A Textbook of

Clinical Pharmacy

for Pharmacy Students

Lesson 3:

Information Technologies in Clinical Pharmacy

ISBN-13: 979-13-992562-4-6

Porto, Portugal.

Published by CIPF, S.L.

Suggested citation: Fernandez-Llimos F. Lesson 3: Information Technologies in Clinical Pharmacy. In: A Textbook of Clinical Pharmacy for Pharmacy Students. Porto: CIPF; 2026.

Available at http://www.cipf-es.org

Lesson 3: Information Technologies in Clinical Pharmacy

Contents

3.1 Information Technologies in Clinical Pharmacy

The integration of information technologies (IT) into clinical pharmacy practice has transformed how pharmacists access, interpret, and apply clinical data to optimize patient care. Drug information centers from 1960s were rooms where the printed information sources were preserved and clinical pharmacists were the professionals who had the knowledge and the skills to search for the information and answer clinical questions from other health care professionals. However, modern clinical decision-making increasingly depends on digital infrastructures capable of managing vast and complex datasets, from electronic health records and computerized physician order entry systems, or bibliographic databases, to clinical decision support tools and point-of-care electronic information sources. These systems enable pharmacists to ensure appropriate therapy monitoring and contribute to multidisciplinary care processes informed by real-time evidence. This chapter explores how these technologies enhance the pharmacist's role in clinical decision-making, focusing on their design, implementation, and impact on the quality and safety of pharmacotherapy.

3.1.1 Clinical information in a Portuguese hospital

In a typical Portuguese public hospital of the Serviço Nacional de Saúde, clinical and medication-related information relevant for pharmacists is distributed across several systems. Typically, hospitals use: SClínico as electronic clinical record, Clinidata as laboratory information system, and a logistics and pharmacy management system such as GHAF (Gestão Hospitalar de Armazém e Farmácia) or Sistema de Gestão Integrada do Circuito do Medicamento (SGICM).

SClínico Hospitalar, developed by Serviços Partilhados do Ministério da Saúde (SPMS) organism depending on the Ministry of Health, is the core **electronic health record (EHR)** that

integrates the former medical and nursing systems and supports all major hospital care settings. It is used primarily by physicians and nurses to document clinical encounters, diagnoses and problems, procedures, clinical notes (SOAP model in some modules), vital signs, and to enter diagnostic requests and medical orders. SClínico, as the national hospital electronic health record in the Portuguese SNS, has been repeatedly criticized in usability and workflow studies for several structural limitations that are directly relevant to clinical pharmacy. Users report high cognitive load and **navigation complexity**, with **information fragmented across multiple screens** and tabs, which makes it difficult and time-consuming to reconstruct a coherent medication history, follow laboratory trends, or understand the temporal sequence of clinical events needed for pharmacotherapeutic decisions. Documentation interfaces are often rigid and poorly aligned with clinical reasoning, leading to **extensive use of free text**, duplication of entries, and **variable data quality**, which in turn hampers secondary uses of data such as automated clinical decision support, quality indicators, or AI applications.

Hospital laboratories usually operate a dedicated laboratory information system (LIS), often based on solutions like **Clinidata**, which manages the entire analytical workflow from electronic test ordering and sample identification to result validation and reporting of structured laboratory values. In practice, this means that numerical and categorical laboratory results (for example, creatinine, liver enzymes, blood counts, microbiology and antibiograms) are generated and stored in the LIS and may be viewed either through direct access to the LIS interface or via summarized integration into clinical views in SClínico, depending on the local technical configuration.

The logistics and pharmacy perspective is covered by systems such as **GHAF or SGICM**, solutions specifically designed to manage the hospital warehouse and pharmacy, including purchasing, stock management, internal distribution, and cost analysis. GHAF provides tools to control the whole hospital supply chain, integrating purchasing processes, stock control, dispensing to wards and services, and providing statistical and analytical reports for management and optimization of resources. But, more important for a clinical pharmacist, these systems are used as **computerized prescription order entry (CPOE)**.

So, it is important to keep in mind that clinical pharmacists, when validating the prescriptions in a Portuguese hospital, should have access to information compiled in several different

Types of data storage in clinical records

Unstructured data correspond to everything written or stored in “free text form”, without predefined fields: narrative progress notes, discharge summaries, reasons for therapy changes, radiology reports in prose, or imaging files. These elements often contain the clinical reasoning, the description of adverse drug reactions, or the contextual details.

Structured data are the “cells in a table” type of information: each item has a predefined place and format (for example, serum creatinine as a numeric field with units, ATC-coded medicines in a medication list, ICD-coded diagnoses in the problem list). Such data can be easily filtered, counted, and used by clinical decision support systems (CDSS).

Semi-structured data sit between these two and are increasingly important for interoperability. They are usually files that look like documents but have tags or a defined structure that software can parse, such as HL7 CDA/C-CDA documents or HL7 FHIR resources containing sections for problems, medications, allergies, and laboratory results.

information systems, not always sufficient interoperability, with crucial information saved as free text (instead of coded fields). Disease and drugs coding standards will be further explored in Pharmacoepidemiology and Pharmacovigilance course (5th year).

