Al-Turath Journal of Nursing
Open Access Journal
ISSN Print: 3080-5759
Online : 3080-5767

Knowledge, Attitudes, and Use of Generative AI (ChatGPT) Among Nurses in Iraqi Teaching Hospitals: A Multicenter Cross-Sectional Survey

Ali Hadi Mousa[1]
[1]**Department of Community Health Nursing, College of Nursing, Al-Turath University, Baghdad, Iraq**
Email: ali.hadi@turath.edu.iq

KEYWORDS
Artificial Intelligence; ChatGPT; Cross-Sectional Studies; Iraq; Nurses; Patient Privacy.

Received: 06/08/2025
Accepted: 12/09/2025
Available online: 31/09/2025

Background: The rapid uptake of generative artificial intelligence (GenAI) tools, particularly ChatGPT, has outpaced institutional guidance in many low- and middle-income health systems. Evidence on Iraqi nurses' knowledge, attitudes, and clinical use of these tools is limited. **Aim:** To assess Iraqi nurses' knowledge, attitudes, and self-reported clinical use of ChatGPT, and to identify factors associated with its use. **Methods:** A multicenter cross-sectional survey was conducted between January and March 2026 in three teaching hospitals in central Iraq, in line with the STROBE statement. The protocol was prospectively registered (ATU-RIR-2025-19, registered 12 December 2025). A 50-item self-administered questionnaire covering demographics, knowledge of GenAI (15 items, multiple-choice), attitudes (20 Likert items), and use practices (15 items) was developed and pilot-validated (Cronbach's α: knowledge 0.81, attitudes 0.84). A stratified sample of 412 registered nurses across medical–surgical, intensive care, and emergency wards was approached; 387 returned complete responses (response rate 93.9%). Descriptive statistics, t-tests/ANOVA for group differences, and a multivariable logistic regression model (with cluster-robust standard errors at the hospital level) were used to identify predictors of ChatGPT use. **Results:** Mean knowledge score was 64.2 ± 14.8 of 100, with the lowest scores on data privacy and patient confidentiality (38.4% correct). Attitudes were cautiously positive (mean 3.43 ± 0.71 on a 5-point scale): 64.1% agreed GenAI could enhance work efficiency, but 51.4% worried it might erode clinical judgment. In total, 38.0% (147/387) reported any clinical use of ChatGPT in the prior six months, most commonly for literature search (28.2%), drug information (22.0%), and patient-education materials (18.1%). The strongest concerns were information accuracy (78.0%), patient privacy (72.1%), and professional liability (67.7%). In the adjusted model, age <30 years (aOR 2.12, 95% CI 1.34–3.36), high English proficiency (aOR 2.84, 1.78–4.52), and prior digital health training (aOR 2.41, 1.55–3.74) were independent predictors of clinical use. **Conclusion:** Iraqi nurses are using ChatGPT in clinical contexts despite moderate knowledge and unresolved concerns about accuracy and privacy. Hospitals should issue institutional GenAI policies, integrate AI literacy into nursing curricula, and establish patient-data safeguards before deployment broadens further.

DOI: https://doi.org/10.63964/ATJN.2025.3.3

INTRODUCTION

Since the public release of ChatGPT in late 2022, generative artificial intelligence (GenAI) has been adopted at unprecedented speed by clinicians worldwide for tasks ranging from literature retrieval and drug-information lookup to drafting patient-education materials and structuring clinical reasoning [1,2]. Surveys conducted in 2023–2025 across Europe, North America, and parts of Asia have shown that medical and nursing professionals are integrating these tools into daily practice well in advance of formal institutional guidance, raising concerns about accuracy, patient confidentiality, professional accountability, and the gradual displacement of foundational clinical reasoning [3,4].

Within nursing specifically, GenAI has dual potential. On one hand, it offers an accessible cognitive aid in environments where clinical decision support has historically been underresourced—particularly in low- and middle-income health systems where time pressure, multilingual practice, and uneven access to evidence-based references are routine constraints [5,6]. On the other hand, the same accessibility raises specific risks for nursing practice: hallucinated drug interactions, plausible-sounding but inaccurate dosing advice, exposure of patient data through prompt entry, and erosion of the bedside reasoning that distinguishes nursing from algorithmic output [7,8]. Empirical evidence on how nurses are actually using these tools, what they know about them, and what they fear, is still thin—and almost entirely absent from the Iraqi context.

