

# Comparing the National Library of Medicine (NLM)'s Medical Text Indexer (MTI) to Human Indexing: A Pilot Study

Eileen Chen, MLIS

*School of Information, University of British Columbia*

Dr. Julia Bullard, MA, MLIS, PhD

*School of Information, University of British Columbia*

Dean Giustini, MLS, MEd

UBC Biomedical Branch Library, Vancouver Hospital



# Background on the Medical Text Indexer (MTI):

MTI was developed in 2002 by the National Library of Medicine (NLM) & Lister Hill National Centre for Biomedical Communications.

*There have been three versions:*

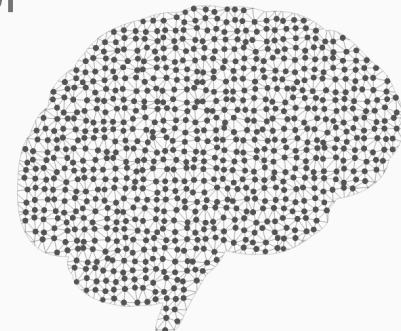


- **MTI (2002)**: term recommender for human indexers
- **MTI First Line (2011-2021)**: semi-automated MTI assisted by human indexers
- **MTI-Auto (2022)**: fully automated indexer (extent of human curation unknown)

*For more information (abstract, data, glossary, etc) on this project, see: <https://osf.io/4k69q/>*

# MTI-Auto (2022):

- Most recent version, but not publicly-available for testing
- *Pattern-based indexing* based on titles & abstracts of papers
- Machine learning used for applying subheadings
- Human reviewers perform quality assurance reviews for '*select citations*' ...



For more information on this project, see: <https://osf.io/4k69q/>

# Aims:

- *To compare MTI with human indexing...*
- Evaluate indexing quality in high vs. low-impact biomedical journals indexed in Medline (PubMed)
- Identify MTI errors, and anomalies in assigning MeSH terms & check tags

For more information on this project, see:  
<https://osf.io/4k69q/>

# Method used for sample:

- *Selected* 20 biomedical articles published in year 2000 (before MTI was created)
- *Identified* key journals from the ‘Abridged Index Medicus’ (AIM) = a journal subset of *Medline (PubMed)*
  - Of 120+ AIM journals, 10 with the highest 2020 Journal Impact Factor (JIF) AND 10 with the lowest JIF were chosen (N=20)
- *Excluded* articles without abstracts or MeSH indexing

# Interactive MTI Tool

*...is a free online tool provided by the NLM. The version of the MTI available is the MTIFL (retired by the NLM in 2021)...*

## Two Output Options:

- 1) *Just the Facts (JTF):*
  - a) Shorter list of ranked terms that represents MTI's final choices
- 2) *Full Listing:*
  - a) Longer list of all terms retrieved by MTI, ranked & explained
  - b) Includes confidence scores, MeSH type, and pathway(s) used to retrieve term

# Results

# Assigned Index Terms – Mean #?

- **MTI and humans** created **more index terms** for high-JIF group than low-JIF group
- Difference was greater for MTI (6.4 terms)
- **Journals with most MTI terms?**
  - *Lancet* (26), *JAMA* (21), *Blood* (21),  
*Annals of Internal Medicine* (21)
- **Journals with least MTI terms?**
  - *Nursing Clinics of North America* (4),  
*Journal of Nursing Administration* (5),  
*Journal of Laryngology and Otology* (7)

	<b>MTI assigned:</b>	<b>Human assigned:</b>
Top 10 JIF journals:	16.6 terms	13.5
Lowest 10 JIF journals:	10.2 terms	11.2
<b>Difference?</b>	<b>6.4 terms</b>	<b>2.3</b>

# Main Headings

- Of a total of **174** main headings used by humans for 20 articles, MTI included:
  - **80** in JTF list
  - **92** in Full Listing
  - Missed **2** altogether.
- In **19** instances, the MTI used an acceptable synonym to a human-indexed term.

The recall rate for relevant terms is high in the Full Listing, but many relevant terms are not ranked highly enough.



## Case Study: MTI vs. Human Terms

- *The MTI missed several major headings*
- *The word “attention” was misinterpreted literally:*  
  
*“The concept of nursing practice models [...] has attracted the attention of nursing administrators in the last decade...” (Upenieks 2000)*

## Article #17: The relationship of nursing practice models and job satisfaction outcomes

Shared Terms	MTI Only	Human Only
<i>Humans (0); Job Satisfaction* (1)</i>	Social Responsibility (2); Climate Change (3); <b>Attention</b> (4)	<b>Models, Nursing*</b> (5); <b>Nursing*</b> (8); <b>Outcome Assessment, Health Care*</b> (31); <b>United States</b> (53)

Italics = check tags; \* = human-indexed major heading; () = MTI rankings in Full Listing

# Check Tag (Age, Sex, Species) Coverage

- Of **72** check tags used by humans across sample (N=20) articles, MTI shared **38**.
- Of remainder, **30** appeared in Full Listing, and **4** were missed altogether.
- MTI used **5** check tags not used by humans, 4 of which were appropriate choices.

The check tag “Aged” was missed in 4 instances.

