Computer Based Neuroradiology Training Using a "Small Worlds" Approach to Differential Diagnosis.

Lisa Hinkley, George du Boulay, Catriona Good, Benedict du Boulay, Briony Teather, Derek Teather, Mike Sharples & Nathan Jeffery

CSRP 540

October 2001

ISSN 1350–3162

UNIVERSITY OF SUSSEX AT BRIGHTON

Cognitive Science Research Papers

Computer Based Neuroradiology Training

Using a "Small Worlds" Approach to Differential Diagnosis.

Lisa Hinkley (1), George du Boulay (2), Catriona Good (2), Benedict du Boulay (3)♦,

Briony Teather (4), Derek Teather (4), Mike Sharples (5) & Nathan Jeffery (4).

(1) Oxford Brookes University, Oxford

(2) Institute of Neurology, National Hospital, London.

♦ *(3)* Author for correspondence*:*

School of Cognitive and Computing Sciences,

University of Sussex, Brighton, BN1 9QH

bend@cogs.susx.ac.uk

(4) De Montfort University, Leicester

(5) University of Birmingham

INTRODUCTION

Radiological diagnosis presents a particularly difficult task as it requires the extraction and interpretation of information on individual patients from complex medical images and reasoning in relation to competing pathologies that often exhibit similar abnormal features in the images. Studies of expertise in diagnostic medical cognition examine differences between practitioners with different levels of experience in terms of their cognitive processes and skills (including hypothesis generation and evaluation, memory performance, diagnostic reasoning and the organisation of clinical knowledge).

According to Kushniruk, Patel & Marley (1) expert physicians organise diagnostic knowledge on the basis of small subsets of diseases and their distinguishing features. They speculate that the process of diagnosis involves the physician rapidly focusing on relatively small sets of logically related diseases (i.e. “small worlds”) and carrying out a limited number of comparisons among these diseases to complete the diagnostic process. It is hypothesised that diseases contained within a small world would typically share certain overlapping features, and this is the basis for their membership of that particular small world. However, the diseases contained within a small world differ in terms of the presence or absence of certain other features, allowing the expert to distinguish between the candidate diseases contained within a small world.

In order to investigate the small worlds hypothesis, Kushniruk et al. re-analysed the verbal protocols collected in the studies carried out by Joseph & Patel (2) and Patel, Evans & Kaufman (3). They were particularly concerned with examining the networks of relationships among the hypotheses and findings generated by experts (endocrinologists solving endocrine cases) and sub-experts (cardiologists solving endocrine cases). The networks produced by the experts were found to contain few elements (i.e. a limited number of hypotheses and findings) and these were tightly connected, displaying a high degree of coherence and relatedness. Furthermore, expert physicians quickly focused on those cues and critical findings (‘critical cues’) that most clearly distinguished among competing diagnoses in the hypothesis set under consideration. In contrast, the sets of hypotheses generated by non-experts (experts working in a different field) often contain large numbers of diagnostic hypotheses, belonging to different disease categories.

Hence, according to the small worlds hypothesis, diagnosis by expert physicians involves the production of a small set of related clinical hypotheses that can explain the signs and symptoms. The experts then look for critical cues contained in the case, which allow them to distinguish between these candidate hypotheses and to produce a “final diagnosis”.

The MEDIATE project is a collaborative venture the initial aim of which (the MR Tutor) was to provide a computer-based training system to help neuroradiology trainees develop the skills required to become experts in the field. In particular, the tutor is designed to:

1. Teach trainees a structured Magnetic Resonance (MR) Image Description Language (IDL). With reference to radiology in general, Rogers, Arkin, Baron, Ezquerra & Garcia (4) discuss the need for a "lexicon of both anatomical landmarks and diagnoses which is understood and accepted by experts and residents alike".
2. Ensure that trainees are exposed to a wide range of cases representing the major brain diseases and gain an indication of the variation in visual features both between and within a disease category. Cases of the same disease can vary tremendously in terms of their MR presentation: Mervis and Pani (5) demonstrated that classification learning is maximised when subjects are first introduced to instances that best (most typically) represent the class under study and are only later exposed to less typical instances.
3. Expose trainees to cases drawn from a variety of small worlds, where each small world consists of cases from diseases that produce similar images, and so help them to discriminate between potentially competing hypotheses.

This paper briefly introduces the Image Description Language used to describe the database of cases and also the statistical method underpinning the display of these cases in an interactive Overview Space. The main part of the paper describes how the small worlds to be displayed in the tutor have been determined.

