



## BACKGROUND

- High quality screening colonoscopy is the hallmark of effective colorectal cancer (CRC) prevention.
- Measuring colonoscopy quality indicators is crucial for quality improvement efforts.
- However, accurately and routinely measuring colonoscopy quality indicators for providers and health systems remains a challenge.

## OBJECTIVES

- To develop a natural language processing (NLP) algorithm that utilizes machine learning tools to automatically and accurately characterize four prespecified quality indicators for screening colonoscopies.
- To validate the performance of the NLP algorithm against manual chart review.

## METHODS

### Study setting:

- Large academic healthcare system that performs >15,100 screening colonoscopies yearly in 6 endoscopy centers.

### Study Population:

- We randomly selected 600 screening colonoscopies performed between 6/2020-2/2021.

### Analyses:

- We trained and developed an NLP algorithm to measure colonoscopy indication (IND), bowel preparation (BP), cecal intubation (CI), and successful cecal intubation (SCI) (Figure 1).
- We compared NLP-derived quality metrics and manual chart review (two physician chart review as gold standard).

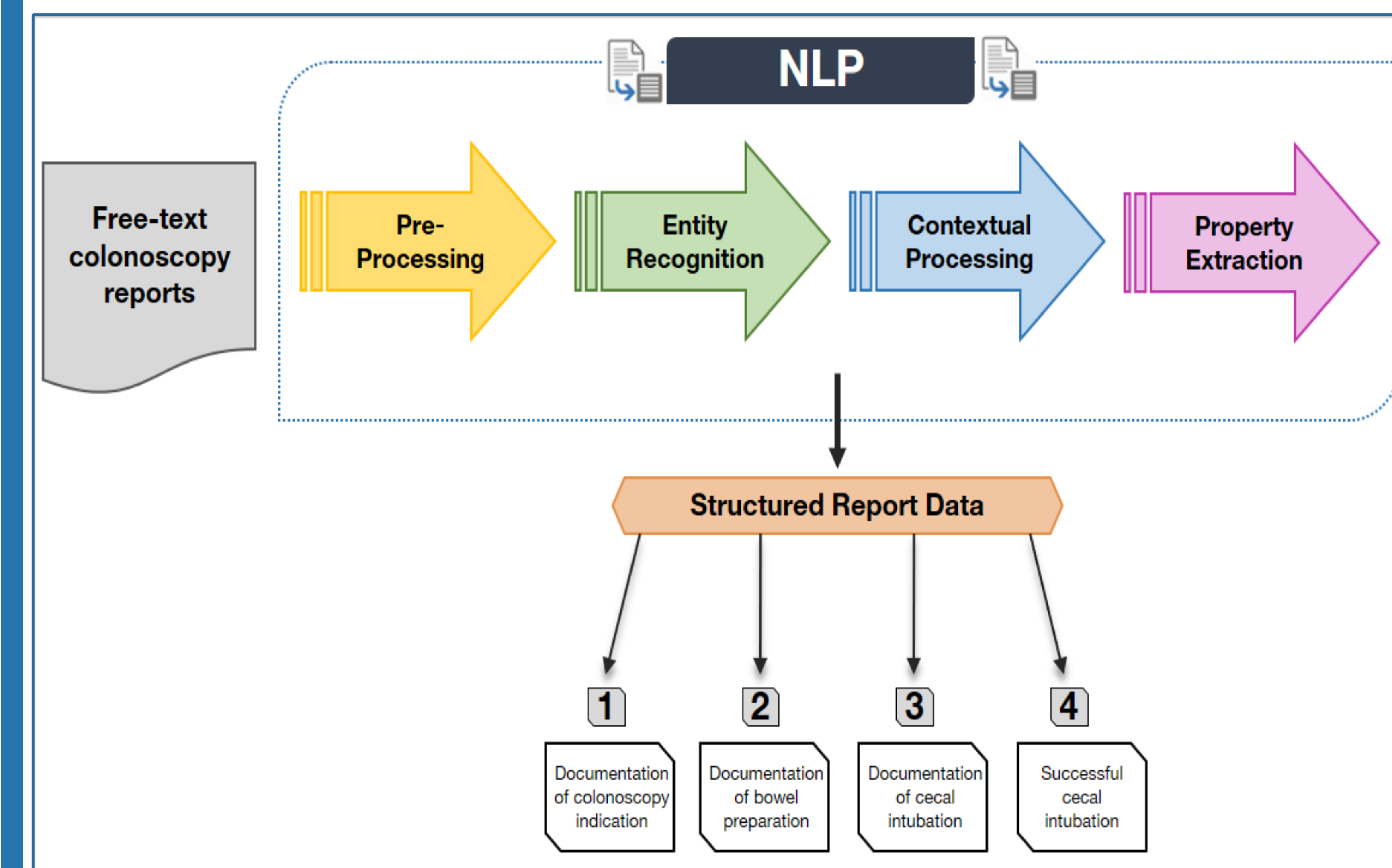
## RESULTS

- NLP sensitivity ranged from 99.3% to 100.0% and specificity ranged from 94.3% to 100.0% for all four studied quality indicators.
- Because of conflicting documentation by endoscopist in the same report, NLP misclassified 2 cases for colonoscopy indication and 1 case for bowel preparation.
- NLP misclassified 12 cases for successful cecal intubation, which were mainly due to either the endoscopist not mentioning the word “cecum” or documenting “terminal ileum” instead of “cecum.”
- NLP had perfect performance (100%) for cecal intubation.

**Table 1: NLP Performance on four colonoscopy quality indicators**

Documentation of colonoscopy indication (IND)			
Manual review	NLP		Total
	“Screening” detected	“Non-screening” detected	
“Screening” detected	314	0	314
“Non-screening” detected	2	284	286
<b>Total</b>	<b>316</b>	<b>284</b>	<b>600</b>
Test characteristics	<b>Sensitivity</b>	99.3%	