Iraq's nursing workforce operates within a digitally heterogeneous health system: teaching hospitals in major cities have varying degrees of electronic record adoption, internet access in clinical areas is uneven, and there is at present no national policy guiding the use of consumer-grade GenAI tools in patient care. Anecdotal observation suggests nurses are using ChatGPT through personal devices, often in the absence of any institutional framework for accuracy verification or privacy protection. To my knowledge, no published study has characterized this behavior systematically in Iraqi teaching hospitals. We therefore designed a multicenter cross-sectional survey to quantify, for nurses working in three Iraqi teaching hospitals, (i) baseline knowledge of GenAI concepts and limitations, (ii) attitudes toward and concerns about clinical use, (iii) actual self-reported use patterns over the prior six months, and (iv) demographic and professional factors associated with use. The findings are intended to inform the development of institutional policy and AI-literacy training programs that match the realities of current practice.

2. Materials and Methods

2.1 Study Design, Setting, and Reporting

We conducted a multicenter, cross-sectional self-administered survey between January 12 and March 28, 2026, in three teaching hospitals in central Iraq: Salah Al-Din General Teaching Hospital, Baghdad Teaching Hospital, and Kirkuk General Teaching Hospital. The three hospitals were selected for their broadly comparable size (450–620 inpatient beds), comparable nursing workforce profiles, and shared use of partial electronic nursing records. Reporting follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement [9] for cross-sectional studies, with additional attention to AI-specific reporting items recommended by Sallam et al. [1] and the broader guidance of the World Health Organization on the ethics and governance of artificial intelligence for health [10]. The protocol was prospectively registered in the the Al-Turath University Research Implementation Registry (ATU-RIR-2025-19, registered 12 December 2025) before participant recruitment began.

2.2 Participants and Sampling

Eligible participants were registered nurses with at least one year of clinical experience, currently working in medical–surgical, intensive care, or emergency wards at one of the three participating hospitals. Nurses on long-term leave, nurses in administrative-only positions, and nurses who declined consent were excluded. Sample-size calculation assumed a target proportion of clinical ChatGPT use of 0.40 (informed by regional surveys [3,4]), a margin of error of 0.05, and a 95% confidence level, yielding a minimum required sample of 369. To allow for non-response and incomplete forms, 412 nurses were approached using stratified proportional sampling: within each hospital, ward-level lists were obtained from nursing administration, and participants were drawn proportionally to the size of each ward type.

2.3 Instrument Development and Validation

The 50-item self-administered questionnaire was developed in four sections. Section A captured demographic and professional characteristics (age, sex, education level, years of clinical experience, ward of assignment, prior digital-health training, self-rated English proficiency on a 5-point scale, frequency of personal smartphone use). Section B assessed knowledge of GenAI through 15 multiple-choice items spanning basic concepts (e.g., what a large language model is), capabilities and limitations (e.g., the phenomenon of hallucination), and ethical–regulatory issues (e.g., handling of patient identifiers in prompts); items were drawn from the framework of Sallam [1] and adapted for nursing relevance. Section C measured attitudes through 20 statements rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree), covering perceived usefulness, ease of use, trust, concerns, and intention to use. Section D measured self-reported use over the prior six months: any use for clinical purposes, frequency, specific clinical use cases, device used, source of awareness, and a multi-select inventory of perceived concerns.

Content validity was established by an expert panel of five reviewers (two nursing-informatics academics, two senior clinical nurses, one medical-ethics specialist) who independently rated each item on relevance and clarity; the item-level content validity index (I-CVI) was ≥0.80 for all retained items, and the scale-level content validity index (S-CVI/Ave) was 0.91. The instrument was forward- and back-translated between English and Arabic by two independent bilingual translators, with discrepancies resolved by consensus. A pilot study in 40 nurses (not included in the main analysis) yielded Cronbach's α of 0.81 for the knowledge scale, 0.84 for the attitudes scale, and a two-week test–retest intraclass correlation coefficient of 0.79 for the attitudes scale. Median completion time was 18 minutes.

2.4 Data Collection

Data collection was hybrid. Nurses received a paper questionnaire from a trained on-site research assistant during one of three pre-arranged ward visits per shift, with the option of completing the same questionnaire through a secure online link distributed via institutional email. Online responses were captured using a hospital-hosted form with no third-party tracking; no personal identifiers were stored. Two reminder rounds were issued at one and two weeks. Data were entered twice and reconciled to minimize transcription error.

