Assessment of Clinical Reasoning: A Script Concordance Test Designed for Pre-Clinical Medical Students

Mary T. Johnson, Ph.D.
Associate Dean, Academic Affairs
Professor, Microbiology & Immunology

Des Moines University
College of Osteopathic Medicine
Outline

• Clinical reasoning: A Conceptual Framework
  • Uncertainty
  • Diagnosis
  • Analytic or non analytic?
  • “Illness Scripts” and Clinical Diagnosis

• Script Concordance Test
  • Principles
  • Applications and Results
Uncertainty

- Healthcare professionals must constantly make decisions in the face of uncertainty.
- Medical students are challenged by ambiguous situations & need practice in this area to become expert clinicians.
Uncertainty

Uncertainties are related to:

• Limited information
• Data subject to more than one interpretation
• New context for an illness, precepts of EBM don’t apply
Diagnosis: A Categorization Task

• Grouping patients’ illnesses according to known attributes

• Allows clinicians to take action

Analytic or non-analytic?

- Non analytic reasoning:
  Fast, unconscious, perceptual-based

- Analytic reasoning:
  hypothetico-deductive model
  Deliberate, reflective, slower
The Hypothetico-Deductive Model

• Think aloud studies
• Initial clues allow a hypothesis to be developed
• Data is collected to affirm or rule out the hypothesis; iterative process
• Both experts and novices do this, but experts take shortcuts

_Elstein, Shulman & Sprafka, Medical Problem Solving, 1990_
Development of Clinical Reasoning

student  
biomedical concepts

novice  
causal reasoning

intermediate  
hypothetic deductive

expert  
pattern recognition
Activation of Relevant Hypotheses

Expert practitioners:

• A non analytic process (usual)
  – Pattern recognition*
  – Memories of previous patients & experiences (spontaneous, unconscious, automatic)

• An analytic process (less usual)
  – Deliberate induction of possible explanations** (logical, conscious, carefully controlled)

* Norman, Medical Education, 2007
**Mamede & Schmidt, Medical Education, 2004
Knowledge Organization

- **Activated hypotheses:**
  Physicians access networks of relevant knowledge

- **Script theory:**
  - How information is processed
  - Knowledge organized for specific tasks
  - Networked knowledge
  - Links between clinical features and diagnostic entities

*Charlin et al. (2007) Scripts and clinical reasoning. Medical Education, 41: 1178*
The Illness Script: A Fit?

EXAMPLE - sinusitis attributes: pain, rhinorrhea, fever...

• If the value is **ACCEPTABLE** → raises the Hypothesis activation level

• Level sufficiently high = Dx Decision

• If the value is **UNACCEPTABLE** → the Hypothesis is rejected (in this case, bloody secretions, bone destruction)
Clinicians constantly evaluate new information for the impact on an activated hypothesis.

Multiple micro-judgments are involved in Clinical Diagnosis.

Each micro-judgment can be assessed to understand a clinician’s reasoning.
Towards Holistic Assessment

• Traditional MCQ Testing
• Direct observation evaluation (Simulation & OCSE; Clerkships)
• ePortfolio: learner reflection and self-assessment
• Script Concordance Test – SCT
  Expert-referenced evaluation of a learner’s clinical reasoning
What is the SCT?

• Method of assessment for clinical data interpretation
  – Examines steps used in clinical reasoning
  – Case-based assessment
  – “Real Life” scenarios allow uncertainty

• Standardized
  – Same stimulus for each learner
  – Objective automated scoring replaces subjective judgment of skilled observers
The Indiana Statewide System for Medical Education

• 9 Sites for preclinical education - variability in formats (PBL, TBL, integrated, mostly “traditional 2 plus 2” curriculum, MD)

• All 320 students: Indianapolis Health Sci Ctr clerkship training (+ emerging regional sites)

SUNY Stony Brook Medical School

• State University of New York (SUNY)
  – Down State, Syracuse, Stony Brook

• Stony Brook University Medical Center
  – 120 students per class all in same pathway, traditional curriculum, clerkships (MD)
Florida State University College of Medicine

• Main Campus in Tallahassee
  – One site for preclinical education (MD)
• Regional campus model for clinical years
  – 120 students each class (some Rural Track)
  – Community physician preceptor model

DMU College of Osteopathic Medicine

• Campus in Des Moines, 221 students/yr
  – One site for preclinical training;
  Yr 2 systems curriculum (DO)
• Community physician preceptor model
  – 60% clerkship rotations are in Iowa;
  90% of students are trained in the Midwest
  – Rural, Global Health, & Academic Medicine
### A Brief Clinical Problem

- Physical sign
- Pre-existing condition
- Laboratory result
- Imaging study

### A Relevant Hypothesis is Posed

### Student Must Make a Decision

### SCT Approach

<table>
<thead>
<tr>
<th>Clinical scenario / stem</th>
<th>If you were thinking of...</th>
<th>And then you find...</th>
<th>This hypothesis becomes...</th>
</tr>
</thead>
<tbody>
<tr>
<td>The hypothesis</td>
<td>New clinical data</td>
<td>-2 -1 0 +1 +2</td>
<td></td>
</tr>
</tbody>
</table>

- 2 = much less probable
- 1 = less probable
0 = neither less or more probable
+1 = more probable
+2 = much more probable
An elderly African-American man presents to the primary care physician with a persistent cough and he complains of recently developing hoarseness in his speech. He has had some mild but persistent back pain for about 6 months. His temperature was normal and he is currently on medication for a sinus infection and chronic high blood pressure.

