Human-AI Collaboration in the Pharmacy: A Cluster Randomized Controlled Trial of Generative AI for Medication Counselling and Adherence

Document date: 9 June 2026

Abstract

Background and Aim: Medication counseling within community pharmacies is crucial for managing chronic diseases, yet significant challenges regarding correctness and completeness remain in Jordan. Although generative artificial intelligence (AI) can be utilized for patient education, there is a lack of research on clinical impact and safety of AI in medication counseling conducted by pharmacists in real-world practice. The aim of this study is to evaluate the effect of pharmacist-supervised AI-assisted medication counseling on the correctness and completeness of counseling information and 30-day medication adherence among patients in Jordanian community pharmacies.

Materials and Methods: This pragmatic, two-arm cluster randomized controlled trial enrolled 136 adult patients across 16 community pharmacies in Jordan (8 clusters per arm). Pharmacists in the intervention arm used a standardized prompt strategy with ChatGPT® to generate counseling drafts, which were then verified and edited before delivery. The control arm provided usual counseling. Co-primary outcomes were correctness and completeness of counseling information (percentage scores based on blinded transcript analysis). Secondary outcomes included 30-day medication adherence (General Medication Adherence Scale [GMAS]), immediate patient understanding, and satisfaction. Data were analyzed using mixed-effects linear and logistic regression models.

Results: Baseline characteristics were well-balanced between arms. The intervention arm achieved significantly higher scores for both correctness (82.4 ± 11.6 vs. 64.9 ± 14.8; adjusted mean difference [AMD]: 16.8; 95% CI: 9.8 to 23.9; $p<0.001$) and completeness (85.7 ± 10.9 vs. 68.3 ± 15.1; AMD: 15.9; 95% CI: 8.7 to 23.1; $p<0.001$). Patients in the intervention arm also demonstrated significantly better 30-day adherence (AMD: 2.0; $p=0.006$; adjusted OR for good adherence: 1.96; 95% CI: 1.04 to 3.70; $p=0.037$). Patient understanding and satisfaction were also significantly improved ($p\leq0.001$). While face-to-face counseling time was longer in the intervention arm (4.7 vs. 3.6 minutes; $p<0.001$), 76.5% of AI-generated drafts required

professional editing or rejection to ensure clinical safety. Qualitative analysis revealed that while AI ensured a high quality floor, the mandatory pharmacist review layer was essential, intercepting four high-severity clinical near misses (5.9%) and simplifying overly technical medical jargon in 10.3% of encounters.

Conclusion: The use of pharmacist-supervised generative AI significantly enhanced the quality of medication counseling and short-term patient adherence in community pharmacy settings. Such evidence validates the idea of human-AI collaboration in which AI acts as a comprehensive scaffold clinical on the condition that pharmacist's supervision remains mandatory.

Keywords: Generative AI, ChatGPT, Community Pharmacy, Medication Adherence, Jordan, Clinical Decision Support.

Introduction

Community pharmacists are extremely high accessible among healthcare practitioners in Jordan, and they act as the first point of contact for any concerns related to medicines, chronic diseases, and drug supply (Abdel-Qader *et al.*, 2024). Nevertheless, studies conducted in Jordan demonstrate that improvements can be made in the counselling and dispensing process in terms of quality and reliability (Abdel-Qader *et al.*, 2021; Hammad *et al.*, 2022; Naser and Sbeat, 2022). This is because studies on community pharmacies have found instances where dispensing errors occurred in their practices. Additionally, they found inconsistencies in the amount of information provided to patients in counselling sessions. Patient satisfaction with the counselling process was also inconsistent (De Rosis, Ferrè and Pennucci, 2022).

The use of generative artificial intelligence (AI) in medicine and pharmacy is gaining more prominence, where these technologies are considered useful tools in relation to medication management processes, communications with patients, and other aspects (Li *et al.*, 2025). According to the International Pharmaceutical Federation, there is room for AI in enhancing the competencies of pharmacists in relation to medication therapy, counselling, and workflow, among others (FIP, 2025). Emerging comparative literature also suggests that large language models may perform relatively well in some drug-counselling tasks while remaining weaker in higher-risk or more complex clinical functions, reinforcing the view that these systems may be most useful when embedded within a pharmacist-supervised workflow rather than used autonomously (Huang *et al.*, 2024; FIP, 2025).

In Jordan, the published literature to date appears to have focused mainly on pharmacists' awareness, perceived usefulness, and concerns regarding generative AI rather than on patient-facing interventional studies. Jordanian pharmacists have generally reported interest in the potential usefulness of ChatGPT® in pharmacy practice, but they have also expressed concerns about privacy, ethics, legal responsibility, and regulatory uncertainty (Abu Hammour *et al.*, 2023). More broadly, pharmacy professionals in the Middle East and North Africa region have reported substantial ethical concerns regarding AI implementation, particularly around data privacy, cybersecurity, transparency, and lack of legal regulation (Hasan *et al.*, 2024). Overall, these results indicate that rather than focusing on whether pharmacists are aware of generative AI technology,

the more important issue is whether the use of these technologies by pharmacists can enhance the counseling services provided to patients.

In Jordan, for instance, medication adherence continues to be an important concern for patients suffering from chronic illnesses and multimorbidity. According to recent Jordanian research, only about 50% of patients with multimorbidity were considered adherent at a high or moderate level, where issues related to burden, cost, and access hindered adherence (Abdel-Qader *et al.*, 2025). At the same time, Jordan continues to face a substantial burden of non-communicable diseases, particularly hypertension, diabetes, and cardiovascular risk conditions that often require long-term pharmacotherapy and repeated counselling (WHO EMRO, 2024). In this context, interventions that improve the correctness, completeness, and clarity of medication counselling may plausibly improve short-term understanding and adherence.

Prior Jordanian community-pharmacy studies have already shown that counselling performance can improve through structured interventions. For example, inhaler counselling quality improved after workshop-based education in Jordanian community pharmacies, and simulated-patient work has shown that structured prompting can increase the amount of information pharmacists provide during counselling encounters (Elayeh *et al.*, 2019; Hammad *et al.*, 2022). Building on that literature, the present study evaluates whether a pharmacist-supervised generative AI workflow can serve as a practical, structured support tool during real counselling encounters.

Accordingly, this cluster randomized controlled trial evaluated the effect of pharmacist-supervised generative ChatGPT®-assisted medication counselling on the correctness and completeness of counselling information delivered in Jordanian community pharmacies and examined whether any improvements were accompanied by better immediate patient understanding**,** patient satisfaction, and 30-day medication adherence among adults receiving chronic medicines.

