

An automated pipeline for analysis of PRO-CTCAE free-text responses with application to the ACCRU RU011201I clinical trial

Patrick Johnson, Sandra Mitchell, Ethan Basch, Lauren Rogak, Yuelin Li, Erik Asmus, David Hillman,
Karen Mittu, Minetta Liu, Amylou Dueck

Background

PRO-CTCAE[®] is a standardized tool that enables patients to self-report symptomatic adverse events in cancer clinical trials.

Part 1. Structured items

	Joint Pain				
	0	1	2	3	4
Frequency	Never	Rarely	Occasionally	Frequently	Almost constantly
Severity	None	Mild	Moderate	Severe	Very Severe
Interference	Not at all	A little bit	Somewhat	Quite a bit	Very much

- ✓ Standard, easy to analyze
- ✗ Predefined, limited

Part 2. Unstructured items

Free text response: “my finger nails are discolored and my throat is sore”

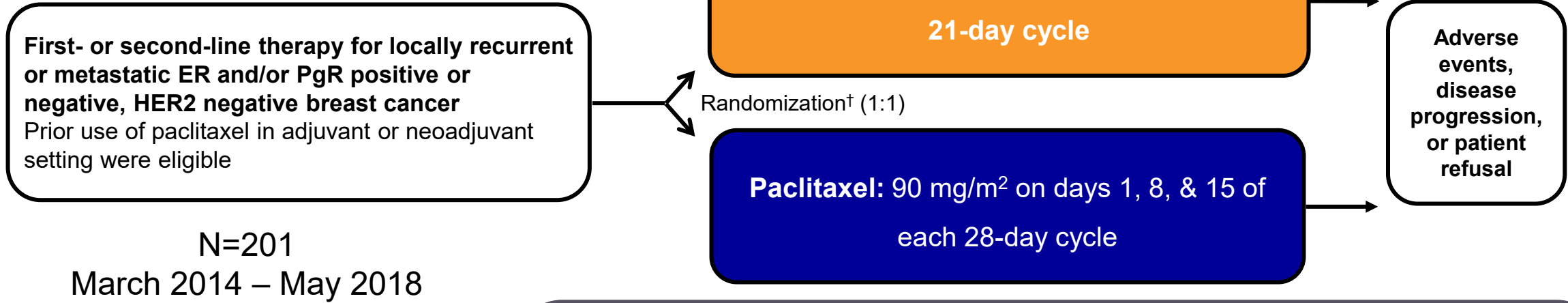


- ✓ Additional items
- ✗ Manual, costly, and time-consuming to structure

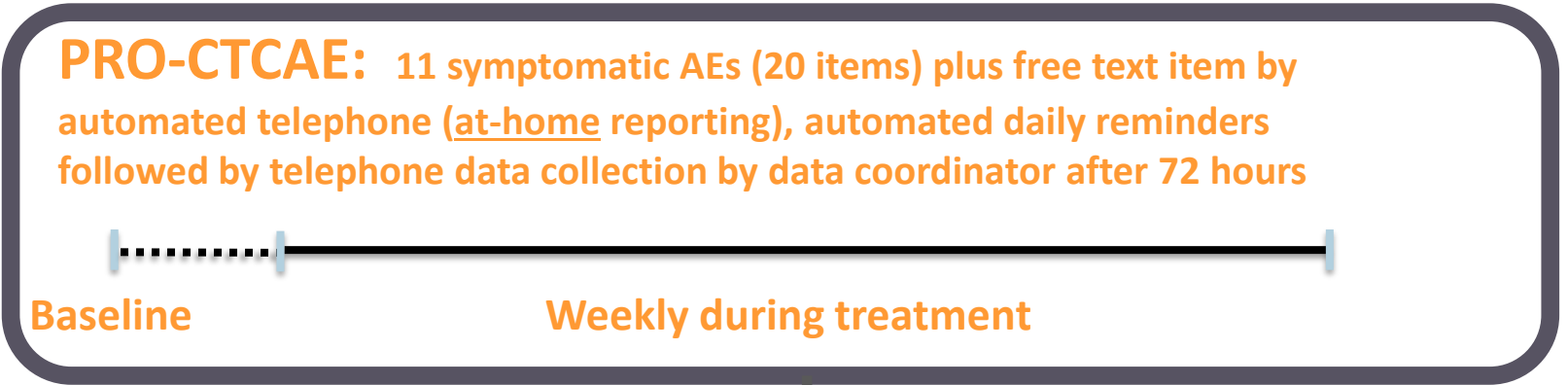
Goal: Can we (at least partially) automate the standardization of free text items?



