Review Article

What Work Has Artificial Intelligence Actually Taken on in Pharmacy? A Scoping Review of Dispensing, Clinical Review, and Medicines Information

Krzysztof Wiśniewski[1]*, Małgorzata Nowak[1], Tomasz Adamczyk[2]

[1]Department of Pharmacy AI Scoping and Review, Faculty of Pharmacy, Warsaw University of Life Sciences, Warsaw, Poland. [2]Department of Clinical Pharmacy and AI Task Analysis, Faculty of Pharmacy, Jagiellonian University, Krakow, Poland.

Abstract

Claims about artificial intelligence in pharmacy frequently combine conceptual proposals, retrospective models, and operational systems, making it difficult to determine what work has actually been assigned to AI and where pharmacist judgment remains necessary. To map AI-supported tasks across dispensing, prescription and clinical review, medicines information, patient-facing services, and administrative pharmacy work while separating technical performance from workflow usefulness, safety, and autonomy. Peer-reviewed Q1 journal articles published from 2017 through 2025 were eligible when they evaluated an AI or hybrid AI component performing or supporting an identifiable pharmacy task and provided extractable evidence on setting, data, reference standard, human involvement, validation, comparator, or outcome. Searches covered MEDLINE/PubMed, Embase, Scopus, Web of Science, CINAHL, International Pharmaceutical Abstracts, IEEE Xplore, and supplementary citation and publisher searching. Evidence was charted by task, AI method, setting, data source, reference standard, pharmacist role, validation level, comparator, outcome, and autonomy, then synthesized using a review-derived task–maturity–oversight framework.

Searches identified 1,266 records. After removal of 472 duplicates, 794 records were screened; 153 reports were sought, 149 full texts were assessed, and 22 studies were included. Evidence concentrated on prescription prioritization, intervention prediction, anomaly detection, pill recognition, targeted patient services, and medicines-information answering. Pharmacists generally retained contextual review, authorization, communication, or final action. AI has primarily taken on bounded tasks that rank, flag, recognize, retrieve, or draft. The evidence does not establish autonomous completion of comprehensive pharmacy work or permit technical accuracy to be treated as proof of safety, benefit, or deployment readiness.

Keywords: Scoping review, Digital pharmacy, Artificial intelligence, Pharmacy practice, Medication safety, Evidence synthesis

Introduction

Artificial intelligence in pharmacy encompasses machine learning, deep learning, natural-language processing, large language models, computer vision, anomaly detection, and hybrid decision-support systems. These methods perform different functions and require different forms of evaluation. Predicting which prescription may require attention, identifying a tablet image, drafting a medicines-information response, and authorizing a medication decision are therefore not equivalent forms of work [1].

Existing reviews suggest that direct evidence remains smaller and more uneven than the visibility of pharmacy-AI claims implies. Evaluations cluster around prescription screening, medication-order prioritization, dispensing-related recognition, patient identification, and information support, while routine-practice studies remain comparatively limited [2, 3]. Publication frequency also does not establish maturity: numerous retrospective models may coexist with little prospective testing, independent validation, sustained implementation, or evidence of patient benefit.

Medication-management AI is also studied in primary care, medical, nursing, informatics, and population-health settings without a clearly defined pharmacy actor [4]. Classifying every medication-related algorithm as pharmacy AI would inflate the evidence base and conceal who created labels,

Address for correspondence: Krzysztof Wiśniewski, Department of Pharmacy AI Scoping and Review, Faculty of Pharmacy, Warsaw University of Life Sciences, Warsaw, Poland.
krzysztof.wisniewski@sggw.edu.pl
Received: 18 July 2025; **Accepted:** 14 November 2025

How to cite this article: Wiśniewski K, Nowak M, Adamczyk T. What Work Has Artificial Intelligence Actually Taken On in Pharmacy? A Scoping Review of Dispensing, Clinical Review, and Medicines Information. Arch Pharm Pract. 2025;16(4):34-42. https://doi.org/10.51847/KoMXgftABY

reviewed outputs, authorized action, communicated recommendations, and retained professional responsibility. Conversely, restricting pharmacy work to product supply would omit medication review, patient outreach, population-health screening, and medicines-information services.

This review therefore asked what work AI actually performed, prioritized, supported, or evaluated in dispensing, prescription and clinical review, medicines information, patient-facing services, and administrative workflows during 2017–2025. It distinguished proposed systems from retrospective testing, prospective evaluation, operational use, and independent validation; mapped retained human involvement; and separated model output from medication action. It did not ask whether AI should replace pharmacists or whether technical performance established universal safety, effectiveness, or readiness for deployment.

Review Design, Questions, and Protocol

Review Design and Protocol

A scoping-review design was selected because the evidence encompassed heterogeneous tasks, technologies, settings, study designs, reference standards, and outcomes. Reporting followed PRISMA-ScR [5], and conduct was informed by updated JBI methodological guidance [6]. Scoping methodology was appropriate for clarifying definitions, mapping evidence distribution, and identifying gaps rather than estimating a pooled intervention effect [7]. Extraction and presentation were organized around prespecified review questions [8].

