

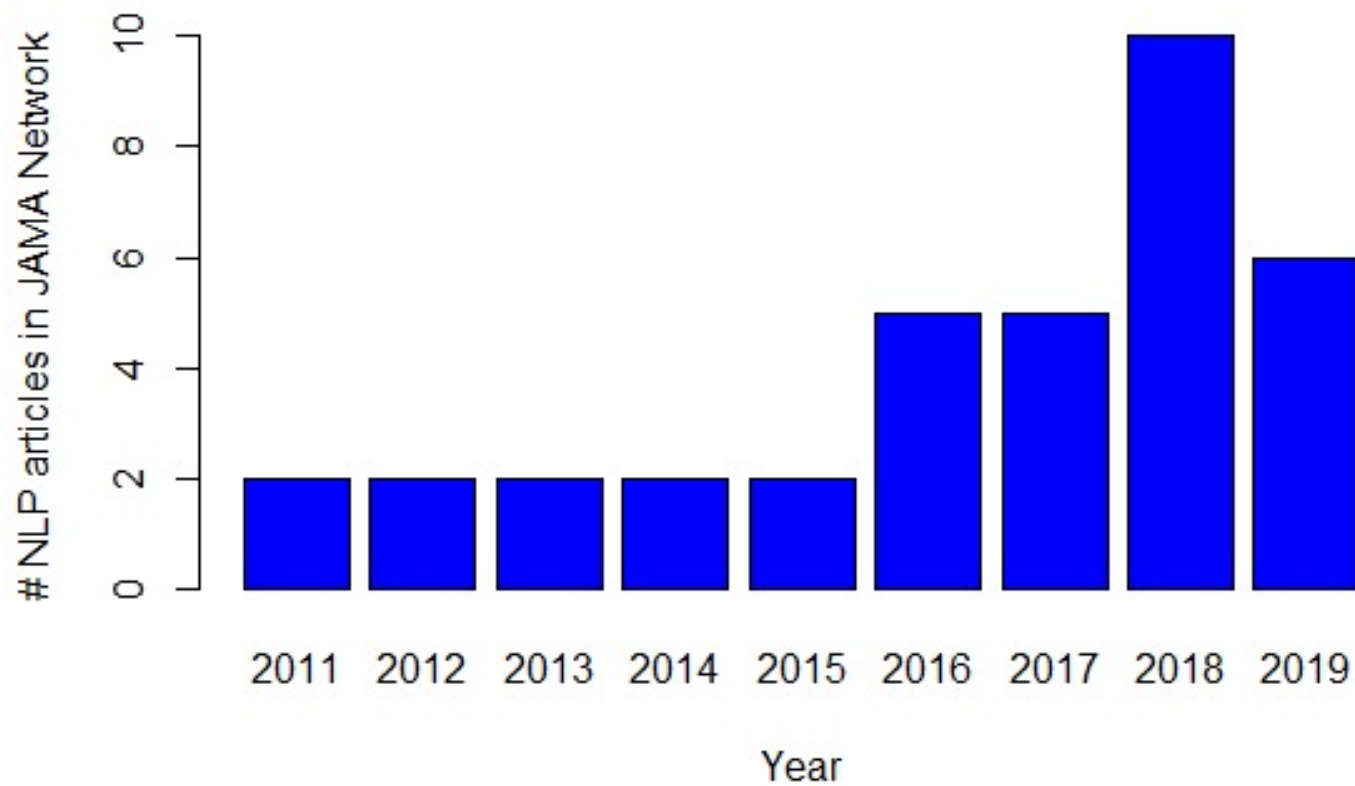
# Biomedical NLP in Practice

Matthew Engelhard

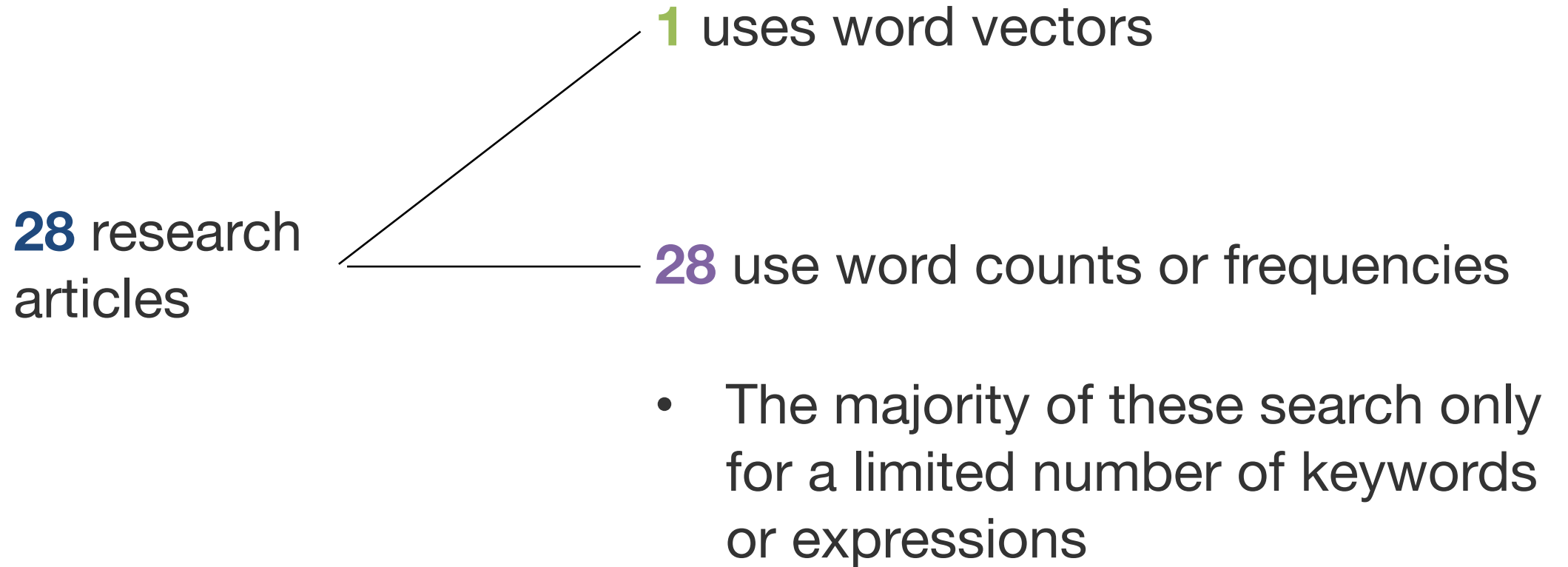
# Bag of Words

...is common and effective for biomedical NLP