Phase III: ACCRU RU0112011



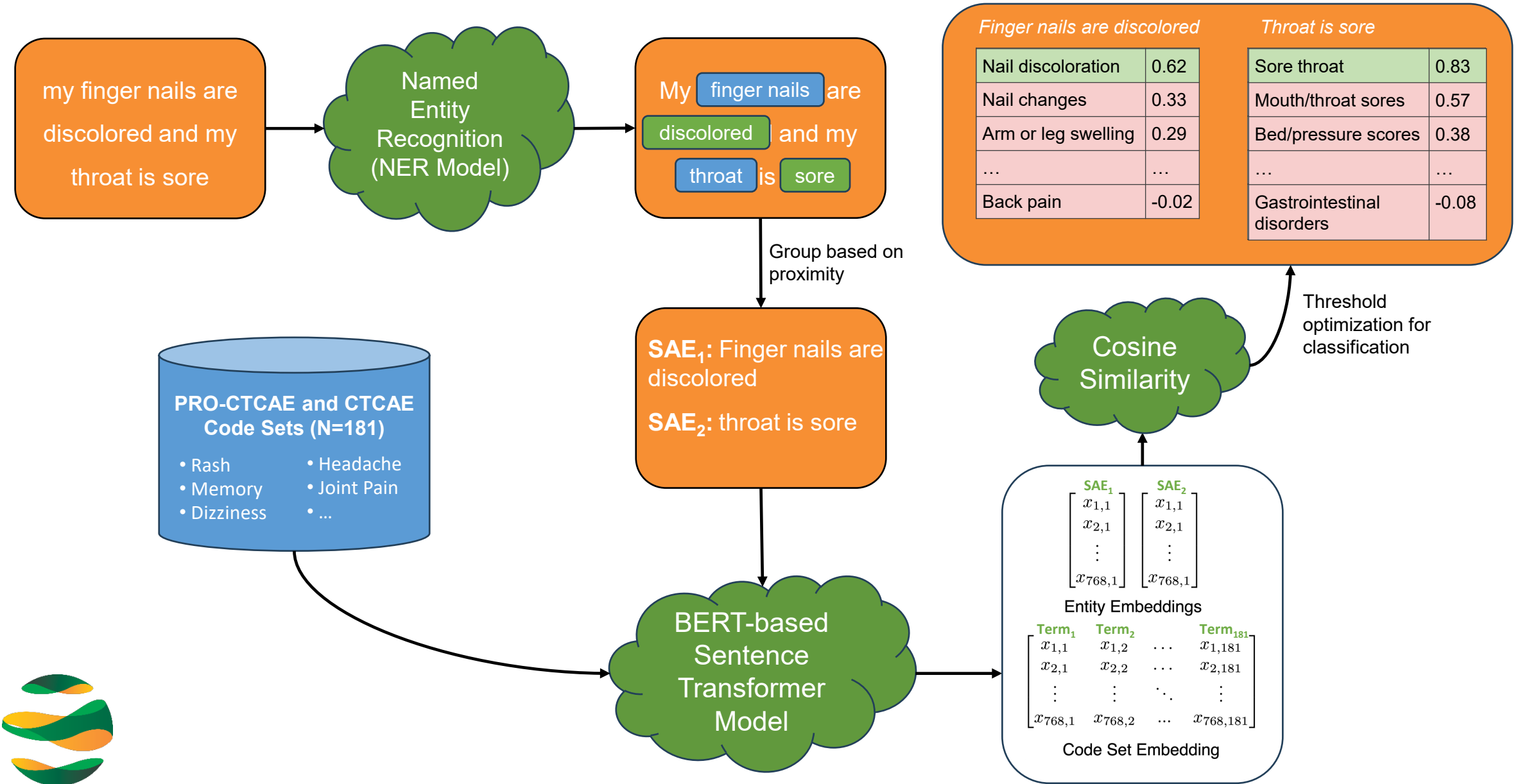
† stratified by:
 prior taxane (yes vs no)
 ER/PgR status (ER or PgR + vs ER & PgR -)
 planned line of chemotherapy (first- vs second-line)