Review Questions Using PCC

The population included pharmacists, pharmacy technicians, pharmacy services, medication orders and records, patients receiving pharmacist-mediated services, and medicines-information questions. The concept was an AI or hybrid system that classified, predicted, ranked, identified, generated, verified, or recommended an output relevant to pharmacy work. Contexts included community, hospital, ambulatory, primary-care, health-system, dispensing, medication-review, medicines-information, and administrative settings.

The primary question was: What pharmacy work has an AI system actually performed or supported? Secondary questions examined the AI method, task, setting, data source, reference standard, human role, validation level, comparator, outcome, autonomy, safety, equity, workflow consequence, and applicability boundary.

Eligibility Criteria

Eligible empirical studies were peer-reviewed Q1 journal articles published from 1 January 2017 through 31 December 2025, had a verifiable DOI, evaluated an AI or hybrid AI component, addressed an identifiable pharmacy task or pharmacist-mediated workflow, and provided extractable evidence relevant to the review questions. Methodological and reporting papers supported explicit methods claims but were not classified as included pharmacy-work studies.

Conference-only papers, unpublished preprints, non-peer-reviewed reports, and studies outside the date window were excluded. Drug discovery, formulation, manufacturing, and general diagnostic AI were excluded unless a pharmacy task and actor were central. Deterministic rules-only systems were excluded unless evaluated as comparators or inseparable elements of a hybrid system. Duplicate reports without unique evidence and papers unable to support a task-level conclusion were also excluded.

Definitions of AI and Pharmacy Work

AI comprised machine learning, deep learning, natural-language processing, large language models, computer vision, anomaly detection, and relevant hybrid systems. Pharmacy work was defined as an activity performed, supervised, interpreted, or acted on by pharmacy personnel or a medicines-information service.

A task was considered "taken on" when the system produced a classification, prediction, ranking, recognition, generated response, verification signal, or recommendation that affected subsequent inspection, communication, or action. Evidence maturity was classified as proposed, retrospectively tested, prospectively evaluated, operationally implemented, or independently validated. Autonomy was coded separately according to whether pharmacist confirmation remained necessary. Clinical usefulness was not inferred from discrimination or accuracy alone.

Eligibility, Definitions, and Information Sources

Databases and Supplementary Sources

The search covered MEDLINE/PubMed, Embase, Scopus, Web of Science Core Collection, CINAHL, International Pharmaceutical Abstracts, and IEEE Xplore for eligible journal articles. Supplementary identification used backward and forward citation searching, publisher records, DOI verification, and related-article searching. Records, reports, and studies were maintained as distinct units [9].

Search Strategy and 2017–2025 Window

Database-specific strategies combined terms for AI methods; pharmacists, pharmacy services, and practice settings; tasks such as prescribing, medication review, intervention, dispensing, pill identification, adherence, anticoagulation, and medicines information; and evaluation terms including validation, workflow, safety, errors, and outcomes. Searches were limited to publications from 2017 through 2025. Full database strategies, dates, platforms, and limits were documented according to PRISMA-S principles [10].

Deduplication and Record Management

Each retrieved record received a unique identifier linked to its source, search date, title, authors, year, DOI, duplicate status, screening decisions, retrieval status, exclusion reason, study-

family identifier, and synthesis role. Deduplication proceeded through exact DOI, normalized DOI, normalized title, title–first-author–year comparison, and manual review of online-first and final versions, corrections, and preprint–journal pairs.

Dual Screening and Conflict Resolution

Two reviewers independently screened titles and abstracts and assessed potentially eligible full texts. Disagreements were resolved through discussion and adjudication. One primary exclusion reason was assigned to each excluded full-text report. AI-specific extraction fields captured intended use, model input, output, human involvement, error handling, workflow integration, and permitted action [11]. Prospective evaluations were additionally examined for prespecified integration, comparator, analysis, and management of unavailable or poor-quality inputs [12].

Search, Selection, Extraction, Appraisal, and Synthesis

Data-Charting Framework

The charting form captured publication year, country, setting, population or unit of analysis, pharmacy task, AI method, data source, reference standard, comparator, pharmacist role, study design, validation level, reported outcome, safety, equity, implementation, and autonomy. Model performance was separated from task performance, workflow consequence, clinical outcome, and patient consequence.

Prediction-model studies were examined using relevant PROBAST domains and TRIPOD+AI reporting concepts [13, 14]. Workflow-facing studies were examined for human–AI interaction and failure handling using DECIDE-AI-informed fields [15]. MI-CLAIM informed extraction of data, model, evaluation, and reproducibility information without being converted into a numerical quality score [16].

Table 1 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for eligibility, definitions, and evidence-extraction framework for the review What Work Has Artificial Intelligence Actually Taken On in Pharmacy? A Scoping Review of Dispensing, Clinical Review, and Medicines Information.

Table 1. Eligibility, definitions, and evidence-extraction framework for the review What Work Has Artificial Intelligence Actually Taken On in Pharmacy? A Scoping Review of Dispensing, Clinical Review, and Medicines Information.