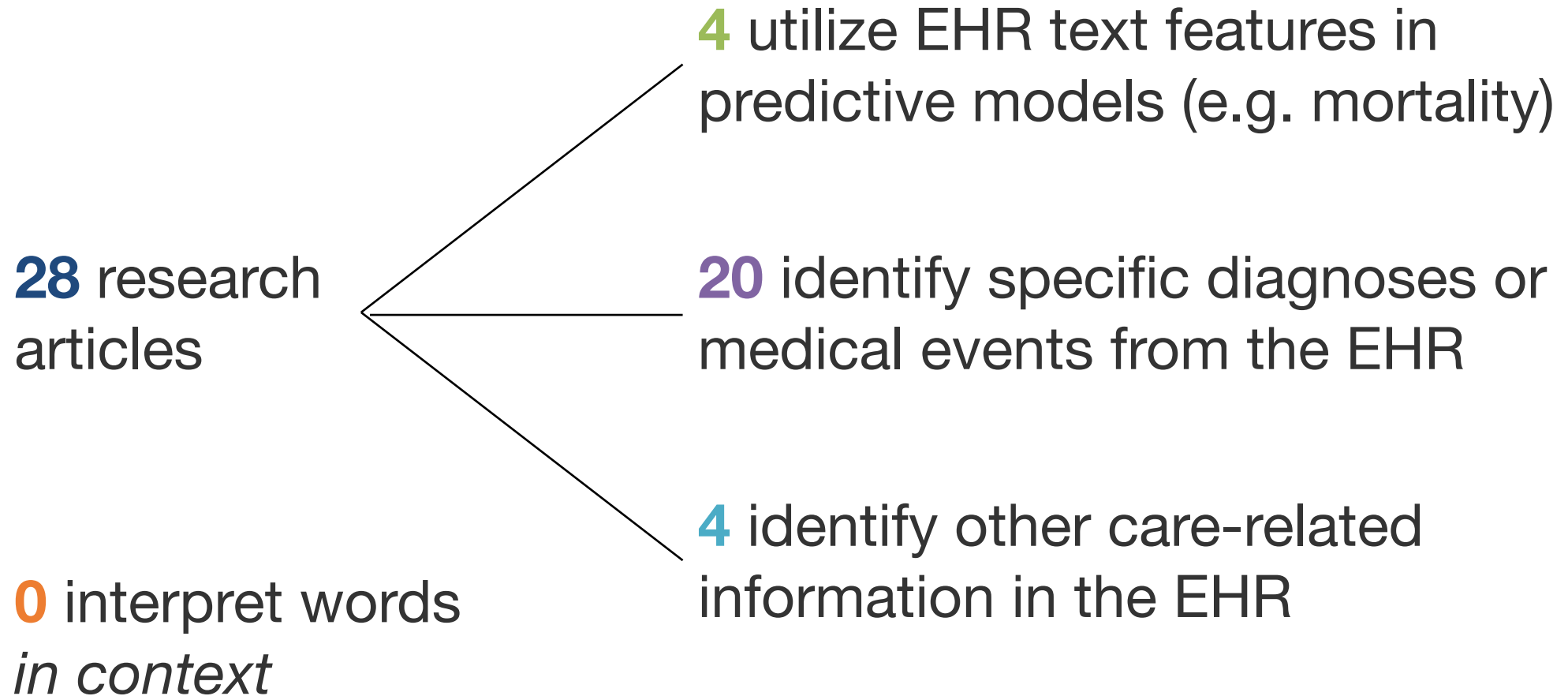
# A Survey of NLP in JAMA...



# A Survey of NLP in JAMA...



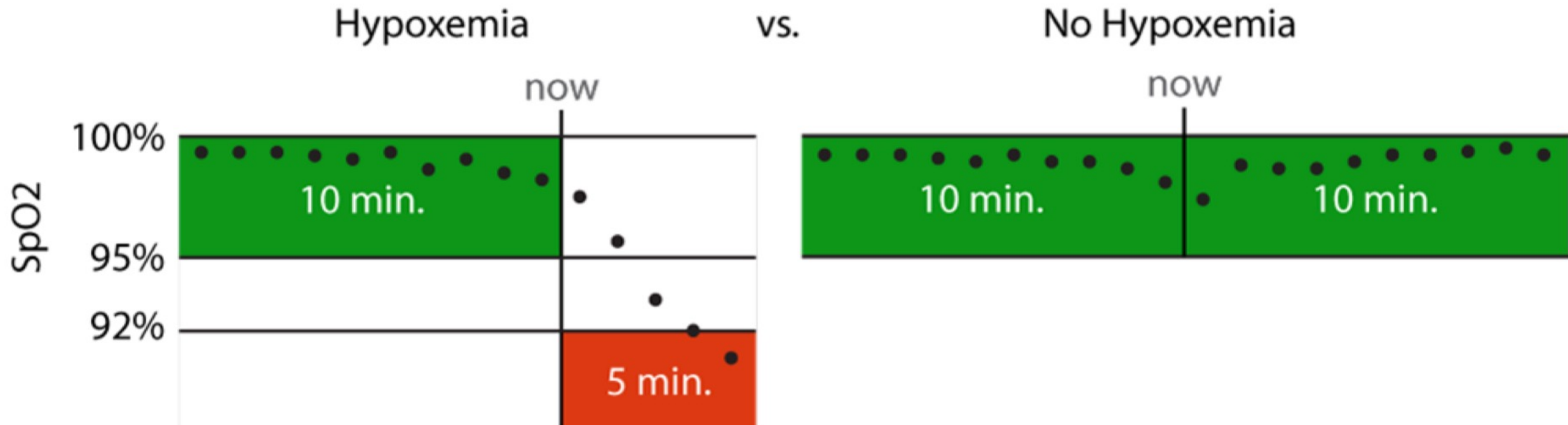
# A Survey of NLP in JAMA...



