**Background:**

NCBI hosts XX database, and each database contains literature from journals pertinent to various sub-domains. Pubmed and PMC are two databases containing a wealth of public health literature, and search tools for those databases were developed first, with others (e..g Scopus, Springerlink) in the pipeline.

Methods

The tool enables several functions: searching for and storing literature in a relationship database; and querying the database.

**Searching for literature and storing in relationship database:**

**Search:**

User-input to the tool includes the database to be searched, search terms, and search term modifiers for advanced searching.

A program is run to prompt the user for search parameters, which are then copied to a text file. When the main program is run, the search parameters are read from the parameter text file.

**Fetch:**

Fetching IDs and XML documents

Parse:

Push:

Papers that have already appeared in previous searches are marked as duplicates, and the parsed data is not pushed to the database twice; they are simply marked as duplicates in the database.

**Database schema:**

The data in fields parsed from the literature are housed in a relational database containing six tables: and ID, Detail, Keyword, Title, Author and Text. The ID table stores the information on a particular search, including a unique Pull ID, the search term(s), the pull time, pull source (e.g., the NCBI database), and pull by (the user).

The Detail table links to the ID table by Pull ID, and contains more information on that piece of literature including the unique ID native to the database (e.g., the PMID), the publication type, and the “store” type – that is, whether the piece of literature has been return already from a previous search and is a duplicate.

Publication type can include your typical research paper as well as conference literature.

**Querying the data:**

Querying to identify novel or emerging findings:

* The publication type field enables users to search specifically for conference literature
* Users can do an evaluation of the search terms used by querying based on the search parameters stored in the