Review element	Operational definition	Eligibility or decision rule	Data field	Rationale	Bias or ambiguity risk	Planned synthesis use
AI system	Machine learning, deep learning, NLP, large language model, computer vision, anomaly detection, or hybrid	Must perform or support an identifiable task	Method and version	Prevents "AI" from replacing task description	Unspecified software or concealed rule logic	Stratification by method and task
Pharmacy work	Activity performed, supervised, interpreted, or acted on by pharmacy personnel	Pharmacy actor, service, or responsibility must be explicit	Actor, setting, workflow point	Separates pharmacy from general medication AI	Medication data may be mistaken for pharmacy work	Mapping of task ownership
Task taken on	Classification, prediction, ranking, recognition, generation, verification, or recommendation	Output must have a defined workflow function	Task, output, subsequent action	Centers demonstrated work	Broad claims may exceed the component studied	Task taxonomy
Evidence maturity	Proposed, retrospective, prospective, operational, or independently validated	Highest level directly demonstrated	Design and validation	Separates publication from maturity	Simulated deployment may appear prospective	Maturity mapping
Human role	Label provider, reviewer, interpreter, authorizer, communicator, or decision-maker	Roles before and after output must be extracted	Oversight, override, authority	Makes human contribution visible	Human work may be underreported	Oversight and autonomy analysis
Reference standard	Pharmacist judgment, intervention, chart review, expert panel, database identity, or outcome	Provenance and adjudication recorded	Label source and adjudication	Defines what performance represents	Historical action may not be ground truth	Interpretation of agreement and error
Validation	Internal, temporal, external, multicenter, prospective, or independent	Development testing is not independent validation	Site, period, sample unit	Supports applicability assessment	Data overlap may inflate performance	Transferability assessment
Outcomes	Model, task, workflow, safety, clinical, patient, equity, or implementation outcome	Outcome levels remain separate	Metric, denominator, uncertainty	Prevents accuracy from becoming benefit	Selective technical reporting	Outcome-level synthesis
Usefulness and autonomy	Meaningful workflow or patient effect; action without pharmacist confirmation	Neither inferred from accuracy alone	Effect, confirmation, authority	Separates performance from decision rights	"Automated" may mean only computational	Demonstrated autonomy

Safety and equity	Harmful output, subgroup performance, representativeness, and access	Extract assessed findings and record nonassessment	Safety and subgroup fields	Identifies consequential evidence gaps	Nonreporting may be mistaken for no risk	Gap and applicability synthesis

Note: Operational definitions and synthesis categories are review-derived unless attributed to a methodological source.

Classification of Pharmacy Tasks

Tasks were classified by their immediate function rather than by broad author labels. Domains comprised dispensing and supply, prescription and clinical review, medicines information, patient-facing services, and administrative or population-health work. Functions included recognition, retrieval, prediction, prioritization, anomaly detection, generation, verification, and recommendation. A separate workflow layer identified whether the output prompted inspection, escalation, communication, authorization, or action.

Synthesis and Visualization

Heterogeneity in tasks, designs, comparators, and outcomes precluded quantitative pooling. Evidence was organized by task, method, setting, reference standard, human role, maturity, outcome, and autonomy. Contradictory findings were retained rather than resolved through vote counting. The task–maturity–oversight framework was treated as an original review-derived synthesis rather than a validated implementation model.

Deviations from Protocol

No deviation altered the review question, publication window, eligibility framework, or evidence classification. Quantitative synthesis was not undertaken because the included studies examined non-equivalent tasks, reference standards, settings, and outcomes.

Search and Selection Results

The searches identified 1,266 records: 214 from MEDLINE/PubMed, 286 from Embase, 331 from Scopus, 198 from Web of Science Core Collection, 71 from CINAHL, 84 from International Pharmaceutical Abstracts, 53 from IEEE Xplore, and 29 from supplementary searching. After removal of 472 duplicates, 794 records underwent title-and-abstract screening. Of these, 641 were excluded and 153 reports were sought for retrieval. Four reports could not be retrieved, leaving 149 full-text reports for eligibility assessment.

A total of 127 full-text reports were excluded: 36 did not examine an identifiable pharmacy-practice task, 14 contained no eligible AI component, 19 were ineligible publication types, nine were outside the 2017–2025 window, three lacked a verifiable DOI, 16 did not meet the journal-quartile criterion, 18 provided insufficient task-level evidence, seven were duplicate reports without unique evidence, and five were non-peer-reviewed. Twenty-two reports representing 22 distinct studies were included. No quantitative synthesis was undertaken.

Figure 1 presents the completed review-selection pathway in a black-and-white PRISMA-ScR flow diagram for evidence selection, documenting identification, deduplication, screening, eligibility, exclusions with reasons, and final inclusion using the reconciled record-accounting values.