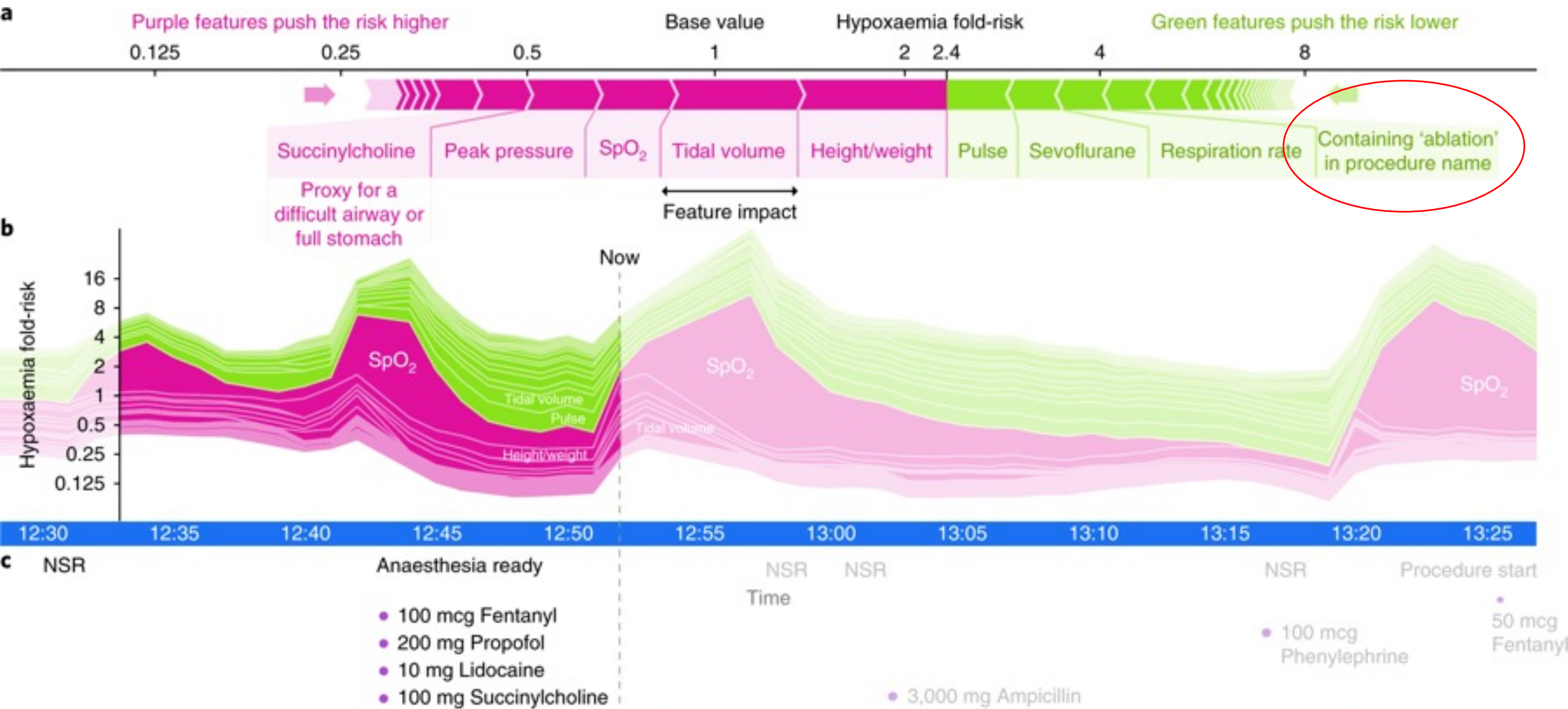
# Hypoxemia Prediction during Surgery

## Real-time Prediction Task:

- hypoxemia (yes/no) in the next 5 minutes
- based on data from the Anesthesia Information Management System
- static features + real-time features collected up to that time point



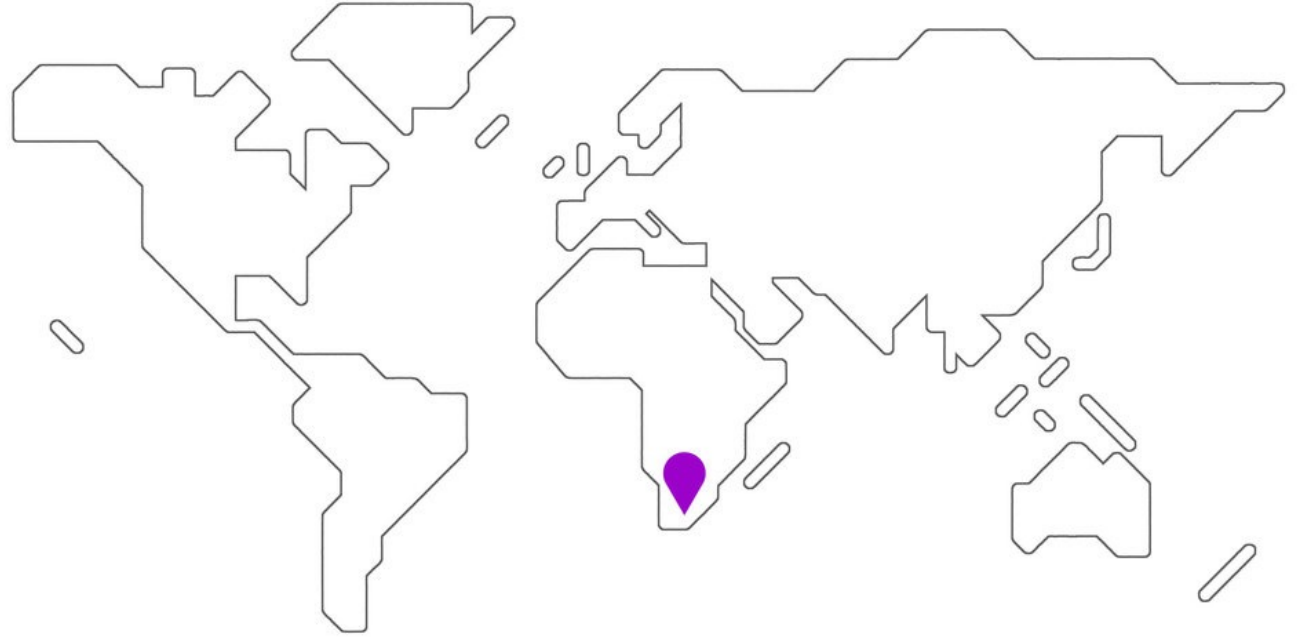
# A majority of features are keyword counts



