RUSH UNIVERSITY

RUSH MEDICAL COLLEGE • COLLEGE OF NURSING • COLLEGE OF HEALTH SCIENCES • THE GRADUATE COLLEGE

Disease Surveillance: The Need For a Robust Natural Language Processor

Julio C. Silva, MD, MPH, FACEP, Marilyn Hallock, MD,
Dino P. Rumoro, DO, FACEP, Shital Shah, PhD, Gillian Gibbs, MPH,
Jamil Bayram, MD, MPH, Michael Waddell, PhD

Funded by: Department of Defense

Telemedicine and Advanced Technology Research Center: W81XWH-06-1-0785

Background

RUSH UNIVERSITY

- Real-time disease surveillance
  - Is critical for rapid diagnosis and public health intervention
  - Needs to be capable of accurately processing both coded and free-text clinical information

GUARDIAN

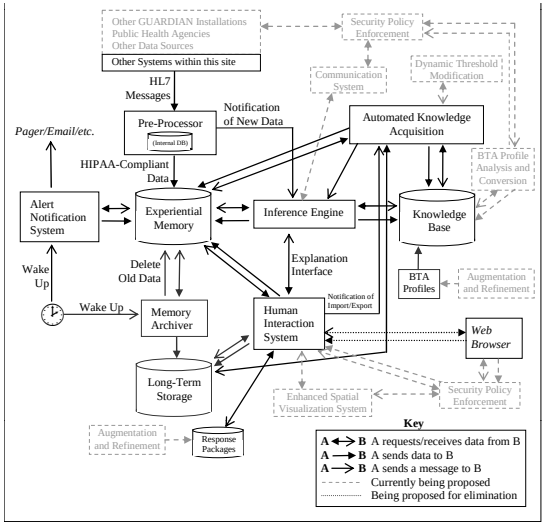
RUSH UNIVERSITY

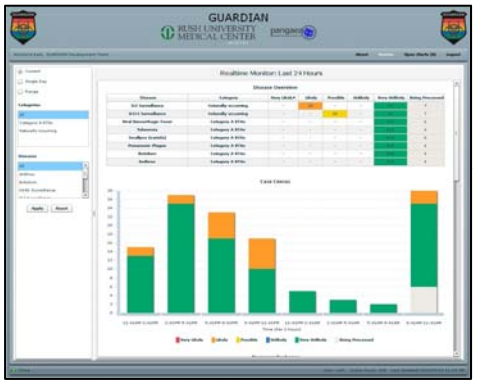
Geographic Utilization of Artificial Intelligence in Real-Time for Disease Identification and Alert Notification

- Real-time, automated, knowledge-based, disease surveillance system

Inputs → *Algorithm/ System Architecture* → *Outputs*

Flow sheet Data •Numeric •Pick Lists •Vital Signs	Chief Complaint •QuickReg •Pick List •Free Text	Nurse's Notes (Pick List)
Medical/ Surgical History (Pick List)	History of Present Illness (Pick List)	Resident's Notes (Free Text)
Attending Notes (Free Text)	Review of Systems (Pick List)	Physical Exam (Pick List)





Objective

RUSH UNIVERSITY

- To compare the GUARDIAN natural language processor (NLP) algorithm's accuracy for influenza-like illness detection to:
  - Physician chart review (gold standard)
  - Unmodified MetaMap Transfer (uMMTx); an open source software component developed by the National Library of Medicine

Methods

RUSH UNIVERSITY

- Retrospective cross-sectional study design
- Emergency department, large urban tertiary care facility
- 1,122 emergency department patients
- Between November 1 - 7, 2009 (influenza season)

Methods

RUSH UNIVERSITY

- Influenza-like illness (ILI) status
  - ILI defined as fever and cough and/or sore throat
  - Fever assessed by vital signs and reported fever
- Patient charts scanned by:
  - GUARDIAN NLP
  - uMMTx
  - Emergency Department Attending Physician
- Discrepancies resolved by second round of physician chart review

Methods

RUSH UNIVERSITY

- GUARDIAN NLP and uMMTx
  - Conduct exact word matches
  - Analyze synonyms
  - Contain large dictionary of terms
  - Differentiate among related terms
- Modifications made to improve accuracy of GUARDIAN NLP algorithm
  - Includes common misspellings
  - Incorporates abbreviations
  - Handles negations
    - 16 word window before and after each negation word