3.1.2 Clinical information in a Portuguese Primary Care (incl. Pharmacies)

In the Portuguese Serviço Nacional de Saúde (SNS), the **Registo de Saúde Eletrónico (RSE)** is the longitudinal electronic health record that aggregates information from different levels of care, including primary care, hospitals, and some private providers, and makes it accessible through the SNS 24 citizen area and professional interfaces. In practical terms, this means that, if properly integrated, the RSE could provide clinical pharmacists with a synthesized view of diagnoses, past encounters, laboratory and imaging reports, vaccination data, and medication information documented in primary care. From a governance perspective, the Portuguese RSE follows national regulations and, increasingly, the framework of the European Health Data Space, which requires that core health data be recorded electronically in standardized, interoperable formats, to support both direct care and secondary uses such as quality monitoring and research.

In primary care, general practitioners in the SNS typically use **SClínico Cuidados de Saúde Primários** and the **Prescrição Eletrónica Médica (PEM)** to register encounters and issue electronic prescriptions for medicines and diagnostic tests. PEM is a certified application developed under the coordination of SPMS that supports fully dematerialized prescriptions, transmission of prescription data to pharmacies, and the inclusion of therapeutic notes and structured alerts on medication safety.

Portuguese community pharmacies almost universally use dedicated pharmacy management software such as **Sifarma** (including Sifarma 2000 and its successors), which integrates dispensing, stock management, billing, and clinical support functionalities. These systems manage the electronic dispensing of PEM prescriptions, record actual dispensing events, generate statistics on medicine use, and incorporate decision support modules such as drug–drug interaction checking or alerts related to reimbursement and regulatory constraints, but they are typically not integrated with hospital information systems.

This fragmentation reinforces the need for pharmacists to actively seek and interpret information across systems, and underlines why future developments in interoperability, standard terminologies, and user-centered design of the RSE and prescribing software are directly relevant for the safety and efficiency of pharmacotherapy in Portugal

3.2 Clinical Decision Support Systems (CDSS)

Clinical decision support systems (CDSS) can be defined as computer-based tools that provide health professionals with patient-specific and knowledge-based information, at appropriate times, to improve clinical decisions, including the safety and quality of pharmacotherapy. In practice, pharmacy-oriented CDSS range from simple drug–drug interaction alerts embedded in dispensing software to complex, integrated systems that combine laboratory data, diagnoses and medication profiles to generate dosing recommendations or identify potentially inappropriate prescriptions.

Classical knowledge-based CDSS are often said to comprise three main components: a knowledge base (e.g. drug interaction rules, dosing guidelines), an inference engine that matches rules with patient data, and a communication mechanism that presents alerts, reminders or

recommendations within the user's prescribing, validation or dispensing environment. Knowledge-based CDSS (the dominant type in current pharmacy systems) rely on explicit rules or clinical guidelines, typically encoded as **IF–THEN statements**, such as "IF creatinine clearance <30 mL/min AND drug = dabigatran THEN suggest dose reduction or alternative."

Non-knowledge-based CDSS use **machine-learning** or other pattern recognition techniques (e.g. neural networks, support-vector machines) to infer predictions directly from data rather than from predefined rules, for example predicting adverse drug events or high-risk patients based on large EHR datasets. Non-knowledge-based CDSS are still less common in routine pharmacy workflows owing to explainability and validation challenges.

From a usability perspective, CDSS can be classified as passive or active according to how forcefully they present information. Passive CDSS require the pharmacist or prescriber to actively consult a resource (for example, opening a drug monograph, a dosing guideline, or a static interaction list embedded in dispensing software), so the system does not interrupt workflow. **Active CDSS generate unsolicited prompts**, such as pop-up alerts for contraindicated therapies, dose range checks, allergy warnings or reminders about missing laboratory monitoring; these are central to medication safety but can contribute to alert fatigue if not well tuned.

Integrated CDSS use structured data from CPOE/EHR (diagnoses, current medication list, renal function, microbiology results) to trigger real-time decision support during prescribing, validation and administration, which is the model most associated with reductions in prescribing errors and improved appropriateness of therapy. **Stand-alone CDSS** include web-based or desktop applications, commercial drug-information platforms, or dedicated clinical calculators that the pharmacist launches separately and often populate manually, which may be easier to deploy but less able to exploit rich patient-level data.

