THE UNIVERSITY OF MELBOURNE
POSTERA CRESCAM LAUDE

THE UNIVERSITY OF UTAH

Department of Veterans Affairs

Automatic Extraction of Skin and Soft Tissue Infection Status from Clinical Notes

Jamie Rhoads – Dept. Dermatology, University of Utah School of Medicine
Lee Christensen – Dept. Internal Medicine, University of Utah School of Medicine
Skylar Westerdahl – Dept. Dermatology, University of Utah School of Medicine
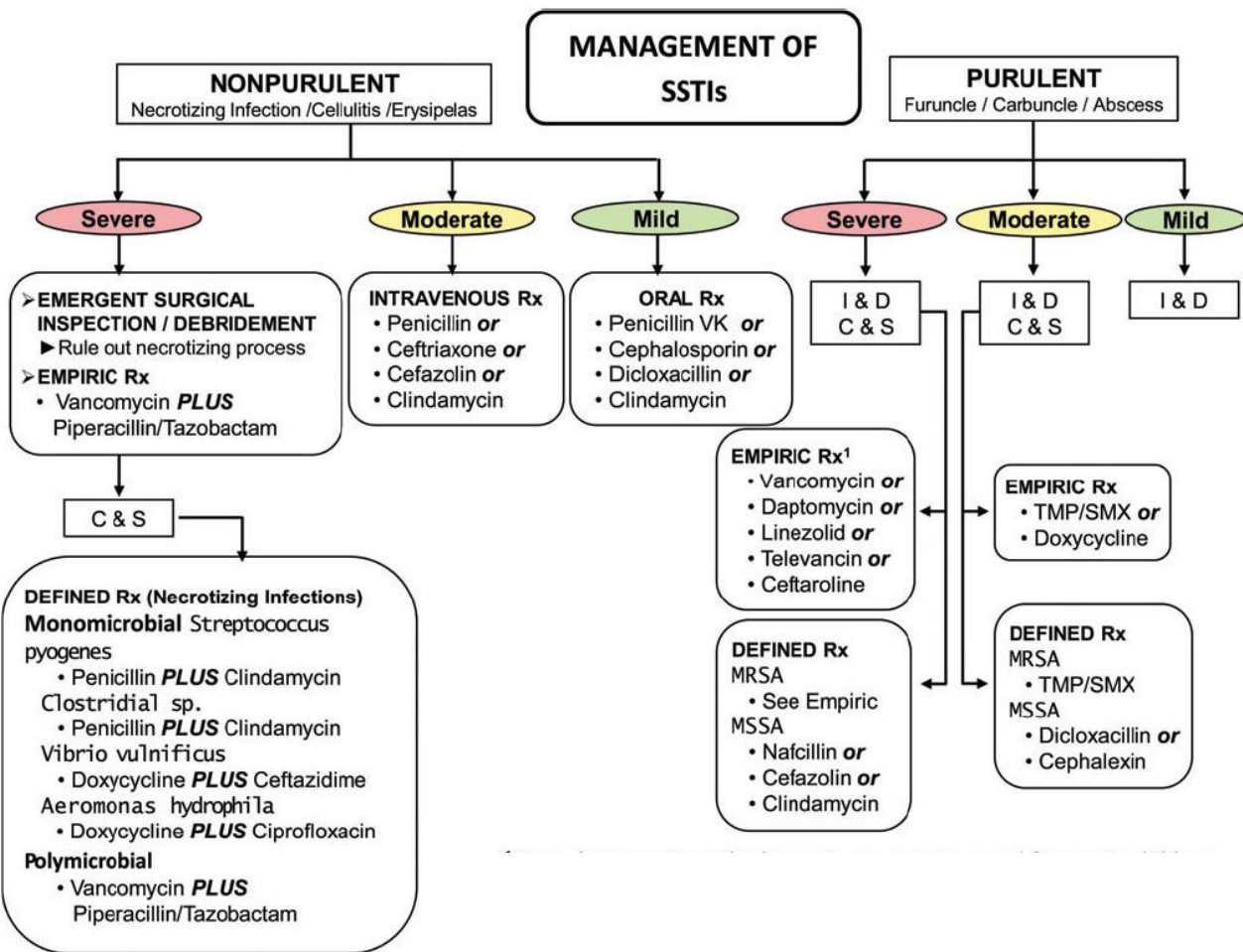Vanessa Stevens – Dept. Internal Medicine, University of Utah School of Medicine
Wendy Chapman – Centre for Digital Transformation of Health, University of Melbourne
Mike Conway – School of Computing & Information Systems, University of Melbourne