Materials and Methods

Study design and reporting framework

This study was a pragmatic, parallel, two-arm cluster randomized controlled trial design, with the community pharmacy as the unit of randomization and the patient encounter as the unit of analysis. Pharmacies were randomized into two parallel arms: The Intervention Arm (AI-assisted counseling) and the Control Arm (usual care). Cluster randomization was appropriate because the

intervention was delivered at workflow level, and pharmacists working within the same pharmacy were likely to influence each other's counselling behavior; therefore, individual randomization would be vulnerable to contamination (Campbell *et al.*, 2012; Hemming *et al.*, 2017).

The protocol was developed in accordance with SPIRIT 2013 and SPIRIT-AI (Chan *et al.*, 2015; Rivera *et al.*, 2020), the intervention was described using the TIDieR framework (Hoffmann et al., 2014), and the final manuscript was reported in line with the CONSORT extension for cluster randomized trials and CONSORT-AI (Campbell *et al.*, 2012; Liu *et al.*, 2020).

Ethical considerations and AI governance

Ethics approval was obtained from the institutional review board in the University of Petra before recruitment begins (**S/22/12/2025)**. Written informed consent was obtained from participating pharmacists and patients before any data collection was initiated (See Appendices K and L).

Given the intervention's AI component, additional safeguards were implemented. No names, phone numbers, national identification numbers, exact addresses, or full medical records were entered into ChatGPT®. Only the minimum necessary clinical context required to generate counselling support was used. Where available, privacy-enhanced or enterprise settings that did not use prompts for model training was preferred. Pharmacists were trained to identify hallucinations and unsafe outputs, and any output conflicting with authoritative drug information was rejected and corrected before counselling was delivered. A log of unsafe or discrepant outputs was maintained. These safeguards were supported by FIP guidance and by published pharmacist concerns regarding privacy, cybersecurity, legal uncertainty, bias, and accuracy in AI use (Abu Hammour *et al.*, 2023; Hasan *et al.*, 2024; FIP, 2025).

Study setting

The trial was conducted in licensed community pharmacies in Jordan, stratified into Jordan's three administrative regions: North, Central, and South between January 2026 and March 2026, to improve external validity and to capture routine counselling practices across different pharmacy contexts. Jordanian community pharmacies are a suitable trial setting because they are highly accessible first-contact sites for medicine-related advice, and prior Jordanian studies have successfully evaluated counselling quality, inhaler education, OTC case management, and patient

satisfaction in this setting (Elayeh *et al.*, 2019; Wazaify *et al.*, 2019; Abdel-Qader *et al.*, 2021; Hammad *et al.*, 2022; Naser and Sbeat, 2022).

Eligibility criteria

Community pharmacies were eligible if they were legally registered in Jordan, provided routine prescription dispensing services, had at least one licensed pharmacist available during recruitment hours, and agreed to participate for the full trial period. Pharmacies that were already using structured AI-assisted counselling tools as part of routine practice were excluded to reduce baseline contamination.

Licensed pharmacists with minimum 2 years of clinical experience working in participating pharmacies were eligible if they provided direct patient counselling and consent to take part in the study. Pharmacists on temporary placement for less than one month, or pharmacists not involved in patient-facing counselling, were excluded.

Eligible patients were adults aged 18 years or older who presented with a new prescription or a refill for a chronic medication requiring counselling within the following classes: antihypertensives, oral antidiabetics, lipid-lowering agents, anticoagulants, or inhaled maintenance therapies. These classes were selected due to their high prevalence in Jordan and the complexity of their associated counseling requirements (Abdel-Qader *et al.*, 2021, 2025; WHO EMRO, 2024). Patients provided informed consent and completed 30-day follow-up. Exclusion criteria included acute infections, psychiatric disorders, oncological conditions, or severe acute illness requiring urgent referral, cognitive impairment precluding informed consent, hearing or communication barriers that prevented interview completion without a caregiver, and inability to provide a follow-up phone number. Chronic medicines were an appropriate focus because medication adherence remains a major problem in Jordan and other similar settings, and pharmacist counselling has been associated with better adherence in randomized evidence syntheses (Kelly *et al.*, 2023; Abdel-Qader *et al.*, 2025).

Recruitment and consent

Pharmacies were approached through professional networks, regional pharmacy directories, and direct site visits. After pharmacy-level consent was obtained, eligible pharmacists were enrolled and baseline data collected. During the patient recruitment phase, consecutive eligible patients

attending participating pharmacies during predefined study sessions were invited to participate. Informed written consent was sought prior to any data collection. In order to reduce potential selection bias, a log of recruitment was kept at each pharmacy regarding the number of patients screened, eligible, enrolled, refused, and lost prior to baseline (Appendix I. Log of screening, enrolment, and recruitment). The flow of pharmacies and participants through each phase of the study is provided in line with CONSORT 2010 for cluster randomized trials (Figure 1).

Figure 1. CONSORT Flow Diagram of the Trial

Intervention development

The intervention was based on a pharmacist-supervised generative-AI counselling support workflow using a study-approved version of ChatGPT®. The intervention was designed to augment, not replace, professional judgement. Such an intervention strategy is compatible with the current guidelines for pharmacists' involvement with AI technologies, emphasizing the need for supervision, management of input/output, visibility of interaction between humans and AI, and risk assessment regarding inaccuracies, privacy concerns, drift, and bias (Liu *et al.*, 2020; Rivera *et al.*, 2020; FIP, 2025).

The intervention relied solely on ChatGPT® Basic/Free plan (OpenAI, 2022) as the only generative AI counseling support system. It was decided to use one model in order to reduce intervention heterogeneity and enhance the interpretability of intervention effectiveness. Such an approach was motivated by recent medical literature that demonstrated better-than-average performance of ChatGPT® over other AI tools in a variety of settings (Sallam *et al.*, 2024; Fattah *et al.*, 2025), although it is true that the performance varies depending on the application.

Prior to conducting the trial, the research team (four academic clinical pharmacists and two experienced community pharmacy practitioners with an average experience of 12 years) developed: (1) standard prompt templates for different classes of medications; (2) medication-specific references derived from reliable drug information sources and package inserts; (3) a counselling checklist and scoring with criteria for accuracy and completeness; and (4) a pharmacist training guide, which includes demonstrations of safe prompting, unsafe prompting, recognizing hallucinations, and proper procedures. All materials were then validated for face and content validity by three professors of clinical pharmacy practice, each with over 15 years of experience

in patient counseling research and health informatics. Modifications were made after validation and pilot-testing with 10 non-study participants.

Pharmacists in intervention pharmacies underwent a training session lasting about 5 hours delivered by the research team, which involved the following elements: (1) aims and workflow of the trial; (2) proper utilization of safe and standardized prompts; (3) complete prohibition of entering any type of identifiable patient data; (4) cross-checking AI prompts with reliable medication references; (5) editing the AI prompt into concise patient counselling; and (6) documentation guidelines for discrepancies, overrides, and suspicious outputs.

