

# Automated Reconciliation of Radiology Reports with Discharge Summaries

Bevan Koopman, Research Scientist

**AUSTRALIAN E-HEALTH RESEARCH CENTRE**



@bevan\_koopman

# The problem of missed limb fractions in the Emergency Room

# An obvious fracture



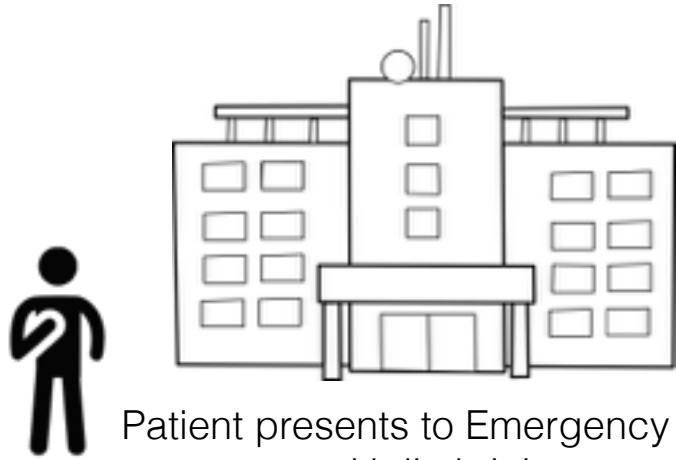
Colles  
fracture

# A subtle abnormality



Defect  
through  
cortex

# The problem of missed limb fractures in the ER



Patient presents to Emergency Room  
with limb injury

# The problem of missed limb fractures in the ER



Xray reviewed  
by ER doctor



Patient presents to Emergency Room  
with limb injury

# The problem of missed limb fractures in the ER



Xray reviewed  
by ER doctor

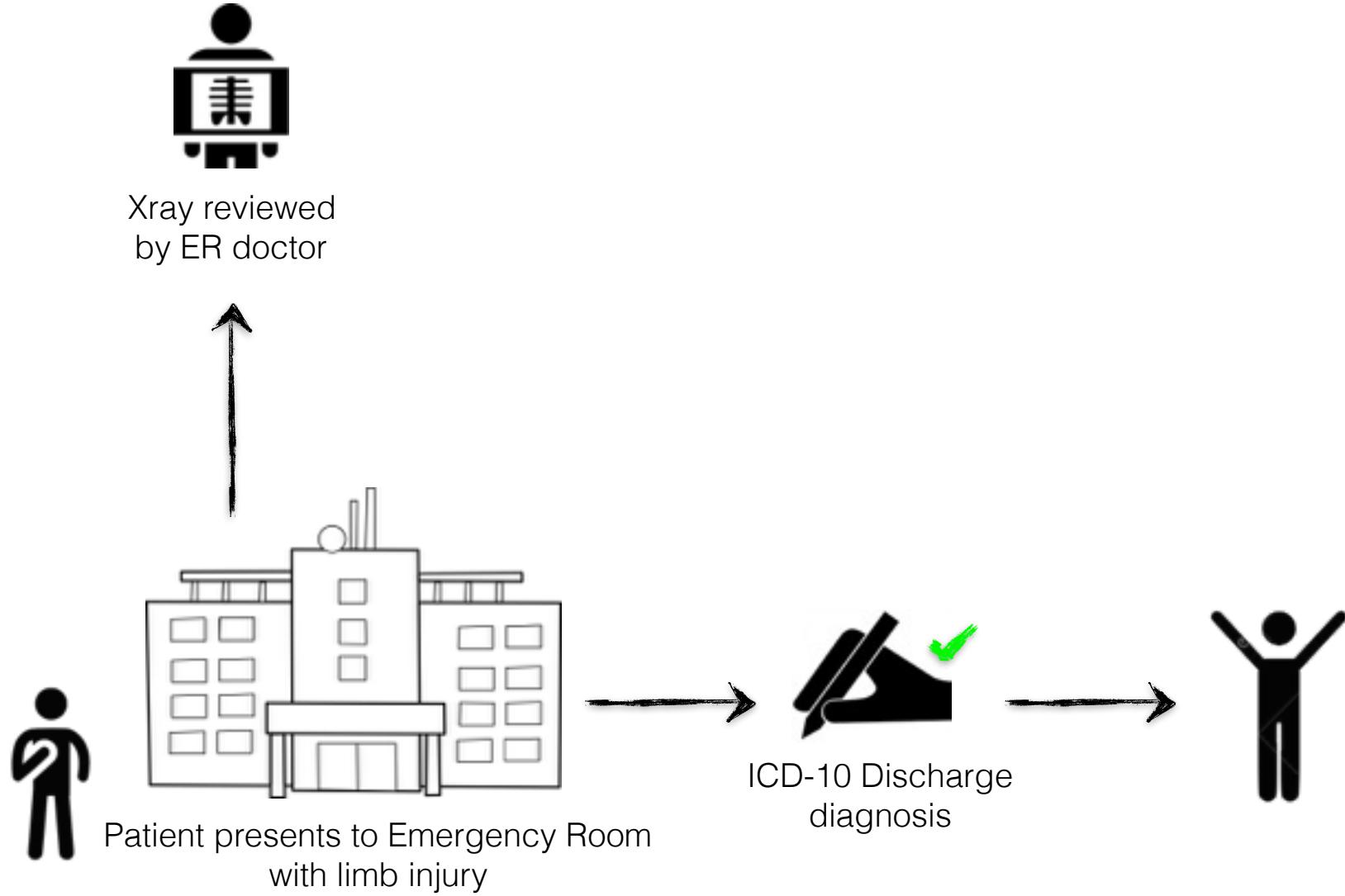


Patient presents to Emergency Room  
with limb injury



ICD-10 Discharge  
diagnosis

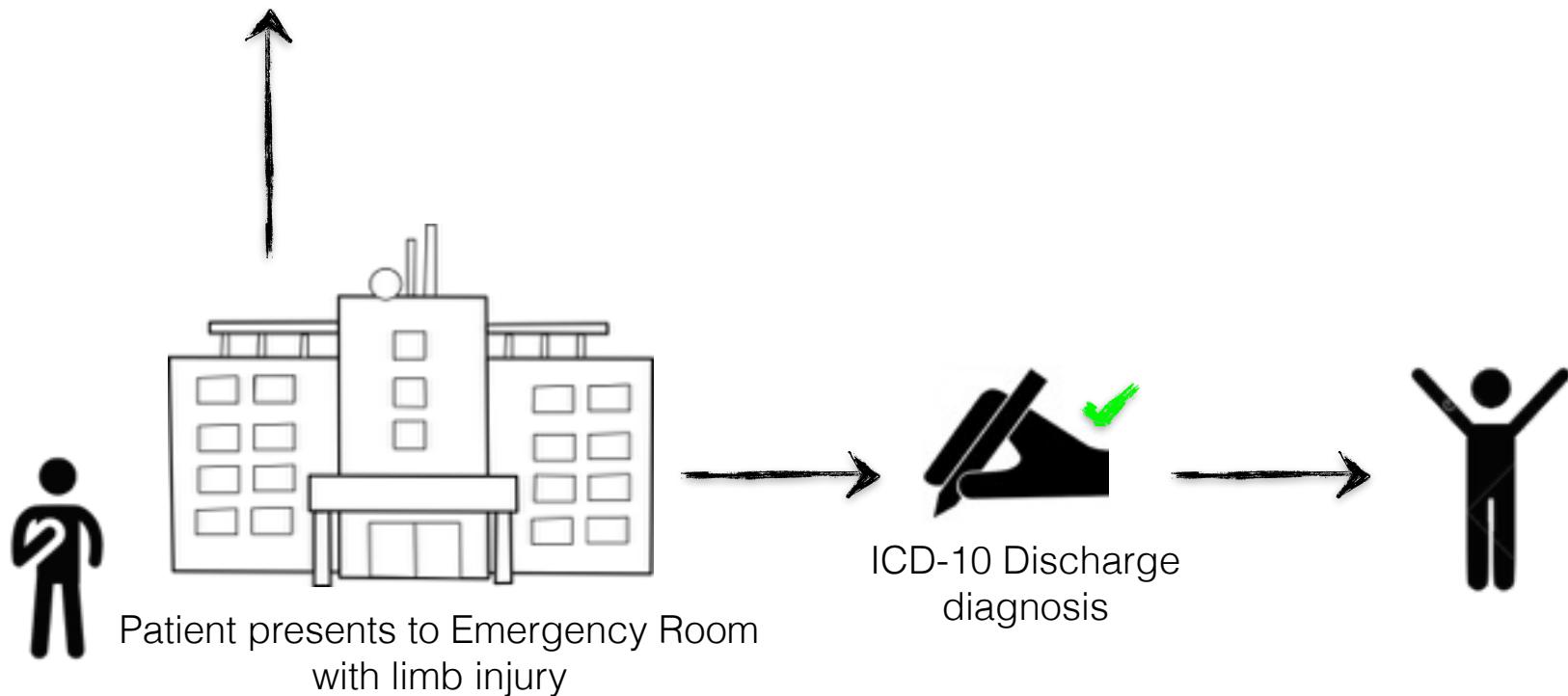
# The problem of missed limb fractures in the ER



