Human Pathology (2017) **61**, 1–8

ELSEVIER

Human PATHOLOGY

www.elsevier.com/locate/humpath

Progress in pathology

Evidence-based pathology in its second decade: toward probabilistic cognitive computing

Alberto M. Marchevsky MD[a,*], Ann E. Walts MD[a], Mark R. Wick MD[b]

[a]*Department of Pathology and Laboratory Medicine, Cedars-Sinai Medical Center, Los Angeles, CA 90048*
[b]*Department of Pathology, Virginia College of Medicine, Charlottesville, VA 22908-0214*

Received 26 June 2016; revised 8 August 2016; accepted 2 September 2016

Keywords:
Evidence-based;
Pathology;
Cognitive computing;
Predictive analytics;
IBM Watson

Summary Evidence-based pathology advocates using a combination of best available data ("evidence") from the literature and personal experience for the diagnosis, estimation of prognosis, and assessment of other variables that impact individual patient care. Evidence-based pathology relies on systematic reviews of the literature, evaluation of the quality of evidence as categorized by evidence levels and statistical tools such as meta-analyses, estimates of probabilities and odds, and others. However, it is well known that previously "statistically significant" information usually does not accurately forecast the future for individual patients. There is great interest in "cognitive computing" in which "data mining" is combined with "predictive analytics" designed to forecast future events and estimate the strength of those predictions. This study demonstrates the use of IBM Watson Analytics software to evaluate and predict the prognosis of 101 patients with typical and atypical pulmonary carcinoid tumors in which Ki-67 indices have been determined. The results obtained with this system are compared with those previously reported using "routine" statistical software and the help of a professional statistician. IBM Watson Analytics interactively provides statistical results that are comparable to those obtained with routine statistical tools but much more rapidly, with considerably less effort and with interactive graphics that are intuitively easy to apply. It also enables analysis of natural language variables and yields detailed survival predictions for patient subgroups selected by the user. Potential applications of this tool and basic concepts of cognitive computing are discussed.

1. Introduction

The term "evidence-based-medicine" (EBM) was first used at McMaster University by Sackett et al in the early 1990s to define a systematic approach to analyze the peer-reviewed literature and extract only valid conclusions as the basis of clinical decisions [1-6]. These investigators described various methodological details that can bias the interpretation of research data and increased awareness that not all data published in the peer-reviewed medical literature are necessarily accurate. They also advocated the need to evaluate the internal validity of study results using "evidence levels." EBM publications also emphasized the concept of "external validity," demonstrating that the results of a study performed on a particular patient cohort are not necessarily valid for patients in other cohorts or applicable to individual patient care. Sackett et al [3] formally defined the term EBM in 1996 as "the conscientious and judicious use of current best evidence from clinical care research in the management of individual patients." EBM advocates have promoted the use of systematic reviews of the literature to gather "best available evidence" and of various statistical tools such as simple Bayesian

* Corresponding author at: Department of Pathology and Laboratory Medicine, Cedars-Sinai Medical Center, 8700 Beverly Blvd, South Tower Rm 8709, Los Angeles, CA 90048.
E-mail address: alberto.marchevsky@cshs.org (A. M. Marchevsky).

http://dx.doi.org/10.1016/j.humpath.2016.09.002
0046-8177/

metrics, odds, positive and negative predictive values, and meta-analysis to analyze this evidence. Although the definition of EBM has emphasized the management of individual patients, the approach has paradoxically stimulated the use of standardized "evidence-based" practices for patient groups. These evidence-based guidelines have been designed to facilitate daily practice for physicians that are too busy to calculate various probabilities for their patients based on data from systematic literature searches. Multiple "expert" teams have attempted to facilitate this task by publishing guidelines that recommend how to diagnose and treat patients with a wide variety of conditions [7-9]. Although these guidelines are based on systematic literature reviews and are labeled evidence-based to emphasize their quality, there is often no "best evidence" to address various aspects of clinical practice. In the absence of such hard evidence, these guidelines often include recommendations formulated with considerable input by "expert opinion" [1,10]. Although the use of evidence-based guidelines has helped standardize practice and has probably improved clinical outcomes, it does not provide specific recommendations for individual patients. Indeed, currently available best evidence from the literature has been collected from patient cohorts that are likely to differ with respect to patient age, comorbidities, therapy, study endpoints, and/or other features.