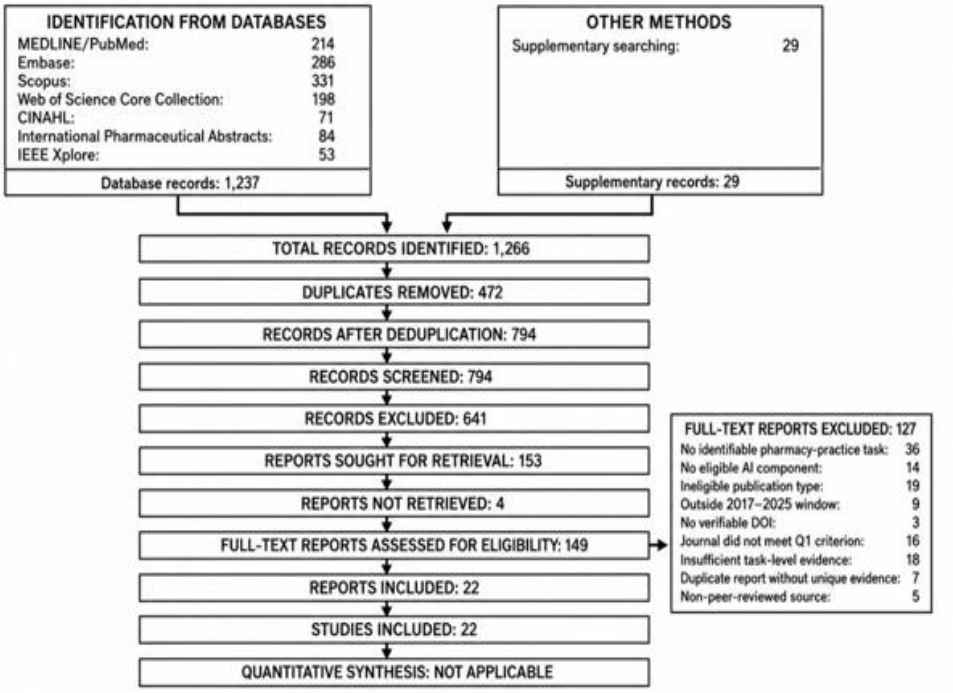

Figure 1. PRISMA-ScR Flow Diagram for Evidence Selection

Publication and Geographic Characteristics

The included evidence represented hospital, ambulatory, primary-care, community-facing, patient-support, and medicines-information contexts. Prescription-risk prioritization was evaluated in hospital medication-review workflows [17], while alert generation was examined through retrospective clinical and cost analysis [18]. Dispensing-related evidence included deep-learning pill identification [19], and operational patient-service evidence included an AI-supported pharmacist adherence program [20]. Most studies remained institution-specific, limiting assumptions about transferability across countries, formularies, prescribing systems, pharmacy roles, and patient populations.

Evidence-Base Characteristics

AI Methods Represented

Supervised machine learning predominated in prescription prioritization and intervention prediction. Deep convolutional networks were used for pill recognition, anomaly-detection methods identified unusual prescriptions or dosing patterns, and large language models were evaluated mainly for medicines-information tasks. Hybrid systems combined learned models with rules or pharmacist review.

Provider-action data were used to predict whether an inpatient medication order would receive pharmacy intervention [21]. Because the outcome represented recorded professional action, it could reflect local workload, documentation, and policy as well as clinical need. Pharmacists' perceptions of atypical medication orders did not always correspond directly with model predictions [22], emphasizing that algorithmic and professional judgments were not interchangeable.

Dispensing and Supply Tasks

Dispensing-related evidence focused mainly on visual recognition. Multicenter testing examined pill-recognition performance across clinical settings and platforms [23]. The demonstrated function was product recognition from an image, not autonomous completion of prescription interpretation, product selection, legal verification, labeling, final checking, supply, counseling, or accountability.

Table 2 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for characteristics and classification of the evidence included in the review What Work Has Artificial Intelligence Actually Taken On in Pharmacy? A Scoping Review of Dispensing, Clinical Review, and Medicines Information.

Table 2. Characteristics and classification of the evidence included in the review What Work Has Artificial Intelligence Actually Taken On in Pharmacy? A Scoping Review of Dispensing, Clinical Review, and Medicines Information.

Evidence category	Study or source design	Population or setting	AI system or intervention	Comparator or reference standard	Outcome or construct	Validation level	Key limitation
Prescription-risk prioritization [17]	Retrospective development and evaluation	Hospital prescriptions	Machine-learning decision support	Pharmacist-identified medication-error risk	Prioritization for review	Local retrospective evaluation	Predicted risk does not establish actual harm or safe exclusion
Prescribing-alert identification [18]	Retrospective clinical and cost analysis	Health-system medication orders	Machine-learning alert system	Chart review and existing decision support	Potential prescribing-error detection	Local retrospective evaluation	Prevented harm may be estimated rather than observed
Pill-image identification [19]	Algorithm-development study	Pill-image dataset	Deep convolutional network	Reference pill identity	Image classification	Internal testing	Controlled recognition is not complete dispensing
Pharmacist-led adherence support [20]	Operational program evaluation	Patients targeted for adherence intervention	AI-supported case targeting plus pharmacist outreach	Program or usual workflow	Adherence-related outcome	Operational evaluation	AI contribution cannot be isolated from pharmacist work
Intervention prediction [21]	Retrospective prediction study	Inpatient medication orders	Machine learning using provider actions	Historical pharmacy intervention	Probability of intervention	Internal validation	Historical action is not an independent harm standard
Atypical-order assessment [22]	Human–AI comparative evaluation	Hospital pharmacists and medication orders	Machine-learning atypicality model	Pharmacist perception	Agreement and divergence	Local comparative evaluation	Professional perception is context-dependent
Multicenter pill recognition [23]	Multicenter real-world evaluation	Clinical images across platforms	Deep-learning pill recognition	Verified pill identity	Recognition across sites	Multicenter validation	Does not establish dispensing autonomy or safety

Note: Evidence categories describe the immediate function evaluated and do not imply equivalent maturity, clinical consequence, or autonomy.

