Clinical Decision Support Historical Perspectives – 2

Component 11 / Unit 3

Scoring and heuristics

- Knowledge is represented as profiles of findings that occur in diseases
- There are measures of importance and frequency for each finding in each disease
- Found to be most "scalable" approach for comprehensive decision support systems
- Examples – INTERNIST-1/QMR, Dxplain, Iliad

History of systems using scoring and heuristics approach

- INTERNIST-1
  - Original approach, aimed to develop an expert diagnostician in internal medicine (Miller, 1982)
  - System originally designed to mimic the expertise of an expert diagnostician at the University of Pittsburgh, Dr. Jack Meyers
  - Evolved into Quick Medical Reference (QMR) where goal changed to using knowledge base explicitly (Miller, 1986)
- DxPlain used principles of INTERNIST-1/QMR but developed more disease coverage (Barnett, 1987)
  - Only system still available: http://www.lcs.mgh.harvard.edu/dxplain.asp
- Iliad attempted to add Bayesian statistics to the approach (Warner, 1989)

INTERNIST-1/QMR knowledge representation

- Disease profiles – findings known to reliably occur in the disease
- Findings – from history, exam, and laboratory
- Import – each finding has a measure of how important it is to explain (e.g., fever, chest pain)
- Properties – e.g., taboos, such as a male cannot get pregnant and a female cannot get prostate cancer

Findings in diseases

- For each finding that occurs in each disease, there are two measures
  - Evoking strength – the likelihood of a disease given a finding
    - Scored from 0 (finding non-specific) to 5 (pathognomic)
  - Frequency – the likelihood of a finding given a disease
    - Scored from 1 (occurs rarely) to 5 (occurs in all cases)

Disease profile for acute myocardial infarction

Dx: MYOCARDIAL INFARCTION ACUTE

Is associated with 134 Finding(s) arranged: (w/References)
1. In Textbook order: History, symptoms, signs, labs
2. By Frequency

	Ev	Fr
Past Medical History...		
Symptoms of Current Illness...		
Chest Pain Substernal At Rest	2	4
Chest Pain Substernal Lasting 20 Minute(s) Or Gtr	3	4
Chest Pain Substernal Unrelieved By Nitroglycerin	3	4
Onset Abrupt	0	4
Chest Pain Substernal Crushing	3	3
Chest Pain Substernal Radiating To Neck And/Or Upper Extremity(ies)	3	3
Chest Pain Substernal Severe	2	3
Abdomen Pain Acute	1	2
Abdomen Pain Epigastrium	1	2
Abdomen Pain Epigastrium Unrelieved By Antacid	1	2
Abdomen Pain Exacerbation With Exercise	1	2
Abdomen Pain Non Colicky	1	2
Abdomen Pain Present	0	2
Abdomen Pain Severe	1	2

INTERNIST-1/QMR scoring algorithm

- Initial positive and negative findings are entered by user
- A disease hypothesis is created for any disease that has one or more of the positive findings entered
- Each disease hypothesis gets a score
  - Positive component based on evoking strengths of all findings
  - Negative component of score based on frequency from findings expected to occur but which are designated as absent
- A diagnosis is made if the top-ranking diagnosis is >80 points (one pathognomic finding) above the next-highest one
  - When diagnosis made, all findings for a disease are removed from the list, and subsequent diagnoses are made
- Performed as well as experts in NEJM clinical cases (Miller, 1982)

Limitations of INTERNIST-1 and evolution to QMR

- Limitations
  - Long learning curve
  - Data entry time-consuming
  - Diagnostic dilemmas not a major proportion of clinician information needs
  - Knowledge base incomplete
- Evolution to QMR (Miller, 1986)
  - Less value in “case” mode
  - More value in knowledge exploration mode, e.g.,
    - Rule diseases in and out
    - Obtain differential diagnoses
    - Link to more detailed information
  - Became commercial product but did not succeed in marketplace

Toward the modern era

- By the late 1980s and early 1990s, it was apparent that
  - Diagnostic process was too complex for computer programs
  - Systems took long time to use and did not provide information that clinicians truly needed
  - “Greek Oracle” model was inappropriate to medical usefulness (Miller, 1990)
- More recently
  - Diagnostic decision support systems less effective than therapeutic systems (Garg, 2005)
  - General failure of AI and ESs to live up to the hype of the 1980s has been acknowledged (Mullins, 2005)
  - But diagnostic error still does continue, and harms patients (Garber, 2007)

Where are we headed now?

- Decision support evolved in the 1990s with recognition of their value within EHR
  - Rules and algorithms most useful in this context
  - Evolution from broad-based diagnostic decision support to narrower therapeutic decision support (covered in following segments)
- AMIA "roadmap" for future provides three "key pillars" (Osheroff, 2006; Osheroff, 2007)
  - Best knowledge available when needed
  - High adoption and effective use
  - Continuous improvement of knowledge and methods

But the quest for diagnostic decision support continues

- Isabel (www.isabelhealthcare.com) – "Second generation" approach uses
  - Natural language processing to map entered text into findings
  - List of differential diagnosis with 30 most likely diagnoses grouped by body system, not probability
- Performance studies
  - Initial development and validation for pediatrics (Ramnarayan, 2006) – reminded of one clinically important case 1 of 8 times
  - Subsequently extended and evaluated in emergency department (Ramnarayan, 2007) – displayed correct diagnosis 95% of time and 90% of time showed "must-not-miss" diagnoses
  - Now expanded to adult internal medicine (Graber, 2008) – pasting in text from NEJM case reports had correct diagnosis suggested in 48 of 50 cases for key text and 37 of 50 cases for all text

Other continuing approaches – "Googling" for a diagnosis?

- Large quantity of text in Google may hold latent knowledge?
  - Found in a case study to make diagnosis of a rare condition (Greenwald, 2005)
  - When text of NEJM cases entered, 15 of 26 had correct diagnosis in top three suggested (Tang, 2006)