<table>
<thead>
<tr>
<th>If you were thinking of...</th>
<th>And then you find...</th>
<th>This hypothesis becomes...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung cancer</td>
<td>Patient quit smoking 5 years ago</td>
<td>-2 -1 0 +1 +2</td>
</tr>
</tbody>
</table>

- 2 = much less probable
- 1 = less probable
0 = neither less or more probable
+1 = more probable
+2 = much more probable
SCT and scoring

• Do the clinical decisions chosen by the learner have *concord* with those of the “Reference Panel”
  – A group of experienced family medicine & generalist practitioners (hospitalists)
  – All answers are recorded
  – Points depend on the number of Reference Panel answers

*Example: 13 panel members*

<table>
<thead>
<tr>
<th></th>
<th>0 : 7 members</th>
<th>1 : 6 members</th>
<th>All other answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/7 = 1 point</td>
<td>6/7 = 0.86 point</td>
<td>0/13 = 0 point</td>
<td></td>
</tr>
</tbody>
</table>
SCT Data?

• Combined institution 3 tier vetting process for questions
• Combined institution expert panel for answer key
• Combined scores for comparison of test validity
• Combined institution 4th year medical student volunteer group
• IRB approval obtained, all 4 institutions
Random Response: 30 (n = 130)

Mean Values for SCT Scores
5-Year Results

Concordance Scores (Percent of Expert Mean)

Experts
4th YR Students
2nd YR, IUSM
2nd YR, SBUMC
2nd YR, FSU
2nd YR, DMU
SCT Validation

• Internal Consistency
  – 75 test items
  – Cronbach’s Alpha = 0.73

• Data Differentiation
  – 2nd year students compared to 4th’s
    • 1 way ANOVA p<0.0001
  – 2nd and 4th years to experts
    • 1 way ANOVA p<0.0001
  – IUSM to SBUMC to FSUCOM to DMUCOM
    • No significant difference p=0.20
Good Reliability With Shorter Testing Time

<table>
<thead>
<tr>
<th>Testing Time (hrs)</th>
<th>MCQ</th>
<th>SCT</th>
<th>Oral exam</th>
<th>Long case</th>
<th>OSCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.62</td>
<td>0.80</td>
<td>0.50</td>
<td>0.60</td>
<td>0.47</td>
</tr>
<tr>
<td>2</td>
<td>0.76</td>
<td>0.85</td>
<td>0.69</td>
<td>0.75</td>
<td>0.64</td>
</tr>
<tr>
<td>4</td>
<td>0.93</td>
<td>0.82</td>
<td>0.86</td>
<td>0.78</td>
<td></td>
</tr>
</tbody>
</table>

Coefficients = Cronbach alpha
Brian Jolly, Monash University 2007
Z-Transform for Student Reports

The SCT is scaled to have an expert mean of 1000 and a standard deviation of 100. The important comparative metric is shown below, where individual results can be seen relative to other groups.

\[
Z_i = \frac{X_i - M}{SD}
\]

where

- \(X_i\) = any particular value in the data set;
- \(M\) = the mean of the data set; and
- \(SD\) = the standard deviation of the data set.

SCT Z-Scores in Quartiles, Class of 2015
Concordance with Expert Mean

1 SD Above Class Mean
- 827

Class Mean: 628
- 740

1 SD Below
- 553
- 336

Class Mean: 132
Conclusions

• 2-School Results published in 2011
  – *Medical Teacher* 33(6):472-7
  – First account of SCT used for preclinical medical student assessment
  – Similar to studies published with “less novice” medical learners (clerks/residents)

• Evidence of validity and reliability
  – Face validity with students remains high

• Costs lower than OSCE/SPAL examinations
  – Complements other assessments; doesn’t replace

• Assesses learner response to clinical ambiguity
Ongoing Assessment of Clinical Reasoning

• IUSM: Assessment of Problem-Solving Competence
  – 2\textsuperscript{nd} year general SCT; 4\textsuperscript{th} year EM rotation

• Convergent Validity with other EM evaluations
  – Med students: $r(266) = 0.28$; $p < 0.01$ with USMLE Step 2, CK Emergency Medicine section
  – Residents: $r(35) = 0.69$; $p < 0.001$ with in-training exam