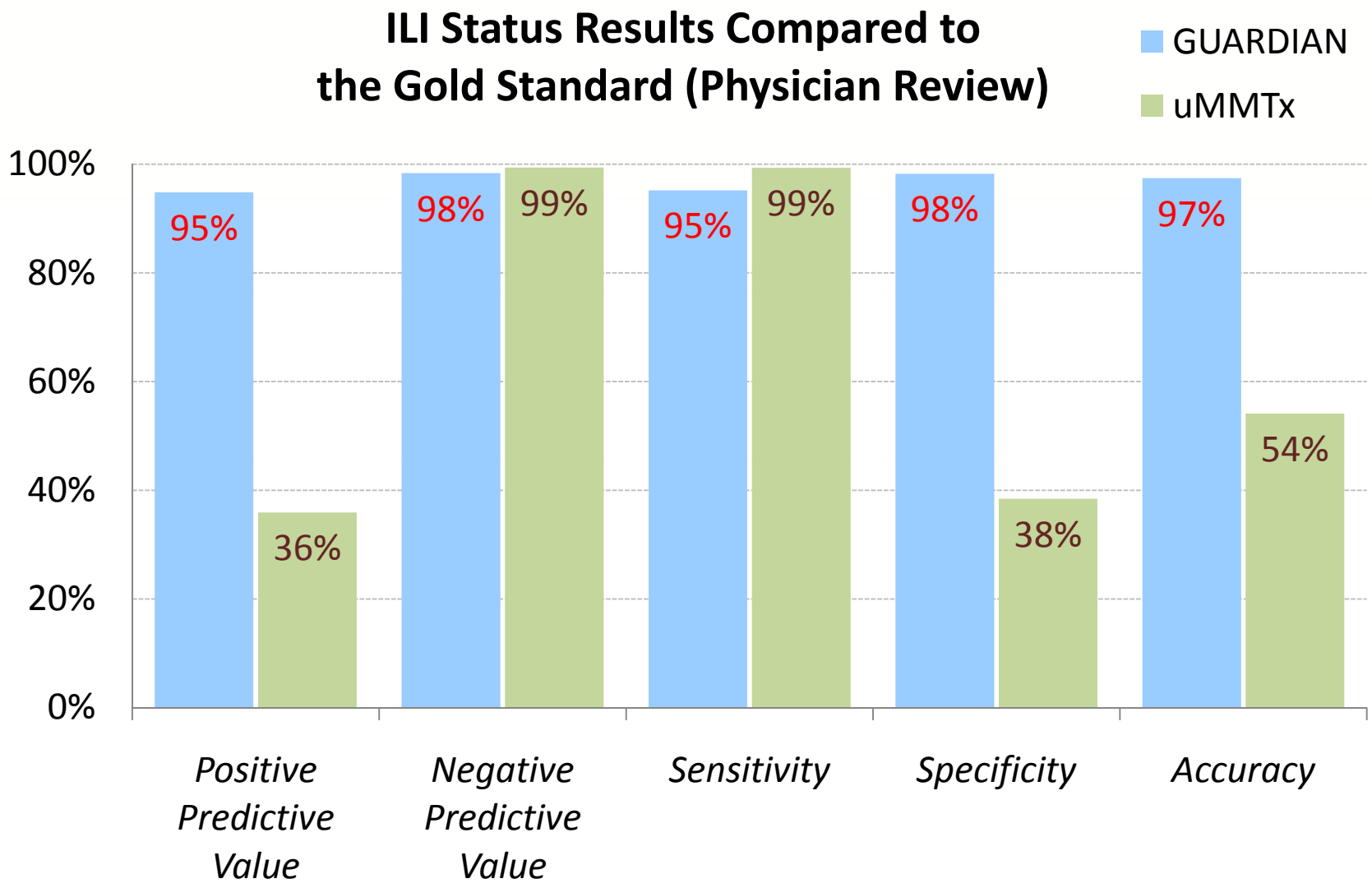
Results

RUSH UNIVERSITY

		ILI Status (Physician Review)	
		Negative	Positive
ILI Status (GUARDIAN NLP)	Negative	818	14
	Positive	15	275
ILI Status (uMMTx)	Negative	320	2
	Positive	513	287