# Global Maternal Health

**Maternal Health HelpDesk:**

**2 million women connected to  
NDoH staff via SMS**



<https://www.praekelt.org>

**Binary Classification: Urgent Message? (Yes/No)**



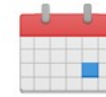
# Suggested Email Responses



Messaging



Health



Appts & Visits



Questionnaires

## Message Center

[ASK A QUESTION](#)

Inbox Sent Messages



Sort by:

Received Date

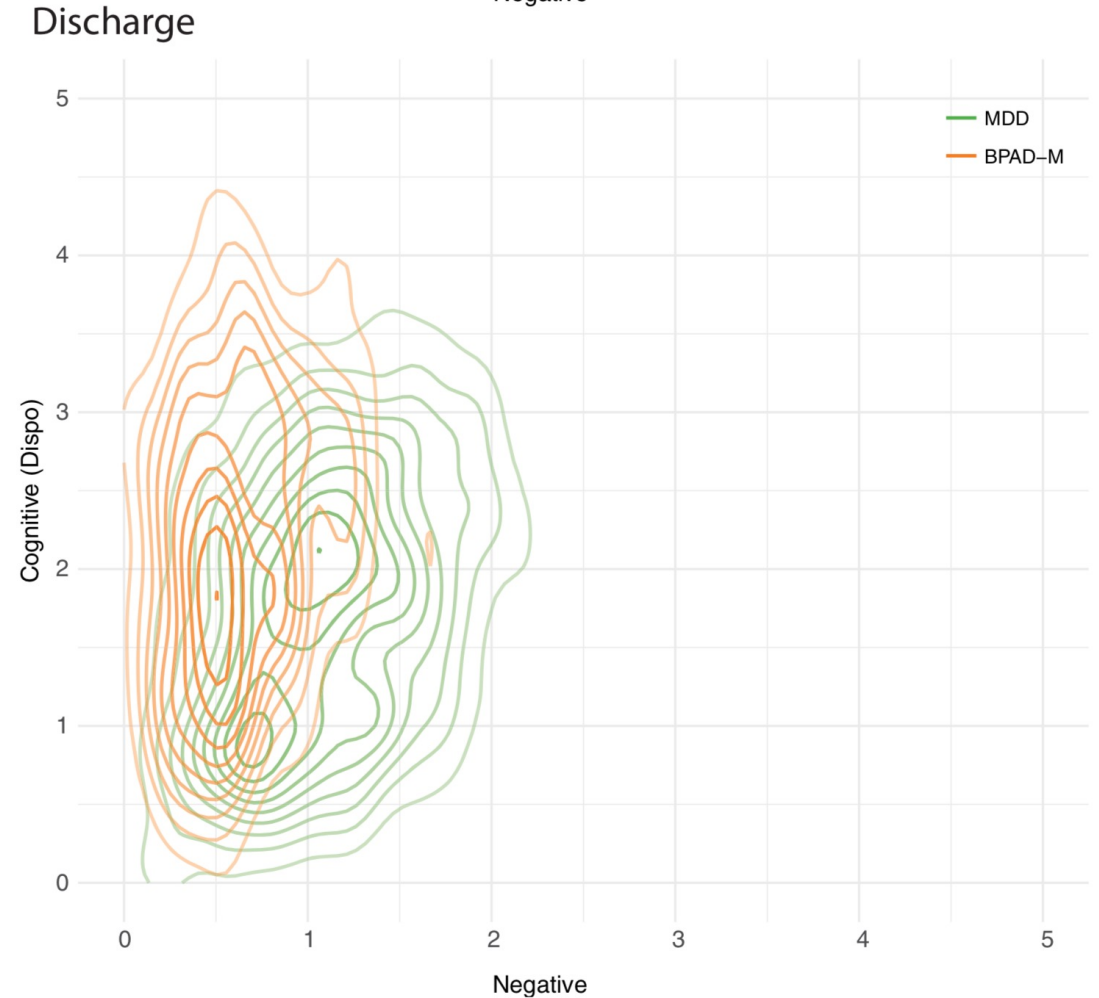
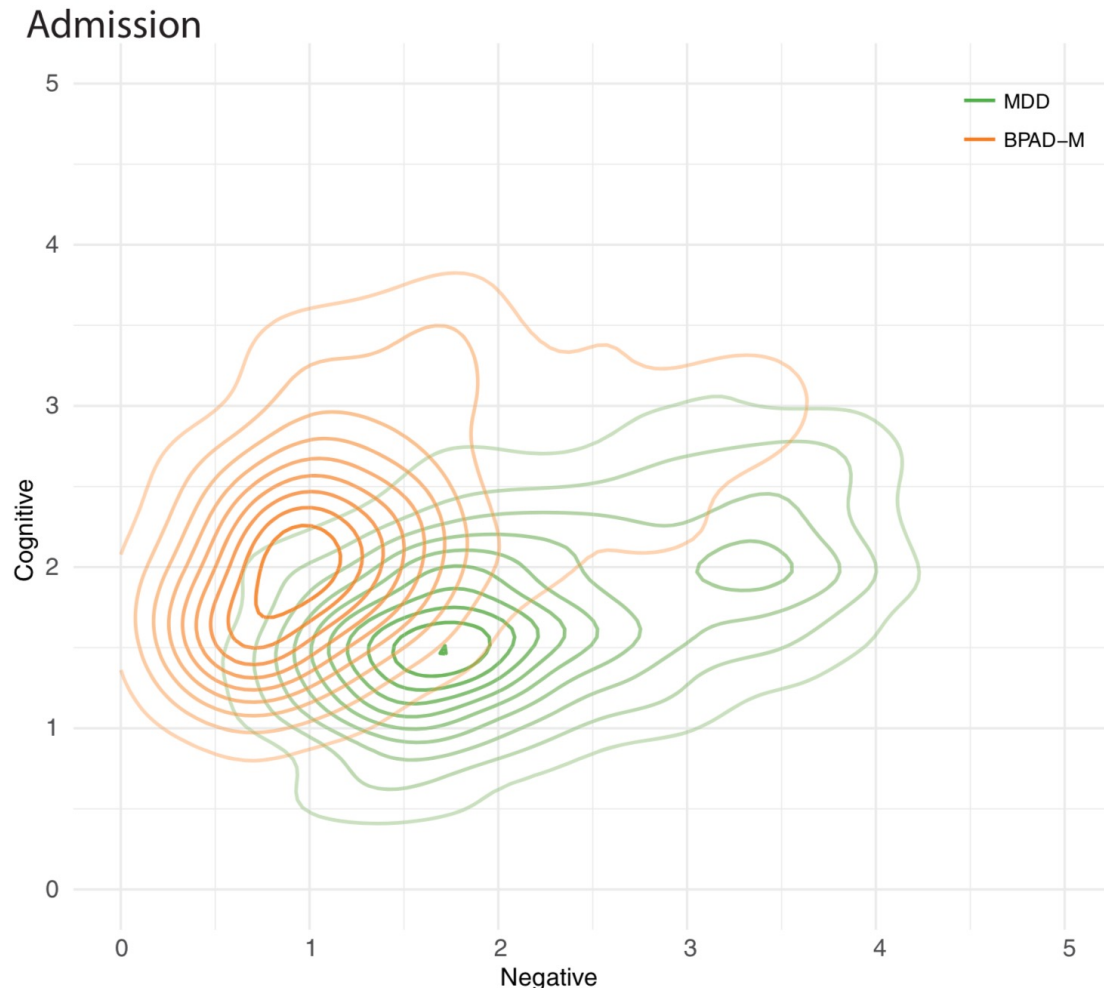