The training process was also supplemented by mock counselling scenarios and a brief competency assessment prior to allowing pharmacists to apply the intervention during actual encounters (Appendix B. Pharmacist Training & Competency Checklist). This is justified by previous Jordanian research involving feedback and training sessions for pharmacy professionals demonstrating significant gains in counselling-based practices, as well as AI recommendations that always stress the importance of user competence and governance prior to any intervention (Elayeh *et al.*, 2019; Hammad *et al.*, 2022; FIP, 2025).

Intervention arm procedures

For all eligible patients in the intervention arm, the pharmacist performed the standard patient assessment and determined which medicine(s) needed counselling. Then, the pharmacist input a prompt in a de-identified format into ChatGPT®. The prompt was a request for an easy-to-understand counselling document with information regarding the indications for the medication, dosage, schedule, route, course, missed doses, possible side effects, important precautions, storage, and advice on taking the medicine as prescribed (Appendix A). The pharmacist ensured that the content generated by the AI was accurate and clear, making corrections where necessary, and then gave verbal counselling to the patient.

The AI output was never provided to the patients without pharmacist evaluation. It is worth noting that pharmacists could also reject the AI output as inaccurate, insufficient, hazardous, and inappropriate altogether. Reproducibility was ensured through documenting the date and time, prompt template employed, acceptance/revision/rejection, and time taken for AI-assisted

preparation. Documentation was especially crucial in this AI experiment since reporting guidelines advocate for clear specification of input, output, human intervention, and failure of performance (Liu *et al.*, 2020; Rivera *et al.*, 2020). Concerns regarding privacy, legal uncertainty, accuracy, and bias have also been reported among pharmacists in Jordan and the wider region, which further supports strict pharmacist supervision and de-identification procedures (Abu Hammour *et al.*, 2023; Hasan *et al.*, 2024).

Control arm procedures

Pharmacies randomized to the control arm continued to provide usual medication counselling according to their standard routine practice, without access to the AI prompt templates or study AI workflow. Control pharmacists used their usual professional references, as would occur in routine care, but they were not trained in or asked to use ChatGPT® during the trial period.

Outcome measures

Co-primary outcomes

Correctness of counselling information: Correctness of counselling information was defined as the proportion of clinically applicable counselling domains communicated accurately and appropriately during the encounter, compared with a medication-specific reference sheet compiled from the British National Formulary (BNF, BNF Edition 89, 2025). For each applicable domain, two blinded senior pharmacist assessors classified the information as correctly informed, misinformed, or uninformed, adapted from published simulated-patient pharmacy studies (Surur *et al.*, 2017). For the primary correctness outcome, domains were scored as 1 if correctly informed and 0 if not correctly informed, with both misinformed and uninformed domains scored as 0 (Surur *et al.*, 2017). A counseling deficiency was defined as any instance where a required clinical domain was provided with inaccurate information (incorrectness). The correctness score was calculated as:

$$\text{Correctness score} = \frac{\text{Number of applicable domains correctly informed}}{\text{Total number of applicable domains}} \times 100$$

This approach was based on published domain-level counselling assessments and allowed objective quantification of counselling accuracy (Surur *et al.*, 2017; Hammad *et al.*, 2022) (Appendix C. Counselling assessment checklist).

Completeness of counselling information: Completeness of counselling information was defined as the proportion of essential and clinically applicable counselling domains that were addressed during the encounter. Each applicable domain was scored as 1 if addressed in any form and 0 if omitted. A counseling deficiency was defined as any instance where a required clinical domain was omitted (incompleteness). The completeness score will be calculated as:

$$\text{Completeness score} = \frac{\text{Number of applicable domains addressed}}{\text{Total number of applicable domains}} \times 100$$

This scoring method followed the item-based percentage scoring approach used in previous pharmacy simulated-patient research, in which each information item provided was awarded one point and the total content was converted to a percentage (Hammad *et al.*, 2022). The assessed domains included indication/purpose, dose, frequency/timing, duration where relevant, route or method of administration, missed-dose advice, major side effects, precautions/interactions/contraindications, storage, adherence advice, and monitoring or warning signs requiring medical review (Surur *et al.*, 2017; Ali, Shimels and Bilal, 2019; FIP, 2021; Hammad *et al.*, 2022) (Appendix C. Counselling assessment checklist).

Outcome assessment procedures

Because the trial also evaluated patient adherence, the preferred assessment method for the co-primary counselling outcomes were audio-recorded real-patient counselling encounters rather than simulated-patient visits. Each consented encounter was audio-recorded and de-identified before scoring. The research evaluation panel consisting of two independent clinical pharmacists, blinded to allocation, scored the transcripts against the same pre-specified medication-specific reference sheet derived from the BNF. Discrepancies were resolved by consensus, and a third senior clinical pharmacist adjudicated unresolved disagreements. To assess inter-rater reliability, at least 20% of transcripts was double-scored and agreement was quantified using weighted kappa or intraclass correlation coefficients, as appropriate.

The decision to use a structured, blinded scoring system was informed by the strong use of simulated-patient and structured-feedback methods in Jordanian community pharmacy research, as well as the broader pharmacy literature recognizing structured simulated-patient methodology as a robust way to evaluate practice quality (Elayeh *et al.*, 2019; Hammad *et al.*, 2022).

Clinical Severity Assessment

Clinical severity was assessed for both incorrectness (the provision of clinically inaccurate information) and incompleteness (the total omission of an essential counseling domain). It was independently assessed by a panel of three senior clinical pharmacists using a 3-point scale adapted from established medication safety literature (Abdel-Qader *et al.*, 2021):

- **Low Severity:** Minor omissions or wording issues unlikely to affect safety or efficacy (e.g., omitting specific storage humidity, using medical jargon, such as myalgia instead of muscle pain).
- **Moderate Severity:** Omissions or errors that could lead to sub-therapeutic effects or patient distress (e.g., incompleteness regarding missed-dose instructions for antihypertensives, or incorrectness regarding administration timing relative to food).
- **High Severity:** Omissions or errors with high potential for significant patient harm (e.g., incompleteness regarding red-flag bleeding signs for anticoagulants, or incorrectness regarding a major drug-drug interaction).

The research evaluation panel reviewed all instances where the co-primary outcome score for a domain was 0 (misinformed or uninformed). Disagreements were resolved through consensus.

Secondary outcomes

Thirty-day medication adherence: The main patient-level secondary outcome was medication adherence at 30 days after the index encounter. Adherence was assessed using the General Medication Adherence Scale (GMAS) via telephone follow-up, because it has been validated in Jordan and captures intentional, unintentional, and cost-related non-adherence dimensions (Naqvi *et al.*, 2018). GMAS total scores were analyzed continuously and, secondarily, dichotomized using validated thresholds where appropriate **(**Appendix G. Thirty-day follow-up case report form).

