

```
graph LR; A[PMC website] --> B[Build query]; B --> C["Music[Title] AND (fMRI[Abstract] OR functional magnetic resonance imaging[Abstract])"]
```

PMC website

Build query

"Music[Title] AND (fMRI[Abstract] OR functional magnetic resonance imaging[Abstract])"

pubget

Format URL
[https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=PMC&term=Music\[Abstract\]+AND+\(fMRI\[Abstract\]+OR+functional+magnetic+resonance+imaging\[Abstract\]\)&usehistory=y](https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=PMC&term=Music[Abstract]+AND+(fMRI[Abstract]+OR+functional+magnetic+resonance+imaging[Abstract])&usehistory=y)

Retrieve journal articles as nested XML

```
<?xml version='1.0' encoding='UTF-8'?>
<article xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:rml="http://www.w3.org/1998/Math/MathML" article-type="systematic-review">
  <?properties open_access?>
  <front>
    <journal-meta>
      <journal-id journal-id-type="nlm-ta">Front Hum Neurosci</journal-id>
      <journal-id journal-id-type="iso-abbrev">Front Hum Neurosci</journal-id>
      <journal-id journal-id-type="publisher-id">Front. Hum. Neurosci.</journal-id>
      <journal-title-group>
        <journal-title>Frontiers in Human Neuroscience</journal-title>
      </journal-title-group>
    </journal-meta>
  </front>
</article>
```

Output tabular data

pmcid	title	keywords	abstract	body
		evidence-based MRI	Despite the variability in	Despite the variability in
7903841	Clinical Anatomy of the Human	mFL		
		mild cognitive decline		
		ASL MRI	We sought to systematically	In this comprehensive
9776573	The Utility of Arthroscopy			
		Microcalcification		
		Magnetic resonance imaging	Up to a pretest probability of 0.5	None of the authors

Output texts ready for annotation

```
1[["text": "P\u011919k\u0105la, Przemys\u0142aw A. and \u0141141azarcz, Dominik P. and Rosa, Mateusz A. and P\u011919k\u0105la, Jakub R. and Baginski, Adam and G\u00f3b\u0142, Alberto and Wojciechowski, W\u0142\u0105d and Tomaszewski, Krzysztof A. and Lafrade, Robert F. and J Sports Med. 2021;30(1):1-10. doi: 10.1093/ncnl/ncn1. Anatomy of the Posterior Meniscocofemoral Ligament of Wrisberg: An Original MRI Study, Meta-analysis, and Systematic
```

labelbuddy

Ground truth

Annotate texts

labelbuddy — review-neuro-meta-analyses.lb

File Preferences Help

Annotate Labels & Documents Import & Export

Set label for selected text:

stopped-here no-access candidate for replication **N studies found** check for n studies found

pmid: 34060940

Luna, Licia P and Sherbaf, Farzaneh Ghazi and Sair, Haris I and Mukherjee, Debraj and Oliveira, Isabella Bezerra and Köhler, Cristiano André Radiology, 2021

The diagram illustrates the process of extracting structured data from a text snippet. On the left, a text snippet is shown: "...We had 19 healthy participants in this study (10 male; age range 20-28...)". An arrow points from this text to a table on the right. Above the table, the text "Train/test/use methods for extracting further data" is displayed. The table has four columns: "count", "females", "count", "males", and "count". The rows of the table are: 19, 10, 18, 67, 9, 7, 9, and an empty row.

pubextract

Train/test/use methods for extracting further data

"...We had 19 healthy participants in this study (10 male; age range 20-28...)"

count	females	count	males	count
19		9		10
10		3		7
18		9		9
67				

custom code

The graph, titled "Analyze data", plots the "Median sample size" (y-axis, 0 to 50) against the "Publication year" (x-axis, 1996 to 2020). It compares four data sources:

- David & al. [2013] (green dotted line)
- Poldrack & al. [2017] (blue dashed line)
- pubextract [in 2023] (orange solid line)
- GPT-3 [in 2023] (red dashed line)

Shaded regions around the lines represent confidence intervals. The median sample size generally increases over time, with a notable sharp increase for GPT-3 around 2012, reaching a median of approximately 30 by 2020.

Figure 1: Median sample size of the data source used in each paper. The graph shows the median sample size (y-axis, 0 to 50) versus publication year (x-axis, 1996 to 2020). Four data sources are compared: David & al. (2013) (green dotted line), Poldrack & al. (2013) (blue dashed line), pubextract (in 2023) (orange solid line), and GPT-3 (in 2023) (red dash-dot line). Shaded areas represent the distribution of sample sizes. pubextract and GPT-3 show a steady increase in median sample size over time, while Poldrack & al. and David & al. show more fluctuation, with Poldrack & al. peaking around 2015 and David & al. peaking around 2011.