Pathology and laboratory medicine has been a late comer to the evidence-based space [10-13]. Nevertheless, interest in publishing systematic literature reviews, developing guidelines that assess the quality of study results using criteria that are appropriate to pathology, and aggregating data from multiple studies using quantitative meta-analysis rather than ad hoc "expert-opinion"–based tables has increased since our 2004 article in *Human Pathology* [10-14]. Our previous article also introduced basic concepts of "medical" decision analysis and described potential applications of artificial neural networks, logistic regression, Bayesian belief networks, and other multivariate statistical methods to diagnose cases or estimate the prognosis of individual patients. However, these methods have not been widely adopted in the pathology literature probably because they are cumbersome to use and/or it is difficult to validate the external validity of results [13,15-18].

Recently, the health care industry has shown increased interest in the application of similar statistical methods combined with data mining techniques [19]. Several analytical tools and software such as SAS Analytics Pro (SAS Institute, Singapore), PEGA Analytics (PEGA, Cambridge, MA), IBM Watson Predictive Analytics (IBM, Armonk, NY), and others that can be used over the Internet have been developed [19-21]. IBM Watson Easy Analytics is a "smart data discovery" service available on the "cloud" that provides a suite of predictive analytics and data visualization software that allows for rapid analysis of research data. In this study, we explore the use of IBM Watson Easy Analytics software to forecast the prognosis of patients with pulmonary carcinoid tumors based on histology and proliferative indices.

2. Materials and methods

To assess the accuracy and potential applicability of IBM Watson Easy Analytics for a pathology study, we used this system to analyze a dataset previously reported in *Modern Pathology* in 2012 by 2 of us (A. E. W. and A. M. M.) [11]. The study was approved by our institutional review board. Sections from 101 pulmonary carcinoid tumors were classified according to the 2004 World Health Organization classification of lung tumors into typical carcinoids (n = 78) and atypical carcinoids (n = 23) [22]. Tumors were immunostained for Ki-67, and Ki-67 indices were measured using an image analysis system whereby 5 randomly selected tumor fields were assessed at ×20 magnification. Tumors were staged according to the seventh edition of American Joint Commission on Cancer [23]. Overall survival information was obtained from hospital records. Follow-up periods ranged from 2 to 142 months (median, 26 months). Data were analyzed with the help of a professional statistician using Fisher exact test, Wilcoxon 2-sample test, Kaplan-Meier survival statistics, Cox proportional hazards models, and receiver operating characteristic curves. The statistical analysis required considerable time, effort, and expense.

3. Results of previous study

The mean Ki-67 indices for typical carcinoids (3.7 ± 4.0) and atypical carcinoids (18.8 ± 17.1) were significantly different ($P < .001$). Survival for patients with typical and atypical carcinoids were also significantly different ($P < .001$). Receiver operator curve analysis suggested that a Ki-67 index cutoff value of 5% provided the best fit for survival. When considered together, World Health Organization (histological) classification was a stronger predictor of overall survival than Ki-67 index. A few patients with typical carcinoids and Ki-67 indices greater than 5 appeared to have worse survival after 5 years. Survival rates by 1 and 5 years, diagnosis, and Ki-67 indices were stratified in a table that showed considerable overlap between various data points. For example, 5-year survival for typical carcinoids and greater than 5% Ki-67 index ranged in 95% confidence interval from 65.1 to 99.1, whereas 5-year survival for atypical carcinoids and greater than 5% Ki-67 index ranged from 52.1 to 91.3.

4. Analysis with IBM Watson analytics

An IBM Watson Easy Analytics free account was accessed using an Internet browser (Chrome; Google, Mountain View, CA). A spreadsheet with the same data used in the previous study was uploaded into IBM Watson Easy Analytics. It included the following data elements: diagnosis (typical or atypical carcinoid), Ki-67 index, survival status (dead or alive), and length of available