

Active Neural Networks to Detect Mentions of Changes to Medication Treatment in Social Media

D. Weissenbacher, S. Ge, A. Klein, K. O'Connor, R. Gross, S. Hennessy, G. Gonzalez-Hernandez
<https://doi.org/10.1101/2020.12.04.20244210>

MOTIVATIONS

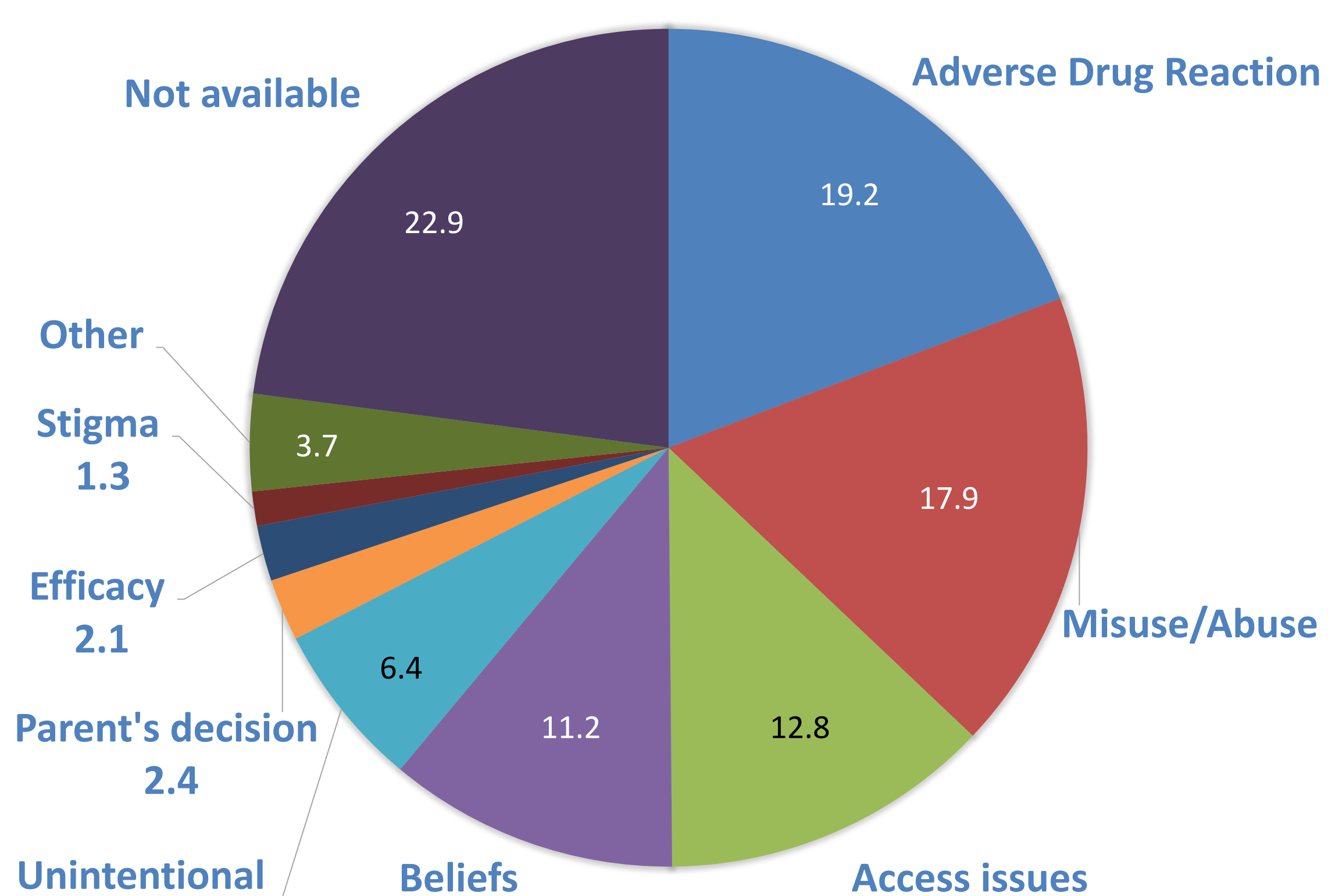
- 50% of patients do not adhere to their treatments, despite serious consequences: worsening health conditions or death
- Understanding reasons for nonadherence** is a condition for successful interventions
- Traditional sources to understand nonadherence are limited, complementary sources such as **Social Media** data can be useful