Filters:

All Messages



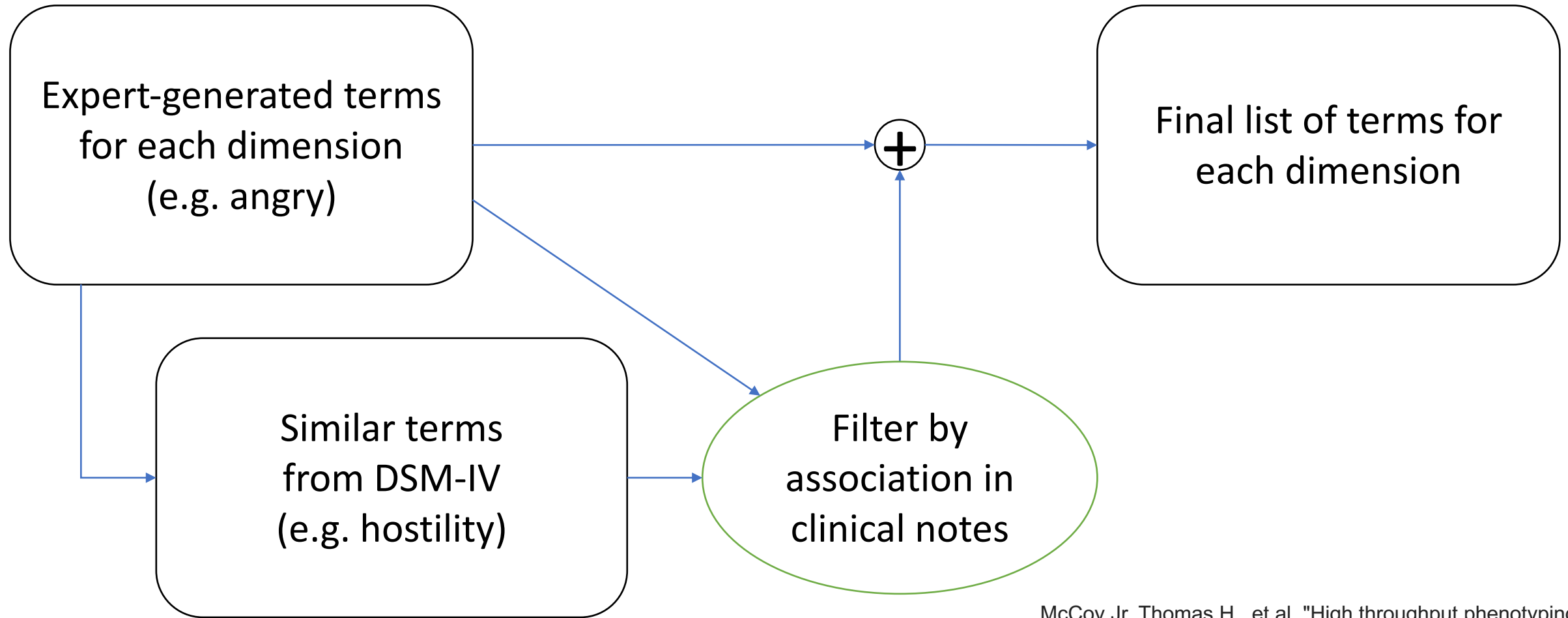
# Trends in psychopathology during stay



McCoy Jr, Thomas H., et al. "High throughput phenotyping for dimensional psychopathology in electronic health records." *Biological psychiatry* 83.12 (2018): 997-1004.

**Figure 1.** Domain comparison contour plots showing change between admission (top) and discharge (bottom). BPAD-M, bipolar disorder–mania; MDD, major depressive disorder.