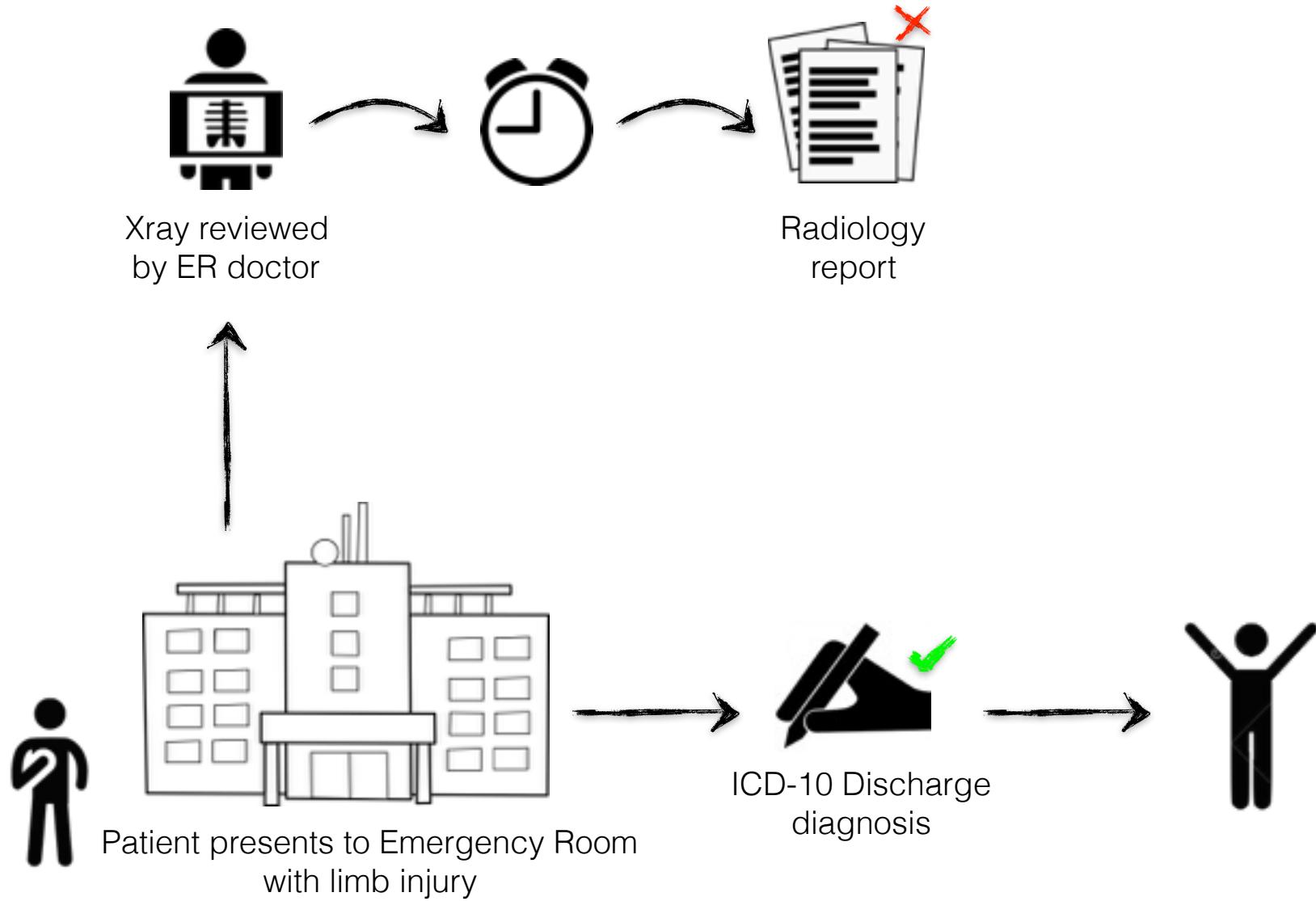
# The problem of missed limb fractures in the ER



Xray reviewed  
by ER doctor



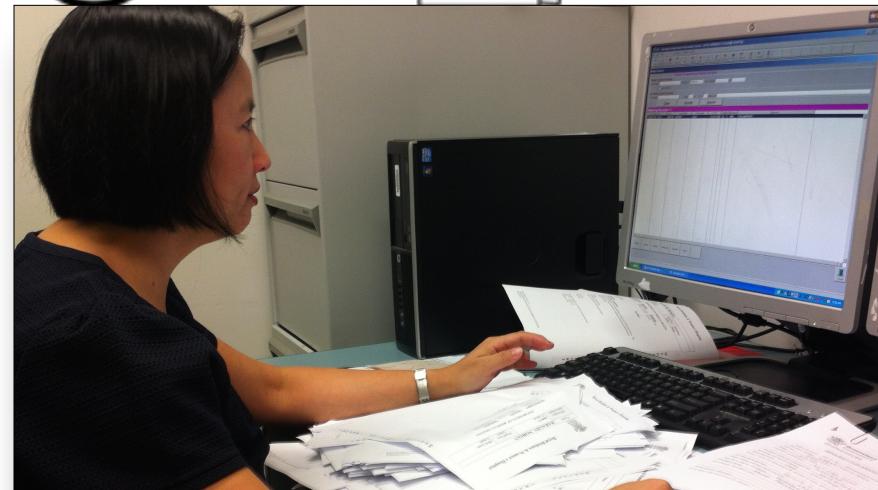
# The problem of missed limb fractures in the ER



