but poses notable challenges that require careful consideration. Our study results reveal significant variability in the prediction accuracy, precision, and recall across the three groups and bleeding severity levels. While anesthesiologists demonstrated high precision in predicting moderate bleeding but had lower consistency in predicting mild and severe bleeding, gynecologists showed moderate performance across all bleeding categories. Still, they struggled with the precision and recall of severe bleeding predictions. ChatGPT4.0, while generally less precise and accurate than the other groups, achieved the highest accuracy in predicting severe bleeding.

One critical issue in clinical practice is the reliance on surgeons' subjective visual estimation, which can lead to significant underestimation or overestimation of blood loss, affecting clinical decision-making[16]. Similarly, a study cohort of laparoscopic surgeries has been found to carry risks of underestimating bleeding, emphasizing the need for more reliable quantitative methods[17]. Interestingly, AI's role as an assistive tool for clinicians presents a nuanced picture. Studies have shown that while AI systems such as GPT-4 outperform clinicians in isolated diagnostic tasks, their integration as diagnostic aids alongside clinicians has not consistently improved performance[18]. This highlights the need for further exploration into how clinicians can be effectively trained to work with AI systems and how these tools can be seamlessly integrated into workflows. For example, in laparoscopic colectomy, AI systems have shown proficiency in real-time detection of bleeding events, achieving remarkable precision[19]. Additionally, machine learning in predicting postpartum hemorrhage has proven effective, with random forest models yielding high accuracy in quantitative assessments[20].

Despite the advancements in medical technology, ethical and legal considerations remain a major barrier to the widespread adoption of AI in clinical practice. For instance, issues surrounding data privacy, potential biases in AI algorithms, and robust regulatory frameworks must be addressed to ensure equitable and responsible use[21]. Furthermore, the psychological and financial burdens stemming from AI's unexplained algorithm could hinder its acceptance among patients and clinicians alike, often referred to as the "black box" problem, raises ethical concerns regarding transparency, patient autonomy, and psychological trust[22].

Limitations

This study has several limitations that could influence its outcomes and interpretations. Firstly, relying on retrospective datasets for analysis introduces inherent biases, including selection bias and inaccuracies in historical records. While AI-assisted tools demonstrated promise in diagnostic accuracy, their'efficacy was evaluated in a controlled environment, which may not reflect real-world challenges. Additionally, the ethical implications of the "black box" nature of AI systems highlight a critical limitation: the difficulty in understanding and validating the reasoning behind AI-driven decisions, which could undermine clinician trust and patient autonomy. Lastly, like the other studies in the literature, this study lacked long-term validation of AI implementation outcomes. While promising short-term results were reported, the absence of long-term outcomes remains an underexplored paradigm of the nature of the ML algorithms.

Interpretation

The findings of this study underscore the transformative potential of AI in healthcare, particularly in enhancing diagnostic accuracy and supporting objective clinical decision-making. However, given the study's limitations, these results should be interpreted cautiously. While ChatGPT demonstrated significant accuracy in hemorrhage prediction, real-time integration into routine practice requires careful consideration of context, data quality, and clinician training.

Generalizability

Despite the study's limitations, our findings provide valuable insights into ChatGPT's potential in healthcare. To enhance generalizability, future research should focus on long-term validating AI systems in varied clinical environments, incorporating diverse patient populations, and addressing disparities in data representation in different languages. By doing so, AI technologies' broader applicability and equity in healthcare can be more effectively realized.

Conclusion

In conclusion, while AI holds the potential to revolutionize healthcare by improving diagnostic accuracy, reducing errors, and optimizing workflows, its integration into clinical practice must be approached with caution. A balance must be struck between leveraging AI's capabilities and addressing its ethical, practical, and educational challenges. Future research should focus on establishing comprehensive ethical guidelines to ensure that AI is a reliable and transparent partner in patient care.

Funding

None.

Acknowledgments

None.

Conflict of interest

The authors declare no conflict of interest.

Availability of data and materials

This published article and its supplementary information files include all data generated or analyzed during this study.