Within this broad taxonomy, several functional CDSS subtypes are particularly relevant to clinical pharmacy practice.[1] Dosing calculators translate guidelines or pharmacokinetic models into patient-specific dose suggestions, for example pediatric weight-based dosing tools that significantly reduce dose calculation errors and shorten the time needed for derivation. Drug–drug interaction checkers and broader drug-utilization review tools automatically screen prescriptions for clinically significant interactions, duplications, contraindications or potentially inappropriate medications, and refined CDSS-assisted interaction checking has been shown to reduce the number of alerts while saving pharmacist time. Finally, therapeutic drug monitoring (TDM) support systems combine measured drug concentrations, dosing history, and patient covariates to propose dose adjustments for narrow-therapeutic-index drugs (e.g. aminoglycosides, vancomycin, some anticonvulsants), thereby standardizing and improving the safety of TDM-guided dosing decisions.

[1] Reis WC, Bonetti AF, Bottacin WE, et al. Impact on process results of clinical decision support systems (CDSSs) applied to medication use: overview of systematic reviews. Pharm Pract (Granada). 2017;15(4):1036.

Alert fatigue and Alert overriding

Alert fatigue is a pervasive and well-documented problem in CDSS. It arises when clinicians are exposed to a high volume of alerts, many of which are low-priority, clinically irrelevant, duplicate, or poorly tailored to the clinical context. Over time, this constant stream of notifications leads users to become desensitized, with alerts perceived more as workflow interruptions than as safety tools, which undermines the original purpose of CDSS and can jeopardize patient safety.

Studies have quantified the extent of alert overriding, particularly for drug–drug interaction (DDI) and drug-allergy alerts. Override rates range from 50% to over 90% of alerts, meaning that most generated warnings are dismissed by professionals. For example, a meta-analysis of DDI alerts found that about 90% of such alerts were overridden, even after local tuning of the CDSS, suggesting that clinicians perceive many alerts as non-actionable in real-world practice.

Multiple factors drive alert fatigue and high override rates. Poor specificity of rule bases, lack of patient-context (e.g. renal function, current lab values), duplicate alerts across systems, and non-stratified severity levels all contribute to a high burden of alerts that clinicians quickly learn to ignore.

- Hussain MI, Reynolds TL, Zheng K. Medication safety alert fatigue may be reduced via interaction design and clinical role tailoring: a systematic review. J Am Med Inform Assoc. 2019;26(10):1141-1149.
- Felisberto M, Lima GDS, Celuppi IC, et al. Override rate of drug-drug interaction alerts in clinical decision support systems: A brief systematic review and meta-analysis. Health Informatics J. 2024;30(2):14604582241263242.

3.3 Artificial Intelligence in Clinical Pharmacy

3.3.1 Terminology in Artificial Intelligence

Artificial Intelligence is used as an umbrella term for computational methods that aim to reproduce specific aspects of human cognitive work, such as pattern recognition and prediction, to support medication-related decisions. Within this broad concept, **Machine Learning** refers to algorithms that learn patterns from data rather than being explicitly programmed for every situation, typically by training models on large datasets derived from electronic health records, pharmacy systems, or laboratory information. **Supervised Learning** denotes methods in which the model is trained on labeled examples (for instance, patients with and without documented adverse drug events) to predict a specific target outcome, whereas **Unsupervised Learning** groups patients or medication profiles into clusters based on similarities in their data without pre-defined labels. **Deep Learning** describes a family of machine-learning techniques, often using **Multilayer Neural Networks**, that are particularly suited to high-dimensional data such as imaging, signals, or very large EHR feature sets, although they tend to be less transparent than simpler models.

Two additional terms are especially relevant for clinical pharmacy applications. **Natural Language Processing** designates methods that enable computers to analyze and extract structured information from free-text sources, such as clinical notes, discharge letters, or

narrative descriptions of adverse drug reactions in the EHR, transforming unstructured text into coded concepts that can be searched, counted, or used as model inputs. **Predictive Analytics** refers to the use of statistical or machine-learning models to estimate the probability of future events for a specific patient (for example, risk of bleeding under anticoagulation, likelihood of hospital readmission, or probability of medication-related harm), often by combining demographic, clinical, laboratory, and medication variables available in routine records.

3.3.2 Artificial Intelligence in Clinical Pharmacy

It is important to distinguish these data-driven AI models from the traditional rule-based clinical decision support systems (CDSS) described earlier in the chapter. Classical CDSS typically rely on explicit, human-authored rules and guidelines encoded as IF–THEN statements, such as "IF creatinine clearance <30 mL/min AND drug = dabigatran THEN suggest dose reduction," and their behavior is fully determined by these predefined rules and structured inputs. By contrast, **AI-based systems, especially those using machine learning, infer patterns from historical data rather than relying solely on pre-specified rules**, allowing them, for example, to predict which patients are most likely to experience an adverse drug event or to benefit from pharmacist review even when no simple rule captures all relevant combinations of variables. In practice, the most robust approaches combine both paradigms: **rule-based CDSS provide transparent, guideline-linked safeguards for well-defined situations, while AI models operate in the background to prioritize patients, flag atypical risk profiles, or summarize complex data, with the clinical pharmacist retaining responsibility for interpreting these outputs and making final decisions**. This complementary model fits a cautious stance in which AI extends the reach and timeliness of clinical pharmacy services without being granted autonomous authority over prescribing or therapeutic decisions.