MATERIALS and METHODS

The Image Description Language (IDL)

The image description language used in the MR Tutor (6,7) is a representation of an expert's diagnostic knowledge in terms of the perceptual cues needed to make an accurate diagnosis. The description language was developed by one of the authors of this paper (with help from a named colleague who made blinded descriptions of sample images for comparison with those of the author). Furthermore, neuroradiologists at two centres have commented on their experimental use of the description language. The terms appear to be understood, but further, more systematic, empirical work is needed to determine whether the tutor is successful in

assisting trainees to use the terminology correctly. The vocabulary is derived from an understanding of the development and histology of lesions but can be applied and understood by radiologists working from images. The particular terms are derived from knowledge of underlying anatomy and histology, but they all refer to visual cues. The language describes both the position and appearance of abnormalities visible in the images. We have accumulated a database of the brain MR images of some 1,200 patients. These have been described by one author, expert in terms of the IDL, allowing trainees to compare their image descriptions with that of an expert (8).

Image Feature Overview Space

The MR-tutor uses a visualisation method to illustrate case typicality and the variation of presentation within and between diseases. The position of cases is computed from their image description with Multiple Correspondence Analysis (MCA) applied to the raw descriptions. Effectively, a multi-dimensional representation (the image description) is reduced to a two-dimensional visualisation, with the data points maximally spread out over the surface. The more similar two cases are, the closer their points will appear in the space. From the image description database we are also able to compute the position of the 'typical case' for each disease, and display it as the disease centroid in the overview space. Cases in physical proximity to this centroid can be classed as highly typical of the disease whereas those further away are less typical. Typicality contours representing the 25th, 50th and 75th percentiles are also contained in the overview space and these allow estimates of case typicality to be made (7). Figure 1 illustrates a typical overview space, in this case constructed from 16 cases of glioma.

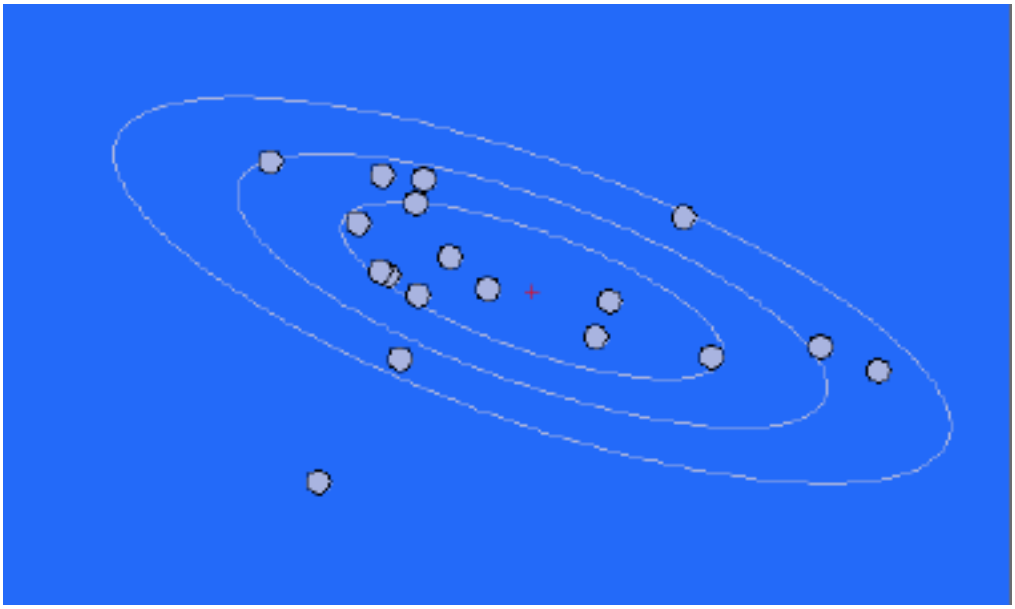

Figure 1. Overview Space for 16 cases of Glioma

Identifying Small Worlds

One of the aims of the MEDIATE project is to develop the MR Tutor to support training in decision making as well as in image analysis and description, using an approach consistent with the small worlds diagnostic hypothesis. This hypothesis suggests that medical experts possess highly structured knowledge that provides various kinds of shortcut to the small set of hypotheses that need to be considered in any situation. From this small world they are able to identify the critical cues that allow them to make a final diagnosis. In keeping with this approach, in decision support mode we envisage the tutor providing diagnostic support at the level of the small world. The tutor will allow the trainee to choose the disease 'small world' under consideration for any particular case (there may be two, three or even more diseases in a small world). The trainee can then view these candidate diseases in the same overview space, allowing them to explore the disease space of each, and just as importantly, the overlap between diseases.