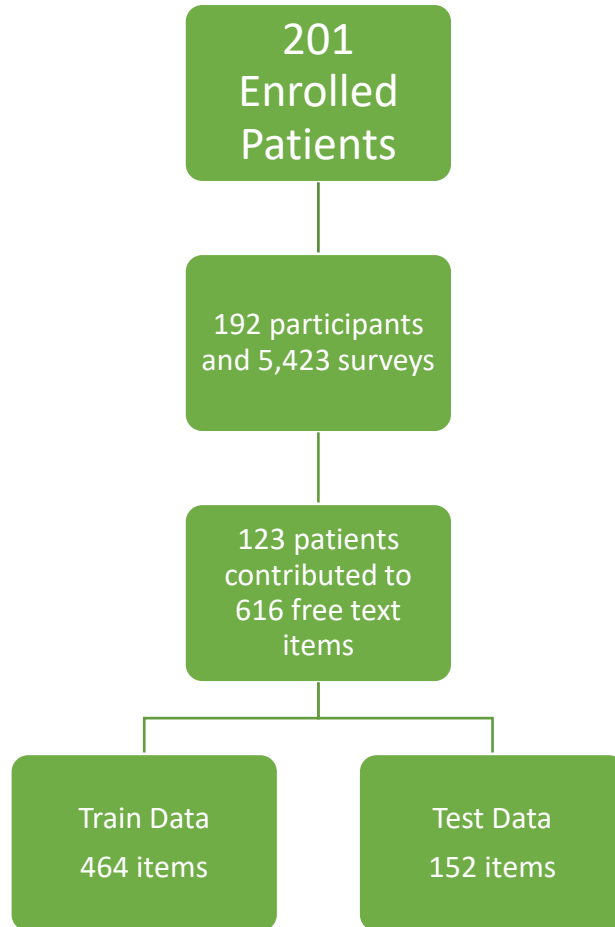
At end of study, two independent coders manually assigned PRO-CTCAE or CTCAE terms to free-text responses for use as reference labels



Pipeline Development



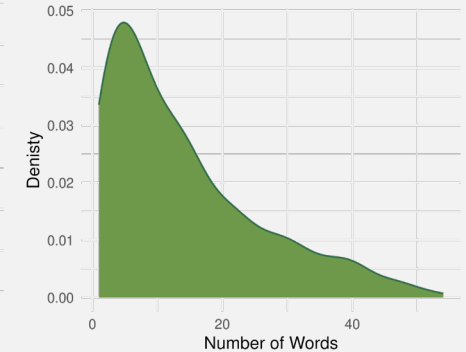
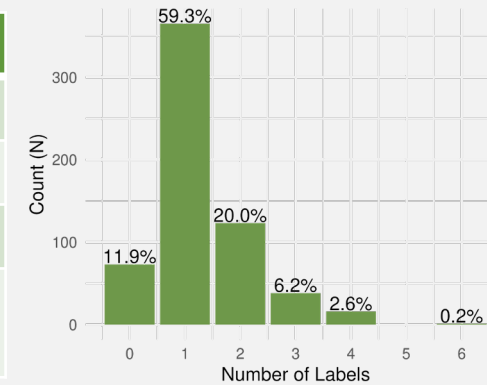
Results



Responses

- 168 total labels, 106 unique labels
- 58.8% PRO-CTCAE terms

Term	Percent
Muscle Cramps	5.7%
Headache	5.3%
Back Pain	3.8%
Numbness or tingling	3.8%



Performance (test data)

Sample-Based Metric	Value
Precision	0.86
Recall	0.74
F1-score	0.75

Conclusion

- Promising results in small proof-of-concept study
- Selectively applying labels in cases of high similarity could allow human coders to focus on more challenging or nuanced survey results
- Approach transferable to other code sets (e.g., MedDRA) and disease groups
- Transferable across institutions and PHI friendly (not reliant on proprietary models or contracts to use ChatGPT or Gemini)

