

Clinical reasoning of Indonesian medical students as measured by diagnostic thinking inventory

Gandes Retno Rahayu¹, Sean McAleer²

Abstract

Introduction: Clinical reasoning skill is one of the most important skills for a good physician. A number of instruments have been developed to measure this skill, including the Diagnostic Thinking Inventory (DTI). Several studies have been carried out to measure its reliability and validity; however, evidence of its construct validity is still limited. This study aims to explore the construct validity of the DTI and to measure the clinical reasoning skills of Indonesian medical students.

Method: The subjects were 1135 medical students and 60 general practitioners. They were asked to complete the Indonesia version of DTI.

Results: Overall reliability of the DTI was .74 and .50 for the flexibility in thinking scale and .70 for the evidence of knowledge structure scale. A one way unrelated ANOVA showed that there were significant differences in the DTI score ($F = 7.097$, $p = .000$), flexibility of thinking subscale ($F = 6.111$, $p = .000$), and the evidence of knowledge structure subscale ($F = 5.306$, $p = .000$) with the scores increasing over the period of medical training and practical experiences. The biggest proportion of subjects in all groups reached the highest level (scored between 171-246/level 5).

Conclusion: This study has shown the construct validity of DTI in a different linguistic context. It also has shown the level of clinical reasoning skills of Indonesian medical students varied with year of study.

Introduction

Clinical reasoning skill is considered as one of the most important skills needed to be a good physician. Earlier paradigms considered clinical reasoning skills as a set of generic skills that incorporate relevant data collection, hypothesis generation, data interpretation and hypothesis evaluation. However the work of Elstein et al. (1978) fostered the idea that the way medical knowledge is structured in the mind of students and physicians is critical to the quality of medical diagnosis. Interest in the way medical diagnostic knowledge is

structured and the role of knowledge structure in medical diagnosis has steadily increased, yielding some clinical reasoning models (Bordage & Lemieux, 1986, 1987; Grant & Marsden, 1987, 1988; Schmidt et al., 1990; Higgs & Jones, 1995). Some research results also have shed light on the clinical reasoning of the novice and the expert. Novice reasoning is characterized by a reliance on biomedical and applied science knowledge. This knowledge base, however, is poorly organized. Expert reasoning is usually characterized by good knowledge structure.

In line with this understanding of the nature of clinical reasoning, a number of methods to assess the process were developed e.g. the Diagnostic Thinking Inventory (Bordage et al., 1990), the Scripts Concordance Test (Charlin et al., 1998; 2000), and Clinical Reasoning Problems (CRP) (Groves, 2002).

The Diagnostic Thinking Inventory (DTI) is a self-report inventory that is mainly based

¹Lecturer in Medical Education,
Department of Medical Education
Faculty of Medicine
Gadjah Mada University, Yogyakarta, Indonesia

²Senior Lecturer,
Centre for Medical Education,
University of Dundee, Scotland, United Kingdom

on research into medical diagnostic and clinical reasoning. The DTI measures two cognitive constructs that emerged from the clinical reasoning research (Lemieux & Bordage, 1986; Bordage & Lemieux, 1987; Grant & Marsden, 1987, 1988), namely flexibility in thinking and evidence for structure in memory. Flexibility in thinking refers to the use of a variety of thinking styles that can be applied during the diagnostic process. Structure in memory refers to the availability of knowledge, stored in memory, during the diagnostic process. It is assumed that availability is a direct consequence of adequate knowledge organization.

Bordage et al. (1990), showed that this inventory has acceptable reliability (alpha coefficient = 0.83), and can discriminate between expert and novice diagnosticians. Groves et al. (2003), suggested that it has the advantage of being independent of knowledge, and is not only applicable to the assessment of clinical reasoning at all levels of expertise but is also able to provide direct insight into the nature of the subject's clinical reasoning process. This understanding could be used as the baseline for developing strategies which support the acquisition of clinical reasoning competency. Sobral (1995, 2002) related the development of clinical reasoning diagnosed by the DTI to learning characteristics, knowledge score and types of curriculum. Others have used it to measure the effect of a specific intervention (Round, 1999), to measure its reliability and validity when used with physiotherapists (Jones, 1997), to assess radiologists' clinical reasoning competency (Peterson, 1999), to measure the concurrent validity of a new clinical reasoning assessment method (Groves, 2002) and to study the clinical reasoning characteristics of diagnostic experts (Groves, 2003). However, the above mentioned studies do not provide evidence on the construct validity of this instrument for use with medical students or physicians. The present study aims:

1. to explore the stability of the construct validity of the DTI in different linguistic contexts.
2. to measure the clinical reasoning skills of medical students in the Faculty of Medicine, Gadjah Mada University, Yogyakarta, Indonesia.

Methods

Context of Study

Gajah Mada University Faculty of Medicine (GMU-FM) is one of the largest universities in Indonesia. Each year the GMU-FM accepts approximately 280 new medical students. From 1990 to 2002 GMU-FM used a hybrid Problem-Based Learning curriculum, but since September 2003 it has been using a full PBL curriculum. There are two programmes within GMU-FM, one is the regular programme which is delivered in the Indonesian language, the other is an international programme where the students come from overseas and the curriculum delivery is in English. The present study used the regular programme.

Subjects

All first year to final (6th) year medical students in GMU SM were included in this study (n=1135). Sixty General Practitioners of varying experience were also recruited.

Measures

The researcher translated the DTI to Bahasa Indonesia (National Indonesian language). Five students who had TOEFL test scores of more than 580 were asked to complete Indonesian and English versions of the DTI. Nunnally & Bernstein (1994) have suggested that parallel forms should be administered at least two weeks apart. One argument for a significant time span between sittings is to minimize memory effect and the tendency to respond to an item based on recalling the earlier response. In this study, the Indonesian version was administered two weeks after the English version.

In this study, each conflicting answer in the Indonesian and English versions was identified. All the questions that yielded different answers were reviewed. New translations for those questions were based on the feedback. The DTI was then distributed to all subjects.

Results

a. Response rate

The response rate for all subject groups is presented in table 1. The overall response rate was 80.75%. The response rate for 1st year students was the highest as it was easier to contact them compared to the other groups. The relatively low rate for 2nd years can be explained by the fact that they rarely attend large group lectures and accessing each one individually was quite difficult in a large medical faculty as Gadjah Mada School of Medicine.

Table 1: The response rate from each subject group

Group	Total number	Returned	Response rate
1 st year medical students	221	221	100 %
2 nd year medical students	202	120	59.4 %
3 rd year medical students	192	162	84.4 %
4 th year medical students	170	147	86.5 %
Clinical students (CS / 5 th and 6 th year medical students)	350	213	60.9 %
General Practitioners 0 - 5 years experience	30	28	93.3 %
General Practitioner > 5 years experience	30	28	93.3 %
Total	1195	919	80.75%

b. Mean scores of DTI and its two subscales

Table 2 shows means for the overall DTI score and its two subscales for each subject groups. A one way unrelated ANOVA showed that there were significant differences in the DTI score ($F = 7.097$,

$p = .000$), flexibility of thinking subscale ($F = 6.111$, $p = .000$), as well as on evidence of knowledge structure subscale ($F = 5.306$, $p = .000$) with the scores increasing over the period of medical training and practical experiences.

Table 2: Mean DTI score and mean subscale scores

Group	Mean (sd)		
	DTI score	Flexibility of thinking	Evidence of knowledge structure
Year 1	165.04 (15.58)	83.55 (8.23)	81.37 (10.20)
Year 2	161.92 (19.82)	82.12 (9.64)	79.61 (11.91)
Year 3	163.24 (16.88)	82.97 (9.10)	80.40 (10.31)
Year 4	168.62 (18.37)	84.51 (9.68)	83.99 (11.53)
Year 5	161.60 (14.30)	81.47 (8.28)	80.11 (8.42)
Year 6	167.37 (16.88)	83.73 (8.71)	83.15 (10.89)
GP 0-5 years	178.15 (17.54)	89.74 (9.86)	88.39 (10.08)
GP > 5 years	178.00 (14.92)	91.15 (10.87)	87.14 (8.68)
Total	165.55 (17.33)	83.61 (9.19)	81.84 (10.70)

c. *Total score according to descriptors of attainment*

In relation to the DTI score, Bordage identified descriptors for five different levels of attainment:

- Level 1 = Poor flexibility and little evidence of structure.
- Level 2 = Some evidence of developing structure and flexibility.
- Level 3 = Evidence of overall developing flexibility and structure.
- Level 4 = Good flexibility of thinking and evidence of structure.
- Level 5 = Excellent flexibility and evidence of structure

The percentage of subjects for each group according to their attainment level is shown in table 3. It is surprising that more

than one-third of year 1 and year 2 medical students scored between 171 and 246 (level 5). It is also of concern that around one-fifth of clinical students were in the lowest position (having poor flexibility of thinking and little evidence of knowledge structure). It shows clearly the "ceiling effect" where the biggest proportion of subjects in all groups reached the highest level (scored between 171-246/level 5). The 2 GP groups scored higher than the others.

d. *Reliability of the DTI*

Overall reliability of the DTI was .74 (α coefficient for internal consistency); for the flexibility in thinking and evidence of knowledge structure it was .50 and .70 respectively. Table 4 shows comparisons with several previous studies.

Table 3: The percentage of subjects for each group according to the DTI attainment level

SCALE	LEVEL	GROUP							GP 0-5 (%)	GP > 5 (%)
		YEAR 1	YEAR 2	YEAR 3	YEAR 4	YEAR 5	YEAR 6			
		(%)	(%)	(%)	(%)	(%)	(%)			
<150		16.1	31.9	18	15	21	19.6	3.7	0	
150-155	1	14.3	9.5	13	6.9	11.4	2	3.7	7.4	
156-160	2	11.5	9.5	12	7.6	11.4	12.7	7.4	7.4	
161-165	3	11.1	7.7	17	14	10.5	10.8	7.4	3.7	
166-170	4	10.1	7.8	11	9.7	20	12.7	22.2	22.2	
171-246	5	36.9	33.6	29	46	25.7	42.2	55.6	59.3	

Table 4: Comparison of reliability across studies

	Bordage et al ,1990 (270 subjects)	Jones, 1997 (48 subjects)	Groves, et al, 2002 (22;33;24;35 subjects)	Present study (919 subjects)
Reliability (α -coefficient)	0.83	0.846	0.66; 0.74; 0.84; 087	0.74

Discussion and Conclusion

In term of clinical reasoning skills the theory indicates that the more experienced the subjects, the better their clinical reasoning skills. The analysis of the DTI scores of the eight groups supports the construct validity of the DTI that has been established in previous studies (Bordage, 1990; Groves et al, 2002).

The overall reliability of the DTI was .74. This coefficient is comparable to the developers' original finding of .83 (Bordage et al., 1990); as well as to previous studies that ranged from .66 to .87 (Groves et al., 2002; Jones, 1997). The reliability of the evidence of knowledge structure subscale was .70. It is also comparable to previous studies (Bordage, 1990). However, the reliability of flexibility in thinking subscale was only moderate (.50). Further analysis showed that if item number 2 (asking about prioritization), number 3 (asking about early interpretation), number 11 (asking about clarification before further data acquisition), and number 16 (asking about data interpretation over acquisition) were deleted the alpha coefficient increases.

The first and second year Indonesian medical students' mean total DTI scores were higher than those reported by Bordage et al., (1990) and Groves et al., (2002). The third year students' mean score is relatively lower than of similar cohorts from previous studies (158.3, Bordage et al., 1990; 168.59 Sobral, 1995; 168.89 Sobral, 2000; 171.1 Groves et al., 2002). No previous studies were identified in the literature that involved fourth year medical students. The fifth year and sixth year medical students' mean score is also lower than in previous studies, while the GPs' score is higher.