<input type="checkbox"/> V Human	<input type="checkbox"/> U Animal
<input type="checkbox"/> W Male	<input type="checkbox"/> Q Mice
<input type="checkbox"/> X Female	<input type="checkbox"/> T Rats
<input type="checkbox"/> A Pregn	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/> J Cats
<input type="checkbox"/> B Inf New (to 1 mo)	<input type="checkbox"/> K Cattle
<input type="checkbox"/> C Inf (1 to 23 mo)	<input type="checkbox"/> L Chick Embryo
<input type="checkbox"/> D Child Pre (2-5)	<input type="checkbox"/> M Dogs
<input type="checkbox"/> E Child (6-12)	<input type="checkbox"/> O Guinea Pigs
<input type="checkbox"/> F Adolesc (13-18)	<input type="checkbox"/> P Hamsters
<input type="checkbox"/> R Young Adult (19-24)	<input type="checkbox"/> S Rabbits
<input type="checkbox"/> G Adult (19-44)	<input type="checkbox"/>
<input type="checkbox"/> H Mid Age (45-64)	<input type="checkbox"/>
<input type="checkbox"/> I Aged (65-79)	<input type="checkbox"/>
<input type="checkbox"/> N Aged (80+)	<input type="checkbox"/>

# Sex Check Tags

**Male (check tag)** was ranked higher in all 6 instances in which *male/female* check tags were used.

Why is there such a bias in the sample?

## Sex Check Tag Rank in MTI Full Listing

Article	Male	Female	Difference (F - M)
1	0*	3rd*	3 ranking places
2	5*	8*	3
13	15	50	35
16	1*	4*	3
18	12*	60*	48
20	9*	70*	61
Mean	7	32.5	25.5

Bolded = included in JTF list of MTI

\* = labelled as a check tag

# Sex Check Tags: MTI vs. Human Differences

Article Title	MTI check tags	Human check tags
Hypertensive emergencies	<i>Pregnancy</i> [0]; <i>Female</i> [1]; <i>Humans</i> [2]	<i>Humans</i>
Application of the Woman Abuse Screening Tool (WAST) and WAST-short in the family practice setting	<i>Humans</i> [0]; <i>Male</i> [1]; <i>Adult</i> [2]; <i>Middle Aged</i> [3]; <i>Female</i> [4]	<i>Adult, Female, Humans, Middle Aged</i>
A comparison of performance on the OMSITE and ABOMS written qualifying examination	<i>Humans</i> [0]; <i>Male</i> [1]; <i>Female</i> [2]	<i>Humans</i>

**MTI can make erroneous assumptions based on populations suggested in abstract.**

# Summary of Findings:

- In sample (N=20), more MeSH terms & accuracy were seen in the **high-JIF articles from 2000**
- High retrieval rates for human-indexed main headings & check tags; however, MTI ranking mechanisms were not consistently reliable
- Check tags reflect a certain bias for **male populations that are not aged**
- *More frequent & accurate use of **medical, operationalizable MeSH terms** than **social and emotional concepts / MeSH***

# Implications:

- Check tag inaccuracy is related to MTI processing abstracts rather than full texts (Mork et al., 2017)
- MTI output would benefit from greater degree of indexer review
- PubMed/Medline end users are encouraged to report problems to [NLM Support Center](#)

# Limitations & Future Research:

## Limitations of this research:

- Our small sample of articles in Medline **does not** yield strong, generalizable findings... & it therefore **cannot** represent all Medline articles
- The Interactive MTI tool we used is dated & *may not* be representative of MTIA (2022) performance

## Future research / directions:

- Monitor & track indexing biases & anomalies
- Collaborate with other scholars, researchers
- Involve indexers / subject experts in projects comparing MTIA indexing to past human indexing

# References

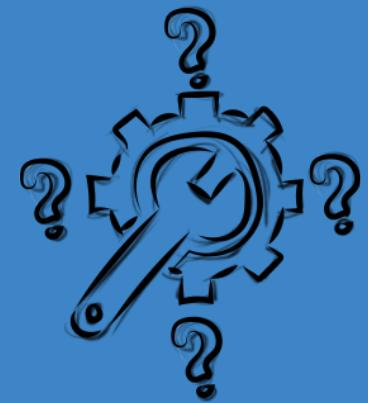
Mork, J., Aronson, A., & Demner-Fushman, D. (2017). 12 years on—Is the NLM medical text indexer still useful and relevant? *Journal of Biomedical Semantics*, 8(1), 8. <https://doi.org/10.1186/s13326-017-0113-5>

National Library of Medicine. (n.d.) *NLM Medical Text Indexer*. Retrieved August 25, 2022, from <https://lhncbc.nlm.nih.gov/ii/tools/MTI.html>

National Library of Medicine. (2017) *Check Tags*. Retrieved August 25, 2022, from [https://www.nlm.nih.gov/bsd/indexing/training/CHK\\_010.html](https://www.nlm.nih.gov/bsd/indexing/training/CHK_010.html)

National Library of Medicine. (2022). Frequently Asked Questions about Indexing for MEDLINE. Retrieved August 25, 2022, from <https://www.nlm.nih.gov/bsd/indexfaq.html#selected>

For a complete bibliography of our project, see <https://osf.io/4k69q/>



# Questions?

Eileen Chen, MLIS  
School of Information, UBC  
[eileen.0415@livemail.tw](mailto:eileen.0415@livemail.tw)

Dr. Julia Bullard, MA, MLIS, PhD  
School of Information, UBC  
[julia.bullard@ubc.ca](mailto:julia.bullard@ubc.ca)

Dean Giustini, MLS, MEd  
UBC Biomedical Branch Library, UBC  
[dean.giustini@ubc.ca](mailto:dean.giustini@ubc.ca)

Our project on OSF <https://osf.io/4k69q/>

