Summary
Literature monitoring is a complicated aspect of pharmacovigilance due to the large amounts of data and range of sources, regional regulatory requirements, and monitoring of local language journals. Even experienced database users find it challenging to build queries that ensure no adverse event is missed while balancing resources needed to review the results. Embase® and its PV Wizard facilitate the building of search strategies with high recall, sensitivity and compliance.
Literature searches in pharmacovigilance: the challenges

As clinical trials involve at most thousands of patients, low-frequency adverse drug reactions (ADRs) are often unknown when a drug reaches the market. In the US alone, ADRs affect around two million people every year, with 100,000 fatalities. Pharmacovigilance (PV) is an essential aspect of drug development that improves patient safety and public health by detecting, assessing, understanding and preventing ADRs. PV activities start upon submission of a marketing authorization application and continue throughout the active authorization period.

Scientific literature is the second-largest source of adverse event reports (AERs) received by the EMA and the fourth-largest source for the FDA. Marketing authorization holders are expected to perform systematic literature searches using global reference databases, such as MEDLINE® and Embase, and local journals from countries where a given medicinal product has marketing authorization. The retrieved information needs to be collated, analyzed and communicated at least once per week. Quality control systems must be in place to allow inspection and audits of the literature monitoring process, as inadequate literature searches can lead to regulatory non-compliance.

Any method of literature monitoring for PV must address the following challenges:

Sources of data: To ensure that no critical safety signal is missed, a comprehensive data search must include information from a broad range of sources: full-text and abstract publications, proceedings from scientific meetings and conferences, systematic reviews and meta analyses, competitor product data, legal documents and patents, dissertations and theses, posters, letters to the editors, clinical trial data, regulatory reports, spontaneous reports, lay press, and unpublished manuscripts including case reports, safety findings and clinical studies. Year by year, unstructured data added to databases such as Embase grows by millions of records.

Regulatory requirements: Globally acting pharmaceutical companies need to consider regulatory frameworks from a range of authorities, including the EMA's Good Pharmacovigilance Practices (1) and the FDA's Guidance for Industry (2). Documents from organizations such as the International Council for Harmonisation must also be taken into account when planning a PV strategy.

Building a perfect search strategy: It is essential to build and maintain search strategies that minimize the risk of missing articles that mention relevant ADRs. As PV searches must be run and reviewed regularly, the workload needs to be kept manageable. Search queries should aim for both high recall to ensure retrieval of all relevant records and high sensitivity to minimize the number of irrelevant records.

Monitoring local language journals: Regulations require systematic reviews of local journals in all countries where a given medicinal product has marketing authorization. In most cases, language is not the biggest challenge, but rather making sure that all relevant journals and articles are included in the search.
Integrating data from the EMA Medical Literature Monitoring Service: In September 2015, the EMA started a Medical Literature Monitoring (MLM) service. It uses Embase, International Pharmaceutical Abstracts (IPA), and the Allied and the Complementary Medicine Database (AMED) to provide quality-controlled literature monitoring for 300 active substances and 100 herbals. Records can be retrieved from the EudraVigilance database.

The goal of the EMA service is to alleviate the burden on marketing authorization holders by enhancing the efficiency of ADR reporting; improve data quality by reducing the number of duplicates; support signal detection activities by national competent authorities and marketing authorization holders; and establish an easy process for compliance with worldwide regulatory requirements.

Due to its limited substance coverage, the MLM service does not relieve pharmaceutical companies from literature monitoring. In addition, non-serious adverse event reports and safety-related topics such as misuse and off-label use are currently not included. Furthermore, EMA standards might deviate from internal case detection and processing standards and company policies.

Inspections and audits: As an essential part of pharmacovigilance, literature monitoring is frequently inspected and audited, even if there is no specific regulatory mandate. Marketing authorization holders must implement a systematic approach to the collection of information about suspected ADRs that is documented in quality-controlled, standard operating procedures. Documentation should comprise search strategies, the source databases used, search dates, search results and details on the quality control checks.
Balancing recall and precision

Having an effective literature monitoring strategy in place is essential for successful pharmacovigilance. But how can the effectiveness of a search strategy be measured? Information science defines precision and recall as quality attributes for search strategies.

Recall determines the percentage of all relevant records present in the search result. At 100% recall, no record reporting an ADR is missed.