The rationale for integrating AI into clinical pharmacy arises from the increasing complexity of pharmacotherapy, the volume and heterogeneity of available data, and the need for timely, **patient-specific risk stratification** and information retrieval. Clinical pharmacists must manage patients with multiple comorbidities, polypharmacy, and dynamic clinical trajectories, while data relevant to medication safety and effectiveness are dispersed across EHR modules, laboratory systems, and free-text documentation. Reviews of AI in clinical pharmacy highlight that machine-learning models and related tools can help identify high-risk patients, detect potential adverse drug reactions or potentially inappropriate medications, and support complex dosing decisions, thereby focusing pharmacists' attention where it is most needed. At the same time, AI-driven systems can continuously monitor large patient populations, triage alerts, and surface succinct, context-specific summaries or recommendations, helping clinical pharmacists to access accurate and up-to-date information at the right moment in the workflow. In this sense, AI is best conceived as an infrastructure that augments the pharmacist's ability to navigate complexity and volume, rather than an independent decision-maker that substitutes professional judgment.[2]

3.3.3 Limitations and risks of AI in Clinical Pharmacy

Machine-learning models for predicting adverse drug events (ADE) in outpatient and EHR-based settings frequently show promising discrimination in development datasets, but several

[2] Alqahtani SS, Menachery SJ, Alshahrani A, Albalkhi B, Alshayban D, Iqbal MZ. Artificial intelligence in clinical pharmacy-A systematic review of current scenario and future perspectives. Digit Health. 2025;11:20552076251388145.

systematic reviews emphasize that **data quality issues and methodological limitations constrain their clinical applicability**. Common problems include class imbalance (few ADEs compared with many non-events), inconsistent outcome definitions, and heterogeneous feature sets, all of which can produce unstable or overly optimistic models that do not transfer well across settings. In addition, many published models **lack robust external validation** and are evaluated only retrospectively in the same or closely related datasets used for model development, so their performance in different populations, health systems, or EHR implementations remains uncertain. Reviews of AI-based ADE prediction and medication alert optimization also underline a persistent gap between good predictive metrics (for example, area under the curve around 0.7–0.8) and demonstrated impact on patient-level outcomes such as reduced ADE incidence, length of stay, or mortality, because very few models have been embedded into clinical workflows and rigorously tested in prospective real-world studies. For clinical pharmacists, this means that AI tools must be interpreted cautiously as decision aids whose outputs require critical appraisal, rather than as validated surrogates for clinical judgment, especially when underlying data sources are incomplete, biased, or poorly standardized.[3]

Beyond technical performance, AI-augmented decision support raises specific risks related to **over-reliance on opaque models, erosion of clinical reasoning, alert fatigue, and accountability**. Complex machine-learning and deep-learning systems often operate as "black boxes," providing risk scores or recommendations without transparent, mechanistic explanations, which can make it difficult for pharmacists to understand why a given patient is flagged and to identify situations where the model may be extrapolating beyond its training data. If such systems are tightly coupled to workflow and presented with a high degree of authority, there is a danger that clinicians gradually defer to the algorithm, weakening their own habit of structured medication review and critical appraisal, particularly under time pressure. At the same time, AI-based efforts to optimize medication alerts must confront the longstanding problem of alert fatigue: although recent reviews suggest that AI can reduce the number of low-relevance alerts generated by CDSS, they also note that most studies report reductions in inappropriate alerts without robust external validation or clear evidence of improved patient outcomes, so the risk remains that yet another layer of algorithmic complexity simply shifts, rather than resolves, the cognitive burden on clinicians.[4]

Finally, legal and ethical analyses stress that introducing AI into medication-related decisions **complicates accountability**, because responsibility is shared among software developers, health institutions, and individual clinicians, while patients may not fully understand the role of algorithms in their care. In this context, **a prudent clinical-pharmacy position is to insist on human-in-the-loop models**, explicit documentation that AI recommendations are advisory, clear governance structures for monitoring performance and bias, and continuous reinforcement of pharmacists' role as the ultimate accountable professionals for pharmacotherapy decisions.

[3] Hu Q, Chen Y, Zou D, He Z, Xu T. Predicting adverse drug event using machine learning based on electronic health records: a systematic review and meta-analysis. Front Pharmacol. 2024;15:1497397.
[4] Hu Q, Li J, Li X, Zou D, Xu T, He Z. Machine learning to predict adverse drug events based on electronic health records: a systematic review and meta-analysis. J Int Med Res. 2024;52(12):3000605241302304.

3.4 PubMed as secondary information source

3.4.1 A brief history of PubMed

To understand the characteristics of modern PubMed, it is useful having an historical perspective on the origins and the evolution of this secondary source for more than 100 years. PubMed is the latest stage in a long evolution of bibliographic control at the US National Library of Medicine (NLM), beginning with index cards, moving to the printed Index Medicus, then to MEDLINE on CD-ROM and proprietary online systems, and finally to a freely accessible web interface known as PubMed.