# Scoring Dimensional Psychopathology

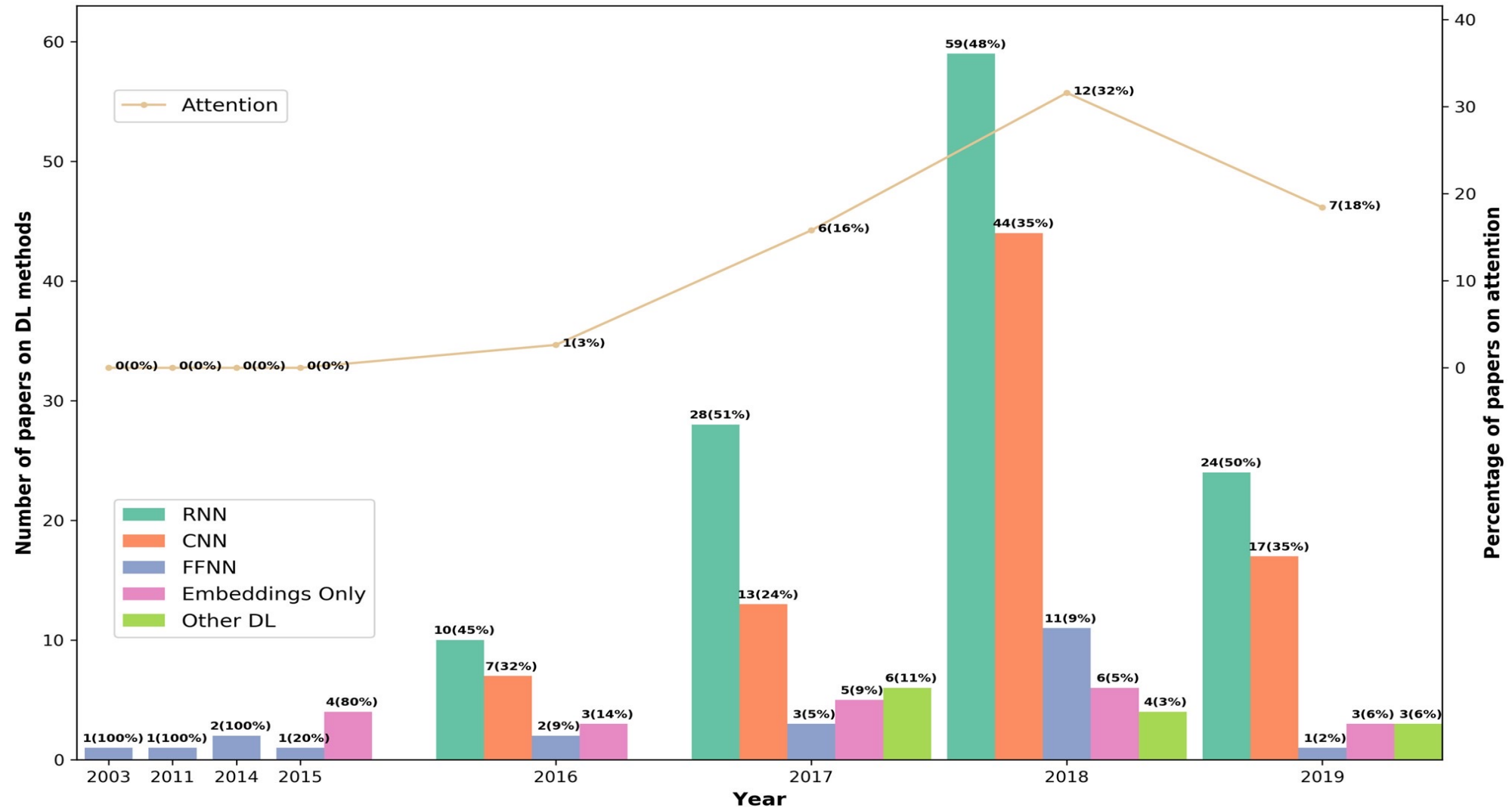


McCoy Jr, Thomas H., et al. "High throughput phenotyping for dimensional psychopathology in electronic health records." *Biological psychiatry* 83.12 (2018): 997-1004.