One way to construct overview spaces containing two or more pathologies and display their overlap requires the selection of sets of confusable diseases, with the overview space for all cases of these confusable diseases giving the joint reference co-ordinate system (7). All cases can then be displayed in the joint co-ordinate system. Such an approach allows the analysis to identify combinations of factors that improve disease separation within a clinically relevant sub-problem hence supporting

diagnosis. Figure 2 presents a joint overview space with fitted typicality contours constructed for three diseases (glioma, infarct and meningioma). The overview space represents an example of the representation of a group of confusable diseases. The space reveals that in the case of the MR description data there is considerable overlap between gliomas and meningiomas, but good separation between infarcts and meningiomas (at least from the descriptions of the T2 weighted images used to construct Figure 2).

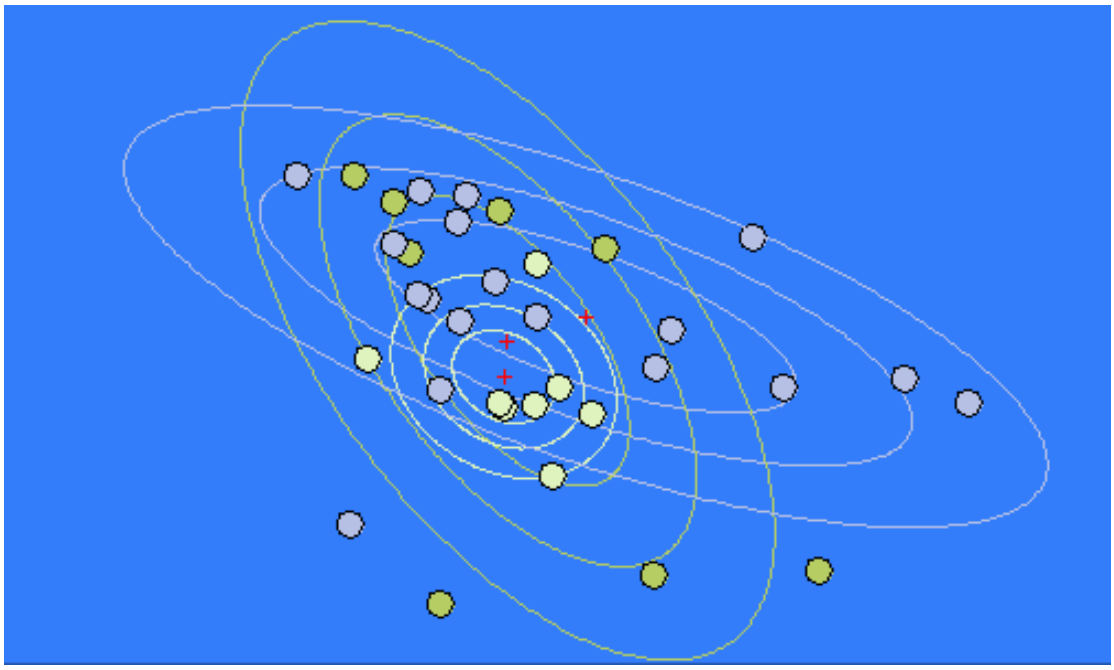

Figure 2. Overview space for case of Glioma, Infarct and Meningioma

Empirical research was needed to establish the differential diagnoses that are problematic, (i.e. diseases that are difficult to distinguish) and therefore represent clinically relevant issues. At present there is only a single group of confusable diseases included in the tutor (the glioma, meningioma & infarct group) and furthermore, this group was based upon the image descriptions of a single expert with empirical work needed to verify his views. The implementation of more confusable subgroups is a necessary condition for a functional decision support system (and training system). This paper describes empirical work undertaken to examine the disease differentiation that expert neuroradiologists find genuinely difficult to make. These groups of confusable diseases can then be implemented in the MR-tutor.

In order to examine what constitutes groups of confusable diseases for experts examining MRI scans of the head, a pro-forma was devised, and distributed to a group of expert neuroradiologists. The pro-forma asked experts to identify a number of scenarios in which examination of MR scans presents differential diagnostic difficulties. They were instructed that the diagnoses should be difficult to make, but at the same time worth making, and asked to list examples of MR based differential diagnostic problems which were not readily solvable and where a decision support tool might be of some assistance. For each differential diagnosis listed by the experts, they were asked to:

6. Rate the difficulty of the differential diagnosis on a scale from 1 - 5 (where 1 was tied to the semantic anchor ‘not normally difficult’, 3 - ‘sometimes difficult’, and 5 - ‘often very difficult’).