# The problem of missed limb fractures in the ER



Xray reviewed  
by ER doctor



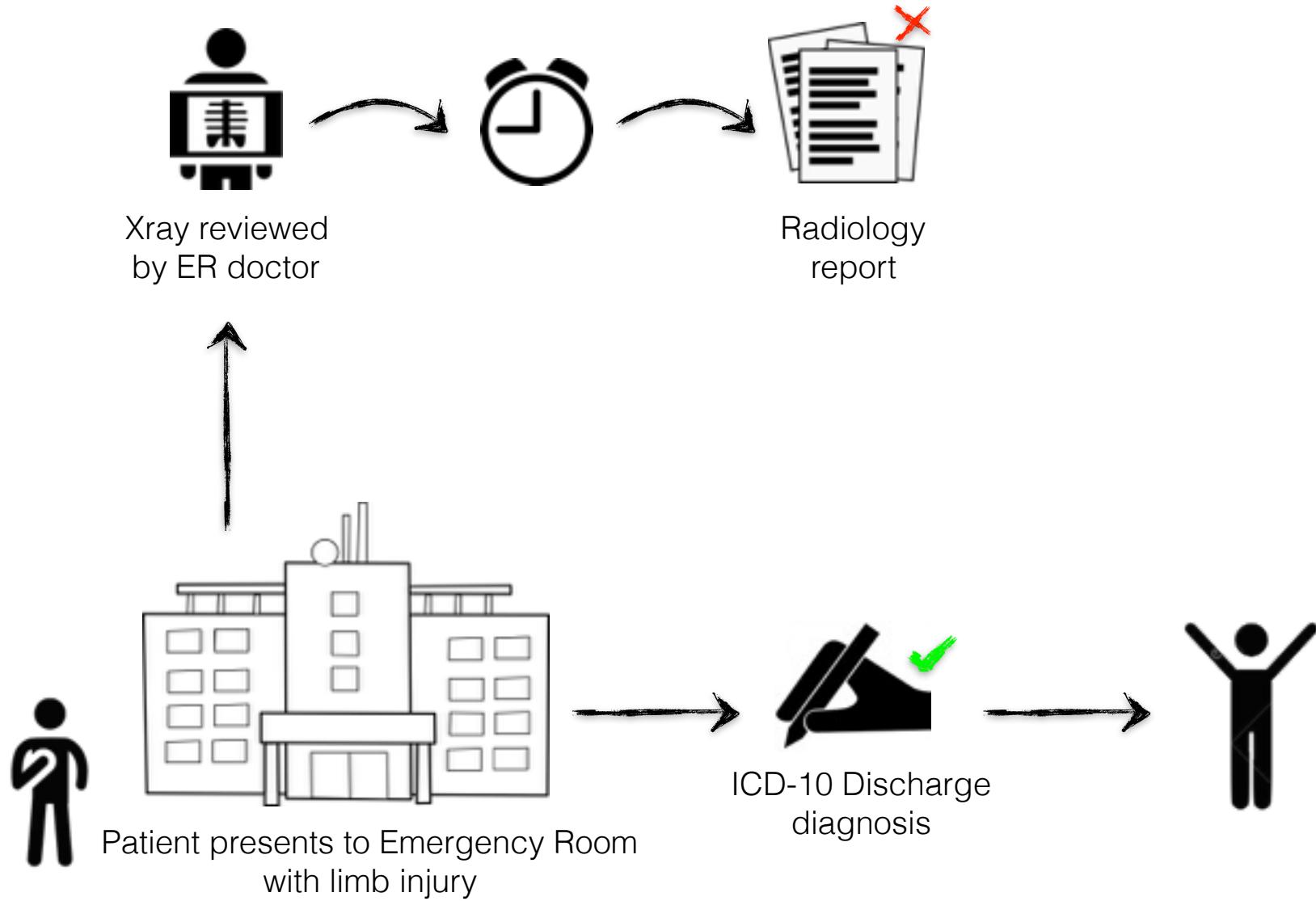
Patient presents to Emergency Room  
with limb injury