AIDH   IMIA

Skin and Soft Tissue Infections (SSTIs)

MANAGEMENT OF SSTIs

NONPURULENT
Necrotizing Infection / Cellulitis / Erysipelas

Severe
- EMERGENT SURGICAL INSPECTION / DEBRIDEMENT
  - ► Rule out necrotizing process
- EMPIRIC Rx
  - Vancomycin **PLUS** Piperacillin/Tazobactam

C & S

DEFINED Rx (Necrotizing Infections)
Monomicrobial Streptococcus pyogenes
- Penicillin **PLUS** Clindamycin

Clostridial sp.
- Penicillin **PLUS** Clindamycin

Vibrio vulnificus
- Doxycycline **PLUS** Ceftazidime

Aeromonas hydrophila
- Doxycycline **PLUS** Ciprofloxacin

Polymicrobial
- Vancomycin **PLUS** Piperacillin/Tazobactam

Moderate
INTRAVENOUS Rx
- Penicillin *or*
- Ceftriaxone *or*
- Cefazolin *or*
- Clindamycin

Mild
ORAL Rx
- Penicillin VK *or*
- Cephalosporin *or*
- Dicloxacillin *or*
- Clindamycin

PURULENT
Furuncle / Carbuncle / Abscess

Severe
I & D
C & S

Moderate
I & D
C & S

Mild
I & D

EMPIRIC Rx[1]
- Vancomycin *or*
- Daptomycin *or*
- Linezolid *or*
- Televancin *or*
- Ceftaroline

DEFINED Rx
MRSA
- See Empiric
MSSA
- Nafcillin *or*
- Cefazolin *or*
- Clindamycin

EMPIRIC Rx
- TMP/SMX *or*
- Doxycycline

DEFINED Rx
MRSA
- TMP/SMX
MSSA
- Dicloxacillin *or*
- Cephalexin

Purulent Infections
Infections of the skin and subcutaneous tissues associated with purulent drainage (pus) or drainable fluid (e.g. abscesses, carbuncles)

Non-Purulent Infections
Infections of the skin which are defined as a diffuse, superficial, spreading infection not associated with purulent drainage or fluid collection

IDSA GUIDELINE

Why not just use structured data?

- Structured data (ICD-9 and -10 codes insufficiently granular)
- Purulent and non-purulent STTIs cannot be captured by ICD codes alone

Purulent Infection
- *There is a fluctuant nodule on the leg*
- *A/P abscess*
- *pt has probable small furuncle of some form on nose*

Non-Pururlent Infection
- *A/P probable cellulitis*
- *Suspected cellulitis of left arm*
- *pt has paroncyhia of the right thumb*

NLP is required to extract purulent vs non-purulent status

⬇ Structured Data

ICD -9 or -10 Code	Description
035, A46	Erysipelas
566, 674.1, H00.03, H60.0, J34.0, K61.0, N48.21, N61.1, N76.6, 091.1	Perianal, breast, eyelid, external ear, nose, or genital abscess
367.01, H05.01, H60.1, N48.22	Cellulitis of orbit, external ear, or genitalia
680	Carbuncle and furuncle
681	Cellulitis and abscess of finger or toe
682	Other cellulitis or abscess

Data and annotation

- Cohort consisted of 2,000 randomly selected US Veterans Affairs patients with an SSTI-related ICD code between 2005 and 2018, yielding 6,576 notes:
    - physician notes
    - physician assistant notes
    - nurse practitioner notes
    - podiatry notes
    - dental notes
- Annotation scheme developed:
    - <u>Mention</u> and <u>document</u> level annotations
    - 400 clinical notes (5 batches of 80)
    - Expert annotators (1 board-certified dermatologist; 1 pharmacist; 1 medical student)
    - Annotation guidelines developed (F-score > 0.9)
- Annotated corpus
    - 6,576 notes (3,288 training; 3,288 test)

Moonstone

- Moonstone rule-based NLP system developed originally for identifying SDoH from clinical notes
- Open-source, Java-based
- System is based around a semantic grammar, a lexical dictionary, a named entity recognition module, and a grammar analysis module
- <u>System advantages</u>: rule-based system allows for rapid development, auditing, and portability
- <u>System disadvantages</u>: requires skill to write rules; complexity of resulting rule set