7. Rate the degree of clinical significance of the differential diagnosis (i.e. the degree to which this distinction affects future treatment of the patient) on a scale from 1-5 (where 1 had the semantic anchor ‘clinically insignificant', 3 - ‘sometimes clinically significant’, and 5 - ‘always clinically significant').

8. Rate the frequency with which this particular differential diagnosis is encountered in clinical practice. Again a scale of 1-5 was used where 1 had the semantic anchor ‘very infrequently’, 3 - ‘occasionally’ and 5 - ‘regularly’.

Subjects

The pro-formas were distributed to expert neuroradiologists at a number of specialist neuroradiology units in the UK and a further two in Barcelona, Spain. Respondents ranged in terms of their experience in MRI imaging in neuroradiology (12-204 months, Mean 98.3 months, SD 64.9 months). 24 completed questionnaires were returned to the research team (6 from Spanish recipients and a further 18 from British neuroradiology specialists). In total 117 problems were identified in the questionnaires (mean 4.88 per questionnaire).

RESULTS

As expected, the experts identified a large number of difficult differential diagnoses. The greatest practicable agreement about disease names and definitions is required in order to progress the studies. However, it was noted that the complexity and variety of disease classification soon obtruded into such an enquiry. Two different, not necessarily incompatible, classifications were investigated. The more detailed is the latest version of the WHO International Classification of Disease (ICD-10) (9), the simpler and more general is the traditional grouping into 10 categories used in neuroradiology training (Developmental, Inflammatory etc., see Table 1). In detail, analysis of the raw data involved examining clusters of responses that appeared to identify similar groups of diagnostic problems. Two expert neuroradiologists (authors of this paper) helped with the analysis which involved identifying similar differential diagnoses in the experts' responses. Firstly all the responses were coded using the WHO International Classification of Disease (ICD-10) and then the experts and a third author looked for overlap in the codes at some level of the ICD hierarchy. By combining overlapping responses, twenty-one disease pairs or groups were identified from the expert questionnaires. Many of the problems were identified by several respondents. However, other problems were idiosyncratic but were deemed sufficiently difficult to be worthy of inclusion. Of the 117 problems identified in the expert questionnaires, 105 were successfully included within this 21-item framework, leaving only 12 outliers. All 21 disease pairs or groups were then named using terms that best described the nature of the items contained within them. Table 2 shows the disease pairs or groups and also examines the extent of the support for each (i.e. how many experts identified a problem that was included under this heading).

Table 1 TRADITIONAL GENERAL PURPOSE CLASSIFICATION

Normal
Developmental
Infectious
Inflammatory
Neoplastic
Vascular
Metabolic
Degenerative
Iatrogenic
Traumatic

Table 2 SMALL WORLDS CONSIDERED WORTH EXPLORING BY EXPERTS

Specified by this number of experts	Diagnosis 1	Diagnosis 2	Enough material in database for preliminary study?
16	**Inflammatory** Inflammatory mass	**Neoplastic** Neoplastic mass	
11	**Vascular** Thrombo-embolic disease	**Inflammatory** Demyelination or encephalomyelitis	Yes
9	**Degenerative** Atrophy	**Degenerative/ Miscellaneous** Hydrocephalus	
8	**Neoplastic** Metastasis	**Neoplastic** Primary tumour	Yes
7	**Iatrogenic** Radiation necrosis	**Neoplastic** Recurrent tumour	
7	**Vascular** Haemorrhage – underlying cause?	**Vascular** Haemorrhage- primary cerebral?	
7	**Infective** HIV related infectious and parasitic diseases		
6	**Neoplastic** Identification of tumours in the pituitary region		
5	**Neoplastic** Glioma	**Vascular** Infarct	Yes
5	**Neoplastic/Infective/Inflammatory** Intra-axial mass	**Neoplastic/Infective/Inflammatory** Extra-axial mass	
4	**Vascular** Vasculitis	**Inflammatory** Demyelination (MS)	
4	**Metabolic/Degnerative/Toxic/Diffuse Inflammatory** Differential diagnosis		
4	**Neoplastic** Differential diagnosis of masses in the Cerebellopontine Angle		
3	**Vascular** Lacunar infarct	**Normality** Prominent Perivascular spaces	
3	**Metabolic** Differential diagnosis of Metabolic/White matter diseases – children		
2	**Vascular** Venous infarct	**Vascular** Arterial infarct	Yes
2	**Vascular** Infarct	**Infective** Herpes simplex encephalitis	
1	**Neoplastic** Low grade glioma	**Infective** Herpes simplex encephalitis	
1	**Vascular** Thrombo-embolic	**Vascular** Vasculitis	
1	**Neoplastic** Meningioma	**Neoplastic** Glioma	Yes
1	**Degenerative** Atypical Parkinsonian Syndromes Dementias		Yes