Immediate patient understanding: Immediately after the counselling encounter, patients completed a brief interviewer-administered understanding assessment based on teach-back principles. They were asked to state, in their own words, the medicine's purpose, how to take it, what to do if a dose is missed, and one important safety or side-effect point **(**Appendix E. Immediate patient understanding assessment (teach-back form)). Responses were scored against

the same encounter-specific medication reference sheet derived from the BNF, producing a total understanding score.

Patient satisfaction: Patients completed a short satisfaction questionnaire immediately after the encounter, covering clarity, usefulness, confidence in the information received, opportunity to ask questions, and overall satisfaction with the counselling process (Appendix F. Patient satisfaction questionnaire). Jordanian studies show that counselling and dispensing procedures are important dimensions of patient satisfaction in community pharmacy, supporting inclusion of this outcome (Naser and Sbeat, 2022).

Counselling time: Total counselling time was measured in seconds from the start of counselling to completion, using audio timestamps. In intervention pharmacies, AI preparation time was also recorded separately to distinguish preparation burden from face-to-face counselling time.

Process and safety outcomes: The study additionally recorded the proportion of encounters in which AI output was fully accepted, edited, or rejected; the frequency and type of detected AI inaccuracies; and any counselling-related incidents (defined as any inaccurate medication-related information that was delivered to the patient during counselling and had the potential to lead to inappropriate medication use, reduced adherence, or minor patient harm, such as transient adverse drug events (Abdel-Qader *et al.*, 2025)) or near misses (defined as any AI-generated counselling statement that contained a clinically inaccurate dose, contraindication, interaction warning, administration instruction, or other medication-related error, but was identified and corrected by the pharmacist before reaching the patient (Abdel-Qader *et al.*, 2021)) identified during the study period. Recording AI failure cases was consistent with AI-trial reporting guidance (Liu *et al.*, 2020) (Appendix H. AI output discrepancy and safety log).

Data collection and quality assurance

Baseline pharmacy data included region, pharmacy type (independent/chain), staffing pattern, daily prescription volume, and internet/device availability. Baseline pharmacist data included age, sex, years of experience, highest qualification, prior AI exposure, and self-rated digital confidence. Baseline patient data included age, sex, education, medication class, new versus refill prescription, number of chronic medicines, and selected health-literacy proxies where feasible **(**Appendix J. Baseline data collection fields).

To standardize scoring and reduce information bias, the research team prepared medication-specific evidence sheets before recruitment starts **(**Appendix D. Medication-specific reference sheet template), piloted the rubric on 10 encounters, and revised unclear criteria before full implementation. Assessors remained blinded to trial arm, and transcripts were coded with anonymous encounter IDs only. A monitoring log documented the protocol deviations, missing data, and follow-up completion.

Sample size determination

The sample size was calculated based on the two co-primary continuous outcomes, namely correctness and completeness of medication counselling. These outcomes were analyzed as percentage scores on a 0–100 scale, because each reflects the proportion of applicable counselling domains delivered accurately and completely. This scoring approach was also consistent with Jordanian community-pharmacy research that reported counselling content as percentage scores (Hammad *et al.*, 2022).

The choice of the target difference (Δ) and variance assumptions was informed by methodological guidance and local pharmacy evidence. A 10-point absolute difference on the 0–100 counselling score was selected as a pragmatic and clinically meaningful improvement for this trial. This value was considered conservative because prior Jordanian community-pharmacy studies have shown poor baseline counselling performance and substantial room for improvement after structured educational interventions (Elayeh *et al.*, 2019; Hammad *et al.*, 2022).

The assumed standard deviation (σ) of 16 points was chosen on a conservative basis from previous Jordanian pharmacy studies. In a simulated-patient study from Jordan reported a spontaneous counselling-content score of 16.2% (SD 15.6) and a total information-content score of 34.4% (SD 16.0) (Hammad *et al.*, 2022). Another Jordanian study found inhaler counselling scores improved from 5.4 (SD 1.6) to 7.8 (SD 0.9) on a 9-point scale and from 4.6 (SD 2.5) to 9.9 (SD 0.6) on a 10-point scale, indicating variability in a similar range when converted to percentages (Elayeh *et al.*, 2019). Therefore, using SD = 16 inflated the variance beyond the observed SDs and was appropriately conservative for planning (Elayeh *et al.*, 2019; Hammad *et al.*, 2022).

Because the trial had two co-primary outcomes, multiplicity was addressed at the design stage by using a Bonferroni-adjusted two-sided alpha (α) of 0.025 for each outcome (Vickerstaff, Omar and Ambler, 2019). The Power (1−β) was set at 80%.

For a two-arm trial with a continuous outcome, the individually randomized sample size per arm was calculated using:

$$n = \frac{2\sigma^2(Z_{1-\frac{\alpha}{2}} + Z_{1-\beta})^2}{\Delta^2}$$

After substituting the parameters, the required sample size under individual randomization was 48.63 patients per arm.

Because this was a cluster randomized controlled trial, the individually randomized sample was inflated using the standard design effect:

$$DE = 1 + (m-1)\rho$$

where (m) is the average number of patients recruited per pharmacy and (ρ) is the intracluster correlation coefficient (ICC). An average cluster size of 10 patients per pharmacy and an ICC of 0.02 were assumed. The assumption of a modest ICC is consistent with methodological guidance for cluster trials, which emphasizes accounting for within-cluster correlation and transparent justification of both cluster size and ICC in the sample-size calculation (Rutterford, Copas and Eldridge, 2015; Hemming *et al.*, 2017). This gave a design effect of 1.18.

After adjustment for clustering, the required sample became 57.38 patients per arm. To allow for 15% loss to follow-up for the 30-day adherence assessment, the final required sample size was 68 patients per arm (136 patients in total).

This sample corresponds to 8 pharmacies per arm and 16 pharmacies in total. This could be achieved by allowing slight variation in cluster sizes, with approximately 9–10 patients recruited per pharmacy across the 16 pharmacies. This approach is acceptable in cluster randomized trials, provided that large imbalances in cluster size are avoided (Hemming *et al.*, 2017).

Randomization, allocation concealment and blinding

Participating pharmacies were randomized in a 1:1 ratio to intervention or control. Randomization was performed by an independent medical statistician using a computer-generated sequence, with stratification by region to improve balance. Variable block sizes were used to reduce predictability. Allocation was concealed until after baseline data collection was completed. Intervention training was then delivered only to pharmacies assigned to the AI-assisted arm.

Blinding of pharmacists was not possible because they knew whether they were using the AI-assisted workflow. However, the following layers of blinding were implemented: transcript scorers for correctness and completeness were blinded to group allocation; the statistician analyzed a masked dataset with anonymized arm labels where feasible; patients were not explicitly told the trial hypothesis comparing AI-assisted with usual counselling, only that the study evaluated medication-counselling processes. These procedures are important because cluster trials involving provider behavior are particularly vulnerable to performance and detection biases if blinding is not addressed carefully (Campbell *et al.*, 2012; Hemming *et al.*, 2017).