METHODS

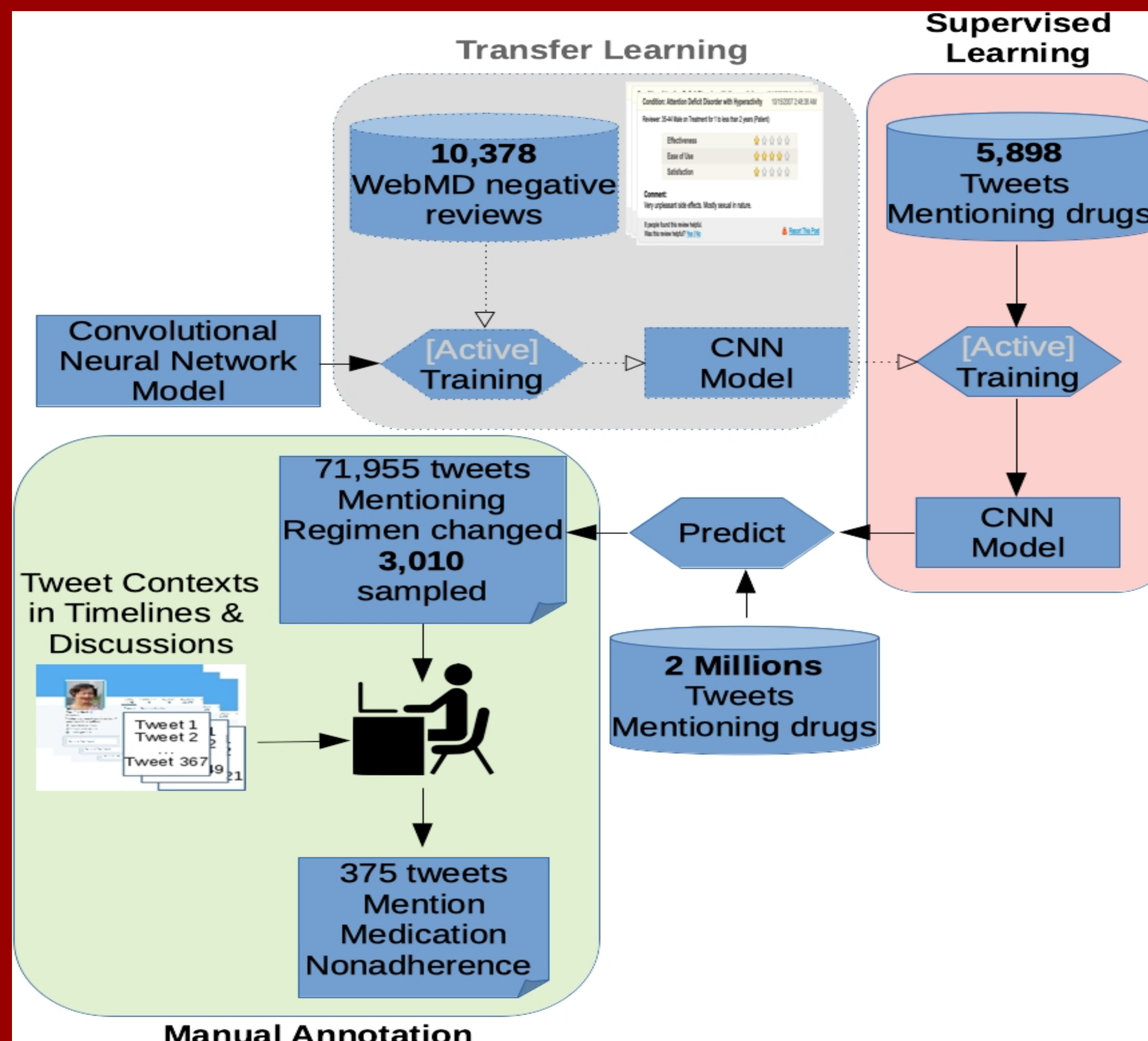
- Two corpora** created:
 - WebMD reviews: 12,000 reviews, ~50% mention a change in medication regimen
 - Twitter corpus: 9,000 tweets, ~8.7% mention a change in medication regimen
- Classifiers** trained on WebMD reviews and knowledge transferred to Twitter corpus, active learning reducing training set
- Manual analysis** of 3010 tweets with a high probability to mention a change in regimen to find nonadherence and the reasons

RESULTS

375 TWEETS MENTIONING NONADHERENCE AND REASONS



From semi-automatic analysis of Twitter posts, we can learn why patients do not adhere to their treatments.