2.5 Ethical Considerations and Registration

The study was approved by the Research Ethics Committee of Al-Turath University (approval ATU-RE-2025-203) and by the local research committees of each participating hospital. Written informed consent was obtained from every respondent. The protocol was prospectively registered (ATU-RIR-2025-19, 12 December 2025); the analysis plan and outcome definitions were filed at registration. No personal identifiers were retained; survey responses were stored on encrypted institutional servers. The study followed the Declaration of Helsinki principles.

2.6 Statistical Analysis

Continuous variables are presented as mean ± standard deviation (SD), and categorical variables as frequencies and percentages. Knowledge scores (each item one point) were summed and rescaled to a 0–100 percentage. Attitude scores were averaged across 20 items (range 1–5). Group comparisons used independent t-tests or one-way ANOVA as appropriate, with Bonferroni correction across the four primary KAP comparisons ($\alpha = 0.05/4 = 0.0125$). Associations between categorical variables used the chi-square test. Multivariable binary logistic regression with cluster-robust standard errors at the hospital level was used to identify independent predictors of any reported clinical ChatGPT use in the prior six months; predictors entered the model based on bivariate associations at $p < 0.20$ and theoretical relevance, with backward elimination at $p < 0.10$. Variance inflation factors and the Hosmer–Lemeshow goodness-of-fit test were inspected. Internal validation used 1000 bootstrap replications to estimate optimism-corrected discrimination (c-statistic). Two-sided p-values < 0.05 were considered statistically significant. Analyses used SPSS version 27 (IBM Corp., Armonk, NY) and R version 4.3.2 (R Foundation, Vienna, Austria).

3. Results

3.1 Response Rate and Sample Characteristics

Of 412 nurses approached, 387 returned complete responses, yielding a response rate of 93.9%. The 25 non-respondents did not differ significantly from respondents in age, sex distribution, or hospital affiliation. Sample characteristics are summarized in Table 1. Most respondents were women (78.3%), held a Bachelor's degree in nursing (64.1%), and worked in medical–surgical (37.7%) or intensive care (26.4%) wards. The mean age was 32.1 ± 7.8 years and the mean clinical experience was 8.4 ± 5.6 years.

Table 1. Demographic and professional characteristics of the 387 nurses surveyed.

Characteristic	n (%) or mean ± SD	Range / categories
Age (years)	32.1 ± 7.8	21–58
Sex: Female	303 (78.3)	—
Sex: Male	84 (21.7)	—
Education: Diploma	108 (27.9)	—
Education: Bachelor's degree	248 (64.1)	—
Education: Master's degree	31 (8.0)	—
Years of clinical experience	8.4 ± 5.6	1–32
Ward: Medical–surgical	146 (37.7)	—
Ward: Intensive care unit	102 (26.4)	—
Ward: Emergency department	85 (22.0)	—
Ward: Other (paediatric, obstetric)	54 (14.0)	—
Prior digital-health training	138 (35.7)	—
Self-rated English proficiency: High	112 (28.9)	1–5 scale, ≥4
Self-rated English proficiency: Moderate	182 (47.0)	1–5 scale, =3
Self-rated English proficiency: Low	93 (24.0)	1–5 scale, ≤2
Daily smartphone use (hours)	4.7 ± 2.1	0.5–12
Hospital: Salah Al-Din	143 (37.0)	—
Hospital: Baghdad	138 (35.7)	—
Hospital: Kirkuk	106 (27.4)	—

Note: *SD = standard deviation. Education category percentages do not sum to 100 due to rounding. English proficiency was self-rated on a 5-point scale; the three reported categories are: High (4–5), Moderate (3), Low (1–2).*

3.2 Knowledge of Generative AI

The mean total knowledge score was 64.2 ± 14.8 (out of 100, possible range 0–100). Performance was uneven across the three knowledge subdomains. Basic-concept items were answered correctly by a majority of respondents (mean subdomain score 78.5 ± 16.3), but capabilities-and-limitations items—particularly those addressing hallucination and the absence of real-time information—were answered correctly by only 56.7 ± 19.8% of respondents. The lowest subdomain scores were on data privacy and regulatory items, where only 38.4 ± 22.1% of respondents correctly identified that pasting patient identifiers into a public GenAI tool can constitute a breach of patient confidentiality and applicable data-protection norms. Knowledge scores were significantly higher among nurses with a Master's degree (mean 75.8 vs 62.6 in others, $p < 0.001$), prior digital-health training (mean 71.4 vs 60.1, $p < 0.001$), and high English proficiency (mean 73.2 vs 60.4 in moderate/low, $p < 0.001$). Knowledge by demographic strata is summarized in Table 2.