References

1. Clarke-Pearson DL, Geller EJ. Complications of hysterectomy. Obstetrics & Gynecology. 2013;121(3):654–673.
2. Meeks GR, Harris RL. Surgical approach to hysterectomy: abdominal, laparoscopy-assisted, or vaginal. Clinical obstetrics and Gynecology. 1997;40(4):886–894.
3. Maresh M, Metcalfe M, McPherson K, Overton C, Hall V, Hargreaves J, et al. The VALUE national hysterectomy study: description of the patients and their surgery. BJOG. An International Journal of Obstetrics & Gynaecology. 2002;109(3):302–312.
4. English EM, Bell S, Kamdar NS, Swenson CW, Wiese H, Morgan DM. Importance of estimated blood loss in resource utilization and complications of hysterectomy for benign indications. Obstetrics & Gynecology. 2019;133(4):650–657.
5. Bonilla DJ, Mains L, Whitaker R, Crawford B, Finan M, Magnus M. Uterine weight as a predictor of morbidity after a benign abdominal and total laparoscopic hysterectomy. The Journal of reproductive medicine. 2007;52(6):490–498.
6. Beck TL, Morse CB, Gray HJ, Goff BA, Urban RR, Liao JB. Route of hysterectomy and surgical outcomes from a statewide gynecologic oncology population: is there a role for vaginal hysterectomy? American Journal of Obstetrics and Gynecology. 2016;214(3):348. e1–348. e9.
7. Catanzarite T, Saha S, Pilecki MA, Kim JY, Milad MP. Longer operative time during benign laparoscopic and robotic hysterectomy is associated with increased 30-day perioperative complications. Journal of minimally invasive gynecology. 2015;22(6):1049–1058.
8. Vree FE, Cohen SL, Chavan N, Einarsson JI. The impact of surgeon volume on perioperative outcomes in hysterectomy. JSLS. Journal of the Society of Laparoendoscopic Surgeons. 2014;18(2):174.
9. Yüzkat N, Soyalp C, Gülhas N. Comparison of the Error Rates of an Anesthesiologist and Surgeon in Estimating Perioperative Blood Loss in Major Orthopedic Surgeries: Clinical Observational Study. Journal of Anesthesia/Anestezi Dergisi (JARSS). 2019;27(4)
10. Stehrer R, Hingsammer L, Staudigl C, Hunger S, Malek M, Jacob M, et al. Machine learning based prediction of perioperative blood loss in orthognathic surgery. Journal of Cranio-Maxillofacial Surgery. 2019/11/01/. 2019;47(11):1676–1681.
11. Kane S, Collins S, Sproat LA, Mangel J. Predictors of transfusion requirement among patients who undergo hysterectomy for benign disease. Journal of Gynecologic Surgery. 2012;28(2):113–115.
12. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G* Power 3. 1: Tests for correlation and regression analyses. Behavior research methods. 2009;41(4):1149–1160.
13. Cantuaria GH, Angioli R, Frost L, Duncan R, Penalver MA. Comparison of bimanual examination with ultrasound examination before hysterectomy for uterine leiomyoma. Obstetrics & Gynecology. 1998;92(1):109–112.

14. Gunasekaran K, Punnagai K, Vijaybabu K. A comparative study of diluents (crystalloid vs colloid) in acute normovolaemic haemodilution (ANH). International Journal. 2014;5(2):64.
15. Landis J. The Measurement of Observer Agreement for Categorical Data. Biometrics. 1977;
16. Gluck O, Mizrachi Y, Kovo M, Divon M, Bar J, Weiner E. Major underestimation and overestimation of visual blood loss during cesarean deliveries: can they be predicted? Archives of Gynecology and Obstetrics. 2017;296:907–913.
17. Sato M, Koizumi M, Inaba K, Takahashi Y, Nagashima N, Ki H, et al. Gynecologists may underestimate the amount of blood loss during total laparoscopic hysterectomy. Obstetrics and Gynecology International. 2018;2018(1):3802532.
18. Goh E, Gallo R, Hom J, Strong E, Weng Y, Kerman H, et al. Large language model influence on diagnostic reasoning: a randomized clinical trial. JAMA Network Open. 2024;7(10):e2440969–e2440969.
19. Horita K, Hida K, Itatani Y, Fujita H, Hidaka Y, Yamamoto G, et al. Real-time detection of active bleeding in laparoscopic colectomy using artificial intelligence. Surgical Endoscopy. 2024;38(6):3461–3469.
20. Wang M, Yi G, Zhang Y, Li M, Zhang J. Quantitative prediction of postpartum hemorrhage in cesarean section on machine learning. BMC Medical Informatics and Decision Making. 2024;24(1):166.
21. Alowais SA, Alghamdi SS, Alsuhebany N, Alqahtani T, Alshaya AI, Almohareb SN, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC medical education. 2023;23(1):689.
22. Xu H, Shuttleworth KMJ. Medical artificial intelligence and the black box problem: a view based on the ethical principle of "do no harm". Intelligent Medicine. 2024;4(1):52–57.