Pressure Ulcer Injury in Unstructured Clinical Notes: Detection and Interpretation

Authors: Mani Sotoodeh, MS, Zelalem H. Gero, MS, Wenhui Zhang, PhD, MS, RN, Roy L. Simpson, DNP, RN, Vicki Stover Hertzberg, PhD, Joyce C. Ho, PhD

S49 - Care Management

Presenter : **Mani Sotoodeh**

Institution : **Emory University**

Twitter: @SotoodehMani

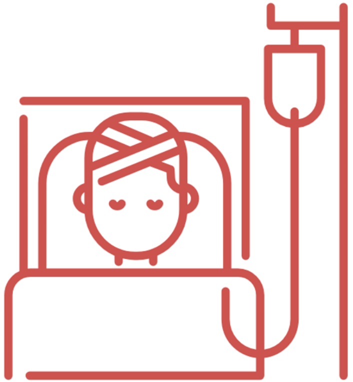
#AMIA2020

Disclosure & Acknowledgement

The authors and their spouse/partner have no relevant relationships with commercial interests to disclose.

This work was supported by NIH grant 1R01LM013323; PIs: V. Hertzberg, Joyce Ho, Roy Simpson.

PUI Detection & Prediction: a necessary task



Over 2.5 million Americans develop PUI each year

Hospital readmissions due to PUI are 75% higher

Reduced quality of life, higher mortality, longer hospital stays

Penalties to the bottom 25% of hospitals as a nursing quality metric

Allocate resources better esp. for intensive and postoperative care units

Manual assessment is inefficient→ Automation

AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.

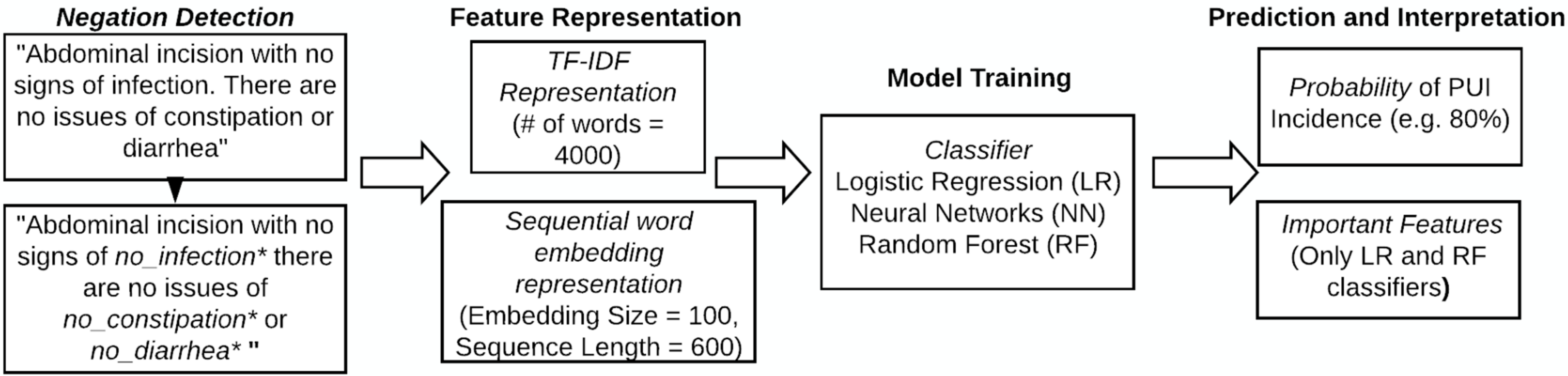
Existing challenges in PUI assessment and PUI detection literature

- Quarterly assessments of one hospital in a single day has many problems
- Missed opportunities to change practices leading to inadequate care
- Inefficiency of risk assessment tools such as the Braden scale
- Most works are on identifying risk factors retrospectively
- Nurse-collected data is not actively mined for patient information

Contributions

- Exploratory analysis of unstructured data (notes) for PUI detection
- Automated PUI detection enables more frequent assessments

Overall framework: PUI detection from notes



Negation detection using NER and Regex

Spacy: Extracting all mentions of named entities including disease, medication, symptoms, and chemicals

NegEX

- Employs defined regular expressions indicating negation
- Limits the scope of the negation phrase
- Filters out phrases that falsely appear to be negation phrases

"showed no evidence of congestive heart failure or pneumonia"

"showed evidence of no congestive heart failure or no pneumonia"

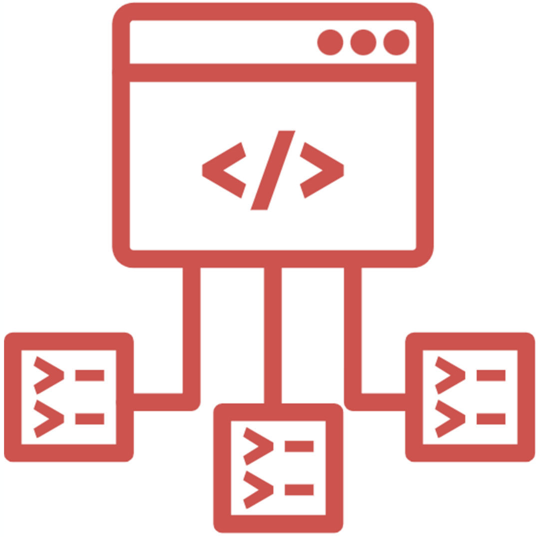
Feature construction and Baseline models