In the 19th- and early 20th-century **card catalogues and cabinet files** were used at all libraries to organize the rapidly growing medical literature. The Surgeon General's Office, later the NLM, compiled these index cards in printed **Index Medicus** in 1879, a monthly bibliographic index of articles from selected biomedical journals that quickly became the standard reference in hospital and academic libraries.

In the early 1960s NLM began to computerize the production of Index Medicus through the Medical Literature Analysis and Retrieval System (MEDLARS), first used for indexing and batch searching from 1963–1964. In 1971, NLM introduced **MEDLINE** (MEDLARS Online), which provided interactive online access to the same core bibliographic data that underpinned Index Medicus, initially via institutional mainframes connected over telephone lines. From the mid-1980s, MEDLINE data began to be distributed on CD-ROM by several commercial vendors, allowing local searching of subsets of the database on microcomputers in medical libraries without continuous network connections. At this time, accessing MEDLINE had an important cost.

PubMed was first released in January 1996 as an experimental web-based interface within NCBI's Entrez system, providing end-user searching of the MEDLINE database over the Internet. The "experimental" label was dropped in April 1997, and on 26 June 1997 NLM and NIH publicly announced free MEDLINE access via PubMed, effectively shifting MEDLINE from a library-mediated, **fee-based resource to a widely accessible public service**.

As Internet access expanded and researchers increasingly used PubMed and other online MEDLINE implementations, subscriptions to the printed Index Medicus declined sharply, leading NLM to cease its print publication after 2004.

3.4.2 Differences between PubMed and MEDLINE

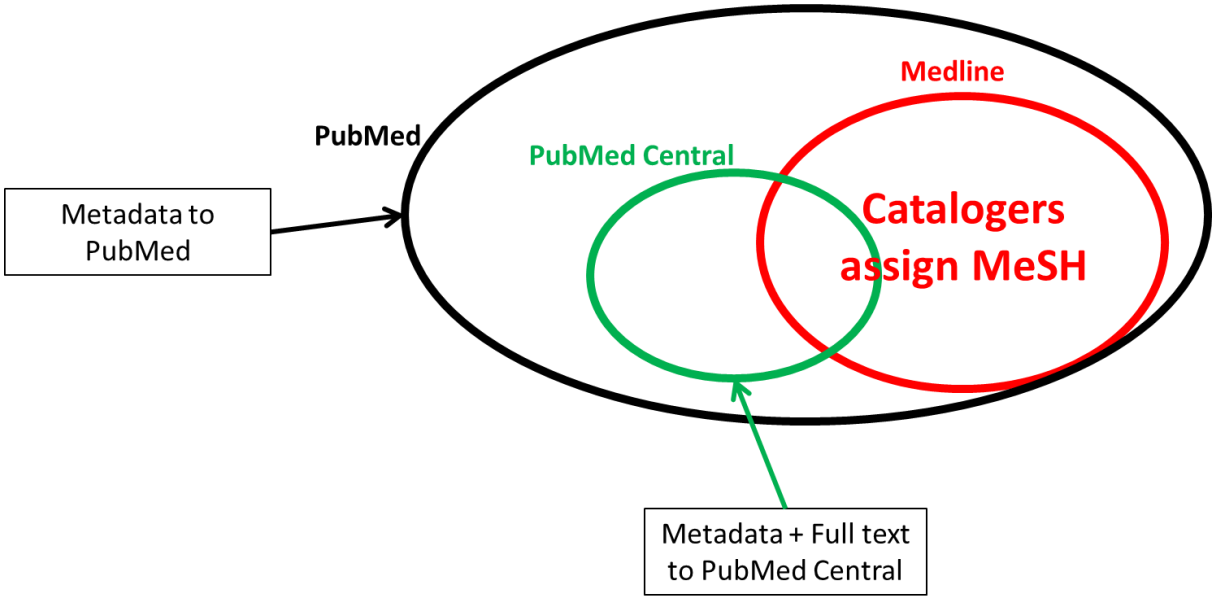
PubMed, **MEDLINE** and **PubMed Central** are closely related NLM resources, but they differ in scope and function: MEDLINE is the main curated bibliographic database, PubMed is the search platform that includes MEDLINE, PubMed Central, and NCBI Bookshelf (full-textbooks, reports and other scholarly documents in the life and health sciences, maintained by NLM).

MEDLINE is NLM's principal bibliographic database of citations and abstracts from selected biomedical and life-science journals, with coverage from the mid-1960s and some earlier material. Journals to be included are selected by the Literature Selection Technical Review

Committee, and **MEDLINE records are indexed with MeSH and other controlled fields**, making it a curated[5] subset of PubMed with high and relatively homogeneous quality standards.

