In October 2013, PubMed launched a relevance sort option for search results, with a short statement in the NLM Technical bulletin[1]. They did not provide additional information about the extent to which this innovation might be useful. For which kind of medical queries or topics would this function be effective and perform best?

Users were left to empirically discover the value and performance of this new display option and make their own judgement. PubMed is the last of the major scientific search tools to offer relevance ranking in the display option. Before 2013 PubMed only computed a score of topical relatedness in connection with the function Related citations.

**Introduction**

In this context our aim was to confront IRS relevance ranking and user relevance ranking. We decided to base our study on the relevance sort option in three major scientific databases on their native platforms.

**Method**

We performed three different subject searches in PubMed using commonly used search features combining free text and controlled vocabulary. The search queries were adapted for Embase and Web of Science (WoS)[2]. For each topic, an expert from our university hospital provided five ‘must read’ articles.

**Results**

All articles were found in the databases, but not all ranked in the top 30 when the relevance display is activated. Thus, in our tests, relevance sorting in all three databases does not overlap the subjective, user-perceived relevance. PubMed scored fairly well in comparison with the two other tools. But our tests are not sufficient to determine if this observation can be seen as a fundamental trend. Relevance judgments in medicine are based on many criteria (study design, clarity, scope, novelty...) that algorithms certainly need further development.

**Background**

Why did NLM wait before introducing this function? Is it because relevance is a fundamental, yet multidimensional process, when integrated into Information Retrieval Systems (IRS)? Two main types of relevance can be distinguished:  

- Algorithmic relevance, or relevance with respect to the query  
- User-perceived relevance, or relevance with respect to a user’s needs and judgement

In our opinion these should appear high in the ranking, preferably in the first 30 results. We then compared these selections with the rankings obtained from the bibliographic search tools, keeping in mind that the databases each have their particular features: coverage, indexing and search settings.

**Conclusion**

Relevance judgments are known to differ between experts, and their judgments may also vary at different times. Relevance is a subjective and even volatile notion. This complex topic is sufficiently challenging to be investigated deeper. Apart from NLM, database producers do not publish their algorithms, probably because they are considered trade secrets. Medical librarians should therefore collaborate to research this topic and compare their findings. They should give feedback to the developers to invite them to further refine the ranking function.

**Bibliography**