To validate a search strategy for its recall rate, a common industry practice is to use a GOLD data set as a standard that contains all relevant records retrieved from the marketing authorization holder’s safety databases in a specific time frame. If the safety database contains a set of records retrieved through literature screening from, for example, 2014 to 2017, data obtained during three months (a different one from each year) are selected and grouped as one GOLD data set.

This GOLD data set is then used as a standard to validate the result obtained with the search strategy. If all data from the GOLD set are retrieved with the search strategy, recall is 100%.

Precision reflects how many of the records in a result set are relevant. Low-precision search strategies yield many irrelevant records, making the triaging process inefficient. Still, in some cases, low precision cannot be increased without the risk of missing a relevant record. The overall goal in literature monitoring is to create a search formula with 100% recall and a precision rate that is as high as possible.

Building a successful search strategy requires the use of comprehensive databases with in-depth drug indexing, a good understanding of metadata fields available in the database (e.g., trade names), and profound search syntax expertise. Knowledge of medicinal product properties and known interactions helps optimization of the search strategy.

When spelled out, PV searches can be hundreds of lines long. Companies typically need to monitor several, sometimes hundreds of drugs in parallel. How can reliable and reproducible literature monitoring be achieved?
The PV Wizard search form is an intuitive query builder that was integrated into Embase to facilitate successful construction of PV search strategies.

Building search queries based on industry best practices with PV Wizard

Embase is recommended by the EMA for the capture of records referring to ADRs. Embase supports PV with over 32 million records from almost 8,300 journals and over 7,000 conference proceedings and abstracts. More than 6,000 records are added daily. In addition, Embase features Emtree®, a thesaurus function for life science topics that provides consistent description of medical information across different types of literature, ensuring comprehensive retrieval of relevant records. Emtree is updated three times per year with information on new drugs, diseases, devices and procedures to ensure the capture of the most recent records.

Having all this data collated and indexed in a single database provides the foundation for literature monitoring, but an efficient search strategy that retrieves all relevant records is essential for successful PV.

The PV Wizard search form (Figure 1) is an intuitive query builder that was integrated into Embase to facilitate the successful construction of PV search strategies. Developed in collaboration with experts in the pharmaceutical industry to reflect best practices of experienced database users, it includes pre-coded search terms to describe PV-relevant content, e.g., adverse drug reactions.

With PV Wizard, the process of building effective strategies is significantly simplified and shortened, making literature monitoring accessible even for novice users. Large companies that need to monitor many substances in parallel can profit from PV Wizard, as can generic drug manufacturers faced with upcoming changes in FDA regulations that will require them to take more responsibility for warning patients about potential product risks.

![PV Wizard - Drug Name](image)

**Figure 1.** The PV Wizard is a dedicated search form integrated into Embase
The PV Wizard allows entry of all search terms in an intuitive step-by-step approach described below. Terms that describe adverse reactions are pre-coded and can be easily selected or de-selected from a drop-down menu, or entered manually.

The Emtree functionality supports the process of finding best search terms and synonyms. Additionally, if a substance is listed in the EMA MLM records, information from this database can be searched directly in the search folder available through Embase.com. Boolean operators are chosen automatically, applying the correct search syntax. The current search strategy and the number of retrieved records can be reviewed at each stage during query building.

Once the search strategy has been validated for recall and precision, email alerts can be set up to automate the search. Email alert history, search results and query details are saved and documented to comply with audit requirements. Figure 2 shows an example of a search query built using PV Wizard. Figure 3 shows the schematic image of the search strategy design.

Application example: Building a search query for paracetamol

At the start of a pharmacovigilance search query, a given medicinal product is searched in combination with subheadings such as “adverse drug reaction”, “drug toxicity”, or “drug interaction”. Expanding the search of its active substance with all alternative and variant names leads to a more comprehensive coverage. Theoretically, this would be sufficient to achieve a 100% recall, but at the cost of very low precision.

To balance recall and precision, PV Wizard contains additional indexing terms that narrow down the search. These include text searches for “adverse effects” or “complications”, information on special situations such as “pregnancy” or “medication error”, and human limits (such as adult vs. children or gender-specific effects). A combination of Boolean and proximity operators yields an intersection containing all possible records that indicate critical safety information for this medicinal product.

