

# Indexing Guide

**Embase Classic is the bibliographic database that covers the Embase backfile from the years 1947 - 1973. Drawn from the 43 Excerpta Medica Abstract Journals published during this timeframe, Embase Classic makes them accessible in electronic format for the first time.**

Embase Classic contains historic index terms and has been enhanced with additional indexing controlled by the current Emtree thesaurus for all 1.8 million records included in the database.

## ***Original indexing***

Historic index terms in Embase Classic are generally different from current Emtree index terms. At the time, indexing of each abstract journal was controlled by the responsible editor within Excerpta Medica. Inconsistencies abounded between the indexes of each section. In 1963, it was decided to synthesize a common thesaurus from the available index terms, which would define both an accepted preferred terminology for all the concepts used, as well as a list of synonyms that would automatically be translated to the preferred terms.

The result was an authority list of 25,000 preferred terms and 50,000 synonyms, which was named MALIMET (the Master List of Medical Terms). Some 25 years later, in 1988, when MALIMET had grown to 250,000 preferred terms it was decided to make the thesaurus more suitable for online use by decreasing the number of preferred terms and imposing a hierarchical structure upon it. A faceted thesaurus structure was used. Emtree, the Excerpta Medica Tree structure, was the result. So although a controlled authority list was used for indexing part of the records included in Embase Classic (from 1963 onwards) many of the original index terms fall outside the current Emtree.

The value of the historic index terms is that they are based on the full-text article. They therefore supplement the information contained in the titles and abstracts of Embase Classic. They are inappropriate however for web -based searching in many cases. In some cases outdated terminology is used, in other cases the terminology does not adhere to Embase's current indexing policy.

## ***Additional indexing***

In order to increase recall and precision searching of the archived literature, additional index terms have been added to each record in Embase Classic by automatic indexing. Also, historic index terms have been mapped to the current Emtree. On average four historic index terms included in each Embase Classic record could be matched with Emtree. These terms are listed as controlled index terms, along with the additional index terms. On average three historic index terms in each record could not be matched with Emtree. These have been used as input for the automatic indexing process. The average number of new index terms included in each Embase Classic record is 11.

## ***Benefits from additional indexing of Embase Classic***

As a result of the additional indexing and mapping of historic terms, Embase Classic provides effective search, retrieval and linkage tools. Benefits include:

- Current Emtree terms can be used to search Embase Classic in combination with Embase
- “Drug terms” and “medical terms” can be searched separately
- Searching of index terms can be limited to “major focus”
- CAS registry numbers are included in records enhanced with drug index terms
- DOIs are included in records for which the full-text source article is available
- online

Note that Emtree “link terms” (subheadings) are not linked in Embase Classic as they are in current Embase. Instead, these terms can be searched either in free text or as regular index terms.

### ***Process and Methodology for additional indexing***

The most important elements of the automatic indexing process include:

- Automatic indexing based on Article Title, Abstract and Historic Index Terms
- Mapping of historic index terms to current Emtree terminology
- Scanning of article titles for old terminology resulting in the addition of old terms to Emtree as synonyms or candidate terms (terms awaiting approval for inclusion)
- Domain knowledge in combination with domain knowledge of the relevant timeframe for optimizing the rules for the automatic indexing process

### ***More information***

Please refer to [www.embase.com](http://www.embase.com)

**Last updated: February 2010**

