

Decision Making By Emergency Room Physicians And Residents: Results From a Clinical Trial



Presented by: Mike Hine, Carleton University

Based on Working Paper by:

Ken Farion, CHEO

Mike Hine, Carleton University

Wojtek Michalowski, University of Ottawa

Szymon Wilk, Poznan University of Technology



To be submitted to: Decision Sciences



Agenda

- Clinical Decision Support Systems (CDSS)
- Expert / Novice Decision Making
- Research Purpose
- Mobile Emergency Triage (MET)
- Methodology
- Analysis
- Results
- Findings
- Implications



CDSS: Background

- “any computer program designed to help health-care professionals make clinical decisions” (Musen, Shahar, & Shortcliffe, 2001)
 - Tools for information management
 - medical information retrieval systems for managing and extracting medical knowledge, and electronic patient record systems (EPRS: Shortcliffe, 1993) for managing patient data
 - Tools for focusing attention
 - remind clinicians about actions that might require attention. Such systems are normally present in the ICU or laboratories
 - Tools for providing patient-specific recommendations
 - assess or advise using patient-specific clinical data



CDSS: Development

- Decision models are based on expert knowledge
 - From past data and research, medical books, clinical practice guidelines (CPG)
- Relying on expert knowledge assumes that clinicians will provide accurate data for ALL required inputs
 - Accurate collection
 - Accurate assessment
- Implies equal level of collection and assessment skills across users



Expert/Novice Decision Making

- In medical institutions, there are different clinician user groups based on expertise (staff physician vs. residents)
- in complex domains such as medicine, it typically takes 10 years of experience before one can be considered an expert (Prietula & Simon, 1989)
- Novice clinicians have weaker information gathering and decision making skills (Johnson & Carpenter, 1986; Mangione et al. 1995)



Research Purpose

- Investigate the usage of a CDSS tool by different clinician user groups (staff physicians and residents)
 - evaluate differences between these two groups
 - Do they collect and assess attributes necessary for system input differently?
 - Do they use similar decision models?
- Addresses a call for a better understanding of real medical personnel making ill structured decision in a naturalistic setting (Klein, 1993) as mediated by technology (Kushniruk, 2001).



MET: Overview

- supports the emergency triage of various presentations of acute pain
- contains clinical modules for handling abdominal pain (Michalowski, Rubin, Slowinski, & Wilk, 2004) and scrotal pain (Michalowski, Wilk, Farion, Pike, Rubin, & Slowinski, 2005)
- The client (PDA) stores clinical data collected during examination by staff physicians and residents and supplies the triage support function.
- Is used directly at the point of care.



MET: Triage Decision

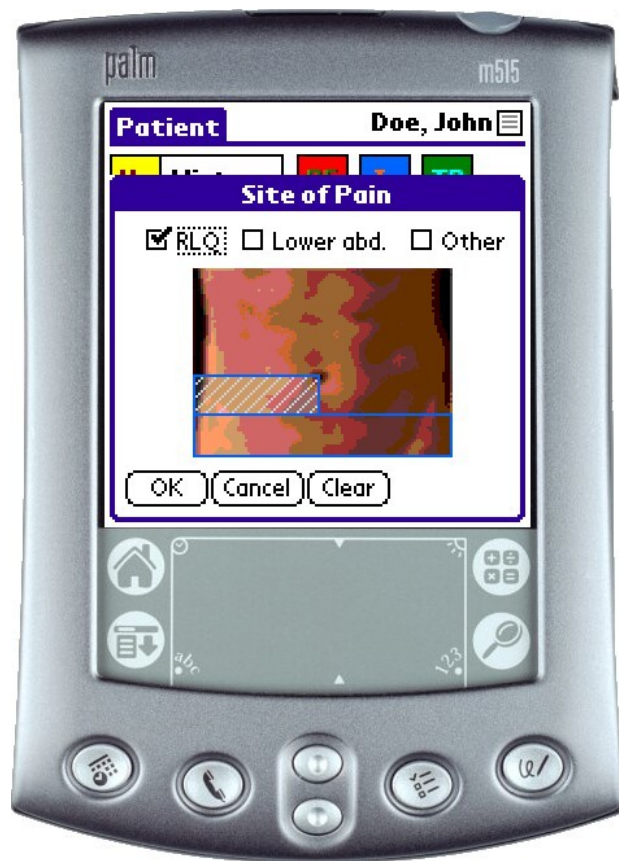
- **Discharge:** patient can be discharged to the care of their regular physician
- **Consult:** surgeon is called because acute appendicitis suspected
- **Follow Up:** further in-hospital evaluation is required