The fact that approximately one-third of the first and second year medical students achieved level 5 was surprising. As explained by Bordage (1990), the DTI can be used in two modes, either in relation to a particular case or in relation to one's general mode of diagnostic thinking. In this experiment the DTI was used in relation to subjects' general modes of diagnostic thinking. The instructions used by the original developers (Bordage et al, 1990) asked the subjects to respond as spontaneously as possible by indicating

how they actually diagnose and not how they think they should diagnose (even for those with little clinical experience). Written instruction was provided very clearly. The researcher also emphasised the instructions orally. However, the first and second year students who have not encountered many patient problems may find difficulties indicating how they actually diagnose. Consequently they probably thought what they should do (try to find the right answer), not what they may actually do in practice. This may have contributed to their high scores. In a formative setting this problem could be handled by asking students to explain their choice.

The fact that approximately one-fifth of clinical students scored below 150 (poor flexibility and little evidence of structure) is a concern. The learning experiences they have undergone for more than four years should have provided sufficient provision to score better.

In conclusion, this study has shown the construct validity of DTI in a different linguistic context. It also has showed the level of diagnostic skills of Indonesian medical students.

Acknowledgement:

We thank to Professor George Bordage who have given permission to use the Diagnostic Thinking Inventory.

References

- Bordage, G., Grant, j. & Marsden, P. (1990) Quantitative assessment of diagnostic ability. *Medical Education*, 24, 413-425.
- Bordage, G. & Lemieux, M. (1986) Some cognitive characteristics of students with and without diagnostic reasoning difficulties. Proceeding of the twenty-fifth Annual Conference on Research in Medical Education of the American Association of Medical College, New Orleans, 185-190.
- Bordage, G. & Lemieux, M. (1987) Structuralism and medical problem solving. *International Semiotic Spectrum*, 7(1), 3-4.
- Charlin, B., et al. (1998) Script questionnaires: their use for assessment of diagnostic knowledge in radiology. *Medical Teacher*. 20(6), 567-571.
- Charlin, B., et al. (2000) The script concordance test: a tool to assess the reflective clinician. *Teaching and Learning in Medicine*, 12:4, 189-195.
- Elstein, A.S., Shulman, L.S., & Sprafka, S.A.(1978) *Medical problem solving: an analysis of clinical reasoning*. Cambridge, MA/London UK, Harvard Univeristy Press.

- Grant, J., & Marsden, P. (1987) The structure of memorised knowledge in students and clinicians; an explanation for diagnostic expertise. *Medical Education*, 21, 98-108.
- Grant, J., & Marsden, P. (1988) Primary knowledge, medical education and consultant expertise. *Medical Education*, 22, 171-172.
- Groves, M., et al. (2002) Assessing clinical reasoning: a method to monitor its development in a PBL curriculum. *Medical Teacher*, 24 (5), 507-515.
- Groves, M., et al. (2003) The clinical reasoning characteristics of diagnostic experts. *Medical Teacher*, 25 (3), 308-313.
- Higgs, J., & Jones, M. (1995) Clinical Reasoning. In: Higgs, J., Jones, M., Eds. *Clinical Reasoning in the Health Professions*. Oxford: Butterworth-Heinemann Ltd.
- Jones, U.F. (1997) The reliability and validity of the Bordage, Grant & Marsden diagnostic thinking inventory for use with physiotherapists. *Medical Teacher*, 19(2), 133-140.
- Peterson, C. (1999) Factors associated with success or failure in radiological interpretation: diagnostic thinking approaches. *Medical Education*, 33, 251-259.
- Round, A.P. (1999) Teaching clinical reasoning-a preliminary controlled study. *Medical Education*, 33, 480-483.
- Schmidt, H.G., Norman, G.R. & Boshuizen, H.P.A. (1990) A cognitive perspective on medical expertise: theory and implication. *Academic Medicine*, 65 (10), 611-621.
- Sobral, D.T., (1995) Diagnostic ability of medical students in relation to their learning characteristics and preclinical background. *Medical Teacher*, 29, 278-282.
- Sobral, D.T. (2000) Appraisal of medical students' diagnostic ability in relation to their learning achievement and self-confidence as a learner. *Medical Teacher*, 22 (1), 59-63.