Moonstone architecture
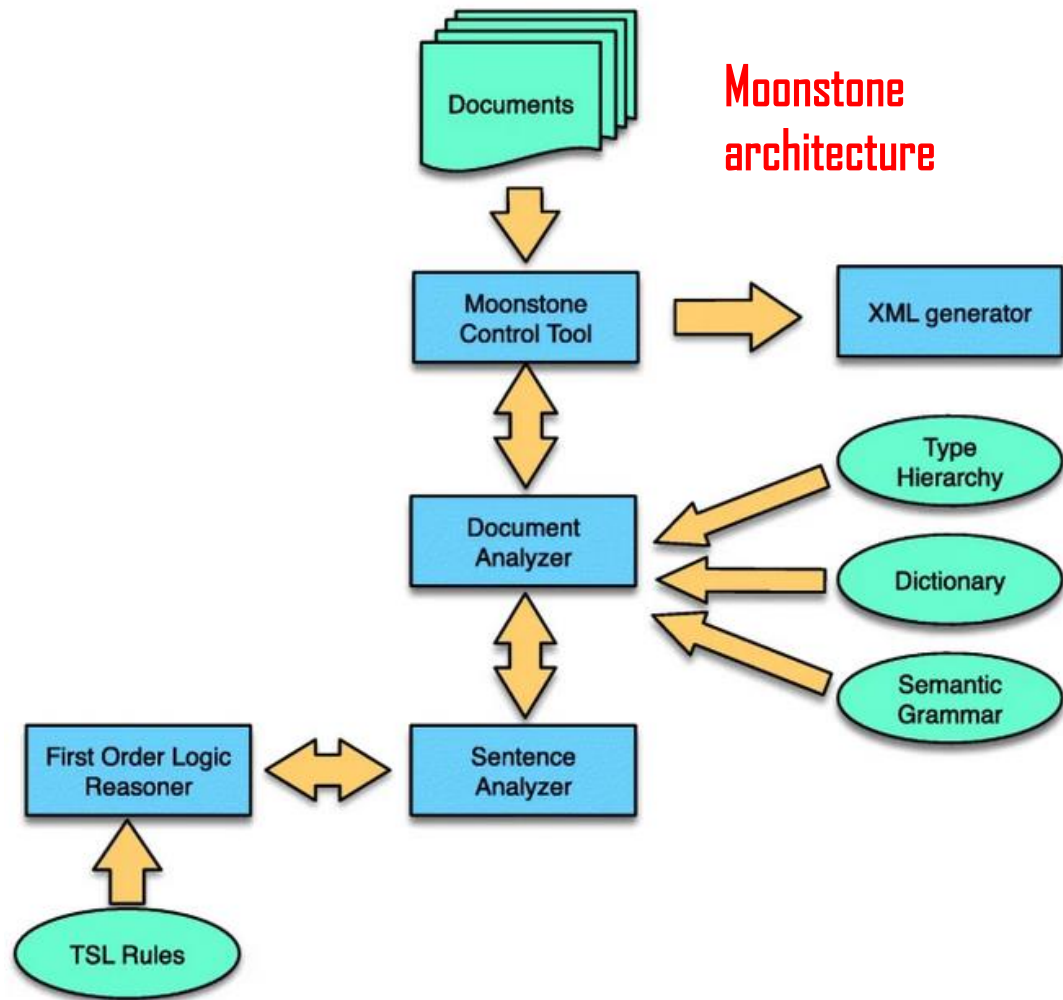
Results – mention level

- F-score for identifying <u>purulent</u> and <u>non-purulent</u> relatively high
- Performance did not degrade substantially compared to training and validation data

Class	Training			Validation			Testing		
	Prec	**Rec**	**F**	**Prec**	**Rec**	**F**	**Prec**	**Rec**	**F**
Purulent	0.77	0.86	0.81	0.71	0.85	0.78	0.7	0.85	0.76
Non-purulent	0.82	0.80	0.81	0.79	0.80	0.80	0.79	0.80	0.80

Results – document level

	Attribute	Training			Validation			Testing		
		Pre	Rec	F	Pre	Rec	F	Pre	Rec	F
SSTI	purulent	0.79	0.85	0.82	0.75	0.88	0.81	0.75	0.90	0.82
	non-purulent	0.91	0.74	0.81	0.89	0.71	0.79	0.91	0.77	0.84

- Relatively high F-score for both purulent and non-purulent document-level classification

Other points
- Also extracted *non-specific SSTIs* and *incision and drainage procedures*
- Hardware requirements limited

Limitations & future work
- Compare to neural NLP methods

Contributions:

1. Evidence that an existing rule-based NLP system – Moonstone – can be adapted to a new, unrelated task
2. The development of an NLP tool that can be utilized for clinical, operational, and epidemiological goals related to automated SSTI type detection

Thanks!

Project leader: Jamie Rhoads, MD, MS
University of Utah
jamie.rhoads@hsc.utah.edu