MET: Decision Attributes

Attribute Code	Attribute Description	Possible Values
Age	Age in number of years	0-6, >= 7 years
Guard	Localized guarding: localized muscle sustained contraction noted when palpating the abdomen	Absent, Present
PainDur	Duration of pain	<=24 hrs, 1-7 and >7 days
PainShift	Shifting of pain	Absent, Present
PainSite	Site of maximal pain	Right lower quadrant (RLQ), lower abdomen, other
PainType	Type of maximal pain	continuous, other
PrevVis	Previous visits in the ER for abdominal pain during the last 48 hours (irrespective of site)	yes, no
RebTend	Rebound tenderness: pain felt at site of maximal tenderness, produced by altering intra-abdominal pressure	absent, present
Sex	Gender	male, female
Tempr	Temperature	<37, 37-39, >= 39 Cel
TendSite	Site of maximal tenderness	RLQ, lower abdomen, other
Vomiting	Vomiting	yes, no
WBC	WBC (white blood cells)	<=4000, 4000-12000, >=12000

MET: Sample Interface





Methodology

- convenience sample of patients was approached 24/7 for approx 6 months in late 2003-early 2004
- Patients arriving at the ED with acute abdominal pain (AP) of less than 10 days duration were considered eligible for enrollment
- To enroll a patient, the clinician retrieved a handheld and identified himself/herself as either resident or staff physician
- Clinician completed the data entry using the structured data collection screens.



Methodology Ctd.

- Clinicians told to only record data for those attributes they felt were relevant to the patient's presentation
- Clinician enters his/her triage prediction
- Where possible, a second observer of the opposite clinician level (i.e., resident, staff physician) independently assessed the patient within one hour of the primary assessment.
- At no point did the clinicians receive a triage recommendation from the system



Participants

- Patients (n=574)
 - 222 seen by both staff physician and resident
 - 235 seen by only staff physician
 - 117 seen by only resident
- Staff Physicians (n=110)
- Residents (n=40)

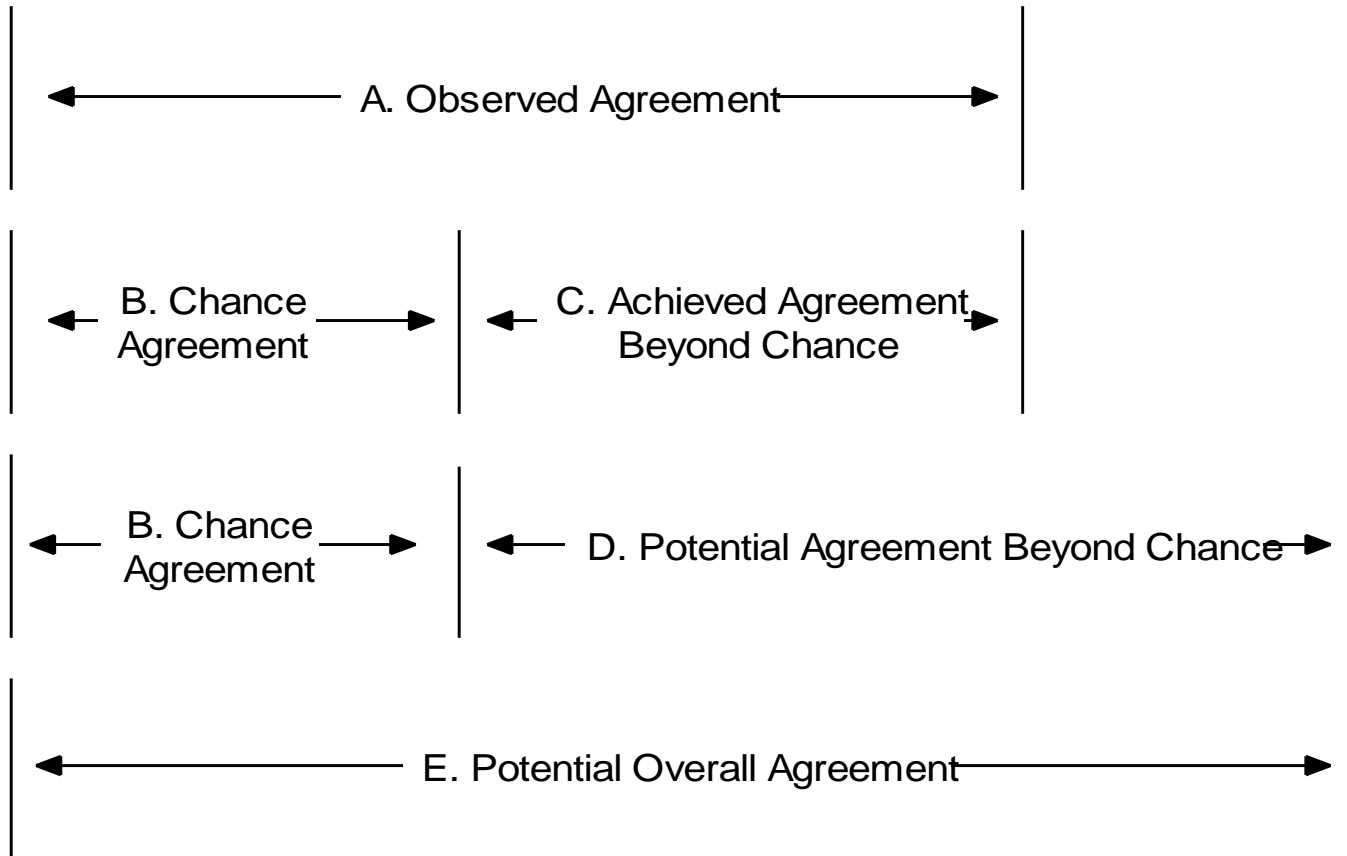


Initial Analysis

- Kappa Statistic
 - Used for interrater reliability
 - Takes into consideration agreement beyond chance
 - Only agreement beyond that expected by chance can be considered “true agreement” (Cohen, 1960)