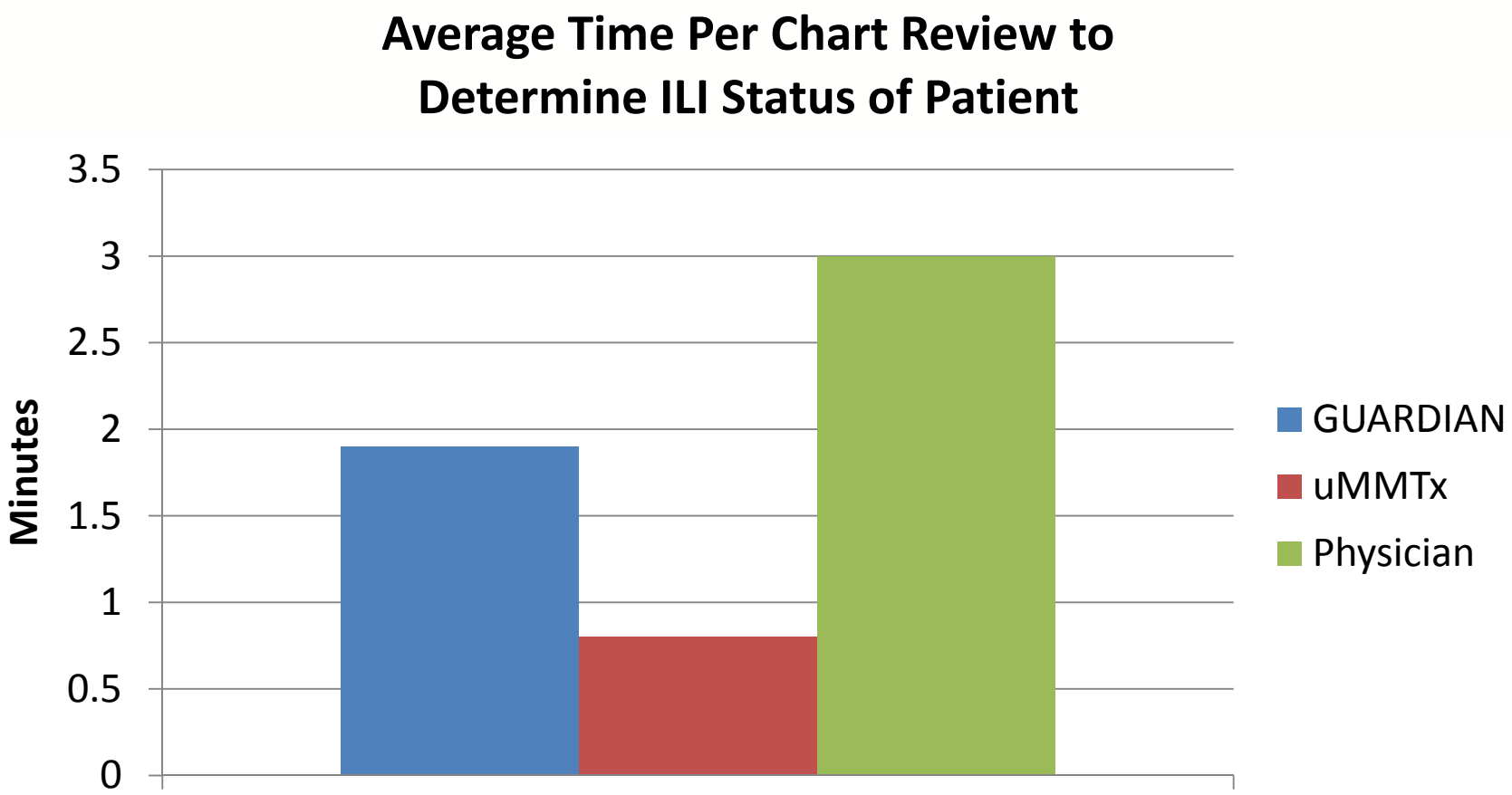
Results

RUSH UNIVERSITY



Results

RUSH UNIVERSITY



Discussion

RUSH UNIVERSITY

- GUARDIAN:
  - Is comparable in accuracy to a licensed physician's manual review of charts
  - Is considerably more accurate than the uMMTx
    - Positive predictive value
    - Specificity
    - Accuracy
  - Utilizes less resources, as measured by time spent reviewing records, than a physician
    - Automated review in real-time

Conclusions

RUSH UNIVERSITY

- A robust NLP component of real-time syndrome surveillance programs can significantly improve accuracy in detecting infectious diseases, which may enhance public health preparedness and response efforts.