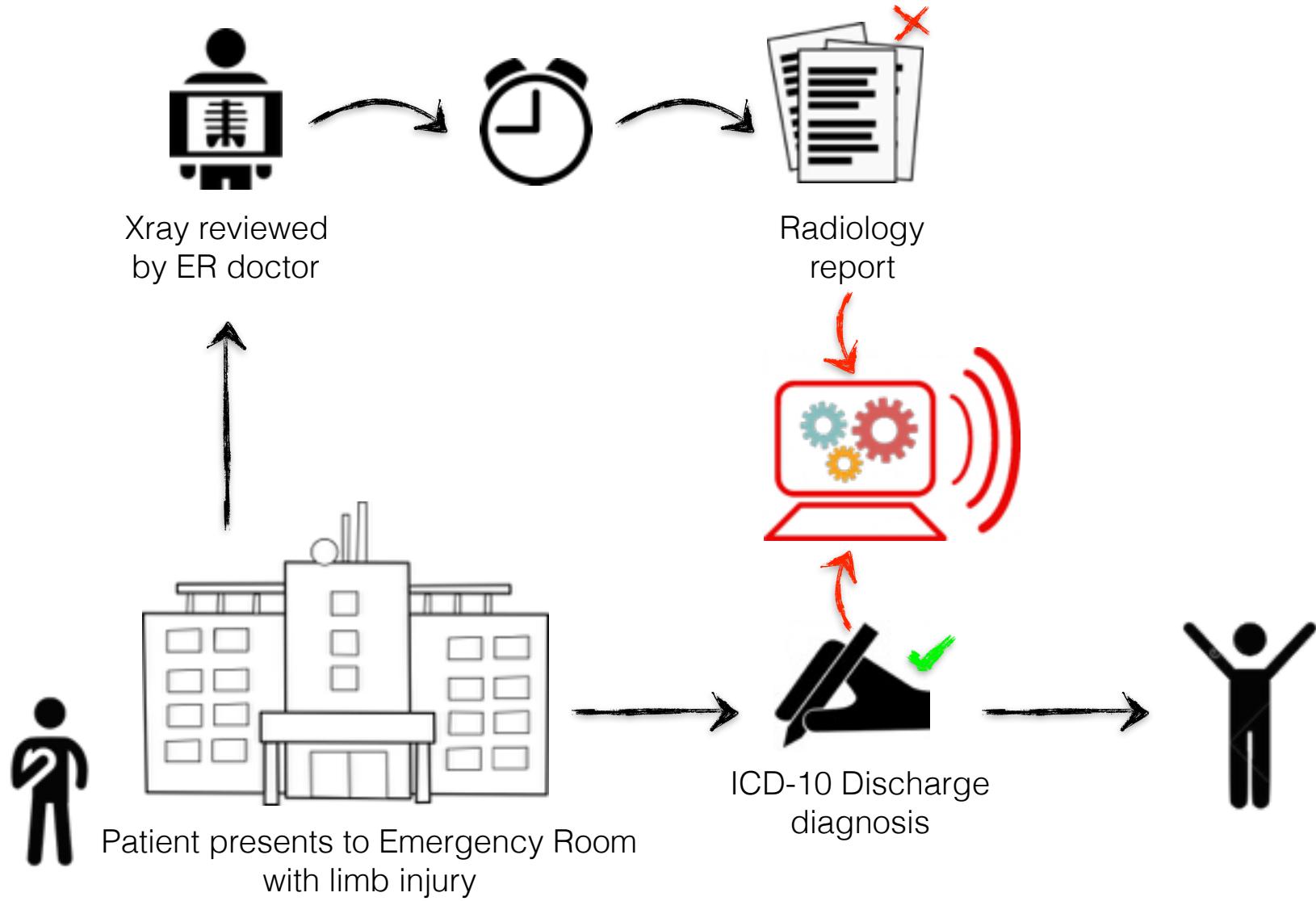
ICD-10 Discharge  
diagnosis



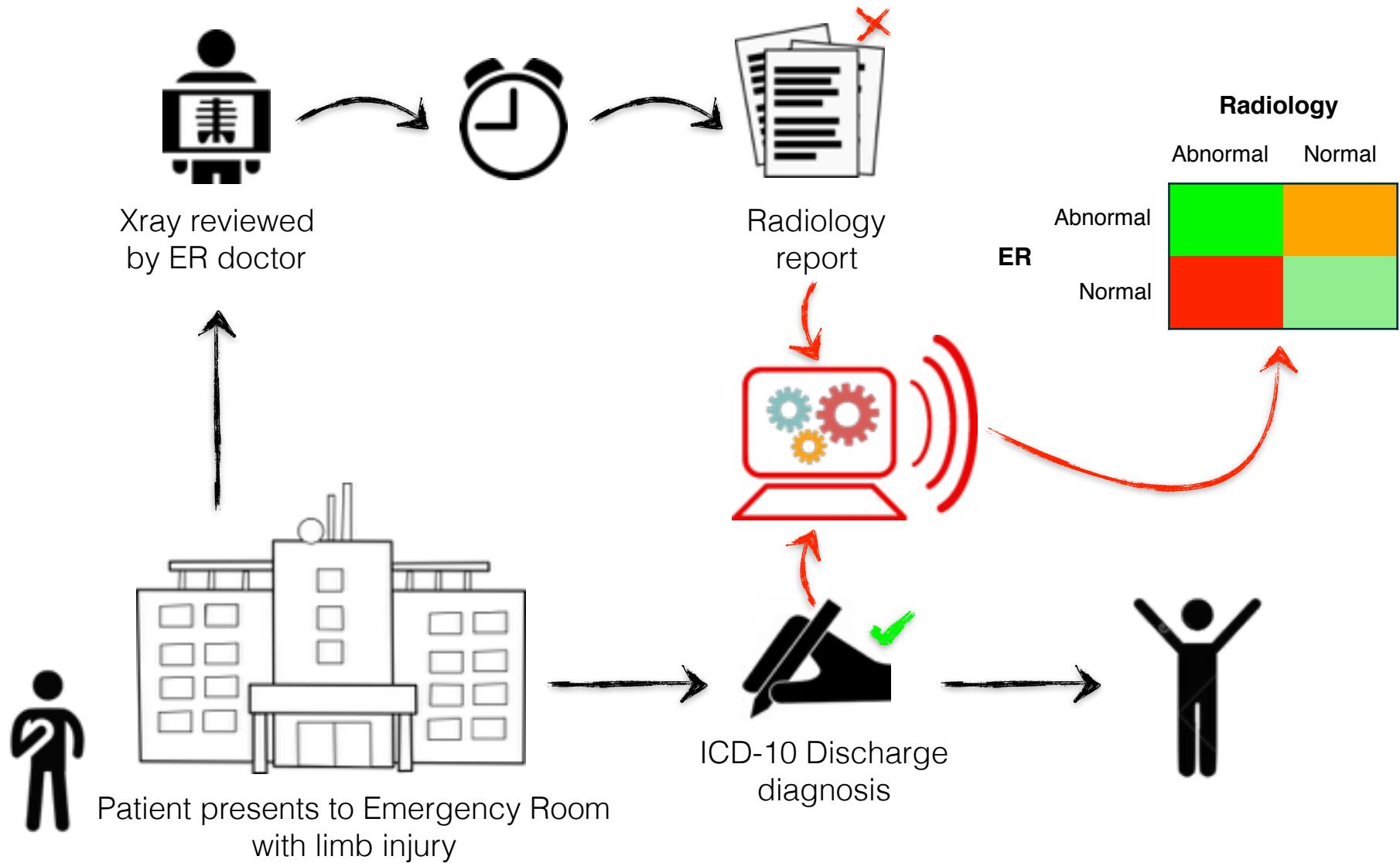
# The problem of missed limb fractures in the ER



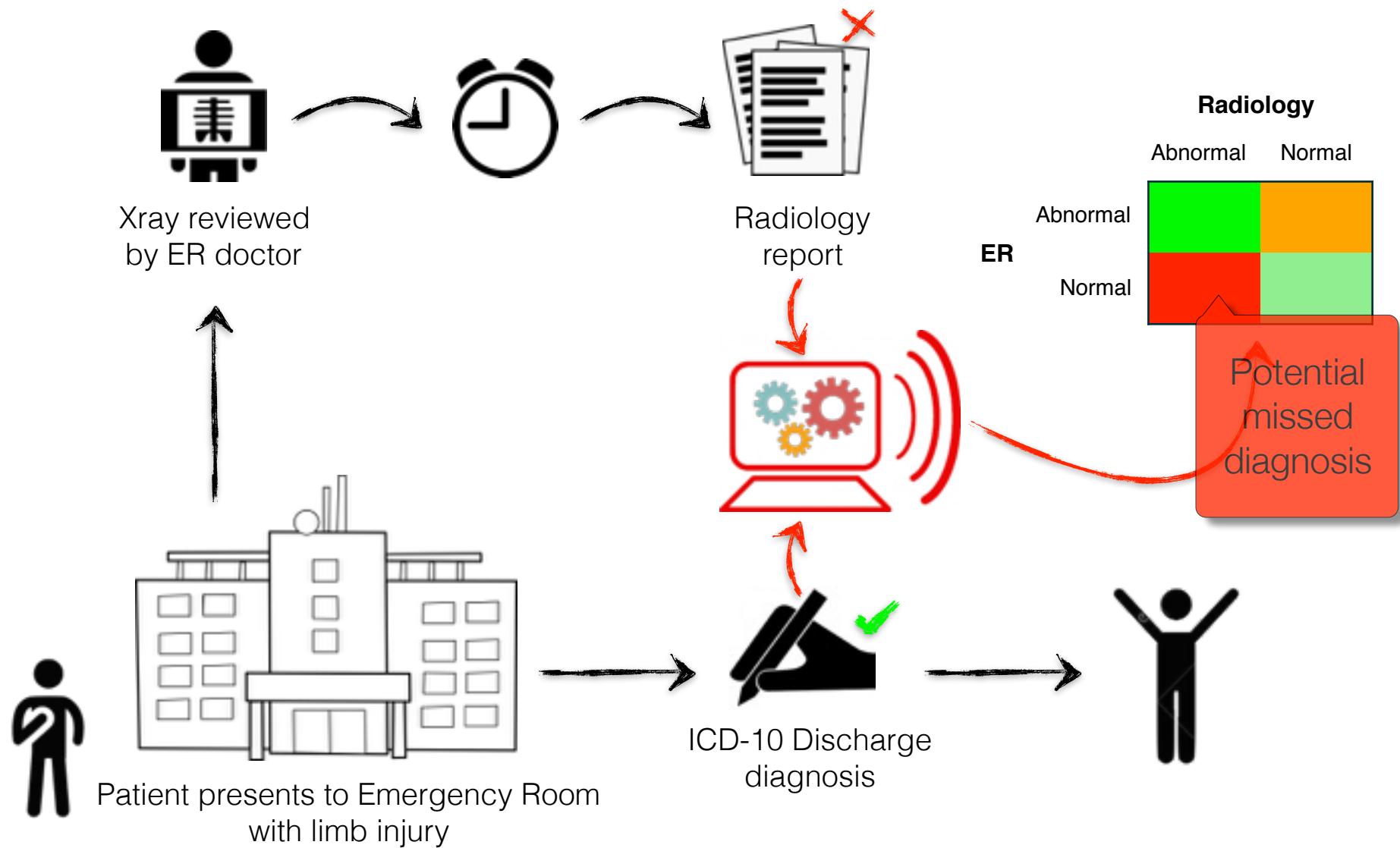
# The problem of missed limb fractions in the ER