	<b>Specificity</b>	100%	
	<b>PPV</b>	100%	
	<b>NPV</b>	99.4%	
	<b>F1 Score</b>	0.996	
<b>Accuracy</b>	99.7%		
Documentation of bowel preparation (BP)			
Manual review	NLP		Total
	BP documented	BP not documented	
BP documented	599	0	599
BP not documented	1	9	1
<b>Total</b>	<b>600</b>	<b>0</b>	<b>600</b>
Test characteristics	<b>Sensitivity</b>	100%	
	<b>Specificity</b>	97.5%	
	<b>PPV</b>	99.8%	
	<b>NPV</b>	N/A	
	<b>F1 Score</b>	0.999	
<b>Accuracy</b>	99.8%		
Documentation of cecal intubation (CI)			
Manual review	NLP		Total
	CI documented	CI not documented	
CI documented	599	0	599
CI not documented	0	1	1
<b>Total</b>	<b>599</b>	<b>1</b>	<b>600</b>
Test characteristics	<b>Sensitivity</b>	100%	
	<b>Specificity</b>	100%	
	<b>PPV</b>	100%	
	<b>NPV</b>	100%	
	<b>F1 Score</b>	1	
<b>Accuracy</b>	100%		
Successful cecal intubation (SCI)			
Manual review	NLP		Total
	SCI documented	SCI not documented	
SCI documented	437	3	440
SCI not documented	9	150	159
<b>Total</b>	<b>446</b>	<b>153</b>	<b>599</b>
Test characteristics	<b>Sensitivity</b>	99.3%	
	<b>Specificity</b>	94.3%	
	<b>PPV</b>	98.0%	
	<b>NPV</b>	98.0%	
	<b>F1 Score</b>	0.987	
<b>Accuracy</b>	98.0%		

**Figure 1: Schematic of natural language processing pipeline**



**Figure 1: The NLP receives input from free-text colonoscopy reports and then outputs the report as structured data that can be used to interpret quality indicators.**

## CONCLUSIONS

- We successfully developed an automated NLP algorithm that is highly accurate and sensitive in determining four priority intraprocedural colonoscopy quality indicators.
- The NLP algorithm is a feasible, efficient, and effective method to consistently track and report colonoscopy quality metrics at the provider and health system level.
- Future directions: application of our NLP algorithm to improve CRC outcomes.

## ACKNOWLEDGEMENTS

UCLA Melvin and Bren Simon Quality Improvement Program  
UCLA Vatche and Tamar Manoukian Division of Digestive Diseases

## CONTACT INFORMATION

Artin Galoosian, MD, MA  
Galoosian@gmail.com

Folasade P. May MD, PhD, MPhil  
FMay@mednet.ucla.edu

#MayLabUCLA  
#UCLAGI