Statistical analysis

All analyses followed the intention-to-treat principle, with patients analyzed according to the allocation of the pharmacy where they were enrolled. Descriptive statistics summarized baseline characteristics by study arm. Continuous co-primary outcomes (correctness and completeness scores) were analyzed using linear mixed-effects models with a random intercept for pharmacy to account for clustering. Fixed effects included study arm and pre-specified covariates, such as region, pharmacy type, medication class, and new versus refill status.

Because the trial had two co-primary outcomes, multiplicity was addressed using a Holm-adjusted familywise alpha. Secondary continuous outcomes (understanding, satisfaction, counselling time, and adherence score) were also analyzed using mixed-effects linear models. Binary secondary outcomes, such as adherent versus non-adherent at 30 days, were analyzed using mixed-effects logistic regression. Effect estimates were reported with 95% confidence intervals and two-sided p values.

The handling of missing outcomes was initially done according to the arm and the reason. In case of missing outcomes due to non-adherence in the 30 days after initiation of the intervention exceeding a low percentage, a sensitivity analysis using multiple imputation on a missing-at-

random assumption was performed. Other types of sensitivity analyses included per-protocol analysis that excluded major deviations from the intervention protocol, as well as subgroup analyses stratified by medication type, whether it is a new prescription or a refill, and region (Rutterford, Copas and Eldridge, 2015; Hemming *et al.*, 2017).

Results

Participant flow and recruitment

A total of 23 community pharmacies were identified for the assessment, of which 16 pharmacies fulfilled the eligibility criteria and were randomly assigned to two arms; 8 pharmacies for the intervention group and 8 pharmacies for the control group. All the randomized pharmacies successfully participated in the study. The number of patients recruited, eligible for enrollment, and selected into each arm was 154, 136, and 68 respectively. Patients in the intervention arm received AI-assisted counseling by the pharmacist, and those in the control group received conventional counseling.

Counselling sessions were audio-recorded for 132 out of 136 patients (97.1%), whereas 126 out of 136 patients (92.6%) returned for the 30-day follow-up. There was a loss to follow-up for the adherence outcome of five patients in both arms.

Baseline characteristics

The demographic details were comparable for the two study arms. The average age of the patients was 56.8 ± 12.4 years in the intervention arm and 55.9 ± 11.8 years in the control arm. Females constituted 52.9% in the intervention arm and 50.0% in the control arm. Antihypertensives, oral antidiabetics, and antihyperlipidemics were the commonly prescribed index drugs in both arms. There was no difference in the percentage of new prescriptions in the two arms (44.1% in the intervention arm vs. 41.2% in the control arm). The average number of chronic drugs used by each patient was 3.1 ± 1.4 and 3.0 ± 1.3 in the intervention and control arms, respectively. The baseline details regarding the participating pharmacists were comparable in the two arms. There were 16 pharmacists in total, with eight in each arm. The average age was 31.8 ± 5.6 years in the intervention arm and 32.4 ± 6.1 years in the control arm. The average duration of experience in

their field was 7.1 ± 4.3 and 7.8 ± 4.9 years, respectively. Prior use of AI tools was reported by 37.5% of pharmacists in the intervention arm and 25.0% in the control arm, while self-rated digital confidence was comparable between groups. Pharmacy-level characteristics, including region, pharmacy type, and prescription volume, were also similar across arms (Table 1).

Table 1. Baseline characteristics of pharmacies and patients by study arm

Characteristic	Intervention arm	Control arm	Total
Pharmacies, n	8	8	16
Independent pharmacy, n (%)	5 (62.5)	5 (62.5)	10 (62.5)
Chain pharmacy, n (%)	3 (37.5)	3 (37.5)	6 (37.5)
Mean daily prescription volume, mean ± SD	84.6 ± 18.2	81.9 ± 17.5	83.3 ± 17.7
Internet/device available for intervention, n (%)	8 (100)	8 (100)	16 (100)
Pharmacists, n	8	8	16
Age, years, mean ± SD	31.8 ± 5.6	32.4 ± 6.1	32.1 ± 5.8
Female, n (%)	5 (62.5)	4 (50.0)	9 (56.3)
Years of professional experience, mean ± SD	7.1 ± 4.3	7.8 ± 4.9	7.4 ± 4.5
Bachelor of Pharmacy, n (%)	6 (75.0)	6 (75.0)	12 (75.0)
PharmD or higher, n (%)	2 (25.0)	2 (25.0)	4 (25.0)
Prior use of AI tools, n (%)	3 (37.5)	2 (25.0)	5 (31.3)
Self-rated digital confidence (1–5), mean ± SD	3.8 ± 0.8	3.6 ± 0.9	3.7 ± 0.8
Patients, n	68	68	136
Age, years, mean ± SD	56.8 ± 12.4	55.9 ± 11.8	56.4 ± 12.1
Female, n (%)	36 (52.9)	34 (50.0)	70 (51.5)
New prescription, n (%)	30 (44.1)	28 (41.2)	58 (42.6)
Refill prescription, n (%)	38 (55.9)	40 (58.8)	78 (57.4)
Number of chronic medicines, mean ± SD	3.1 ± 1.4	3.0 ± 1.3	3.0 ± 1.3

Antihypertensives, n (%)	24 (35.3)	23 (33.8)	47 (34.6)
Oral antidiabetics, n (%)	18 (26.5)	19 (27.9)	37 (27.2)
Lipid-lowering agents, n (%)	14 (20.6)	13 (19.1)	27 (19.9)
Anticoagulants, n (%)	6 (8.8)	7 (10.3)	13 (9.6)
Inhaled maintenance therapy, n (%)	6 (8.8)	6 (8.8)	12 (8.8)

Co-primary outcomes

The intervention arm demonstrated higher correctness and completeness scores than the control arm. The mean correctness and completeness scores were 82.4 ± 11.6 and 85.7 ± 10.9 in the pharmacist-supervised AI-assisted counselling group compared with 64.9 ± 14.8 and 68.3 ± 15.1 in the usual-counselling group, respectively (Table 2). This is evidenced by the 16.8 points in correctness and 15.9 points in completeness, where the AI-assisted group consistently covered domains that were neglected in usual care.