# The problem of missed limb fractures in the ER



# The problem of missed limb fractures in the ER



# Automatic classification of abnormalities in radiology reports

# Three different radiology datasets

<b>Dataset</b>	<b>Description</b>	<b>#Reports</b>	<b>Normal</b>	<b>Abnormal</b>
RBWH	Royal Brisbane & Womens' Hospital (adult)	1,480	58%	42%
RCH	Royal (Brisbane) Childrens' Hospital (child)	498	66%	34%
GCH	Gold Coast Hospital (adult child 38%)	400	62%	38%

## Gold Standard by three ER physicians:

- ER Registrar
- Senior Medical Officer
- Adjudicated by Consultant

# Extracting features from radiology reports

- Token (i.e., a single word)
- punctuation
- token stem; e.g, “fractur”
- token negation; e.g., “no fractur”
- token stem bi-gram
- token stem tri-gram
- SNOMED-CT concept
- SNOMED-CT concept bi-gram

# Train Support Vector Machine (SVM)

# Train Support Vector Machine (SVM)

- How effective is a classifier trained on a mixture of hospital reports?

# Train Support Vector Machine (SVM)

- How effective is a classifier trained on a mixture of hospital reports?
- How does effectiveness differ when we consider each hospital dataset in isolation?

# Train Support Vector Machine (SVM)

- How effective is a classifier trained on a mixture of hospital reports?
- How does effectiveness differ when we consider each hospital dataset in isolation?
- How robust are the methods across different hospitals (train one, test on another)?

# Train Support Vector Machine (SVM)

- How effective is a classifier trained on a mixture of hospital reports?
- How does effectiveness differ when we consider each hospital dataset in isolation?
- How robust are the methods across different hospitals (train one, test on another)?
- In comparison with keyword spotting baseline.

(Thomas et al., 2005; Waghlikar et al., 2013).

# Classification Results (F-measure)

		Testing			
		GCH <sup>CA</sup>	RBWH <sup>A</sup>	RCH <sup>C</sup>	All
Training	GCH <sup>CA</sup>	0.95	0.80	0.88	
	RBWH <sup>A</sup>	0.84	0.89	0.85	
	RCH <sup>C</sup>	0.87	0.81	0.95	
	GCH+RBWH				0.88
	GCH+RCH			0.80	
	RBWH+RCH	0.84			
All					0.92

**C** = Children; **A** = Adults.

33% avg. improvement over keyword baseline

# Classification Results (F-measure)

Training	Testing			
	GCH <sup>CA</sup>	RBWH <sup>A</sup>	RCH <sup>C</sup>	All
GCH <sup>CA</sup>	0.95	0.80	0.88	
RBWH <sup>A</sup>	0.84	0.89	0.85	
RCH <sup>C</sup>	0.87	0.81	0.95	
GCH+RBWH				0.88
GCH+RCH			0.80	
RBWH+RCH	0.84			
All				0.92

**C** = Children; **A** = Adults.

33% avg. improvement over keyword baseline

# Classification Results (F-measure)

Training	Testing			
	GCH <sup>CA</sup>	RBWH <sup>A</sup>	RCH <sup>C</sup>	All
GCH <sup>CA</sup>	0.95	0.80	0.88	
RBWH <sup>A</sup>	0.84	0.89	0.85	
RCH <sup>C</sup>	0.87	0.81	0.95	
GCH+RBWH				0.88
GCH+RCH			0.80	
RBWH+RCH	0.84			
All				0.92

**C** = Children; **A** = Adults.

33% avg. improvement over keyword baseline

# Classification Results (F-measure)

		Testing			
		GCH <sup>CA</sup>	RBWH <sup>A</sup>	RCH <sup>C</sup>	All
<sup>CA</sup> GCH		0.95	0.80	0.88	
<sup>A</sup> RBWH		0.84	0.89	0.85	
<sup>C</sup> RCH		0.87	0.81	0.95	
GCH+RBWH				0.88	
GCH+RCH			0.80		
RBWH+RCH		0.84			
All					0.92

<sup>C</sup> = Children; <sup>A</sup> = Adults.

33% avg. improvement over keyword baseline

# Classification Results (F-measure)

Training	Testing			
	GCH	RBWH	RCH	All
GCH	0.95	0.80	0.88	
RBWH	0.84	0.89	0.85	
RCH	0.87	0.81	0.95	
GCH+RBWH				0.88
GCH+RCH		0.80		
RBWH+RCH	0.84			
All				0.92

C = Children; A = Adults.

33% avg. improvement over keyword baseline

## Classification Results (F-measure)

Able to identify limb abnormalities with high degree of accuracy.

Robust across different hospital reports.

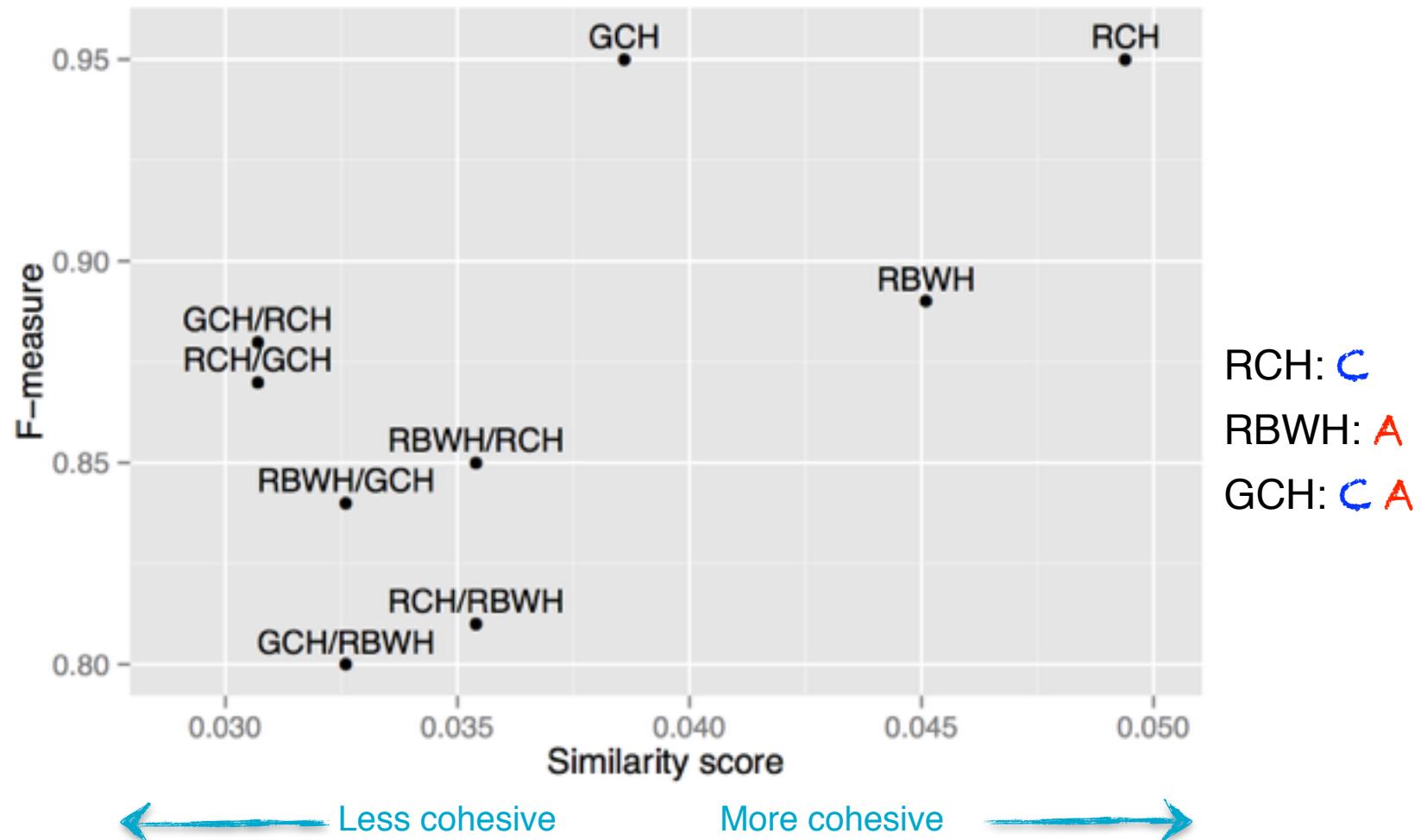
More training data does not equate to better effectiveness.

