

# Lumbar Puncture Success Rate in the Pediatric Emergency Department

Mara Alexeev, MD, MPH<sup>1</sup>,   
@MaraAlexeev

mara.alexeev@childrens.harvard.edu

Assaf Landschaft, M.Sc.<sup>2</sup> Al Ozonoff, PhD, CPPS<sup>3</sup>,<sup>4</sup> Lise E. Nigrovic, MD, MPH<sup>5</sup>  
Marvin B. Harper, MD<sup>5</sup> Jennifer Kupiec, MPH<sup>5</sup> Amir Kimia, MD<sup>5</sup>

<sup>1</sup> Department of Pediatrics, Boston Children's Hospital  
<sup>2</sup> Boston Children's Hospital  
<sup>3</sup> Precision Vaccines Program, Boston Children's Hospital  
<sup>4</sup> Department of Pediatrics, Harvard Medical School  
<sup>5</sup> Division of Emergency Medicine, Boston Children's Hospital

## Background

The frequency of lumbar punctures (LPs) procedures has declined over the past two decades with uncertain impacts on lumbar puncture competence. Unsuccessful LPs can be challenging to identify using available administrative or self-reported data.

## Objectives

1. Our goal was to identify the proportion of attempted LPs that were ultimately unsuccessful.

## Methods

We performed a retrospective cross-sectional study of all ED encounters for children 21 years or younger who presented between 2017-2019 to a tertiary care pediatric center. We defined an unsuccessful LP as no CSF obtained, CSF culture only, or CSF RBC  $\geq 10,000$  cells/mm<sup>3</sup>.

Failure rates of lumbar punctures are much higher than previously published data.

Natural language processing tools captured attempted LPs not captured by structured data.



We developed a Natural Language Processing (NLP) model to identify these procedures from unstructured data (clinical note text), using a sample of manually reviewed notes from one year of data. We randomly selected 70% of the notes as a training set, and held aside the remaining 30% of notes as a validation/testing set. We compared NLP model classification to the manual chart review and calculated the sensitivity and specificity of the classifier for identifying notes describing an LP attempt. We then applied the NLP model to all ED notes during the study period and reviewed the output manually.

We also used unstructured data (e.g. physician orders, laboratory results, billing codes) to flag cases for manual review.

## Results

In the validation cohort, the NLP model identified 118 of the 122 LPs identified by medical records review: sensitivity 96.6% (95% CI 93.4-99.9%) and specificity 97.9% (95% CI 96.4-99.3%). The NLP model was then applied to 117,981 ED encounters during the study period and the NLP classifier output provided 2916 notes for manual review, of which we identified 1,094 notes describing an LP attempted. Screening an additional 713 cases flagged by structured data, identified an additional 19 cases, for an overall 1113 ED encounters with LP attempts (0.63% of encounters) (see figure 1.)

Unsuccessful LPs occurred in 226/1113 (20.3%). Of these, fluid was not obtained in 104 (9.3%), LP sent for culture only in 53 (4.8%), and LP with  $\geq 10,000$  CSF RBC cells/mm<sup>3</sup> in 69 (6.2%).

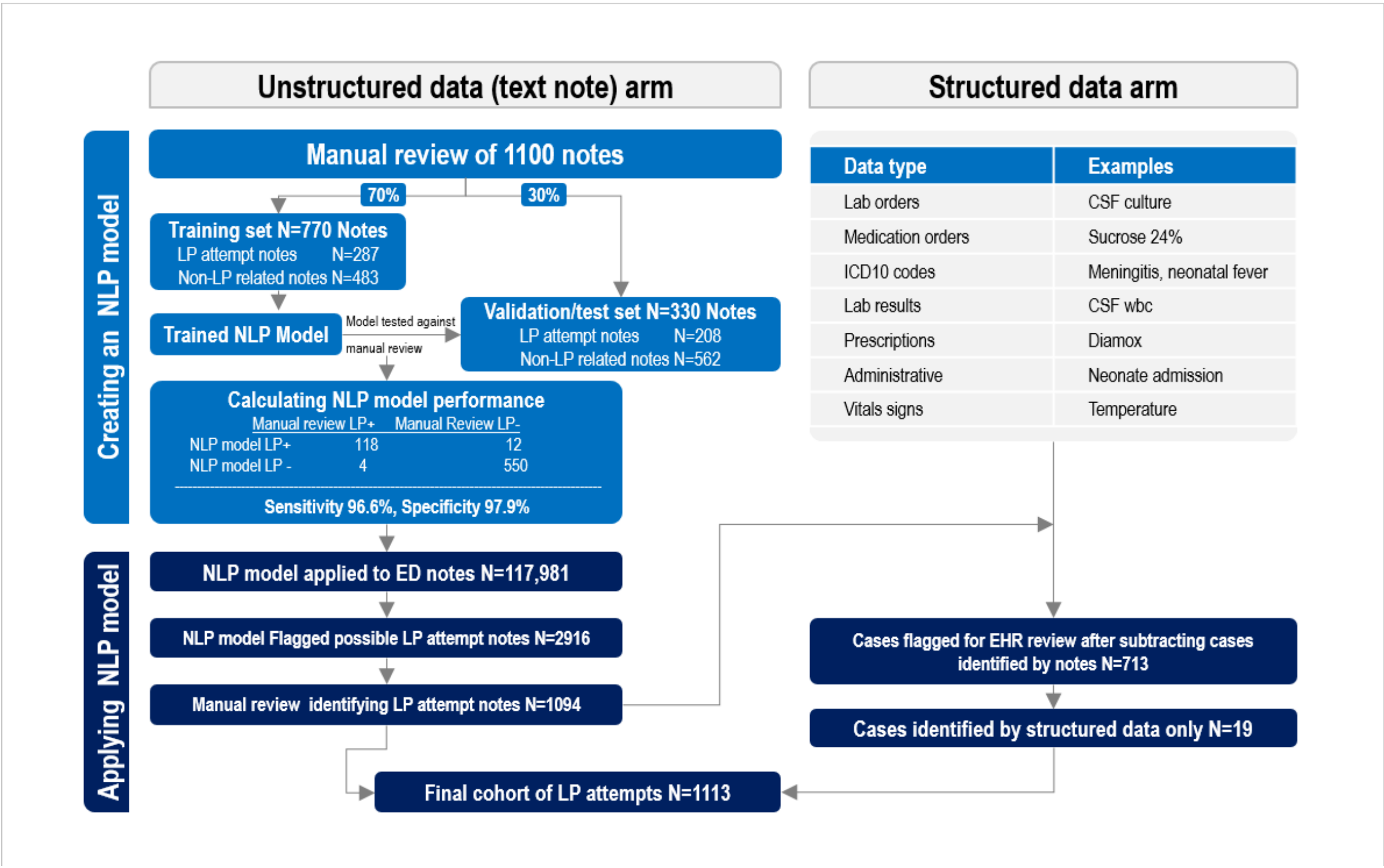


Figure 1: Review process