Table 2. Co-primary outcomes by study arm

Outcome (0–100 scale)	Intervention arm (n=68)	Control arm (n=68)	Adjusted Mean Difference (95% CI)*	P-value
Correctness score	82.4 ± 11.6	64.9 ± 14.8	16.8 (9.8 to 23.9)	< 0.001
Completeness score	85.7 ± 10.9	68.3 ± 15.1	15.9 (8.7 to 23.1)	< 0.001

**Adjusted for pharmacy-level clustering and covariates (age, sex, and new vs refill status).*

Comparison of counseling deficiencies and clinical severity

As shown in Table 3, the qualitative sub-analysis of the 136 counseling encounters revealed clinical risk profile differed significantly between the two study arms. Regarding completeness, the control group was primarily characterized by high-to-moderate severity, with 82.4% (n=56) of patients receiving no instructions on how to manage a missed dose. In contrast, the intervention arm's main deficiency was low-severity (Alert Compliance Bias: 4.4%, n=3), where strict adherence to the AI-generated domains occasionally led the pharmacist to overlook patient-reported context.

Regarding correctness, the control group was primarily characterized by high severity with 11.8% (n=8) of encounters involving complex drug-drug interactions, where pharmacists relied on potentially outdated clinical recall. In the intervention arm, the AI initially produced four high-severity inaccuracies (5.9%). The remaining deficiencies in the intervention arm were of low clinical severity.

Table 3. Qualitative analysis of counseling deficiencies by category and clinical severity (N=136 encounters)

Primary Domain	Study Arm	Deficiency Type	Frequency (n, %)	Clinical Severity	Outcome	Clinical Scenario Example
Completeness	Control	Omission	56 (82.4%)	Moderate	Reached Patient	During Lisinopril/ Valsartan dispensing, the pharmacist correctly explained the daily timing but failed to provide a backup plan for a missed dose, leaving the patient vulnerable to therapeutic fluctuations.
Completeness	Control	Omission	10 (14.7%)	High	Reached Patient	For a patient starting an oral anticoagulant (Warfarin or Apixaban), the pharmacist provided instructions on dosage but failed to mention signs of internal bleeding (e.g., dark stools), representing a major safety gap that reached the patient.
Completeness	Intervention	Alert Compliance Bias	3 (4.4%)	Low	Corrected by Pharmacist	For a patient receiving Atorvastatin, the pharmacist was so focused on completing the AI-generated domains that they dismissed the patient's spontaneous comment about starting a grapefruit-heavy diet, leading to a loss of patient-specific dietary counseling.
Correctness	Control	Interaction Error	8 (11.8%)	High	Reached Patient	When dispensing Simvastatin to a patient already on Clarithromycin, the pharmacist failed to identify the major drug-drug interaction risk. This high-severity clinical error reached the patient because the spontaneous counseling workflow lacked a secondary verification barrier.

Correctness	Intervention	Near Miss	4 (5.9%)	High	Intercepted by Pharmacist	AI suggested Gliclazide was strictly contraindicated in a patient with a minor sulfonamide antibiotic allergy (a common clinical misconception).
Correctness	Intervention	Technical Wording	7 (10.3%)	Low	Edited by Pharmacist	For Amitriptyline, the AI used the term Xerostomia instead of Dry mouth
Correctness	Intervention	Local Context	4 (5.9%)	Low	Edited by Pharmacist	AI suggested Entresto (Sacubitril/Valsartan) as a first-line agent for hypertension in patients with heart failure—a suggestion that was clinically sound but ignored local insurance formulary restrictions and costs.

Secondary outcomes

At 30 days, adherence scores were higher in the intervention arm than in the control arm (Table 4). The mean GMAS score was 29.1 ± 3.8 in the intervention group and 26.8 ± 4.5 in the control group. In adjusted mixed-effects analysis, the intervention was associated with an adjusted mean difference of 2.0 points (95% CI: 0.6 to 3.4; p = 0.006).

When adherence was categorized, 73.0% of participants in the intervention group met the predefined threshold for good adherence compared with 57.1% in the control group. The adjusted odds of good adherence were higher in the intervention arm (adjusted OR 1.96, 95% CI: 1.04 to 3.70; p = 0.037).

Patients in the intervention arm achieved better immediate understanding scores after the counselling encounter. The mean understanding score (range 0–4) was 3.5 ± 0.7 in the intervention group and 2.8 ± 0.9 in the control group. The adjusted mean difference was 0.68 points (95% CI: 0.34 to 1.02; p < 0.001).

Patient satisfaction was also higher among participants who received pharmacist-supervised AI-assisted counselling. The mean satisfaction score (possible range 5–25) was 22.8 ± 2.4 in the intervention arm and 20.6 ± 3.1 in the control arm. Adjusted analysis showed a mean difference of 2.1 points (95% CI: 0.9 to 3.3; p = 0.001).

Mean face-to-face counselling time was longer in the intervention group (4.7 ± 1.1 minutes) than in the control group (3.6 ± 0.9 minutes). The adjusted mean difference was 1.0 minute (95% CI: 0.5 to 1.5; p < 0.001). In the intervention arm, the mean additional AI-preparation time before verbal delivery was 1.2 ± 0.4 minutes.

Table 4. Secondary outcomes by study arm

Outcome	Intervention arm	Control arm	Adjusted Effect Estimate (95% CI)*	P-value
30-day Adherence (GMAS), mean ± SD	29.1 ± 3.8	26.8 ± 4.5	2.0 (0.6 to 3.4)[1]	0.006
Good adherence (threshold), n/N (%)	46/63 (73.0)	36/63 (57.1)	1.96 (1.04 to 3.70)[2]	0.037

Immediate Understanding (0–4), mean ± SD	3.5 ± 0.7	2.8 ± 0.9	0.68 (0.34 to 1.02)[1]	< 0.001
Patient Satisfaction (5–25), mean ± SD	22.8 ± 2.4	20.6 ± 3.1	2.1 (0.9 to 3.3)[1]	0.001
Face-to-face counselling time, min, mean ± SD	4.7 ± 1.1	3.6 ± 0.9	1.0 (0.5 to 1.5)[1]	< 0.001
AI preparation time, min, mean ± SD	1.2 ± 0.4	—	—	—

**Adjusted for pharmacy-level clustering and baseline covariates (age, sex, and prescription status). [1]Adjusted Mean Difference; [2]Adjusted Odds Ratio.*

Data regarding the pharmacist-supervised workflow revealed that the AI output required frequent professional intervention. Among the 68 intervention-arm encounters, AI-generated content was fully accepted without modification in 16 encounters (23.5%), edited before delivery in 48 encounters (70.6%), and rejected and not used in 4 encounters (5.9%). The most frequently recorded AI-related issues were omitted counselling points, overly technical wording, and incomplete missed-dose advice (Table 5).

