Natural Language Processing: An Automated Alternative to Determining Group A Streptococcal Testing

Brian R. Lee, PhD, MPH; Alaina Linafelter, PharmD, BCPPS; Alaina Burns, PharmD, BCPPS; Allison Burris, MD; Heather Jones, RN, MSN, CPNP; Jarrod Dusin, MS, RD, LD; Rana E. El Feghaly, MD, MSCI

Children's Mercy Kansas City; University of Missouri Kansas City School of Medicine

Background

- 12 million pediatric ambulatory visits for acute pharyngitis each year
- Rapid antigen detection test (RADT) for Group A *Streptococcus* is commonly used
- An estimated 40-60% of these RADT are considered inappropriate

Data Collection

- Sample 1: 10% random selection of patients ≥ 3 y/o seen at 5 emergency department/urgent care sites between 4/2018 – 9/2018 and received RADT
- Manual chart review of provider notes to document symptoms:
    - Sore throat
    - Viral (conjunctivitis, rhinorrhea, cough, diarrhea, hoarse voice, viral exanthema)
- Inappropriate RADT was 2+ viral symptoms *or* lack of sore throat
- Sample 2: pharyngitis patients seen in March 2019 at 1 urgent care site using same inclusion and chart review process as Sample 1

Natural Language Processing (NLP)

- R *tidytext* package used to perform NLP on the same manual chart review sample
- Symptom key words flagged
- Algorithm developed to distinguish negation (e.g., "reports cough" vs. "denies cough")
- Sensitivity/specificity of NLP calculated using manual review as gold standard

Results

Sample 1

- Of 720 patients, 320 (44.4%) had inappropriate RADT based on manual chart review
- Inappropriateness from NLP was 44.9% with high sensitivity (88.4%) and high specificity (90.0%)

Table 1: Symptom-specific NLP sensitivity and specificity compared to Sample 1 manual chart review

| Symptom	Frequency	Sensitivity	Specificity
Sore throat	533	92.9%	92.5%
Diarrhea	47	83.0%	98.8%
Cough	239	94.5%	96.5%
Hoarse voice	14	63.6%	100.0%
Conjunctivitis	10	50.0%	100.0%
Rhinorrhea	238	86.1%	95.3%

Sample 2

- Of 382 patients, 174 (45.7%) had inappropriate RADT based on manual chart review
- NLP showed high sensitivity (92.0%) and good specificity (76.8%)

Table 2: Select symptom-specific NLP sensitivity and specificity compared to Sample 2 manual chart review

Symptom	Frequency	Sensitivity	Specificity
Sore throat	268	91.3%	94.6%
Cough	180	94.9%	86.2%
Rhinorrhea	184	90.4%	74.8%

Figure 1: NLP-derived trendline of inappropriate RADT, June 2018-April 2020
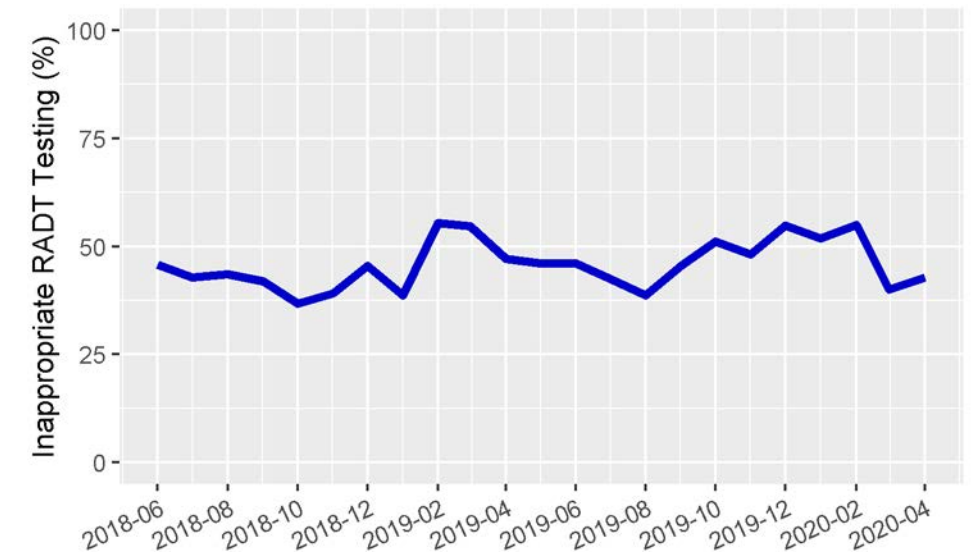
Conclusions

- NLP demonstrated high validity compared to manual chart review
- An NLP approach requires significantly less dedicated time compared to manual chart review (minutes vs. hours)
- NLP less susceptible to subjectivity compared to clinical abstractors
- Once implemented, NLP easily allows for monitoring long-term trends

LOVE WILL.

KU SCHOOL OF MEDICINE
The University of Kansas

UMKC School of Medicine

Children's Mercy KANSAS CITY