Prescription and Clinical-Review Tasks

Prescription-review systems mainly determined which orders should receive attention. A hybrid system was evaluated for ruling selected prescriptions out of comprehensive medication review [24]. Its practical significance depended not only on model performance but also on the consequences of incorrectly deprioritizing an order and on the safeguards retained in the local workflow.

Canceled or self-intercepted medication orders were also used as signals for potential ordering error [25]. This enabled learning from routinely collected behavior, but cancellation remained an imperfect proxy because orders may be stopped for administrative, formulary, timing, or preference-related reasons.

A hospital model predicted pharmaceutical interventions from prescription-review data [26]. Such systems may organize workload by directing attention to cases more likely to prompt intervention, but they do not determine whether an intervention is correct, accepted, clinically consequential, or transferable to another institution.

Medicines-Information Tasks

Medicines-information applications generated or retrieved explanatory content rather than ranking structured medication orders. Evaluations examined correctness, completeness, reference quality, clinical applicability, and potential risk. Interpretation required attention to model version, prompt construction, language, observation date, source transparency, and the distinction between fluent output and accountable advice.

Patient-Facing and Administrative Tasks

An algorithmic population-health intervention identified patients with atrial fibrillation who might require anticoagulation review [27]. Pharmacist assessment excluded many algorithm-identified cases after considering clinical circumstances, and the intervention did not increase anticoagulant initiation. The system therefore demonstrated worklist generation while also showing why algorithmic case finding could not be equated with a final recommendation.

Deep-learning pill identification could assist patients, caregivers, or pharmacy personnel in recognizing tablets against a reference database [28]. Recognition did not establish indication, dose, authenticity, ownership, suitability, or safe administration.

Large-language-model responses were also compared with drug-information answers produced or assessed by clinical pharmacists [29]. The system performed a drafting or answering function, but pharmacists retained responsibility for clarification, evidence checking, individualized interpretation, and communication of uncertainty.

Level of Autonomy Actually Studied

Most systems had low or conditional autonomy. They flagged, ranked, identified, or drafted; pharmacists or other clinicians reviewed outputs before consequential action. Pharmacists also frequently supplied labels, defined intervention outcomes, adjudicated errors, or served as comparators. Human contribution therefore occurred both upstream and downstream of the model.

No included evidence established autonomous completion of comprehensive medication review, dispensing, medicines-information practice, or patient-specific therapeutic decision-making. “Automated” generally described computational execution rather than independence from professional confirmation or authority.

Evaluation Settings and Outcome Types

One medicines-information evaluation compared ChatGPT responses with answers produced by a drug-information center [30]. This offered an accountable professional comparator but did not establish that one center represented every possible expert response or institutional standard.

Search-engine-integrated chatbots were evaluated for the quality and risks of drug information [31]. Their performance depended on the query, platform, access date, retrieval behavior, and underlying sources. A real-world analysis of ChatGPT likewise identified variable performance and potential risks in medicines-information use [32]. These findings concerned particular systems and observation periods rather than a stable property of all language models.

Across task domains, the most common outcomes were accuracy, discrimination, sensitivity, specificity, agreement, and intervention yield. Fewer studies assessed pharmacist workload, review time, override, recommendation acceptance, prescribing change, adherence, patient outcome, safety events, subgroup performance, equity, maintenance, or implementation burden.

Evidence Gaps

Independent external validation remained uncommon. Many models learned from single-site historical actions, embedding local documentation and intervention practices into the target. Calibration, temporal drift, subgroup performance, missing data, alert fatigue, override behavior, and the consequences of false reassurance were inconsistently assessed.

Evidence was sparse for community-pharmacy counseling, medicines reconciliation, transitions of care, controlled-drug governance, supply disruption, pharmacovigilance, long-term patient outcomes, and equity. Nonassessment of subgroup performance could not be interpreted as evidence that performance or access was equitable.

Discussion: Principal Findings and Interpretation

Principal Findings

AI had taken on bounded components of pharmacy work rather than complete professional roles. It classified pill images, identified unusual orders, estimated intervention probability, prioritized review, generated medicines-information responses, and constructed patient worklists. Renal-function data were used to identify patients at risk of inappropriate dosing [33], pharmacy inquiry records supported prediction of dose-related questions [34], and unsupervised methods flagged potential overdose or underdose prescriptions [35].

These functions shared a common structure: the system narrowed, ordered, recognized, or drafted information, after which a pharmacist determined whether the output was relevant and what action was justified. The AI contribution could be operationally meaningful without becoming equivalent to accountable medication decision-making.

Difference between Automation Claims and Demonstrated Work

Broad descriptions such as "automated medication review" concealed the narrower tasks actually evaluated. Across development and daily-practice studies, AI was used principally to prioritize or filter prescriptions for pharmacist review rather than to replace the pharmacist's final clinical judgment [17, 24, 26]. Pill recognition represented one component of dispensing, and response generation represented one component of medicines-information practice.

Technical performance and clinical usefulness were distinct evidential layers. High discrimination could indicate that historical pharmacist actions were predictable without showing that the model improved review quality, prevented harm, reduced inequity, or remained safe after implementation.