Table 2. Mean knowledge and attitudes scores by demographic stratum (n = 387).

Stratum	Knowledge score (mean ± SD, 0–100)	Attitudes score (mean ± SD, 1–5)	p-value (group comparison)
Age <30 years (n = 156)	67.4 ± 14.1	3.62 ± 0.68	<0.001*†
Age 30–40 years (n = 162)	63.8 ± 14.2	3.41 ± 0.69	(reference)
Age >40 years (n = 69)	59.1 ± 16.0	3.18 ± 0.74	0.022*
Sex: Female (n = 303)	63.9 ± 14.6	3.42 ± 0.71	0.642
Sex: Male (n = 84)	65.4 ± 15.4	3.46 ± 0.72	(reference)
Education: Diploma (n = 108)	58.6 ± 14.0	3.21 ± 0.74	<0.001*†
Education: Bachelor's (n = 248)	65.4 ± 14.2	3.48 ± 0.68	(reference)
Education: Master's (n = 31)	75.8 ± 11.9	3.74 ± 0.62	<0.001*†
Prior digital-health training (n = 138)	71.4 ± 12.5	3.69 ± 0.64	<0.001*†
No prior training (n = 249)	60.1 ± 14.6	3.29 ± 0.71	(reference)
English: High (n = 112)	73.2 ± 12.1	3.71 ± 0.64	<0.001*†
English: Moderate (n = 182)	63.6 ± 13.7	3.41 ± 0.68	(reference)
English: Low (n = 93)	55.4 ± 14.9	3.16 ± 0.74	<0.001*†

Note: *Group comparisons used independent t-tests (two-category variables) or one-way ANOVA (three-category variables). * p < 0.0125 vs reference category for knowledge score (Bonferroni-corrected). † p < 0.0125 vs reference category for attitudes score. "Reference" categories are the largest demographic group within each stratum.*

3.3 Attitudes and Concerns

Overall attitudes toward GenAI were cautiously positive: the mean composite attitude score was 3.43 ± 0.71 on the 5-point scale. A majority of nurses agreed or strongly agreed that ChatGPT could enhance their work efficiency (64.1%) and assist with patient-education materials (59.4%). However, only 22.5% trusted GenAI-generated medical information as much as a textbook, and 51.4% expressed concern that prolonged reliance might erode independent clinical judgment. Concerns were widespread and largely overlapping (Table 3): the leading concerns were information accuracy (78.0%), patient privacy and data leakage (72.1%), professional liability (67.7%), absence of institutional policy (63.6%), deskilling (54.5%), cultural appropriateness of advice (40.8%), and English-language limitations (37.7%).

3.4 Patterns of ChatGPT Use

In total, 38.0% of nurses (147/387, 95% CI 33.1–43.0%) reported any clinical use of ChatGPT in the prior six months; 12.1% reported weekly or more frequent use. Use was overwhelmingly conducted on personal devices (smartphone 91.0%, personal laptop 25.3%); only 14.0% of users had ever accessed ChatGPT from a hospital workstation. The most common reported clinical use cases were literature search (28.2%), drug-information lookup (22.0%), patient-education material drafting (18.1%), differential reasoning support (14.0%), discharge instructions (11.1%), and care-plan drafting (9.8%) (Table 3). Notably, 23.1% of users (34/147) acknowledged including some patient-related information in prompts—most commonly age, sex, and primary symptom—although only one respondent reported ever including a name or hospital identifier.

3.5 Independent Predictors of Clinical Use

Adjusted associations from the multivariable logistic regression are presented in Table 4. After adjustment for all candidate predictors and with cluster-robust standard errors at the hospital level, three factors remained independently associated with any clinical ChatGPT use in the prior six months: age <30 years (adjusted OR 2.12, 95% CI 1.34–3.36), high English proficiency (aOR 2.84, 1.78–4.52), and prior digital-health training (aOR 2.41, 1.55–3.74). Master's-level education and ICU posting showed positive but not significant associations. Female sex and total years of clinical experience were not associated with use. The model showed acceptable discrimination (apparent c-statistic

0.78, 95% CI 0.73–0.83; optimism-corrected 0.76 by 1000 bootstrap replications) and adequate calibration (Hosmer–Lemeshow χ^2 = 7.14, p = 0.521). Variance inflation factors were all below 1.8.