PubMed Central is a digital **repository of full-text biomedical and life-science articles**, maintained by NLM/NCBI, that provides free access to complete articles deposited by journals and authors (e.g. under open-access policies or funder mandates). PMC articles have a corresponding PubMed citation, meaning that PubMed functions as the discovery layer (citation and abstract), whereas PMC provides the full text and long-term archiving.

PubMed is the freely accessible **search interface** that retrieves MEDLINE plus additional NLM and publisher-supplied records, including in-process and "ahead of print" citations, some non-MEDLINE journals, older non-indexed citations, and records derived from PubMed Central and NCBI Bookshelf.



3.4.3 Specificity and Sensitivity

In the context of literature searching, **sensitivity** and **specificity** describe how a search strategy balances the trade-off between finding as many relevant records as possible and minimizing the retrieval of irrelevant material, paralleling their use in diagnostic test accuracy. **Sensitivity** refers to the ability of a search to retrieve **ALL** the relevant articles in the database, whereas **specificity** reflects the ability to retrieve **ONLY** records that are relevant to the question posed.

By analogy with diagnostic testing, a **highly sensitive search** strategy is one that has a low "false negative" rate: **very few relevant articles are missed**. In practical terms, high sensitivity is achieved by using broad search terms, extensive synonym lists (combined with OR), generous truncation, and minimal use of restrictive filters. For example, when designing a search filter for randomized controlled trials in PubMed, "high-sensitivity" filters are deliberately constructed to capture almost all trials at the cost of bringing in many non-trial records. This approach is usually

[5] Curated means that the contents have been deliberately selected, checked and organized by experts, with human oversight to ensure quality, consistency and relevance.

preferred in systematic reviews of clinical interventions, where the risk of missing clinically important evidence outweighs the extra screening burden.

Specificity is often sometimes discussed as precision in the information-retrieval literature. A highly specific search yields fewer irrelevant ("false positive") results, meaning that the titles and abstracts presented to the reviewer are more likely to be on-topic. Specificity is increased by using more precise terms (for example, controlled vocabulary such as MeSH headings), by restricting the strategy to key fields (e.g. title/abstract instead of all fields), and by adding additional concepts with AND or using carefully chosen methodological filters.

In real searches, sensitivity and specificity exist in tension: making a strategy more sensitive (by broadening terms or adding synonyms) tends to reduce its specificity, whereas narrowing the strategy to increase specificity usually reduces sensitivity. For clinical pharmacy and other applied disciplines, the appropriate balance depends on the purpose of the search. For a quick clinical query where time is limited and missing a small number of marginally relevant papers is acceptable, a more specific strategy focusing on precise MeSH terms and phrase searching may be appropriate. In contrast, for systematic reviews, guideline development or health-technology assessments, searchers deliberately favor high-sensitivity strategies and accept low specificity, supplementing them with rigorous screening and, where appropriate, post-hoc text-mining or prioritization tools to manage the large result sets.

Sensitivity and Specificity vs. Precision and Recall

In the search and information-retrieval context, recall corresponds to sensitivity as used in diagnostic testing, and precision is the analogue of a "rule-in" property that is closely related to, but not identical with, specificity.

Recall and **sensitivity** are mathematically defined in the same way: both are the proportion of relevant (or "positive") items that are successfully identified by the system among all relevant items present. In information retrieval, recall is therefore the fraction of all relevant documents in the database that the search retrieves, just as sensitivity is the fraction of all patients with the disease who test positive in diagnostic testing.

Precision, on the other hand, is defined as the proportion of retrieved items that are relevant among all the items retrieved by the search. In diagnostic test terminology this is analogous to the **positive predictive value**, and it increases as the number of false positives decreases. **Specificity**, by contrast, is the proportion of truly negative items correctly identified as negative among all negatives (true negatives plus false positives), a quantity that depends explicitly on the pool of negatives. Because in large document collections the number of non-relevant items (true negatives) is enormous and usually not enumerated, information-retrieval research commonly focuses on precision and recall rather than specificity; in that setting, "high precision" plays a role conceptually similar to "high specificity", in that both describe situations where few non-relevant items are labeled as relevant, even though the underlying denominators differ.

3.4.4 Features for an efficient PubMed search strategy

Efficient PubMed searching relies on understanding a small set of syntactic tools that control how the system interprets a query[6]. These include phrase searching with quotation marks, logical structuring with Boolean operators and parentheses, truncation symbols, and field tags that tell PubMed exactly where to look in each record.

Quotation marks are used for phrase searching, which is crucial when a concept is conventionally expressed as a fixed expression. When you type clinical pharmacy without quotation marks, PubMed's automatic term mapping may **split the words and map them separately** to MeSH terms, text words and journal titles, potentially retrieving articles where <<"clinical">>[7] and <<"pharmacy">> **appear far apart in the record**. By contrast, <<"clinical pharmacy">> **enforces adjacency and order**, so PubMed looks specifically for that phrase, usually in the title, abstract, or MeSH terms; this tends to reduce the number of records but increases their specificity. Quotation marks should be used also for single words to avoid any **automatic term mapping**.