Figure 2 presents the original conceptual synthesis for evidence map of AI tasks across dispensing, clinical review, and medicines information, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

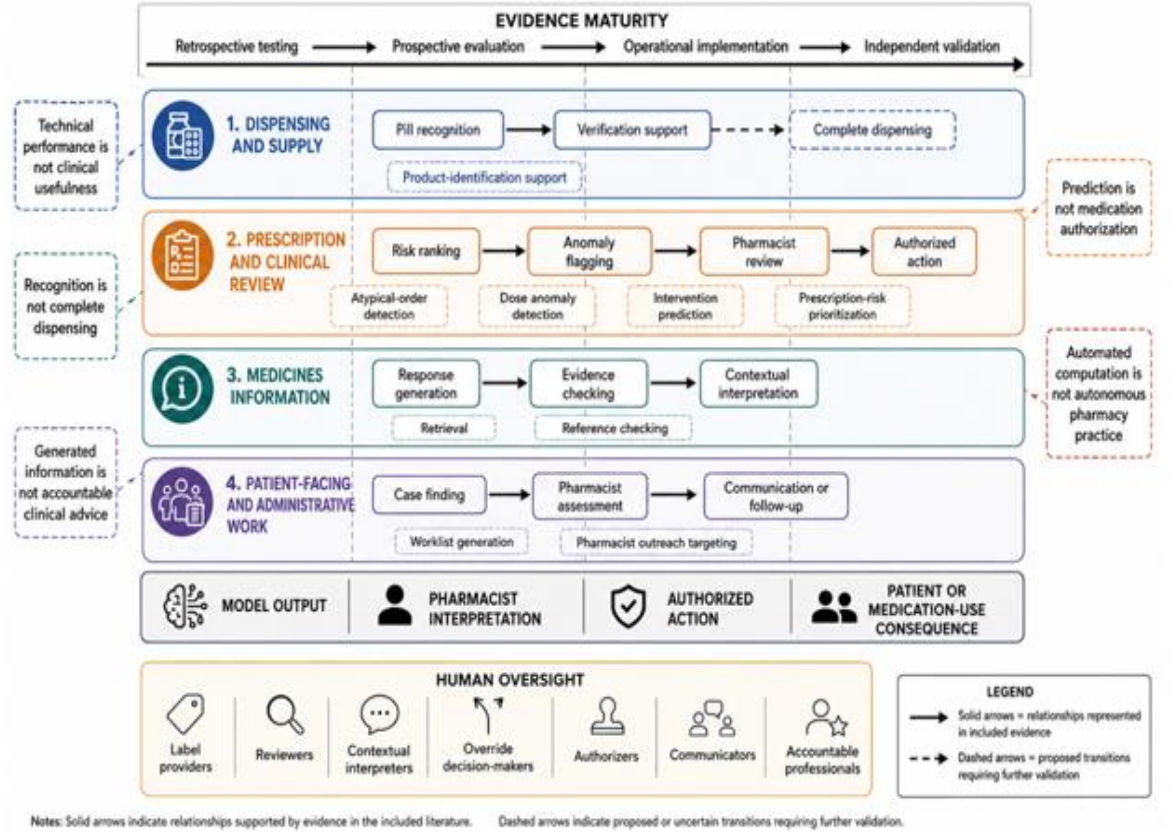

Figure 2. Evidence Map of AI Tasks across Dispensing, Clinical Review, and Medicines Information

Implications for Pharmacy Practice

Evaluation should remain task-specific. Prescription-ranking tools require assessment of missed high-consequence cases, alert burden, pharmacist response, and workflow effect. Recognition tools require testing under realistic image conditions and explicit management of uncertainty. Medicines-information systems require evaluation of factuality, evidence provenance, currency, harmful omissions, and clarification needs.

Evidence should not be transferred between tasks. Ability to prioritize orders does not validate counseling; pill recognition does not validate dispensing; and response generation does not confer authority to make patient-specific treatment decisions. Implementation remains conditional on local validation, professional oversight, escalation, governance, and monitoring.

Implications, Strengths, and Limitations

Implications for Research and Reporting

Future studies should define the unit of pharmacy work before reporting model performance. Drug-specific dose-correction prediction illustrates how rare outcomes, class imbalance, and local treatment practices constrain interpretation [36]. Reports should identify who created labels, who reviewed outputs, which actions were permitted, how uncertainty was communicated, and what occurred when the model and pharmacist disagreed.

Clinical-pharmacy evaluations of ChatGPT also demonstrate the need to distinguish factual correctness from applicability, completeness, and patient-specific safety [37]. Version control is essential because performance may change materially between model generations [38]. Model and interface versions, prompts, access dates, retrieval settings, language, and source availability should therefore be reported.

Priority designs include temporal and external validation, prospective silent evaluation, workflow trials, independent replication, and mixed-method human-factors assessment. Outcomes should extend beyond discrimination to calibration, workload, time, override, recommendation acceptance, prescribing change, detected and missed harm, patient experience, equity, sustainability, and performance under data or model drift.

Strengths and Limitations

A principal strength was the task-centered approach, which distinguished what the system computed from what pharmacy personnel interpreted or authorized. The review also separated technical outcomes from workflow, clinical, safety, patient, equity, and implementation outcomes and treated autonomy as an independent construct.