DISCUSSION

The aim of the reported study was to examine those groups of diseases (small worlds) neuroradiologists find difficult to distinguish when examining MR images of the head. These results, it was hoped, would allow us to focus the aims and objectives of a decision support/training system, and in so doing increase its utility. The study was successful in that it allowed the identification of a series of 21 pairs or groups of diseases that neuroradiologists claim may, under certain conditions, prove difficult to discriminate. However, it should be noted that the identification of confusable diseases and groups of related conditions, so much desired as goals, turns out in practice, to be subject to imprecision and vagaries of pathological nomenclature. The experts in our survey, when asked to list significant differential diagnostic problems, understandably often used common neurological parlance. Names such as 'white matter disease' and 'vascular disease' denote groups for which no agreed pathologies have been listed and no boundaries drawn, though all neuroradiologists can name some of the pathological entities that all would agree would lie within the named curtilage. A decision support system needs to be practical: a teaching system needs to minimise imprecision whilst accepting unsharpness of definitions that cannot be avoided. Sometimes an imprecise term like 'low grade glioma' may be preferred to the very detailed histological terms provided in the WHO ICD codes.

In order to exploit the categories in Table 2, we determined the ICD coding of every individual pathology in the archive of 1200 cases that forms the database of the decision support and tutorial system and we compared these ICD codes to the 21 problem diagnoses cited by the expert neuroradiologists. This voluminous data is available though not published in the present paper.

Bearing in mind evident subsets of more general problems identified by the experts, a further 16 small worlds might be found with sufficient archived data for preliminary study, (see Table 3). So the 21 small worlds identified by the experts as worthy of attention compare with 16 for which material seems to be available for experimental study in the present MEDIATE archive. A number of these small

worlds identified will be made available from the database of the MR Tutor, with the aim of providing decision support training in difficult cases. Hence, for instance, when a trainee (or indeed an expert) has a difficult case thought to be either a glioma or an infarct (sometimes a difficult differential diagnosis identified in five questionnaires), inputting its image description will show where it lies in relation to the glioma and infarct clusters derived from the archived material. This may suggest that the target case is of one or the other pathology or perhaps neither and further trials with other clusters may provoke reconsideration of the possibilities. Examining the overview space of any two or more diseases may throw light on the power of individual image descriptors under different pre-suppositions.

ICD 10 was chosen because of its extensive use, but in spite of its richness and flexibility of choices between broad and narrow levels of classification, new knowledge has already left it behind for some categories of neurological disease. For the degenerative dementias, for example, the four following categories would be better and will now be employed:

a. Dementia of the Alzheimer's type, McKhan, G et al., 1984, (10),
b. Frontotemporal lobar degeneration, Neary, D et al., 1998, (11),
c. Dementia with Lewy-bodies, McKeith, I et al., 1996, (12),
d. Vascular Dementia, Roman, GC et al., 1993, (13)

How should the compiler of a reference or teaching archive of cases cope with the inescapable fact that classifications inevitably expand and change? Parts of the archive that do not change may have a long and useful life. For others, if the future users can be made aware of the details of the disease classification adopted when the archive was new, it may well be possible to make some allowances for obsolescence. It would be good practice to record as much detail as possible when inserting information into the diagnosis section of the database in case that might subsequently permit reclassification of the case in the light of nosological advances.

ICD10 includes much expanded detail about brain tumour types that was little known or used while the archive was being built. There are now additional lines of

classification. In adding new cases it is the intention to record the greatest detail included in the case records.

The list of significant differential diagnoses (small worlds) suggested by the experts' answers to the questionnaire originates from their practical experience. It is based partly on image appearance and partly on clinical presentation or expectations prompted by referring neurologists or neurosurgeons. Many of the small worlds bridge across diverse aetiologies. Systematic knowledge and teaching in medicine on the other hand begins with a classification of disease under about 10 headings (Table 3). It could be useful also to place the identified small worlds, when practicable, within the structure of this time-honoured systematic classification of disease. Indeed, in quoting the diagnostic codes from ICD10 one is already building bridges between the practical work of image interpretation and the systematic scaffolding of nosology.

