

Martin Gerken, Volker Buschmann, Monika Lelgemann  
HTA-Zentrum in der Universität Bremen

### Hintergrund

Systematische Evidenzdarlegungen wie sie z.B. für Leitlinien und HTA-Berichte benötigt werden erfordern als Grundlage eine optimale Suchstrategie. Für eine solche ist dabei zwischen Genauigkeit (precision) und Vollständigkeit (recall) abzuwägen.

Vordefinierte, validierte Filter zur methodischen Auswahl der Literatur sind etabliert und stehen zur Verfügung [1,2].

Jenseits methodischer Vorgaben sind problemspezifische inhaltliche Filter zu definieren. Typische Unklarheiten wie die unscharfe Definition einer Erkrankung oder die nicht ausreichend präzise oder fehlerhafte Verschlagwortung sind zu beachten.

Die für die Entwicklung erforderlichen probatorischen Suchen, in denen Kombinationen verschiedener Suchbegriffe getestet werden, sind von Hand selbst mit „History“-Funktion sehr zeitaufwendig.

### Matrixsuche / Stand der Dinge

Die Matrixsuche ist eine Methode des database minings und meint die Datenbanksuche nach zwei Listen von Suchbegriffen und die Darstellung des Zusammentreffens als Häufigkeitsmatrix. Dabei wird jeder Suchbegriff aus der ersten Liste mit jedem Suchbegriff aus der zweiten Liste mit logischem UND verknüpft, in der Datenbank gesucht und mit der Trefferanzahl als Kreuztabelle dargestellt.

Zur automatisierten Literatursuche gibt es ein Online-Tool [3] welches die Matrixsuche in der frei zugänglichen Datenbank MEDLINE über Pubmed implementiert. Dieses Werkzeug wurde bisher hauptsächlich im Bereich Genetik und Bioinformatik eingesetzt, um z.B. eine Matrixsuche zwischen Genorten und Erkrankungen durchzuführen.

### Matrixsuche / Weiterentwicklung

Entwicklung eines webbasierten Werkzeugs mit verbesserter Funktionalität zur automatisierten Suche in Medline (über Pubmed) durch Einführung einer dritten „Suchdimension“

#### Suchmaske

The screenshot shows the XSearch Pubmed web interface. It features three input fields for search filters: 'Hier Suchbegriff #1 (Zeilen) eingeben', 'Hier Suchbegriff #2 (Spalten) eingeben', and 'Hier Suchbegriff #3 (Tabelle) eingeben'. A 'Suchen' button is located at the bottom right of the filter area. Above the filters, there is a message: 'Die drei Suchfilter werden kreuzweise gegeneinander logisch verknüpft und die Trefferzahl in der 3x3 Matrix Datenbank ausgegeben. Bitte nur maximal drei "Suchen"-Kнопк drücken, da Sucher langsam eine Vorwarnung.' Below the filters, there is a 'Suchen' button.

#### Literatur

- [1] Haynes RB et al. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ*. 2005 May 21;330(7501):1179. Epub 2005 May 13.
- [2] <http://www.ncbi.nlm.nih.gov/entrez/query/static/clinical.shtml>
- [3] <http://pubmatrix.grc.nia.nih.gov/> (Becker KG et al. PubMatrix: a tool for multiplex literature mining. *BMC Bioinformatics*. 2003; 4: 61)
- [4] <http://www.meshpubmed.org>
- [5] <http://www.gpubmed.org> (Doms A, Schroder M: GoPubMed: exploring PubMed with the Gene Ontology. *Nucleic Acids Res*. 2005 Jul 1;33(Web Server issue):W783-6)
- [6] <http://www.chilibot.net/> (Chen H, Sharp BM Content-rich biological network constructed by mining PubMed abstracts. *BMC Bioinformatics*. 2004; 5: 147)

### Weiterentwicklung / praktische Anwendung

Orientierende Datenbankrecherche zum Thema Asthma bronchiale bei Kleinkindern mit folgenden Fragen / Problemen:

- unscharfe Krankheitsdefinition
- unklare Verlässlichkeit der Indexierung nach Altersgruppen
- Abschätzung der Studienzahl in Abhängigkeit vom Publikationstyp bzw. den verwendeten methodischen Suchfiltern gewünscht.

#### Ergebnistabelle

The screenshot shows a search results table with columns for 'Preschool Child', 'Child OR Infant', and 'Child OR Infant'. The table lists various search terms and their corresponding counts. Annotations include: 'RCT spezifisch' pointing to the first row, 'RCT sensitiv' pointing to the second row, 'Systematic Review' pointing to the third row, and 'Trefferzahl' pointing to the total count in the last row. A yellow starburst highlights the 'Trefferzahl'.

Suchbegriff	Preschool Child	Child OR Infant	Child OR Infant
randomized controlled trial[Publication Type] OR randomize[Title: Abstract] AND crossover[Title: Abstract] AND trial[Title: Abstract]	669	1214	1872
Asthma OR Wheeze OR Cough	790	1222	1812
Asthma OR Wheeze OR Cough OR Respiratory Hypersensitivity OR Bronchial Hyperresponsivity OR Bronchial Spasm OR Respiratory Sounds	839	1867	2165
Arthritis OR Wheeze OR Cough	562	1148	1216
Asthma OR Wheeze OR Cough OR Respiratory Hypersensitivity OR Bronchial Hyperresponsivity OR Bronchial Spasm OR Respiratory Sounds	5299	11931	11931
Systematic[SR]	189	424	503
Asthma	191	491	509
Asthma OR Wheeze OR Cough	215	551	584
Asthma OR Wheeze OR Cough OR Respiratory Hypersensitivity OR Bronchial Hyperresponsivity OR Bronchial Spasm OR Respiratory Sounds	224	572	605
XSearch-01.03.2007			1642

### Praktische Anwendung / Ergebnisse

- Darstellung des Suchergebnisses in Tabellenform
- HTML Format ermöglicht einfache Übernahme
- Automatisierte Verlinkung zur Quell-Datenbank
- Export und Speichern über Quell-Datenbank
- Werkzeug steht im Internet frei zur Verfügung <http://www.hta.uni-bremen.de/xsearch>

### Anwendungsoptionen

Mit dem weiterentwickelten Online-Tool lassen sich vorbereitende Arbeiten zur Entwicklung von Suchstrategien optimieren, der Zeitaufwand wird relevant reduziert.

Die automatisierte Suche ermöglicht die vereinfachte Ermittlung der Trefferzahlen bei Kombination und Variation verschiedener Suchbegriffe.

Im Rahmen systematischer Evidenzaufbereitung ermöglicht die Webanwendung die Aufwandsabschätzung.

Dieses Verfahren ergänzt:

- ontologie-basierte Suchhilfen (z.B. MeshPubmed [4], GoPubMed [5])
- prädefinierte methodische Filter (z.B. „clinical queries“ [2])
- automatische Relationsanalysen (z.B. Chilibot [6]).

### Kontaktadresse

Martin Gerken  
Interdisziplinäres Zentrum für Health Technology Assessment in der Universität Bremen  
martin.gerken@uni-bremen.de