Table 5. AI process and safety outcomes in the intervention arm (N=68)

Process/safety variable	n (%)
Total intervention encounters	68 (100)
AI output accepted without modification	16 (23.5)
AI output edited before delivery	48 (70.6)
AI output rejected and not used	4 (5.9)
Any AI-related discrepancy identified	22 (32.4)
Omitted counselling point	9 (13.2)
Overly technical wording	7 (10.3)
Incomplete missed-dose advice	4 (5.9)
Unsupported statement	2 (2.9)
Serious patient safety incidents	0 (0.0)

Inter-rater reliability and sensitivity analyses

Among the double-scored counselling transcripts, inter-rater agreement was high. Weighted kappa values ranged from 0.79 to 0.88 across individual counselling domains, and the intraclass correlation coefficient for total correctness and completeness scores exceeded 0.90. Sensitivity analyses excluding participants with missing 30-day follow-up data yielded effect estimates similar in direction and magnitude to the main analyses. A per-protocol analysis also produced findings consistent with the primary intention-to-treat results.

Discussion

In this cluster randomized controlled trial, pharmacist-supervised generative AI-assisted medication counselling improved both co-primary outcomes, with higher correctness and completeness scores than usual counselling. The intervention was also associated with better 30-day medication adherence, immediate patient understanding, and patient satisfaction, although it required modestly longer face-to-face counselling time. At the same time, most AI-generated outputs required pharmacist editing and a small proportion were rejected outright, while no serious patient safety incidents were observed. Overall, these findings suggest that the value of the intervention lay not in replacing pharmacist judgement, but in supporting pharmacists to deliver more structured, accurate, and comprehensive counselling while maintaining professional oversight.

The improvements in counselling correctness and completeness are consistent with previous Jordanian research showing that counselling performance in community pharmacies can improve when pharmacists are supported by structured educational or assessment tools. In Jordan, Elayeh *et al.* (2019) showed that workshop-based education improved inhaler counselling performance in community pharmacies, while Hammad *et al.* (2022) demonstrated that information provision in simulated-patient encounters improved after prompting and structured assessment. The present study extended that literature by suggesting that pharmacist-supervised generative AI may function as a practical, encounter-level prompting tool that helps pharmacists systematically cover key counselling domains and reduce omissions during real patient care, rather than relying entirely on memory or habitual practice.

Analysis of the encounters showed that the two study arms differed not only in counselling scores but also in the severity and nature of counselling deficiencies. In the control arm, the most important problems were moderate- and high-severity omissions that reached patients, especially failure to provide missed-dose advice and key safety warnings. This represents a major clinical vulnerability in chronic disease management, where therapeutic consistency is paramount. This aligns with prior Jordanian studies, which identified that spontaneous counseling often neglects critical safety information (Abdel-Qader *et al.*, 2021; Hammad *et al.*, 2022). In contrast, the intervention arm was dominated by low-severity wording issues and a small number of high-severity near misses that were intercepted by pharmacists before patient exposure, further reinforces the necessity of the human-in-the-loop model. This distinction is critical: because the AI-assisted workflow incorporated a mandatory verification phase, pharmacists were able to intercept and correct AI-generated hallucinations before information reached the patient. This aligns with established medication-safety frameworks, where near misses represent errors that are detected and mitigated before causing harm, often reflecting the presence of effective safety barriers within the system (Abdel-Qader *et al.*, 2021). This suggests that pharmacist-supervised generative AI improved counselling not only by increasing correctness and completeness, but also by reducing the likelihood that clinically important information gaps would reach the patient.

These results on adherence and knowledge were also quite reasonable from both behavioral and clinical standpoints. According to a systematic review and meta-analysis conducted by Kelly *et al.* (2023), counseling by pharmacists can help in improving adherence to medications; hence, it can be assumed that the results obtained for this measure are clinically sound. Similarly, the increased levels of medication knowledge observed in the counseling group reflected the very purpose of the teach-back approach, which aimed to find out if patients have understood key facts about their medications and can state them independently. Recent research on pharmaceutical counseling also suggests employing teach-back as a component of counseling for increasing patients' knowledge. Consequently, the higher level of counseling accuracy and completeness could have resulted in increased knowledge first and subsequently adherence (O'Mahony *et al.*, 2023).

The study provided a clear link between the 16.8 points in correctness and 15.9 points in completeness and the 1.96 adjusted odds of good adherence. By ensuring that high-severity safety points (like anticoagulant bleeding signs or antidiabetic hypoglycemia warnings) were consistently

addressed, the intervention likely reduced patient anxiety and increased self-efficacy. When patients leave the pharmacy with a complete safety map for their medication, the likelihood of adherence-related complications is logically reduced.

The results on the process and safety were arguably among the most critical outcomes of the study. In terms of the latter, only a small percentage of AI-generated outputs was adopted unaltered, while the vast majority of them had to be edited by pharmacists and some of them even discarded prior to patient use. This trend supports the assertion that AI must be employed as a complementary tool, not as an independent source for counseling, a perspective that aligns well with the FIP's recommendations on the appropriate usage of AI in pharmacy practice (FIP, 2025). For instance, in their investigation of ChatGPT® and Gemini®'s ability to respond to general drug inquiries, Pornwattanakavee *et al.* (2025) observed that both LLMs could handle numerous drug queries but had inconsistent performance and reliability with regards to accuracy. From a Jordanian perspective, these results are highly relevant due to the research conducted which demonstrated that dispensing errors and managing drug interactions pose safety concerns for pharmacists (Abdel-Qader *et al.*, 2021, 2025).

This intervention has been linked to increased encounter times (about 2.3 minutes added for each patient). The time spent is a major concern in the fast-paced working conditions in Jordanian community pharmacies; nevertheless, the substantially increased patient satisfaction indicated that patients appreciated the longer discussion. More studies should be conducted in order to investigate integrated enterprise or artificial intelligence systems that can help lower the time spent on preparations.

There are several strengths of the current study. Firstly, the design of cluster randomization was suitable for the study since the intervention was carried out at the workflow level of the pharmacists, and there was likely to be contamination if an individual randomization were used. Secondly, the intervention was standardized using pre-set prompts, drug-specific guidelines, and training of the pharmacists, which ensured that the intervention was reproducible. Thirdly, the study did not focus on assessing the performance of the AI system alone; instead, it focused on the human-controlled workflow of using AI, which is more realistic since AI systems are most likely controlled by humans in pharmacy practice.

However, several limitations need to be highlighted as well. First, the research was conducted among Jordanian community pharmacies targeting certain categories of chronic medications, which might affect the generalizability of the findings across different settings, acute-care interactions, and other therapeutic areas. Second, a single workflow involving one model of AI was used for the intervention, implying that the conclusions drawn from the study cannot be extrapolated to any other model or new versions. Third, adherence was measured during a relatively short period of 30 days that allowed to capture patients' short-term behavior but not their long-term adherence. Fourth, certain secondary outcomes including patient understanding and satisfaction were measured with the help of instruments developed for this research as opposed to the standardized ones that have been used for many years. Finally, pharmacists could not be blinded to treatment allocation, while blinding to transcripts and masking statistical analyses could not eliminate all forms of potential bias.