**C** = Children; **A** = Adults.

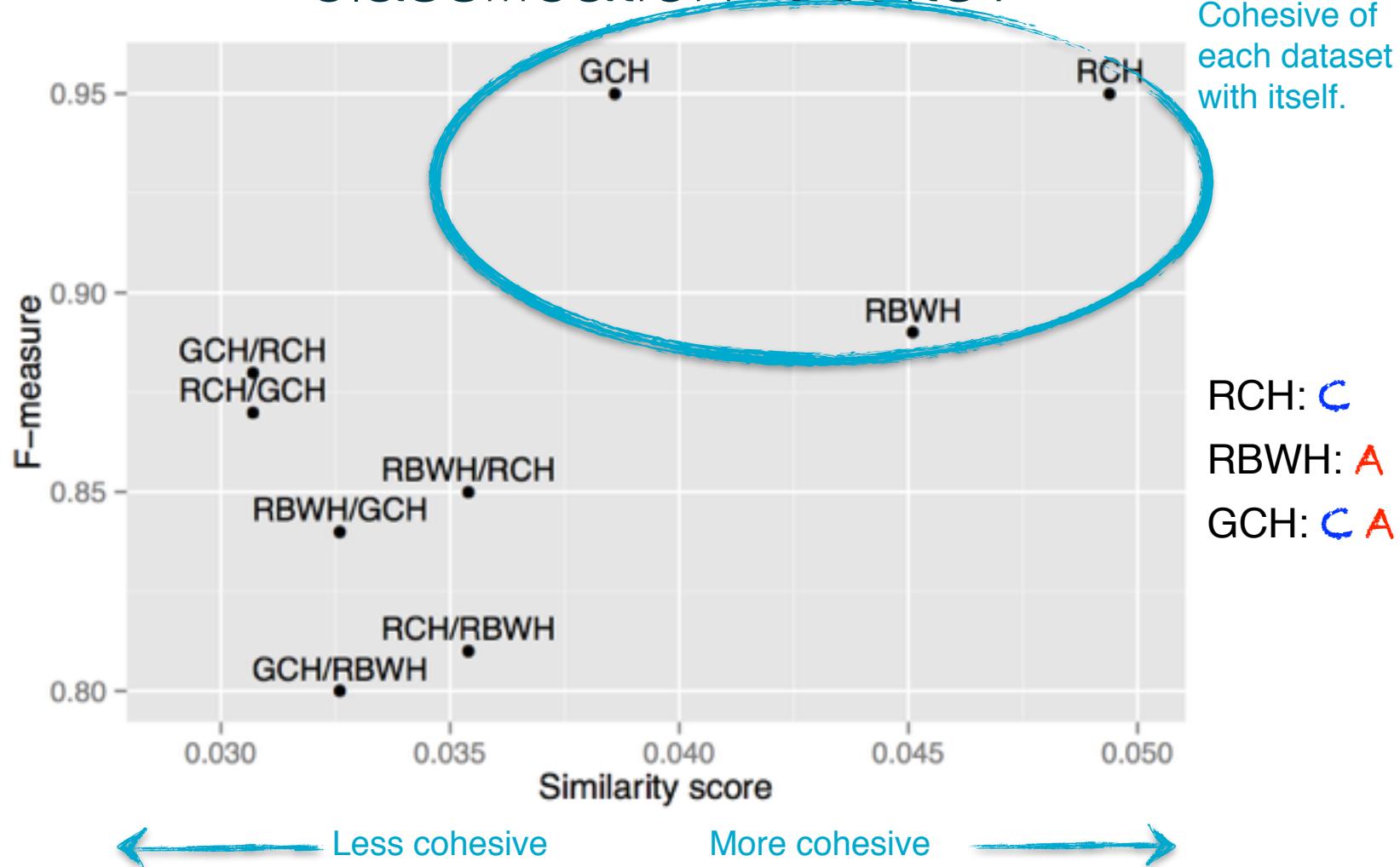
33% avg. improvement over keyword baseline

# How does report similarity affect classification results?

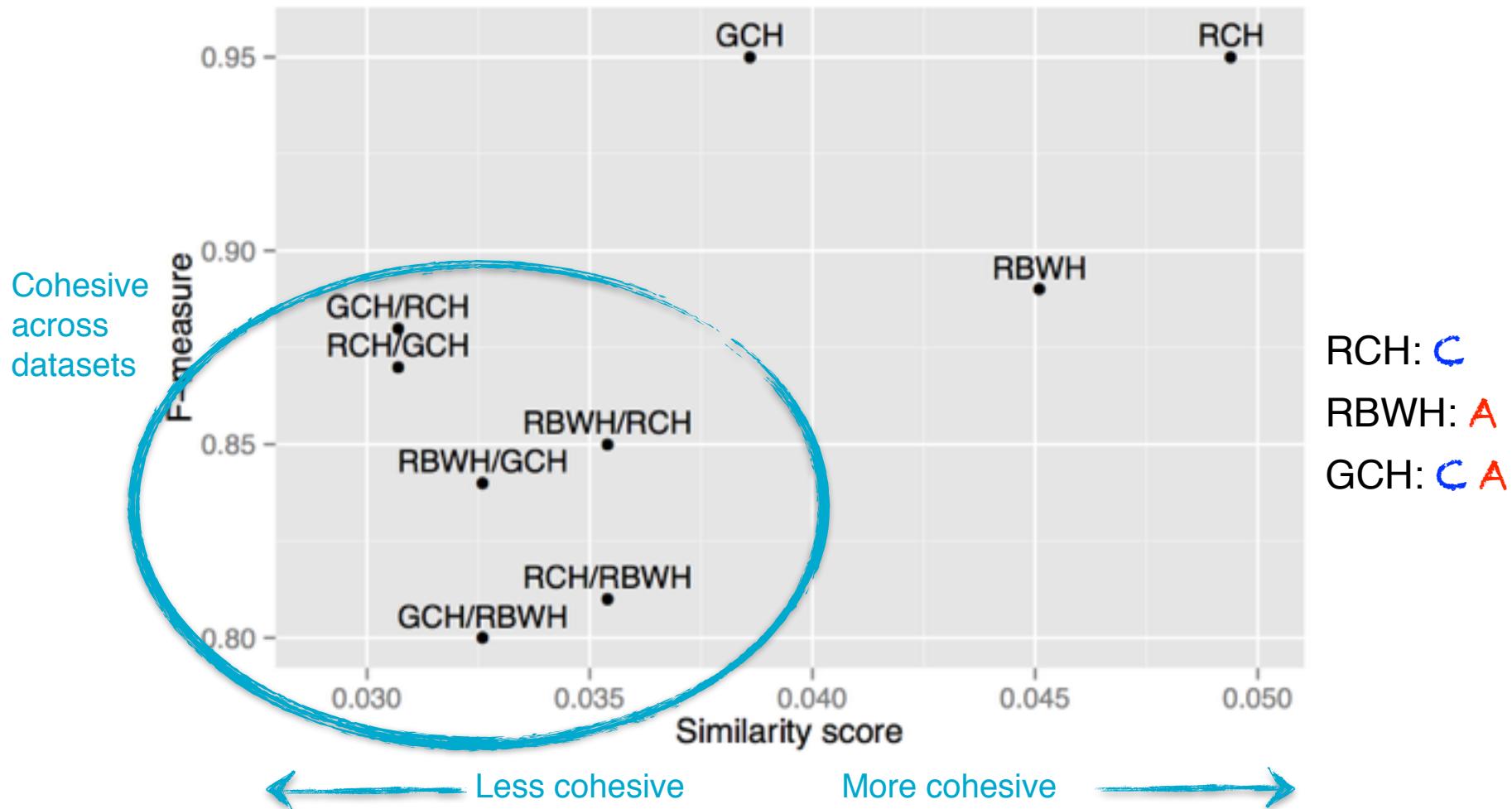
# How does report similarity affect classification results?



