

Expert system support using a Bayesian belief network for the classification of endometrial hyperplasia

M. L. Morrison¹, W. G. McCluggage^{2*}, G. J. Price¹, J. Diamond¹, M. R. M. Sheeran², K. M. Mulholland², M. Y. Walsh², R. Montironi³, P. H. Bartels⁴, D. Thompson⁴ and P. W. Hamilton¹

¹ Quantitative Pathology Laboratory, Cancer Research Centre and Centre for Health Care Informatics, Queen's University, Belfast, UK

² Department of Pathology, Royal Group of Hospitals Trust, Belfast, UK

³ Institute of Pathological Anatomy and Histopathology, University of Ancona, Ancona, Italy

⁴ University of Arizona, Tucson, Arizona, USA

Received: 25 September 2001

Accepted: 25 February 2002

Abstract

Accurate morphological classification of endometrial hyperplasia is crucial as treatments vary widely between the different categories of hyperplasia and are dependent, in part, on the histological diagnosis. However, previous studies have shown considerable inter-observer variation in the classification of endometrial hyperplasias. The aim of this study was to develop a decision support system (DSS) for the classification of endometrial hyperplasias. The system used a Bayesian belief network to distinguish proliferative endometrium, simple hyperplasia, complex hyperplasia, atypical hyperplasia and grade 1 endometrioid adenocarcinoma. These diagnostic outcomes were held in the decision node. Four morphological features were selected as diagnostic clues used routinely in the discrimination of endometrial hyperplasias. These represented the evidence nodes and were linked to the decision node by conditional probability matrices. The system was designed with a computer user interface (Cytolnform) where reference images for a given clue were displayed to assist the pathologist in entering evidence into the network. Reproducibility of diagnostic classification was tested on 50 cases chosen by a gynaecological pathologist. These comprised ten cases each of proliferative endometrium, simple hyperplasia, complex hyperplasia, atypical hyperplasia and grade 1 endometrioid adenocarcinoma. The DSS was tested by two consultant pathologists, two junior pathologists and two medical students. Intra- and inter-observer agreement was calculated following conventional histological examination of the slides on two occasions by the consultants and junior pathologists without the use of the DSS. All six participants then assessed the slides using the expert system on two occasions, enabling inter- and intra-observer agreement to be calculated. Using unaided conventional diagnosis, weighted kappa values for intra-observer agreement ranged from 0.645 to 0.901. Using the DSS, the results for the four pathologists ranged from 0.650 to 0.845. Both consultant pathologists had slightly worse weighted kappa values using the DSS, while both junior pathologists achieved slightly better values using the system. The grading of morphological features and the cumulative probability curve provided a quantitative record of the decision route for each case. This allowed a more precise comparison of individuals and identified why discordant diagnoses were made. Taking the original diagnoses of the consultant gynaecological pathologist as the 'gold standard', there was excellent or moderate to good inter-observer agreement between the 'gold standard' and the

results obtained by the four pathologists using the expert system, with weighted kappa values of 0.586–0.872. The two medical students using the expert system achieved weighted kappa values of 0.771 (excellent) and 0.560 (moderate to good) compared to the 'gold standard'. This study illustrates the potential of expert systems in the classification of endometrial hyperplasias.

Copyright © 2002 John Wiley & Sons, Ltd.

Keywords: endometrium; endometrial hyperplasia; decision support; expert system; Bayesian belief network; inter-observer variation; intra-observer variation