

# A Novel Tool for Translational Research Discovery

PAUL FONTELO, MD, MPH<sup>1</sup> AND FANG LIU, MS<sup>1</sup>

The Institute of Medicine's Clinical Research Roundtable highlighted two translational blocks that prevent discoveries of promising drugs, diagnostic technologies, or therapeutic procedures that are generated by basic research from getting translated into clinical applications.<sup>1</sup> Success is not the norm, even for research with the most promising results.<sup>2</sup> For basic research that ultimately leads to patient applications, the interval is long, often more than 10 years.<sup>2</sup> Several explanations for the lag were proposed: funding sustainability through the tedious discovery to application process, arduous challenges of validating and ensuring that the product is safe and effective, competing research priorities, and a lack of trained translational investigators and participants to carry on research.<sup>1,2</sup>

A set of filters and limiters, specifically selected to find translational research publications, may be useful in finding promising past and recent novel studies on interventions for specific disease processes

Another explanation could simply be a failure to discover novel research. There are now more than 18 million citations in MEDLINE/PubMed from a subset of about 5,200 select journals. The total number of science journals that may contain potential discoveries is manifold more.

The National Library of Medicine (NLM) at the National Institutes of Health has undertaken a project to facilitate the discovery of published translational science research cited in MEDLINE/PubMed. This project is based on the underlying hypothesis that creating a set of filters and limiters, specifically selected to find translational research publications, may be useful in finding promising past and recent novel studies on interventions for specific disease processes.

The project development began by creating a database of more than 6,800 term combinations (eg, "future therapy," "novel application," and "potential therapeutic target") using vocabulary from known translational articles. These terms were used to search PubMed using E-Utilities. The database was continuously modified to optimize time needed to search and maximize yield of translational publications retrieved. The term combinations were reorganized to a current final list of about 50 search limiters derived through repeated searching and ranking of terms.

## Searching for New Discoveries

The user initiates the search using any "intervention" (drug, chemical, target molecule, gene locus, test, etc) or a specific disease process or condition (autism, asthma, cataract, obesity) or narrows the search using both (prostate cancer AND statins, thrombotic thrombocytopenic purpura AND ADAMTS-13, diabetes AND

thiazolidinedione, etc). The search results list up to 100 of the most recent citations on each page depending on the publication date selected. Links are provided to earlier citations if more than 100 are found.

Using RxNorm (<http://www.nlm.nih.gov/research/umls/rxnorm>), a standardized nomenclature for clinical drugs (14,119 terms) and Medical Subject Heading (MeSH) classification of diseases (34,593 terms), the retrieved PubMed IDs (PMID) are sorted into "Interventions" and "Diseases." Abstracts that are not in either one of the two headings are listed separately. PMIDs may be duplicated in the "Disease" or "Intervention" category. Search terms, interventions, and diseases are highlighted in the abstract.

## Usefulness

Preliminary results are promising. As intended, this tool retrieves publications that are considered translational studies. Information retrieval measurements such as precision and recall are ongoing, but from multiple random reviews of retrieved abstracts, they are expected to be high. The more significant measure will be the usefulness of this resource to the translational science community and beyond. In the future, a suggestion will be submitted to NLM's MeSH Unit on creating an assigned indexing term, a "tag" for future translational research. Comments and suggestions on improving the tool are always welcome. Access the transitional science search tool on the Web at <http://tscience.nlm.nih.gov>.

CTS

## Acknowledgements

*This research was supported by the Intramural Research Program of the National Institutes of Health, National Library of Medicine, and Lister Hill National Center for Biomedical Communications.*

## REFERENCES

1. Sung NS, Crowley WF Jr, Genel M, Salber P, Sandy L, Sherwood LM, Johnson SB, Catanese V, Tilson H, Getz K, Larson EL, Scheinberg D, Reece EA, Slavkin H, Dobs A, Grebb J, Martinez RA, Korn A, Rimoin D. Central challenges facing the national clinical research enterprise. *JAMA*. 2003; 289: 1278-1287.
2. Contopoulos-Ioannidis DG, Ntzani E, Ioannidis JP. Translation of highly promising basic science research into clinical applications. *Am J Med*. 2003; 114(6): 477-84.

<sup>1</sup>Office of High Performance Computing and Communications, Lister Hill National Center for Biomedical Communications, National Library of Medicine, Bethesda, Maryland, USA.

Correspondence: P Fontelo ([fontelo@nlm.nih.gov](mailto:fontelo@nlm.nih.gov))

DOI: 10.1111/j.1752-8062.2009.00139.x