# How does report similarity affect classification results?



# How does report similarity affect classification results?



# How does report similarity affect classification results?

Effectiveness not correlated with similarity.

Variance in training data provides model with a mix of reporting styles.

Overall difference are small - models are robust across different hospitals

← Less cohesive      More cohesive →

A  
A

# Reconciliation with Discharge Diagnosis

# Reconciliation



Radiology report  
classification



ICD-10 Discharge  
diagnosis

S63.0: Dislocation of wrist  
S22.3: Fracture of rib  
...

# Reconciliation



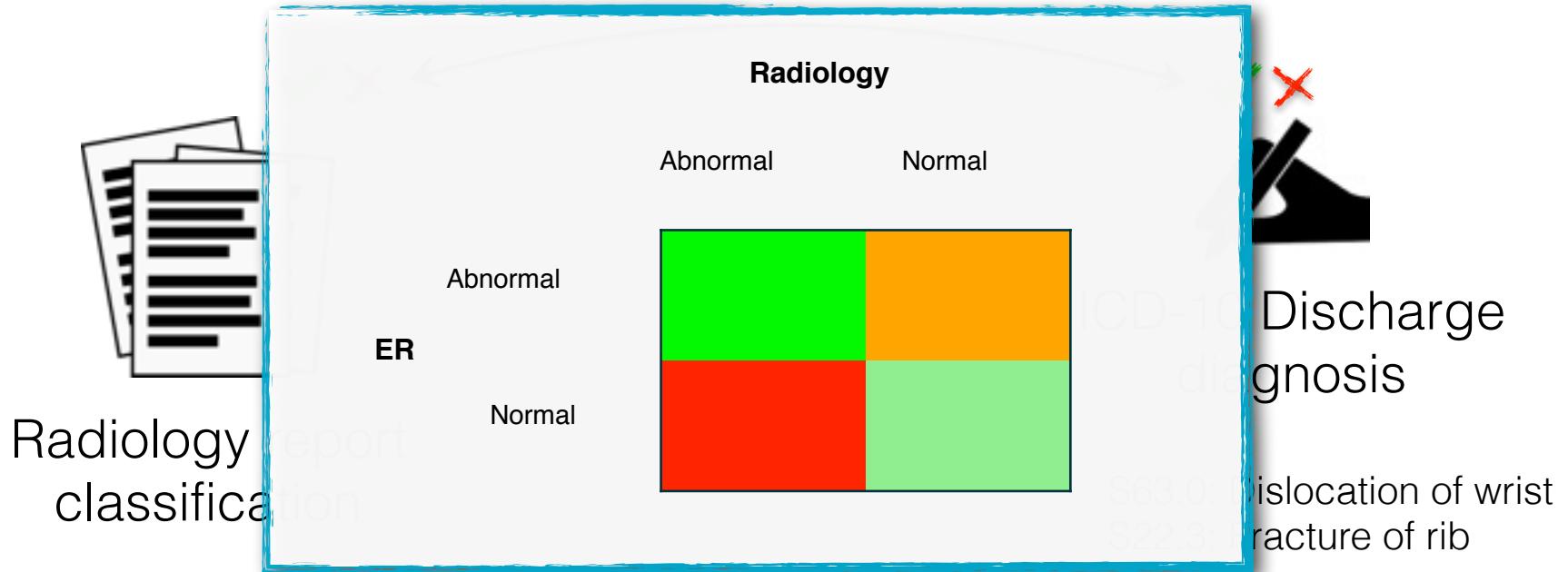
Radiology report  
classification



ICD-10 Discharge  
diagnosis

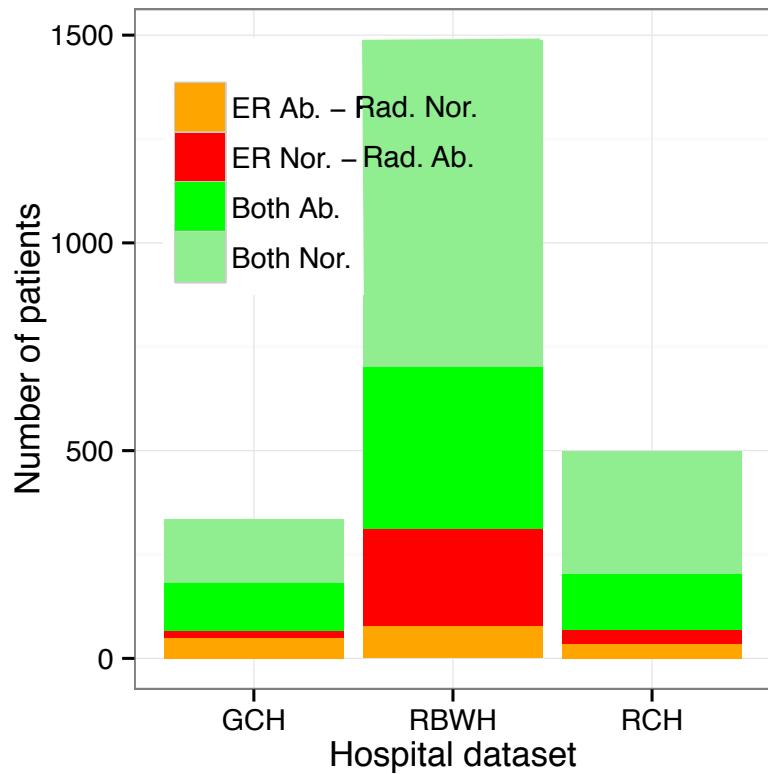
S63.0: Dislocation of wrist  
S22.3: Fracture of rib  
...