Boolean operators define the logical relationships between concepts. The AND narrows the search by requiring that all connected terms appear in the same record, for example "clinical pharmacy" AND "anticoagulants" retrieves articles that discuss both clinical pharmacy and anticoagulants. The OR broadens the search by accepting any of the listed synonyms or related terms, such as <<warfarin OR acenocoumarol OR phenprocoumon>>, which is useful when you want comprehensive coverage of a pharmacological class. The NOT excludes a term, for example <<"warfarin" NOT "animals">>, though in evidence-based searching the NOT operator must be used cautiously to avoid inadvertently discarding clinically relevant studies.

Parentheses control the order in which PubMed processes Boolean operators and allow you to group synonyms into concept blocks, mirroring the PICO[8] or similar frameworks used in clinical pharmacy. Without parentheses, PubMed evaluates Boolean operators from left to right, which can change the meaning of the query; for instance, <<"hypertension" OR "diabetes" AND "clinical pharmacy">> will not behave the same as <<("hypertension" OR "diabetes") AND "clinical pharmacy">>. In practice, you typically first group synonyms with OR, such as first <<("hypertension" OR "high blood pressure")>> and second <<("adherence" OR "compliance" OR "persistence")>>, and then combine these sets with the AND to express a structured question like <<(hypertension OR "high blood pressure") AND (adherence OR compliance OR persistence) AND "community pharmacy">>, which retrieves articles focusing on medication adherence in hypertensive patients in community pharmacy settings. See below How to Build a Query.

Truncation characters increase recall by allowing one stem to represent multiple word endings. In PubMed the truncation symbol is the asterisk (*). Since 2024, new uses of asterisks in PubMed were released, allowing using them in the middle of a word, as well as using more than one.[9] For example, anticoagul* retrieves anticoagulant, anticoagulants, and anticoagulation, so a search like "atrial fibrillation" AND anticoagul* will cover a wide range of anticoagulation-related articles without enumerating each variant; similarly, pharmaco*inet* captures pharmacokinetic and

[6] Query is the common term to call a "search strategy".
[7] <<...>> were used here to frame the search queries and should not be used in actual searches.
[8] PICO: acronym meaning: Population, Intervention, Comparator, Outcome.
[9] https://www.nlm.nih.gov/pubs/techbull/mj24/mj24_pubmed_wildcard.html

pharmacokinetics. However, if the stem is too short (4-letters are required), truncation may pull in irrelevant variants. For instance, stat* would retrieve statistics, status, and other unrelated words.

Field tags provide fine-grained control by telling PubMed exactly which part of the record (which field) should be searched, and they are particularly powerful when combined with the tools above. The [TIAB] tag restricts a term to the title and abstract (e.g. anticoagul*[TIAB]), which can be useful when you want to focus on articles where a concept is central enough to appear in these visible fields, while [MH] restricts to MeSH headings (e.g. "Anticoagulants"[MH]), leveraging NLM's controlled vocabulary and indexing to group related synonyms under one concept. Other tags such as [AU] for author (e.g. Garcia DA[AU]), [TA] for journal title abbreviation (e.g. "Clin Pharmacol Ther"[TA]), or [DP] for publication date (e.g. "2019/01/01"[DP]:"2024/12/31"[DP]) allow you to refine searches further.

Proximity operators allow the user to specify that multiple words must appear close to each other in the title or in the title/abstract fields, within a maximum distance defined by the searcher (the ":~N" part of the syntax). In practice, proximity searching uses the structure <<"adverse reaction"[TIAB:~N]>>, where the field is usually [TI] or [TIAB], and N is the maximum number of words that may appear between them; so, adverse and reaction may appear in any order, with at most two words in between.

3.4.5 Medical Subject Headings (MeSH)

Medical Subject Headings (MeSH) are a **controlled, hierarchically organized vocabulary** created and maintained by the NLM to describe the content of biomedical publications in a standardized way. As a controlled vocabulary, MeSH assigns one term to each concept and links it to synonyms and closely related expressions, so that heterogeneous author language such as "heart attack", "myocardial infarct", or "MI" is normalized under the single MeSH heading "Myocardial Infarction", thereby improving consistency and recall when searching MEDLINE.

Only records that are part of MEDLINE receive MeSH terms during the indexing process, whereas non-MEDLINE articles in PubMed (for example, some PubMed Central-only journals, preprints, and "ahead of print" records) are not assigned MeSH descriptors. This distinction is important for clinical pharmacy literature searches, because a strategy relying solely on MeSH headings will systematically miss a proportion of PubMed records, particularly very recent articles not yet indexed and citations from journals that are not selected for MEDLINE.