# Deep Learning

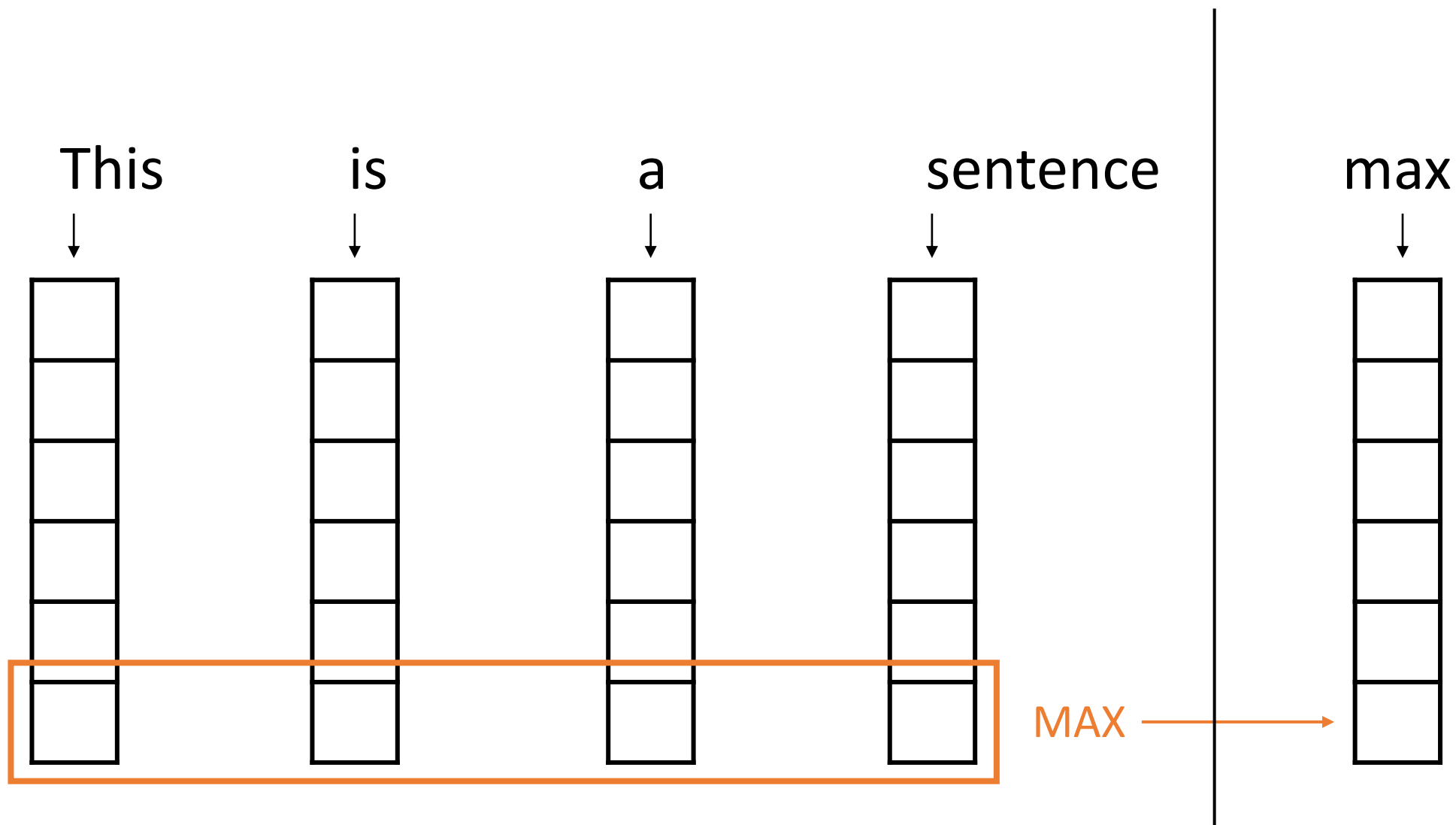
...is on the rise for biomedical NLP

# Rise of Deep Learning NLP in the Clinical Literature

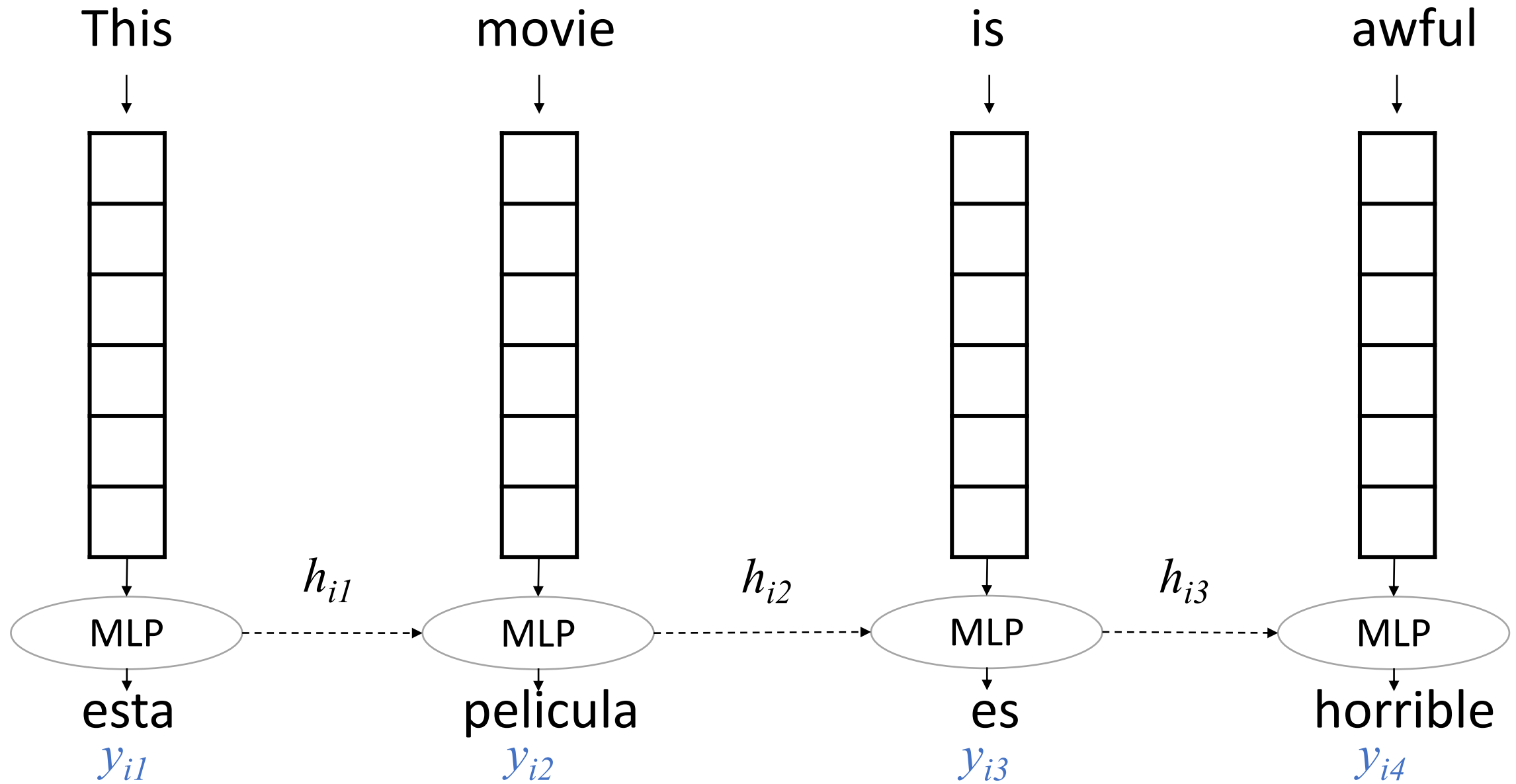


# “Embeddings Only” and feedforward neural network (i.e. MLP)

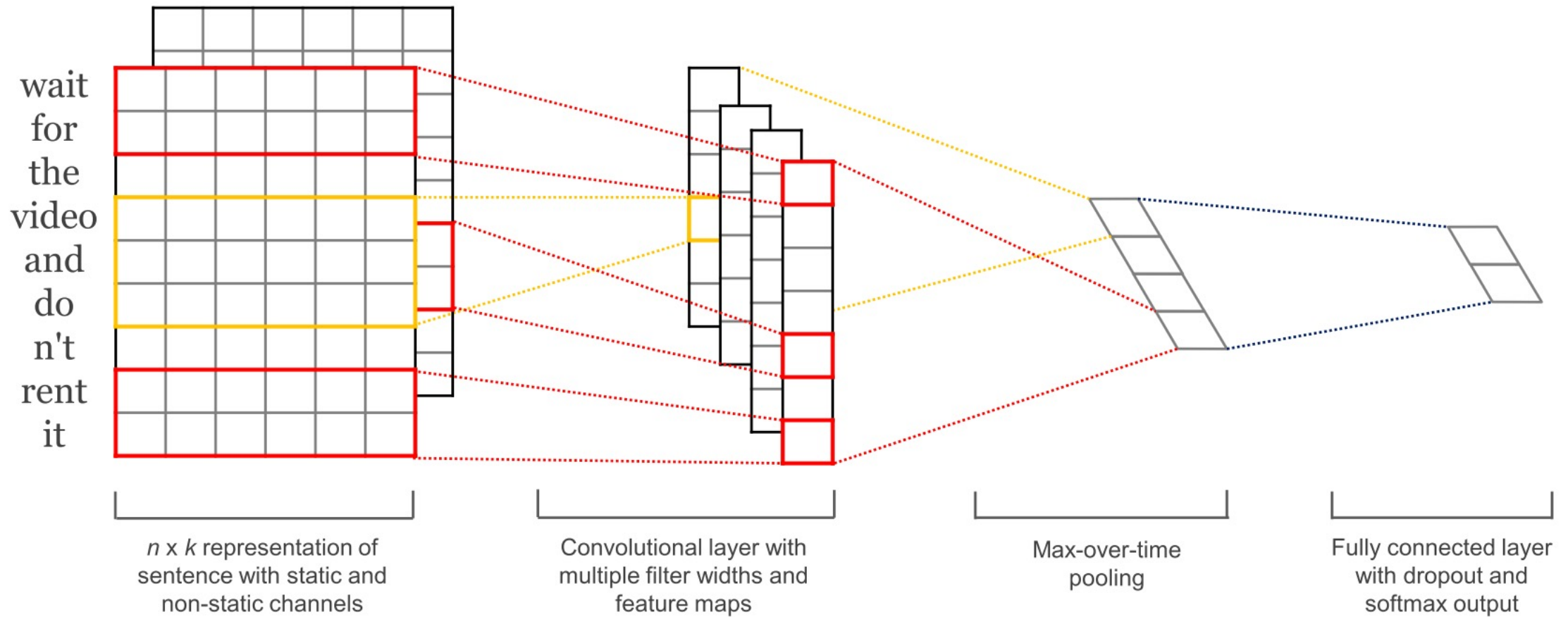
## Simple Word Embedding Model



RNN: transfer *relevant* information about earlier words



# CNNs for NLP





# SOTA in 2021: re-purpose deep, pre-trained NLP models.

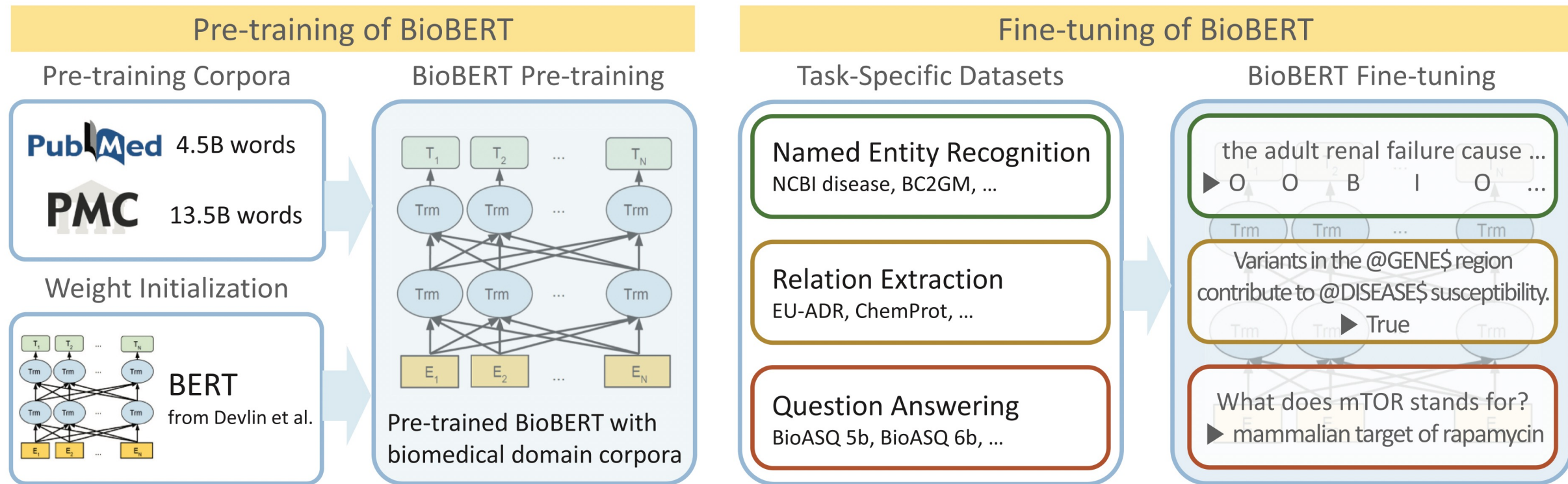


Fig. 1. Overview of the pre-training and fine-tuning of BioBERT

Lee J, Yoon W, Kim S, Kim D, Kim S, So CH, Kang J. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. Bioinformatics. 2020 Feb 15;36(4):1234-40.

# Named Entity Recognition + Simple Models

...is often the most practical approach

# The Old(er) Way

- Unified Medical Language System (UMLS)
- Apache cTAKES
- Rules-based systems to extract medical concepts from free text
- Can then build predictive models based on presence or absence of specific medical concepts

# Same Idea (Med Concept Recognition), Newer Tech

**Choose Sample Text**

The patient is a 30-year-old female with a long history of insulin dependent diabetes, type 2; coron...

**Text annotated with identified Named Entities**

The patient is a 30-year-old female with a long history of **insulin dependent diabetes, type 2** ; **coronary artery disease** ; **chronic renal insufficiency** ; **peripheral vascular disease** , also secondary to **diabetes** ; who was originally admitted