# Reconciliation

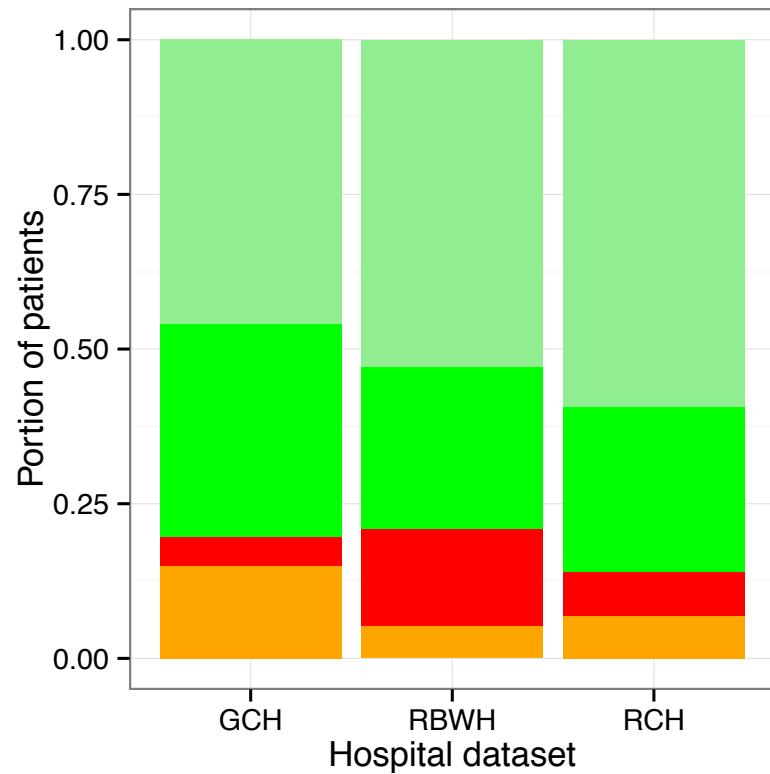


# Breakdown of diagnosis reconciliation

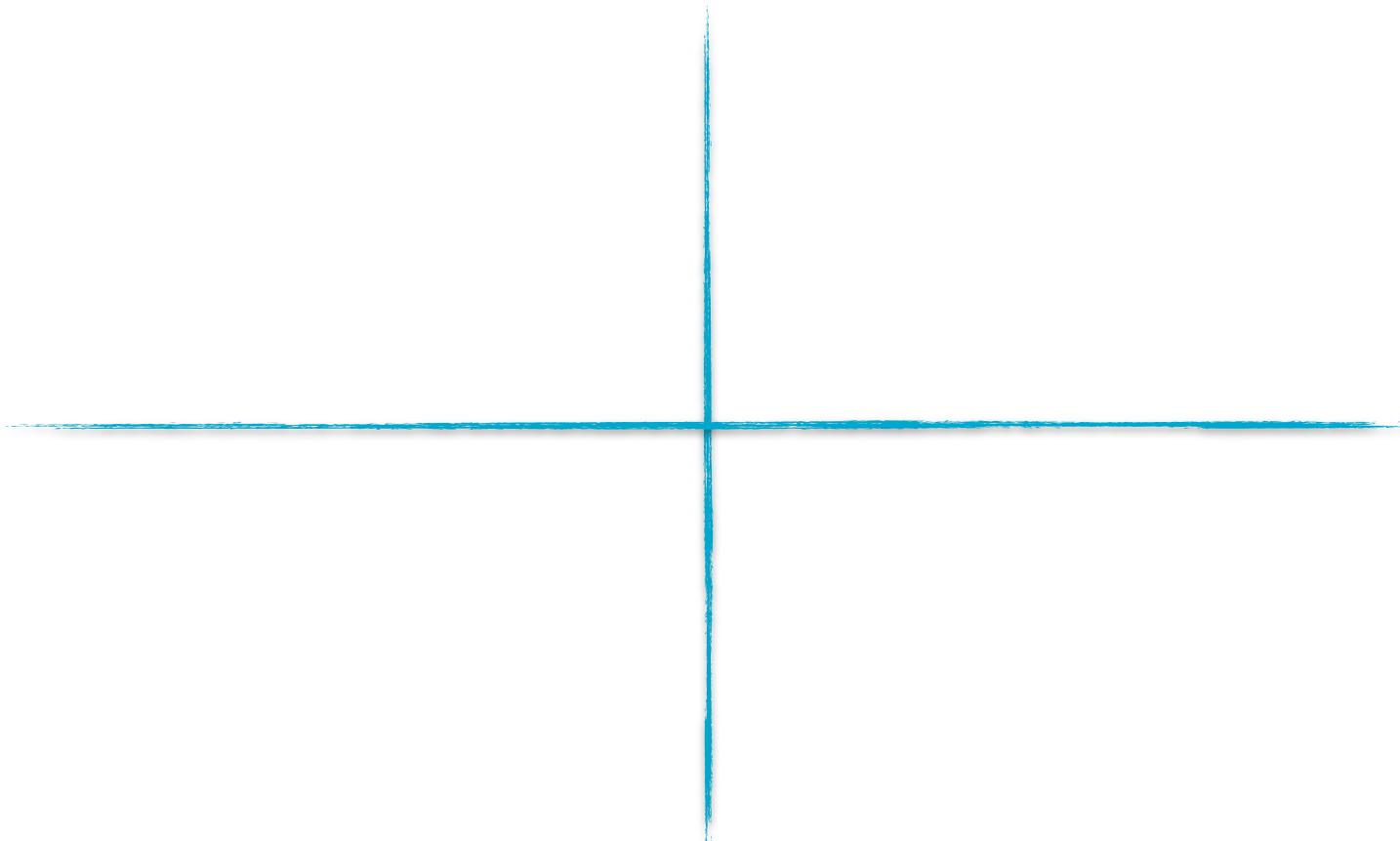
Absolute



Normalised



# 4 categories of red 'flagged cases'



# 4 categories of red 'flagged cases'

Related ED diagnosis

ER ICD: *Crush injury*

ER ICD: *Laceration*

ER ICD: *Strain or sprain*

# 4 categories of red 'flagged cases'

Related ED diagnosis

ER ICD: **Crush injury**

ER ICD: **Laceration**

ER ICD: **Strain or sprain**

Unrelated ED diagnosis

ER ICD: **Self-harm**

ER ICD: **Congestive heart failure**

# 4 categories of red 'flagged cases'

Related ED diagnosis

ER ICD: **Crush injury**

ER ICD: **Laceration**

ER ICD: **Strain or sprain**

Unrelated ED diagnosis

ER ICD: **Self-harm**

ER ICD: **Congestive heart failure**

More general ED diagnosis

ER ICD: **Injury of the lower foot.**

# 4 categories of red 'flagged cases'

Related ED diagnosis

ER ICD: **Crush injury**

ER ICD: **Laceration**

ER ICD: **Strain or sprain**

Unrelated ED diagnosis

ER ICD: **Self-harm**

ER ICD: **Congestive heart failure**

More general ED diagnosis

ER ICD: **Injury of the lower foot.**

Missed ED diagnosis

ER ICD: **No injury found**

ER ICD: **Patient did not wait.**

# Final messages

- Can effectively identify limb abnormalities from radiology reports using NLP+ML
- Robust across different hospital's reports
- Classification results can be reconciled against discharge diagnosis
- Potentially undiagnosed limb fractures can be identified
- Saving doctors time by only reviewing “flagged” cases
- Extend beyond limb abnormalities to cancers and pathology reports



**Queensland Government**  
Queensland Health

## Australian e-Health Research Centre, CSIRO

Dr. Bevan Koopman  
Dr. Anthony Nguyen  
Mr. John O'Dwyer  
Dr. Amol Wagholtikar

Royal Brisbane Hospital  
Royal Brisbane Children's Hospital  
Gold Coast Hospital

Dr. Kevin Chu  
Dr. Ibrahim Mahmoud  
Dr. Kim Lai  
Dr. Gerben Keijzers



Funding by Queensland Emergency Medicine Research  
Foundation Grant, EMPJ-11-158-Chu- Radiology.  
No disclosure.