Text Vector representations

- TF-IDF representation
- Sequential Word Embedding Representation

Baseline models

- *Logistic Regression*
- *Random Forest*
- *Neural Networks*



Data Split

68% training, 12% validation & 20% test, 30 runs

MIMIC III dataset , Problem setup

Open data, 49K hospital admissions, stays of people younger than 20 are removed

PUI-positive from an ICD-9 perspective

- [707, 707.1, 707.2, 707.3, 707.4, 707.5, 707.6, 707.7, 707.9, 707.11, 707.21, 707.22, 707.23, 707.24, 707.25]

PUI-positive from a notes perspective

- [Pressure Ulcer Prevention, Skin Surveillance, Decubitus Ulcers, Impaired Tissue Integrity, Impaird Skin integrity, Bed Sores, Pressure Ulcer, Pressure sore]

Stays indicating PUI in **both** (*neither*) sources constitute → **+** (-) class

Each **+** sample is matched with **4 -** ones closest to it in age, gender, length of stay

Total number of samples: 3,589 **Positive samples**: 856 (23.8%)

Results: Negation Detection boosts prediction

Classifier	Average Test AUC (*SD*)		Average Test F1 (*SD*)	
	Negation-Aware	w/o Negation Detection	Negation-Aware	w/o Negation Detection
Neural Networks (NN)	**0.8462** (*0.0169*)	0.8440 (*0.0161*)	**0.6252** (*0.0291*)	0.6189 (*0.0302*)
Logistic Regression (LR)*	**0.9022** (*0.0120*)	0.8720 (*0.0155*)	**0.6905** (*0.0188*)	0.6455 (*0.0248*)
Random Forest (RF)	**0.9533** (*0.0086*)	0.9530 (*0.0071*)	**0.7887** (*0.0226*)	0.7862 (*0.0219*)

Results : Words created by negation detection contribute greatly to classification

Classifier	Found only in	Indicating PUI's	Words in the Set (importance)
LR	Negation-aware notes	Absence	mso (-0.3182), groundglass (-0.2953), swanganz (-0.2915), preoperative (-0.2730), no_ectopy (-0.2632), no_edema (-0.2560), independent (-0.2531), no_sob (-0.2137), no_pneumothorax (-0.2107), number (-0.1918), no_pulmonary (-0.1881)
LR	Untouched notes	Absence	ganz (-0.3563), mso4 (-0.3431), lat (-0.3431), ward (-0.2286), lima (-0.1443), hyperthermia (-0.1169), pepcid (-0.1156), neoplasm (-0.1043), Sao2 (-0.1023), pyrexia (-0.1006)
RF	Negation-aware notes	Presence or Absence	no_wound (0.0016), apply (0.0011), multipodus (0.0011), swanganz (0.0008), sch (0.0005), clip (0.0004), [no_skin, no_pneumothorax, unit, no_infection] (all 0.0003)
RF	Untouched notes	Presence or Absence	lat (0.0007), ptitle (0.0006), ganz (0.0005), name (0.0004), [followup, numeric, lastname, identifier] (all 0.0003), fi02, defined] (all 0.000[illegible])

Results: Prominent features (words) overlap with medical literature on PUI

AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.

Classifier	Type of Words (or notes)	Indicating PUI's	Most Medically Meaningful Keywords in the Set (importance)
LR	Only no_ .... words	Presence	no_wound (0.2951), no_erythema (0.2303), no_skin (0.1629), no_infection (0.1420), no_obstruction (0.1403), no_ct (0.1291), no_secretions (0.0850), no_lesions (0.0728)
LR	Only no_ .... words	Absence	no_edema (-0.2560), no_stool (-0.1241), no_pain (-0.0923), no_diuresis (-0.0744), no_hemorrhage (-0.0677), no_bleed (-0.0589), no_bleeding (-0.0523)
LR	Only Negation-aware	Absence	swanganz (-0.2915), no_edema (-0.2560), independent (-0.2531), pacu (-0.1765), bloodtinged (-0.1314), no_stool (-0.1241)
RF	Only no_ .... words	Presence or Absence	no_wound (0.0016), [no_skin, no_tube] (0.0003), [no_infection, no_blood, no_insulin, no_erythema, no_hypotension] (all 0.0002), [no_diuresis, no_abscess, no_pain, no_stool, no_edema, no_gtt] (all 0.0001)
RF	Only Negation-aware	Presence or Absence	no_wound (0.0016), multipodus (0.0010), [no_skin , no_infection , no_tube, no_insulin] (all 0.0003), sedate (0.0002)

Using notes is promising, possible next steps

AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.

Findings of the current work

- Detection of PUI for quality assurance purposes using notes of the care team
- Negation Detection improves models' predictive performance
- Most contributing keywords have considerable overlap with medical literature

Potential extensions and directions

- Further improvement by hyperparameter tuning & more complex models
- ConText, an extension for more accurate identification of medical terms
- Predict PUI prior to its occurrence and detecting its stage

Thank you!

Email me at: msotood@emory.edu