Table 3 ADDITIONAL SMALL WORLDS FOR WHICH ARCHIVE MATERIAL MAY BE SUFFICIENT FOR PRELIMINARY STUDIES

Metastasis / Glioma

Metastasis / Lymphoma

Infarct / Lymphoma

Meningioma / Lymphoma

Meningioma / Infarct

Metastasis / Infarct

Glioma / Lymphoma

Metastasis / Meningioma

Glioblastoma / Astrocytoma grades 3 & 4

Benign glioma / Malignant glioma

Oligodendroglioma / Benign glioma

Oligodendroglioma / Glioblastoma

Oligodendroglioma / Malignant glioma

Malignant oligodendroglioma / Benign oligodendroglioma

Pituitary adenoma / Craniopharyngioma

Cavernoma / Aterio-venous malformation

CONCLUSIONS

The MR Tutor is a novel computer-based training system for neuroradiologists. It is currently implemented in Java and makes use of about 40 cases from the archive of 1200 cases (from which at least 16 small worlds may be constructed) with the potential for major expansion. It aids observation, description, interpretation and analysis in a systemic and interactive way. The multiple correspondence analysis technique allows the display of the multi-dimensional spatial relationships between descriptions of lesions in individual cases compared with groups of known diagnoses. It thus tutors trainees in the art of differential diagnosis and highlights the variability both within and between disease categories.

ACKNOWLEDGEMENTS

We thank the British Neororadiological Society and Dr Rosemary Tate for assistance in contacting experts and the experts for their time. The work is supported by Engineering Physical Sciences Research Council grants GR/L53588 and GR/L94598.

REFERENCES

1. Kushniruk AW, Patel VL, Marley AJ. Small worlds and medical expertise: implications for medical cognition and knowledge engineering. Int. J. of Med. Informatics. 1998; 49: 255-271.
2. Joseph GM, Patel VL. Domain knowledge and hypothesis generation in diagnostic reasoning. Med. Decision Making. 1990; 10: 31-46.
3. Patel VL, Evans D, Kaufman D. A cognitive framework for doctor-patient interaction. In: Evans D, Patel V eds. Cognitive Science in Medicine: Biomedical Modelling. 1989; MIT Press, Cambridge, MA, pp. 253-307.
4. Rogers E, Arkin, RC, Baron, M, Ezquerra, N, Garcia, E. Visual Protocol

Collection for the Enhancement of the Radiological Diagnostic Process. In Proceedings of the First Conference on Visualisation in Biomedical Computing. Los Alamitos, 1990; CA: IEEE Computer Society Press, 208-215

5. Mervis CB, Pani JR . Acquisition of basic object categories. Cognitive Psychology, 1980 Oct, 12:4, 496-522
6. Du Boulay GH, Teather BA, Teather D, Higgott MA, Jeffery NP. Standard Terminology for MR image description. Proc XV Symposium Neuroradiologicum, Kumamoto 25 September 1994. Mutsumasa Takahashi, Yukunori Korogi, Ivan Moseley Eds Neuroradiology Supplement 1995; 37: 32-34
7. Teather D, Teather BA, Jeffery NP, du Boulay GH, du Boulay B, Sharples, M. Statistical Support for Uncertainty in Radiological Diagnosis. Method Inform Med 2000; 39:1-6.
8. Sharples M, Jeffery NP, du Boulay, B, Teather BA, Teather D, du Boulay GH. Structured computer-based training in the interpretation of neuroradiological images. Int. J. Med. Informatics, 2000; 60:263-289.
9. ICD – 10. International Statistical Classification of Diseases and related health problems. Tenth Review, World Health Organisation; Geneva. 1992; reprinted 1996.
10. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadian E. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. Neurology, 1984; 34: 939-944
11. Neary D, Snowden JS, Gustafson L, et al. Frontotemporal lobar degeneration: A concensus on clinical diagnostic criteria. Neurology 1998; 51:1546-1554
12. McKeith I, Galasko D, Kosaka K, et al. Concensus guidelines for the clinical and pathological diagnosis for dementia with Lewy bodies (DLB). Neurology 1996; 47:1113-1124,
13. Roman GC, Tatemichi TK, Erkinjuntti T, et al. Vascular dementia: Diagnostic criteria for research studies. Report of the NINDUS-AIREN International Workshop.Neurology 1993; 43:250-260