Conclusion

In conclusion, the current findings suggest that pharmacist-supervised generative AI-assisted medication counselling can improve the accuracy, completeness, and patient-centered effectiveness of counselling in Jordanian community pharmacies, while remaining dependent on pharmacist review for safety. The study therefore supports a model of human-AI collaboration in pharmacy practice in which AI is used to structure and strengthen counselling, but pharmacists remain responsible for clinical verification, contextual judgement, and safe delivery of information. Future studies should examine longer follow-up periods, broader therapeutic areas, and comparative workflows using different models under the same reference-constrained conditions.

References

Abdel-Qader, D. H., Al Meslamani, A. Z., Lewis, P. J., & Hamadi, S. (2021). Incidence, nature, severity, and causes of dispensing errors in community pharmacies in Jordan. International journal of clinical pharmacy, 43(1), 165-173. https://doi.org/10.1007/s11096-020-01126-w

Abdel-Qader, D. H., et al. (2024). A comprehensive analysis of public satisfaction: Community pharmacists' pandemic preparedness in Jordan. Journal of Applied Pharmaceutical Science, 14(8), 160-168.

Abdel-Qader, D. H., et al. (2025). Drug–Drug interaction management among pharmacists in Jordan: A national comparative survey. Pharmacy, 137. https://doi.org/10.3390/pharmacy13050137

Abu Hammour, K., et al. (2023). ChatGPT in pharmacy practice: A cross-sectional exploration of Jordanian pharmacists' perception, practice, and concerns. Journal of Pharmaceutical Policy and Practice, 16(1), 115.

Ali, S., Shimels, T., & Bilal, A. I. (2019). Assessment of patient counseling on dispensing of medicines in outpatient pharmacy of Tikur-Anbessa Specialized Hospital, Ethiopia. Ethiopian journal of health sciences, 29(6), 727.

Campbell, M. K., et al. (2012). Consort 2010 statement: Extension to cluster randomised trials. BMJ, 345.

Chan, A.-W., et al. (2015). SPIRIT 2013 Statement: Defining standard protocol items for clinical trials. Revista Panamericana de Salud Pública, 38, 506-514.

Elayeh, E. R., et al. (2019). Use of secret simulated patient followed by workshop based education to assess and improve inhaler counseling in community pharmacy in Jordan. Pharmacy Practice (Granada), 17(4).

Fattah, F. H., et al. (2025). Comparative analysis of ChatGPT and Gemini (Bard) in medical inquiry: A scoping review. Frontiers in digital health, 7, 1482712.

FIP, I. P. F. (2021). Medication review and medicines use review: A toolkit for pharmacists

Colophon.

FIP, I. P. F. (2025). An artificial intelligence toolkit for pharmacy: An introduction and resource guide for pharmacists. (March).

Hammad, E. A., et al. (2022). Feasibility of using simulated patients for onsite structured practice feedback in Jordanian community pharmacy settings. Saudi Pharmaceutical Journal, 30(3), 225-229.

Hasan, H. E., et al. (2024). Ethical considerations and concerns in the implementation of AI in pharmacy practice: A cross-sectional study. BMC Medical Ethics, 25(1), 55.

Hemming, K., et al. (2017). How to design efficient cluster randomised trials. BMJ, 358.

Huang, X., et al. (2024). Evaluating the performance of ChatGPT in clinical pharmacy: A comparative study of ChatGPT and clinical pharmacists. British journal of clinical pharmacology, 90(1), 232-238.

Kelly, W. N., et al. (2023). Association of pharmacist intervention counseling with medication adherence and quality of life: A systematic review and meta-analysis of randomized trials. Journal of the American Pharmacists Association: JAPHA, 63(4), 1095-1105. https://doi.org/10.1016/j.japh.2023.04.024

Li, L., et al. (2025). Comparative analysis of generative artificial intelligence systems in solving clinical pharmacy problems: Mixed methods study. JMIR Medical Informatics, 13, e76128.

Liu, X., et al. (2020). Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. The Lancet Digital Health, 2(10), e537-e548.

Naqvi, A. A., et al. (2018). Development and validation of a novel General Medication Adherence Scale (GMAS) for chronic illness patients in Pakistan. Frontiers in pharmacology, 9, 1124.

Naser, A. Y., & Sbeat, B. S. A. (2022). Satisfaction with community pharmacies services in Jordan: A cross-sectional study. Saudi Pharmaceutical Journal, 30(11), 1646-1651.

O'Mahony, E., et al. (2023). Development and evaluation of pharmacist-provided teach-back medication counselling at hospital discharge. International Journal of Clinical Pharmacy, 45(3), 698-711.

Pornwattanakavee, S., et al. (2025). Effectiveness of ChatGPT, Google Gemini, and Microsoft Copilot in answering Thai drug information queries: Cross-sectional study. JMIR AI, 4, e79751.

Rivera, S. C., et al. (2020). Guidelines for clinical trial protocols for interventions involving artificial intelligence: The SPIRIT-AI extension. The Lancet Digital Health, 2(10), e549-e560.

De Rosis, S., Ferrè, F., & Pennucci, F. (2022). Including patient-reported measures in performance evaluation systems: Patient contribution in assessing and improving the healthcare systems. The International Journal of Health Planning and Management, 37, 144-165.

Rutterford, C., Copas, A., & Eldridge, S. (2015). Methods for sample size determination in cluster randomized trials. International journal of epidemiology, 44(3), 1051-1067.

Sallam, M., et al. (2024). Human versus artificial intelligence: ChatGPT-4 outperforming Bing, bard, ChatGPT-3.5 and humans in clinical chemistry multiple-choice questions. Advances in Medical Education and Practice, 857-871.

Surur, A. S., et al. (2017). Self-reported and actual involvement of community pharmacists in patient counseling: A cross-sectional and simulated patient study in Gondar, Ethiopia. Pharmacy Practice (Granada), 15(1).

Vickerstaff, V., Omar, R. Z., & Ambler, G. (2019). Methods to adjust for multiple comparisons in the analysis and sample size calculation of randomised controlled trials with multiple primary outcomes. BMC medical research methodology, 19(1), 129.

Wazaify, M., et al. (2019). Assessing insomnia management in community pharmacy setting in Jordan: A simulated patient approach. PLOS ONE, 14(12), e0226076.

WHO EMRO (World Health Organization Regional Office for the Eastern Mediterranean). (2024). Jordan implements WHO HEARTS in primary health care to strengthen management of cardiovascular diseases and related risks. https://www.emro.who.int/jor/jordan-infocus/jordan-implements-who-hearts-in-primary-health-care-to-strengthen-management-of-cardiovascular-diseases-and-related-risks.html

Supporting Information Captions

S1 File. Anonymized Dataset

S2 File. Appendices