Table 3. ChatGPT use practices and prevalence of concerns (n = 387).

Item	n (% of 387)	95% CI
Any clinical use of ChatGPT in past 6 months	147 (38.0)	33.1–43.0
— Frequency: weekly or more often	47 (12.1)	9.0–15.7
— Frequency: monthly	73 (18.9)	15.1–23.1
— Frequency: less than monthly	27 (7.0)	4.7–9.9
Use case: literature search	109 (28.2)	23.7–32.9
Use case: drug-information lookup	85 (22.0)	17.9–26.4
Use case: patient-education materials	70 (18.1)	14.4–22.2
Use case: differential reasoning	54 (14.0)	10.7–17.7
Use case: discharge instructions	43 (11.1)	8.2–14.6
Use case: nursing care plans	38 (9.8)	7.0–13.2
Device: personal smartphone	352 (91.0)	87.7–93.7
Device: personal laptop	98 (25.3)	21.0–30.0
Device: hospital workstation	54 (14.0)	10.7–17.7
Concern: accuracy of information	302 (78.0)	73.6–82.0
Concern: patient privacy / data leakage	279 (72.1)	67.4–76.4
Concern: professional liability	262 (67.7)	62.9–72.3
Concern: lack of institutional policy	246 (63.6)	58.6–68.3
Concern: deskilling	211 (54.5)	49.5–59.6
Concern: cultural appropriateness	158 (40.8)	35.9–45.8
Concern: English language barrier	146 (37.7)	32.9–42.7

Note: *Use cases and concerns were multi-select items; percentages reference the full sample (n = 387) and therefore do not sum to 100. Confidence intervals were computed using the Wilson score method.*

Table 4. Multivariable logistic regression of factors associated with any clinical ChatGPT use in the prior six months (n = 387).

Predictor	Adjusted OR	95% CI	p-value
Age <30 years (vs ≥30)	2.12	1.34 to 3.36	0.001*
High English proficiency (vs moderate/low)	2.84	1.78 to 4.52	<0.001*
Prior digital-health training	2.41	1.55 to 3.74	<0.001*
Master's degree (vs Bachelor's/Diploma)	1.68	0.96 to 2.93	0.069
ICU posting (vs medical–surgical)	1.42	0.91 to 2.21	0.122
Emergency posting (vs medical–surgical)	1.28	0.79 to 2.07	0.314
Years of clinical experience (per year)	0.96	0.92 to 1.00	0.064
Female sex	0.78	0.49 to 1.24	0.296
Daily smartphone use (per hour)	1.18	1.05 to 1.32	0.005*

Note: *OR = odds ratio; CI = confidence interval. Adjusted OR estimates are from a single multivariable binary logistic regression model with cluster-robust standard errors at the hospital level, fitted in R 4.3.2 with the sandwich and lmtest packages. * $p < 0.05$. Apparent c-statistic 0.78 (95% CI 0.73–0.83); optimism-corrected c-statistic 0.76 by 1000 bootstrap replications. Hosmer–Lemeshow goodness-of-fit $\chi^2 = 7.14$, $p = 0.521$. All variance inflation factors <1.8.*

4. Discussion

This is, to our knowledge, the first multicenter survey of Iraqi nurses' knowledge, attitudes, and clinical use of generative AI. Three findings stand out. First, knowledge was uneven: while basic concepts were well understood, knowledge of data-privacy and regulatory implications was poor, with fewer than 4 in 10 respondents recognizing that pasting patient identifiers into a public GenAI tool can constitute a confidentiality breach. Second, attitudes were cautiously positive but accompanied by widespread concern, particularly around accuracy and privacy. Third, despite these concerns, more than a third of nurses had already used ChatGPT for clinical purposes in the prior six months, almost entirely on personal devices and outside any institutional framework.

