

Submitter Name: Mustafa Hakan Gunturkun
Submitted email: hakan@uthsc.edu
PI Name: Hao Chen
PI email: hchen@uthsc.edu

RatsPub: a webservice aided by deep learning to mine PubMed for addiction-related genes

Mustafa Hakan Gunturkun¹, Efraim Flashner², Tengfei Wang¹, Megan K. Mulligan², Robert W. Williams², Pjotr Prins², Hao Chen¹

¹Department of Pharmacology, Addiction Science and Toxicology, ²Department of Genetics, Genomics and Informatics, University of Tennessee Health Science

Interpreting and integrating results from genetic and epigenetic studies on addiction phenotypes require a comprehensive survey of the extant literature. Most often this is conducted by *ad hoc* queries of the PubMed database. Here, we introduce RatsPub, a literature mining web service that searches user-provided gene symbols in conjunction with a set of systematically curated keywords related to addiction, as well as results from human genome-wide association studies (GWAS). We have organized over 300 keywords into seven categories forming an ontology (Suggestions on additional keywords are welcome). The literature search is conducted by querying the NIH PubMed server using a programming interface, which is followed by retrieving abstracts from a local copy of the PubMed archive. The main results presented to the user are individual sentences containing the gene symbol, organized by the keywords they also contain. These sentences are presented through an interactive graphical interface or as tables. GWAS results are displayed using a similar method. All results are linked to the original abstract in PubMed. In addition, a convolutional neural network is employed to distinguish sentences describing systemic stress from those describing cellular stress. The automated and comprehensive search strategy provided by RatsPub facilitates the integration of new discoveries from addiction omic studies with existing literature and improves analysis and modeling in addiction biology. RatsPub is free and open source software. The source code of RatsPub and the link to a running instance is available at <https://github.com/chen42/ratspub>.