to an outside hospital for what appeared to be **acute paraplegia** , lower extremities. She did receive a course of **Bactrim** for 14 days for **UTI** . Evidently, at some point in time, the patient was noted to develop **a pressure-type wound** on the sole of her left foot and left great toe. She was also noted to have **a large sacral wound** ; this is in a similar location with **her previous laminectomy** , and this continues to receive daily care. The patient was transferred secondary to inability to participate in full physical and **occupational therapy** and continue **medical management** of **her diabetes** , the sacral decubitus, **left foot pressure wound** , and associated **complications of diabetes** . She is given **Fragmin** 5000 units subcutaneously daily, **Xenaderm** to **wounds** topically b.i.d., **Lantus** 40 units subcutaneously at bedtime, **OxyContin** 30 mg p.o. q.12 h., **folic acid** 1 mg daily, **levothyroxine** 0.1 mg p.o. daily, **Prevacid** 30 mg daily,

# Conclusions

- Bag of words techniques already go beyond the majority of “NLP” found in the medical literature
- However, deep learning is rising in popularity. Much like in computer vision, we can now use transfer learning to adapt a pre-trained model to the problem at hand. It is best to start with a model pre-trained on a biomedical corpus.
- Alternatively, we can take a more ad hoc approach: identify medical concepts with off the shelf deep learning tech, then build simple models on top of the identified medical concepts