The 38.0% prevalence of any clinical use is broadly consistent with regional surveys of healthcare professionals in 2023–2025 [3,4,11], although direct comparison is difficult because most published studies have surveyed physicians or medical students rather than nurses, and few have separated clinical use from educational or personal use. Iraqi nurses appear to be using GenAI in ways that mirror international peers—predominantly for literature search, drug information, and patient-education drafting—rather than for direct clinical-reasoning support, suggesting that current use is closer to a personal study aid than a bedside decision tool. This pattern reduces, but does not eliminate, the immediate clinical risk.

The most actionable finding concerns privacy. Nearly a quarter of users (23.1%) acknowledged including patient-related information—most commonly demographics and primary symptom—in their prompts. Combined with the low correct-response rate on data-privacy items (38.4%), this signals a clear knowledge–behavior gap that institutional policy can address quickly. International guidance has converged on three minimum requirements for clinical GenAI use: prohibition of patient-identifier entry into general-purpose tools, deployment of locally hosted or contractually compliant alternatives where clinical use is desired, and routine staff training [10,12,13]. None of these is in place at the participating hospitals at the time of writing, and the survey suggests that the absence of policy itself is a top-five concern among nurses (63.6%).

Independent predictors of use—younger age, higher English proficiency, and prior digital-health training—match patterns reported elsewhere [3,14,15]. The English-proficiency association is particularly relevant in an Arabic-first health system: GenAI tools currently perform less reliably in clinical Arabic than in English, and nurses with weaker English are simultaneously less likely to use these tools and more likely to be exposed to potentially lowquality Arabic outputs when they do. Cultural appropriateness was already a flagged concern for 40.8% of the sample, and English-language barriers for 37.7%; these point toward a plausible inequity dimension of GenAI uptake that deserves longitudinal follow-up [16].

From a practice and policy perspective, the implications cluster into three areas. Educationally, structured AI literacy—combining basic concepts, hallucination awareness, prompt-engineering hygiene, and explicit privacy rules—should be added to nursing continuing-education programs; the strong association of prior digital training with higher knowledge and more confident use suggests the return on training is high. Institutionally, hospitals should issue brief, enforceable GenAI policies covering acceptable-use cases, prohibited inputs, device boundaries, and a documented escalation route for concerns. From a research perspective, longitudinal designs are needed to track whether early use generalizes into bedside reasoning, whether knowledge gaps narrow with policy implementation, and whether equity gaps between English-proficient and Arabic-only nurses widen or close as Arabic-language GenAI matures [17,18].

Limitations should be acknowledged. First, all use measures were self-reported and may be subject to social-desirability bias in either direction—under-reporting where nurses fear sanction, over-reporting where they wish to appear technologically adept. Second, the cross-sectional design cannot establish causation between predictors and use. Third, the three participating hospitals are large urban teaching institutions; rural and small-district Iraqi hospitals may show different patterns, and the findings should not be extrapolated without caution. Fourth, no objective measurement of actual prompt content was performed; the privacy-related figures rely on respondent recall. Fifth, the questionnaire was investigator-developed and adapted from international frameworks; despite acceptable internal consistency and content validity, formal external validation in another Arabic-speaking nursing population is warranted. Finally, GenAI tools are evolving rapidly; the picture six months from now may differ from the one captured here.

5. Conclusions

Iraqi nurses surveyed across three teaching hospitals demonstrated moderate baseline knowledge of generative AI, cautiously positive attitudes, and meaningful clinical use—approximately 38% had used ChatGPT for at least one clinical purpose in the prior six months, almost entirely on personal devices and outside any institutional framework. Use was independently predicted by younger age, higher English proficiency, and prior digital-health training. The most acute gap is around data privacy: knowledge of confidentiality implications was the weakest knowledge subdomain, while concerns about privacy were the second most prevalent. Hospitals should respond by issuing clear institutional GenAI policies, integrating AI literacy into nursing continuing education, and—where clinical use of GenAI tools is to be encouraged—deploying locally hosted or contractually compliant alternatives that do not require nurses to use personal devices. Further research should track use longitudinally, evaluate the impact of policy implementation, and address the equity dimension exposed by the English-proficiency gradient.