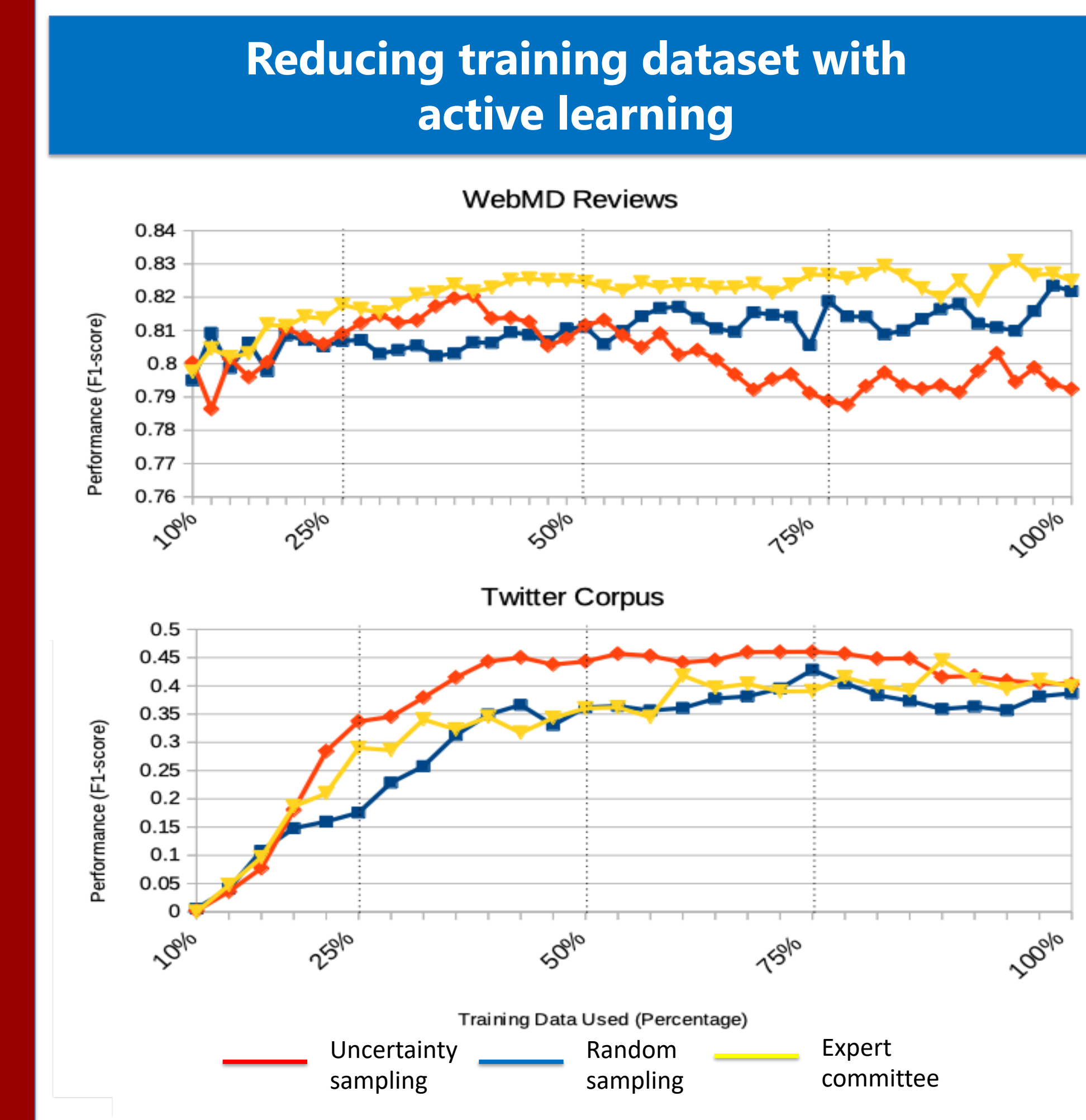
Challenge and solutions

- Mentions of Nonadherence are rare: 0.68% in tweets mentioning statins [Golden et al. 2020]
- We searched for changes in medication regimen and not only for nonadherence
- We reduced the number of training examples using Transfer and Active learning

Classification of changes in medication regimen

Systems performance on the Twitter Corpus test data
(2,300 tweets, 208+/2152-, IAA = 0.649 κ)

Systems	P	R	F1
Without Transfer			
Regular expressions	45.3	41.4	43.2
Random sampling	50.8	34.1	40.4
Uncertainty sampling	39.3	56.2	45.5
Expert Committee (5 CNNs)	51.7	34.7	41.3
With Transfer			
Random sampling	53.6	40.9	46.2
Uncertainty sampling	46.5	52.4	48.4
Expert Committee (5 CNNs)	56.2	45.8	50.4



Acknowledgements and Contacts

This work was partially supported by National Library of Medicine (NLM), grant number R01LM011176. The content is solely the responsibility of the authors and does not necessarily represent the official view of NLM.

Contacts
 Laboratory website: <http://healthlanguageprocessing.org>
 Davy Weissenbacher: dweissen@penncmedicine.upenn.edu, @davy_weiss
 Graciela Gonzalez: gragon@penncmedicine.upenn.edu, @gracielagon