Limitations included heterogeneous terminology, task definitions, reference standards, and validation designs. Historical pharmacist action was frequently both training label and evaluation target. Rapid model change complicated comparisons of language-model studies. The evidence map therefore describes the distribution and boundaries of available work rather than establishing comparative effectiveness across non-equivalent technologies.

Conclusion

Artificial intelligence has taken on identifiable but bounded components of pharmacy work. The most frequently demonstrated functions were prescription prioritization, anomaly detection, intervention prediction, pill recognition, algorithmic case finding, and generation of medicines-information responses. These functions generally reduced, reordered, recognized, or drafted information rather than completing an entire professional task.

Pharmacists remained central as label providers, reviewers, contextual interpreters, authorizers, communicators, and accountable decision-makers. Algorithmic worklists could contain contextually inappropriate cases, recognition tools addressed only one component of dispensing, and fluent medicines-information responses still required evidence checking and clinical interpretation.

Pharmacy AI should therefore be described according to its immediate task, evidence maturity, human oversight, comparator, and outcome. Technical accuracy should not be treated as proof of medication safety, patient benefit, transferable workflow improvement, professional equivalence, or autonomous authority. Stronger evidence will require external validation, prospective workflow testing, consequence-sensitive outcomes, explicit allocation of human and algorithmic roles, equity assessment, version control, and postimplementation monitoring.

ACKNOWLEDGMENTS: None
CONFLICT OF INTEREST: None
FINANCIAL SUPPORT: None
ETHICS STATEMENT: None

References

1. Nelson SD, Walsh CG, Olsen CA, McLaughlin AJ, LeGrand JR, Schutz N, et al. Demystifying artificial intelligence in pharmacy. Am J Health Syst Pharm. 2020;77(19):1556-70. doi:10.1093/ajhp/zxaa218
2. Ranchon F, Chanoine S, Lambert-Lacroix S, Bosson JL, Moreau-Gaudry A, Bedouch P. Development of artificial intelligence powered apps and tools for clinical pharmacy services: A systematic review. Int J Med Inform. 2023;172:104983. doi:10.1016/j.ijmedinf.2022.104983
3. Hatzimanolis J, Riley B, El-Den S, Aslani P, Zhou J, Chaar BB. Applications of artificial intelligence in current pharmacy practice: A scoping review. Res Social Adm Pharm. 2025;21(3):134-41. doi:10.1016/j.sapharm.2024.12.007
4. Damiani G, Altamura G, Zedda M, Nurchis MC, Aulino G, Heidar Alizadeh A, et al. Potentiality of algorithms and artificial intelligence adoption to improve medication management in primary care: A systematic review. BMJ Open. 2023;13(3). doi:10.1136/bmjopen-2022-065301
5. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. Ann Intern Med. 2018;169(7):467-73. doi:10.7326/M18-0850
6. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth. 2020;18(10):2119-26. doi:10.11124/JBIES-20-00167
7. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol. 2018;18(1):143. doi:10.1186/s12874-018-0611-x
8. Pollock D, Peters MDJ, Khalil H, McInerney P, Alexander L, Tricco AC, et al. Recommendations for the extraction, analysis, and presentation of results in scoping reviews. JBI Evid Synth. 2023;21(3):520-32. doi:10.11124/JBIES-22-00123

9. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ. 2021;372. doi:10.1136/bmj.n71
10. Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. Syst Rev. 2021;10(1):39. doi:10.1186/s13643-020-01542-z
11. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK; SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. Lancet Digit Health. 2020;2(10). doi:10.1016/S2589-7500(20)30218-1
12. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ; SPIRIT-AI and CONSORT-AI Working Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: The SPIRIT-AI extension. Lancet Digit Health. 2020;2(10). doi:10.1016/S2589-7500(20)30219-3
13. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med. 2019;170(1):51-8. doi:10.7326/M18-1376
14. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385. doi:10.1136/bmj-2023-078378
15. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924-33. doi:10.1038/s41591-022-01772-9
16. Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco M, et al. Minimum information about clinical artificial intelligence modeling: The MI-CLAIM checklist. Nat Med. 2020;26(9):1320-4. doi:10.1038/s41591-020-1041-y
17. Corny J, Rajkumar A, Martin O, Dode X, Lajonchère JP, Billuart O, et al. A machine learning-based clinical decision support system to identify prescriptions with a high risk of medication error. J Am Med Inform Assoc. 2020;27(11):1688-94. doi:10.1093/jamia/ocaa154
18. Rozenblum R, Rodriguez-Monguio R, Volk LA, Forsythe KJ, Myers S, McGurrin M, et al. Using a machine learning system to identify and prevent medication prescribing errors: A clinical and cost analysis evaluation. Jt Comm J Qual Patient Saf. 2020;46(1):3-10. doi:10.1016/j.jcjq.2019.09.008
19. Wong YF, Ng HT, Leung KY, Chan KY, Chan SY, Loy CC. Development of fine-grained pill identification algorithm using deep convolutional network. J Biomed Inform. 2017;74:130-6. doi:10.1016/j.jbi.2017.09.005
20. Worrall C, Shirley D, Bullard J, Dao A, Morrisette T. Impact of a clinical pharmacist-led, artificial intelligence-supported medication adherence program on medication adherence performance, chronic disease control measures, and cost savings. J Am Pharm Assoc (2003). 2025;65(1):102271. doi:10.1016/j.japh.2024.102271
21. Balestra M, Chen J, Iturrate E, Aphinyanaphongs Y, Nov O. Predicting inpatient pharmacy order interventions using provider action data. JAMIA Open. 2021;4(3). doi:10.1093/jamiaopen/ooab083
22. Hogue SC, Chen F, Brassard G, Lebel D, Bussières JF, Durand A, et al. Pharmacists' perceptions of a machine learning model for the identification of atypical medication orders. J Am Med Inform Assoc. 2021;28(8):1712-8. doi:10.1093/jamia/ocab071
23. Ashraf AR, Somogyi-Végh A, Merczel S, Gyimesi N, Fittler A. Leveraging code-free deep learning for pill recognition in clinical settings: A multicenter, real-world study of performance across multiple platforms. Artif Intell Med. 2024;150:102844. doi:10.1016/j.artmed.2024.102844
24. Levivien C, Cavagna P, Grah A, Buronfosse A, Courseau R, Bézie Y, et al. Assessment of a hybrid decision support system using machine learning with artificial intelligence to safely rule out prescriptions from medication review in daily practice. Int J Clin Pharm. 2022;44(2):459-65. doi:10.1007/s11096-021-01366-4
25. King CR, Abraham J, Fritz BA, Cui Z, Galanter W, Chen Y, et al. Predicting self-intercepted medication ordering errors using machine learning. PLoS One. 2021;16(7). doi:10.1371/journal.pone.0254358
26. Johns E, Guendouz A, Dal Mas L, Beck M, Alkanj A, Gourieux B, et al. Using machine learning to predict pharmaceutical interventions during medication prescription review in a hospital setting. Am J Health Syst Pharm. 2025;82(22):1238-48. doi:10.1093/ajhp/zxaf089
27. Wang SV, Rogers JR, Jin Y, DeiCicchi D, Dejene S, Connors JM, et al. Stepped-wedge randomised trial to evaluate population health intervention designed to increase appropriate anticoagulation in patients with atrial fibrillation. BMJ Qual Saf. 2019;28(10):835-42. doi:10.1136/bmjqs-2019-009367
28. Heo J, Kang Y, Lee S, Jeong DH, Kim KM. An accurate deep learning-based system for automatic pill identification: Model development and validation. J Med Internet Res. 2023;25. doi:10.2196/41043
29. Huang X, Estau D, Liu X, Yu Y, Qin J, Li Z. Evaluating the performance of ChatGPT in clinical pharmacy: A comparative study of ChatGPT and clinical pharmacists. Br J Clin Pharmacol. 2024;90(1):232-8. doi:10.1111/bcp.15896
30. Triplett S, Ness-Engle GL, Behnen EM. A comparison of drug information question responses by a drug information center and by ChatGPT. Am J Health Syst Pharm. 2025;82(8):448-60. doi:10.1093/ajhp/zxae316
31. Andrikyan W, Sametinger SM, Kosfeld F, Jung-Poppe L, Fromm MF, Maas R, et al. Artificial intelligence-powered chatbots in search engines: A cross-sectional study on the quality and risks of drug information for patients. BMJ Qual Saf. 2025;34(2):100-9. doi:10.1136/bmjqs-2024-017476
32. Morath B, Chiriac U, Jaszkowski E, Deiß C, Nürnberg H, Hörth K, et al. Performance and risks of ChatGPT used in drug information: An exploratory real-world analysis. Eur J Hosp Pharm. 2024;31(6):491-7. doi:10.1136/ejhpharm-2023-003750
33. Kaas-Hansen BS, Leal Rodríguez C, Placido D, Thorsen-Meyer HC, Nielsen AP, Dérian N, et al. Using machine learning to identify patients at high risk of inappropriate drug dosing in periods with renal dysfunction. Clin Epidemiol. 2022;14:213-23. doi:10.2147/CLEP.S344435
34. Cho J, Lee AR, Koo D, Kim K, Jeong YM, Lee HY, et al. Development of machine-learning models using pharmacy inquiry database for predicting dose-related inquiries in a tertiary teaching hospital. Int J Med Inform. 2024;185:105398. doi:10.1016/j.ijmedinf.2024.105398
35. Nagata K, Tsuji T, Suetsugu K, Muraoka K, Watanabe H, Kanaya A, et al. Detection of overdose and underdose prescriptions—An unsupervised machine learning approach. PLoS One. 2021;16(11). doi:10.1371/journal.pone.0260315
36. Sato H, Kimura Y, Ohba M, Ara Y, Wakabayashi S, Watanabe H. Prediction of prednisolone dose correction using machine learning. J Healthc Inform Res. 2023;7(1):84-103. doi:10.1007/s41666-023-00128-3
37. Fournier A, Fallet C, Sadeghipour F, Perrottet N. Assessing the applicability and appropriateness of ChatGPT in answering clinical pharmacy questions. Ann Pharm Fr. 2024;82(3):507-13. doi:10.1016/j.pharma.2023.11.001
38. He N, Yan Y, Wu Z, Cheng Y, Liu F, Li X, et al. Chat GPT-4 significantly surpasses GPT-3.5 in drug information queries. J Telemed Telecare. 2025;31(2):306-8. doi:10.1177/1357633X231181922