Cell population data-driven acute promyelocytic leukaemia flagging through artificial neural network predictive modelling

Authors: Rana Zeeshan Haider,[A] Tahir Shamsi[A] and Ikram Uddin Ujjan[B]

Introduction

A targeted and timely offered treatment can be of benefit for patients with acute promyelocytic leukaemia (APML). The current study made use of potential morphological and immature fraction related parameters (cell population data) generated during complete blood cell (CBC) counting, through artificial neural network (ANN) predictive modelling for early flagging of APML cases.

Materials and methods

We collected classical CBC items along with cell population data (CPD) from the haematology analyser at diagnosis of 1,067 patients with haematological neoplasms. For morphological assessment, peripheral blood films were examined. Statistical and machine learning tools, including principal component analysis (PCA), helped in the evaluation of predictive capacity of routine and CPD items. Then ANN predictive modelling driven by the selected CBC items was developed to identify the hidden trend by increasing the predictive accuracy of these parameters in differentiation of APML cases.

Results and discussion

We found a characteristic triad, based on a lower (53.73) platelet count (PLT) with a decreased/normal (4.72) immature fraction of platelet (IPF) with the addition of significantly higher (65.5) deoxyribonucleic acid / ribonucleic acid content-related neutrophil (NE-SFL) parameter in patients with APML compared with groups of patients with other haematological neoplasms. On PCA, APML showed exceptionally significant variance for PLT, IPF and NE-SFL. Through training of ANN predictive modelling, our selected CBC items successfully classified the APML group from non-APML groups at a highly significant (0.894) AUC value with a low (2.3%) false prediction rate. Practical results of using our ANN model were found to be acceptable, with values of 95.7% and 97.7% for the training and

testing datasets respectively. We propose that the triad of PLT, IPF and NE-SFL could potentially be used for early flagging of APML cases in the haematology–oncology unit.

Conclusion

ANN modelling driven by CBC items is a novel approach that substantially strengthens the predictive potential of CBC items, allowing clinicians to be confident of the typical trend raised by these studied parameters. ∎

Conflicts of interest

None declared.

Authors: [A]National Institute of Blood Disease, Karachi, Pakistan;
[B]Liaqat University of Medical and Health Sciences, Jamshoro, Pakistan