Figure 2. An example of a search query for paracetamol built using PV Wizard
Figure 3. Schematic image of a search strategy designed using PV Wizard

Drug names: terms for drug names and variants, searched combining subheadings such as “adverse drug reaction”, “drug toxicity”, “drug interaction”

Adverse drug reactions: terms describing adverse drug reactions

Special conditions: terms describing special situations such as “pregnancy” or “medication error”

Human limit: terms describing human limits (e.g., age or gender)

Results: A manageable set of records with a high recall rate
For example, with PV Wizard, a search query for paracetamol can be created within just a few minutes:

**Step 1.** Access PV Wizard via Embase.com. Enter the product name, here “paracetamol”, into the **Drug name** field. Important subheadings are pre-selected for search in combination with the drug names. Subheadings are Emtree terms that are used as concept qualifiers to provide additional information about context, such as “adverse drug reaction”, “drug interaction” or “drug toxicity”.

![PV Wizard - Drug Name](image)

**Step 2.** Click **Next step** to view the suggested list of alternative drug names for paracetamol is suggested. This list can be added to the search query as a complete set, or can be edited by de-selecting individual names or adding missing ones. Emtree facilitates manual editing by suggesting a list of all possible synonyms.

![PV Wizard - Alternative Drug Names](image)
Step 3. Click Next step to see the Adverse drug reaction term window, which shows a set of pre-coded search terms developed in collaboration with industry experts. It can be edited further if necessary.

Step 4. Click Next step to add special situation terms from a pre-developed list.

Step 5. Click Next step to go to the optional field where human limit search terms can be included. The search can also be narrowed down to a particular time frame on this page.
These five steps are automatically combined with Boolean operators. During query building, Embase displays the current search details, and calculates the number of records in real-time. The full list of search results is displayed, providing a good starting point. As mentioned above, it is important to validate the query by reviewing the search terms for recall and precision.

Once the query design is completed, a regular email alert allows automation of the search. The email alert history can be exported and audited. As Emtree is updated three times per year, the search strategy should be reviewed regularly to ensure it consistently produces a record set meeting its pre-determined quality attributes.

Optionally, the “EMA’s MLM searches” feature allows users to search for information on the 300 chemical and 100 herbal active substance groups that are included in the EMA medical literature monitoring service.

User feedback

A manager of PV and risk management at a large US generics manufacturer discussed the impact of Embase and PV Wizard for routine literature monitoring work.

With over 150 products, the PV and risk management group has a large workload. The manager and team of three PV specialists perform adverse event searches but also answer clinical questions from colleagues inside the company, e.g., providing information required for regulatory procedures, clinical trials and other safety-related situations.

The time this team invests in literature searches is wasted if an important adverse event or information needed for a clinical summary is missed. While choosing the right database is seen important, creating the search string that ensures no important article is missed is the primary challenge.

PV Wizard saves them a significant amount of time. While it previously took them two weeks to create a single search string, this process is greatly facilitated by the pre-defined parameters and search terms. Even though the company only operates in the US market, they frequently review and adapt queries from the EMA MLM service for their products. Overall, the PV Wizard gives the PV and risk management team more confidence with literature monitoring and compliance.

Summary

Literature screening is an important part of pharmacovigilance. Building a high-quality search strategy is essential, but also very challenging. It requires knowledge, experience and skills to ensure that no adverse event report is missed. Regulatory authorities recommend using Embase as the most comprehensive biomedical reference database. With the introduction of the PV Wizard as a new Embase search tool, building an effective and compliant search strategy has now become significantly easier and faster.

PV Wizard gives the PV and risk management team at a large US generics manufacturer more confidence with literature monitoring and compliance.
References


Embase

Embase helps customers improve literature monitoring for adverse events with the world's most comprehensive biomedical literature database.

Learn More

To learn more about Elsevier’s R&D Solutions, visit elsevier.com/rd-solutions/pharma-and-life-sciences or contact your nearest Elsevier office.

ASIA AND AUSTRALIA
Tel: + 65 6349 0222

JAPAN
Tel: + 81 3 5561 5034

KOREA AND TAIWAN
Tel: +82 2 6714 3000

EUROPE, MIDDLE EAST AND AFRICA
Tel: +31 20 485 3767

NORTH AMERICA, CENTRAL AMERICA AND CANADA
Tel: +1 888 615 4500

SOUTH AMERICA
Tel: +55 21 3970 9300