Kappa Statistic





Results: Kappa

Attribute	Kappa	N	Agreement Quality
Localized guarding	0.31	217	Fair
Rebound tenderness	0.45	215	Moderate
Previous visit	0.48	220	Moderate
Type of pain	0.48	220	Moderate
Site of pain	0.51	222	Moderate
Shifting of pain	0.52	216	Moderate
Site of tenderness	0.57	187	Moderate
Duration of Pain	0.83	221	Very Good
Vomiting	0.89	220	Very Good
WBC	0.91	39	Very Good
Temperature	0.95	216	Very Good

< 0.20	Poor
< 0.40	Fair
< 0.60	Moderate
< 0.80	Good
to 1	Very good



Further Analysis

- Interested in which attributes were significant in predicting triage decision for both residents and staff physicians
- Collapsed triage decision making outcomes of ‘discharge’ and ‘follow-up’ into single category
 - Allowed us to isolate the ‘consult’ category which is reflective of suspecting acute appendicitis
- Logistic Regression with two category triage decision and dependent variable and all other collected attributes as independent variables.



Results: Logistic Regression

Logistic regression analysis for residents (n = 294 patients)

Variable	Wald Statistic	p-value	Critical DEFF
Age	0.251	0.617	0.065
Gender	3.159	0.076	0.823
Pain Duration	0.325	0.850	0.085
Pain Site	0.153	0.926	0.040
Pain Type	1.833	0.176	0.477
Vomiting	0.005	0.944	0.001
Previous Visit	0.053	0.818	0.014
Temperature	1.327	0.515	0.346
Tenderness Site	9.971	0.007**	2.597
Localized Guarding	13.469	0.000***	3.508
Rebound Tenderness	8.164	0.004**	2.126
Pain Shifting	2.222	0.136	0.579

*p < 0.05

**p < 0.01

***p < 0.001



Results: Logistic Regression

Logistic regression analysis for staff physicians (n = 385 patients)

Variable	Wald Statistic	p-value	Critical DEFF
Age	1.013	0.314	0.264
Gender	1.260	0.262	0.328
Pain Duration	0.614	0.736	0.160
Pain Site	6.862	0.032*	1.787
Pain Type	6.879	0.009**	1.791
Vomiting	4.674	0.031*	1.217
Previous Visit	3.604	0.058	0.939
Temperature	2.312	0.315	0.602
Tenderness Site	3.194	0.203	0.832
Localized Guarding	7.662	0.006**	1.995
Rebound Tenderness	16.005	0.000***	4.168
Pain Shifting	2.985	0.084	0.777

*p < 0.05

**p < 0.01

***p < 0.001



Results: Design Effect

- Sample: Multiple physicians and multiple residents each of whom may have seen more than one patient.
- Cluster is formed by a physician and the multiple patients they saw, and by each resident and the multiple patients they saw.
- Calculate a “critical DEFF” adjust the statistic in question to the point where it was no longer significant at .05.
- critical DEFF was calculated for the Wald statistic using the following formula:

$$- \text{Critical DEFF} = \frac{\hat{W}}{c^2}$$



Results: Design Effect

	Staff: Consult	Res: Consult
Age	0.314	0.617
Gender	0.262	0.076
Duration of Pain	0.736	0.850
Site of Pain	0.032	0.926
Type of Pain	0.009	0.176
Vomiting	0.031	0.944
Previous Visit	0.058	0.818
Temperature	0.315	0.515
Tenderness Site	0.203	0.007
Localized Guarding	0.006	0.000
Rebound Tenderness	0.000	0.004
Shifting of Pain	0.084	0.136



Main Findings

- physicians and residents recorded different values for many of the predefined decision making attributes while examining the same patients
- Attributes that required a physical examination had relatively low Kappas
 - These attributes are ‘difficult to assess’
- Physicians and residents had some different significant attributes in their decision making models
- Significant attributes in both decision making models included those that are difficult to assess



Implications

- Additional support for assessing/recording difficult attributes
 - Imprecise/uncertain information
 - Uncertain option
 - Confidence factors
 - Explanations to assist in difficult assessments
 - Searchable case base (indexed through additional information)
- Different embedded decision making models based on different user group?



Comments / Questions...