6. References

[1] Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare. 2023;11(6):887. doi:10.3390/healthcare11060887

[2] Lee P, Bubeck S, Petro J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. N Engl J Med. 2023;388(13):1233-1239. doi:10.1056/NEJMsr2214184

[3] Temsah MH, Aljamaan F, Malki KH, et al. ChatGPT and the future of digital health: a study on healthcare workers' perceptions. JMIR Med Educ. 2023;9:e49877. doi:10.2196/49877

[4] Aljamaan F, Temsah MH, Altamimi I, et al. Reference hallucination score for medical artificial intelligence chatbots: development and usability study. JMIR Med Inform. 2024;12:e54345. doi:10.2196/54345

[5] Buchanan C, Howitt ML, Wilson R, Booth RG, Risling T, Bamford M. Predicted influences of artificial intelligence on the domains of nursing: scoping review. JMIR Nurs. 2020;3(1):e23939. doi:10.2196/23939

[6] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56. doi:10.1038/s41591-018-0300-7

[7] Nashwan AJ, Abujaber AA. Harnessing the power of ChatGPT in nursing education: a critical reflection. Cureus.

2023;15(11):e49998. doi:10.7759/cureus.49998
[8] McGrow K. Artificial intelligence: essentials for nursing. Nursing. 2019;49(9):46-49. doi:10.1097/01.NURSE.0000577716.57052.8d
[9] von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 2007;370(9596):1453-1457. doi:10.1016/S0140-6736(07)61602-X
[10] Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. J Am Med Inform Assoc. 2020;27(3):491-497. doi:10.1093/jamia/ocz192
[11] Abdaljaleel M, Barakat M, Alsanafi M, et al. A multinational study on the factors influencing university students' attitudes and usage of ChatGPT. Sci Rep. 2024;14:1983. doi:10.1038/s41598-024-52549-8
[12] Liu J, Wang C, Liu S. Utility of ChatGPT in clinical practice. J Med Internet Res. 2023;25:e48568. doi:10.2196/48568
[13] Liaw SY, Tan JZ, Lim S, et al. Artificial intelligence in virtual reality simulation for interprofessional communication training: mixed method study. Nurse Educ Today. 2023;122:105718. doi:10.1016/j.nedt.2023.105718
[14] Dave T, Athaluri SA, Singh S. ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. Front Artif Intell. 2023;6:1169595. doi:10.3389/frai.2023.1169595
[15] Wang C, Liu S, Yang H, Guo J, Wu Y, Liu J. Ethical considerations of using ChatGPT in health care. J Med Internet Res. 2023;25:e48009. doi:10.2196/48009
[16] Pailaha AD. The impact and issues of artificial intelligence in nursing science and healthcare settings. SAGE Open Nurs. 2023;9:23779608231196847. doi:10.1177/23779608231196847
[17] Goh E, Gallo R, Hom J, et al. Large language model influence on diagnostic reasoning: a randomized clinical trial. JAMA Netw Open. 2024;7(10):e2440969. doi:10.1001/jamanetworkopen.2024.40969
[18] Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. Nature. 2023;620(7972):172-180. doi:10.1038/s41586-023-06291-2
[19] Chow JCL, Sanders L, Li K. Impact of ChatGPT on medical chatbots as a disruptive technology. Front Artif Intell. 2023;6:1166014. doi:10.3389/frai.2023.1166014
[20] Choudhury A, Shamszare H. Investigating the impact of user trust on the adoption and use of ChatGPT: survey analysis. J Med Internet Res. 2023;25:e47184. doi:10.2196/47184
[21] Booth RG, Strudwick G, McBride S, O'Connor S, Solano López AL. How the nursing profession should adapt for a digital future. BMJ. 2021;373:n1190. doi:10.1136/bmj.n1190
[22] Almagharbeh WT, Alfanash HA, Alnawafleh KA, et al. The impact of AI-based decision support systems on nursing workflows in critical care units. Int Nurs Rev. 2024. doi:10.1111/inr.13011
[23] Nashwan AJ, Abujaber AA, Choudry H. Embracing the future of physician-patient communication: GPT-4 in gastroenterology. Gastroenterol Endosc. 2023;1(3):132-135. doi:10.1016/j.gande.2023.07.004
[24] Sezgin E, Sirrianni J, Linwood SL. Operationalizing and implementing pretrained, large artificial intelligence linguistic models in the US health care system: outlook of generative pretrained transformer 3 (GPT-3) as a service model. JMIR Med Inform. 2022;10(2):e32875. doi:10.2196/32875
[25] Iftikhar L. DocGPT: impact of ChatGPT-3 on health services as a virtual doctor. EC Paediatr. 2023;12:45-55. doi:10.31080/ECPE.2023.12.01215