Conceptually, MeSH is structured as a set of tree branches, in which broad descriptors sit at higher levels and more specific concepts appear as narrower terms beneath them, allowing users and indexers to navigate from general to specific topics. PubMed takes advantage of this hierarchical structure through the **Automatic Explosion** feature: when a search is performed with a MeSH heading, the system also retrieves all records indexed with any of its more specific descendants in the tree, so a search for Myocardial Infarction will automatically include Anterior Wall Myocardial Infarction, Inferior Wall Myocardial Infarction, and other narrower infarction terms unless this behavior is explicitly turned off.

All MeSH Categories
Diseases Category
Cardiovascular Diseases
Heart Diseases
Myocardial Ischemia
Myocardial Infarction
Anterior Wall Myocardial Infarction
Inferior Wall Myocardial Infarction
MINOCA
Non-ST Elevated Myocardial Infarction
Shock, Cardiogenic
ST Elevation Myocardial Infarction

The **MeSH Database**, accessible directly from the PubMed interface[10], is the principal tool to identify appropriate subject headings for a given concept and to inspect their position in the hierarchy. When a term is entered in the MeSH Database, PubMed displays the preferred MeSH descriptor, its definition, synonyms or entry terms, subheadings, and its tree locations.

Because MeSH indexing is concept-based rather than word-based, searching with MeSH retrieves articles about a subject even when the exact wording does not appear in the title or abstract, which is particularly useful in fields with variable terminology, such as “adverse drug reactions” or “medication adherence”. In practice, optimal PubMed strategies in clinical pharmacy usually combine MeSH headings (taking advantage of their controlled vocabulary and automatic explosion in MEDLINE records) with carefully chosen text words and proximity operators, so that both indexed MEDLINE citations and non-MEDLINE PubMed records are captured in a transparent and reproducible way.

3.4.6 How to Build a Query

A typical advanced query in a PubMed search strategy should combine text words and MeSH, for example <<(“direct oral anticoagulants”[TIAB] OR DOAC*[TIAB]) AND “Atrial Fibrillation”[MH] AND (“Medication Adherence”[MH] OR “adherence”[TIAB] OR “compliance”[TIAB])>>, which yields a more precise and transparent strategy than relying solely on free text. To build this query, a systematic approach helps.

Clinical pharmacists can use one of several acronyms to structure a query and frame clinical questions. The most used acronym is PICO, which stands for population, intervention, comparator, and outcome. Sometimes it is expanded to include study design (PICOS) or time frame (PICOT). There are many other acronyms that fit better with different types of studies or research questions. Some examples are:

- SPIDER: Sample, Phenomenon of Interest, Design, Evaluation, Research type.
- PEO (qualitative use): Population, Exposure (or Experience), Outcome (understood as impact/experience rather than clinical endpoint).
- PPhTS: Participants, central Phenomenon, Time, Space (for some qualitative/public health questions).
- SPICE: Setting, Perspective (or Population), Intervention/Interest, Comparison, Evaluation.

[10] https://www.ncbi.nlm.nih.gov/mesh

- ECLIPSE: Expectation, Client group, Location, Impact, Professionals, Service.
- 5Ws+H: Who, What, Where, When, Why, How.

The purpose of using these acronyms is to break down the research question into smaller, independent pieces of information.

When creating a query, it is helpful to break the problem down into smaller parts. Clinical pharmacists can use these acronyms or simply divide the question into logical parts. One helpful method is to construct a table with these parts as the rows and two columns: one for free-text words and one for MeSH terms. For example, if you want to search for articles about studies that evaluated medication adherence in patients with atrial fibrillation who were prescribed DOACs (direct oral anticoagulants), one initial step could be to create a table with three rows covering the concepts of DOACs, adherence, and fibrillation, as well as two columns for free text and MeSH terms. Then, populate the table with different words in the first column and MeSH terms in the second column that represent each concept.

	Free-text [TIAB]	MeSH [MH][11]
DOACs	• "direct oral anticoagulants" • "DOAC*"	• "Rivaroxaban" • "Dabigatran"
Adherence	• "adherence" • "compliance"	• "Medication Adherence"
Fibrillation	• "Fibrillation"	• "Atrial Fibrillation"

Once the table is populated, each row should be combined with an OR Boolean operator. Finally, all the rows should be combined with an AND Boolean operator (use parentheses to group each row).

("direct oral anticoagulants"[TIAB] OR "DOAC*"[TIAB] OR "Rivaroxaban"[MH] OR "Dabigatran"[MH])

AND

("adherence" [TIAB] OR "compliance" [TIAB] OR "Medication Adherence" [MH])

AND

("Fibrillation" [TIAB] OR "Atrial Fibrillation" [MH])

Then, the final query could be:

("direct oral anticoagulants"[TIAB] OR "DOAC*"[TIAB] OR "Rivaroxaban"[MH] OR "Dabigatran"[MH]) AND ("adherence" [TIAB] OR "compliance" [TIAB] OR "Medication Adherence" [MH]) AND ("Fibrillation" [TIAB] OR "Atrial Fibrillation" [MH])

[11] Some other Supplementary Concepts could be used here, but they are out of the scope of this chapter.