An Artificial Intelligence Approach to Thrombophilia Risk

João Vilhena, Universidade de Évora, Évora, Portugal
Henrique Vicente, Universidade de Évora, Évora, Portugal
M. Rosário Martins, Universidade de Évora, Évora, Portugal
José Grañeda, Hospital do Espírito Santo de Évora, Évora, Portugal
Filomena Caldeira, Hospital do Espírito Santo de Évora, Évora, Portugal
Rodrigo Gusmão, Hospital do Espírito Santo de Évora, Évora, Portugal
João Neves, Drs. Nicolas & Asp, Dubai, UAE
José Neves, Universidade do Minho, Braga, Portugal

ABSTRACT

Thrombophilia stands for a genetic or an acquired tendency to hypercoagulable states, frequently as venous thrombosis. Venous thromboembolism, represented mainly by deep venous thrombosis and pulmonary embolism, is often a chronic illness, associated with high morbidity and mortality. Therefore, it is crucial to identify the cause of the disease, the most appropriate treatment, the length of treatment or prevent a thrombotic recurrence. This work will focus on the development of a diagnosis decision support system in terms of a formal agenda built on a Logic Programming approach to knowledge representation and reasoning, complemented with a computational framework based on Artificial Neural Networks. The proposed model has been quite accurate in the assessment of thrombophilia predisposition (accuracy close to 95%). Furthermore, the model classified properly the patients that really presented the pathology, as well as classifying the disease absence (sensitivity and specificity higher than 95%).

KEYWORDS

Artificial Neuronal Networks, Decision Support System, Degree of Confidence, Knowledge Representation and Reasoning, Logic Programming, Quality of Information, Thrombophilia, Venous Thromboembolism

INTRODUCTION

Thrombophilia is a genetic or acquired tendency to hypercoagulable states that increase the risk of arterial and often, venous thrombosis. Venous ThromboEmbolism (VTE) is represented by two main clinical events: Deep Venous Thrombosis (DVT) and Pulmonary Embolism (PE) which often constitute a unique clinical picture in which PE follows DVT (Previtaly, Bucciarelli, Passamonti, & Martinelli, 2011). VTE is a multifactorial disease, often asymptomatic and associated with high morbidity and mortality (East & Wakefield, 2010).

The two most common factors for inherited (genetic) thrombophilias are the V Leiden mutation and the prothrombin G20210A gene mutation. Other mutations or polymorphisms associated with increased risk of thrombosis are methylenetetrahydrofolate reductase 677C (Lim & Moll, 2015). People can have one abnormal gene (heterozygous state) or two abnormal genes (homozygous state).
Less common genetic thrombophilia include deficiencies of the blood clotting proteins called protein C, protein S, and antithrombin (Lim & Moll, 2015; Previtaly et al., 2011).

Secondary hypercoagulable states are a frequently acquired VTE in patients with underlying systemic diseases or clinical conditions, known to be associated with an increased risk of thrombosis (e.g., malignancy, pregnancy, and use of oral contraceptives, myeloproliferative disorders, hyperlipidemia, diabetes mellitus and abnormalities of blood vessels). The most common acquired thrombophilia is antiphospholipid antibody syndrome (Previtaly et al., 2011).

The incidence of VTE is estimated at 56-160/100.000 people/year (East & Wakefield, 2010) and is strongly age-dependent, rising nearly 1% per year in old age (Rosendaal, 2005). Studies show that about 70% of patients presenting the first episode of VTE are over 60 years old, and the rate of recurrence is higher when the first episode of VTE occurs before 60 years old (Pernod et al., 2009). Moreover, a non-O blood group is associated with an increased risk of proximal deep vein thrombosis of the lower limbs with or without pulmonary embolism. The addition of inherited thrombophilia increases the thrombotic risk conferred by non-O group alone by almost 3-fold (Spiezia et al., 2013). Environmental or acquired VTE risk factors also include previous immobility, surgery, obesity, smoking habits and pregnancy or postpartum status (Anderson & Spencer, 2003; Enga et al., 2012; Goldhaber, 2010). Furthermore, the long distance travel is often associated with an up to 4-fold increased risk of VTE compared to non-travelers (Bartholomew, Schaffer, & McCormick, 2011).

Follow-up of patients for prolonged periods after an initial DVT or PE has revealed that VTE often is a chronic illness requiring lifelong prevention strategies. Thus, it is crucial to identify the cause, the most appropriate treatment, how long the treatment should be or how to prevent a thrombotic recurrence (Sinescu, Hostiuc, & Bartos, 2011). In the case of patients with a high risk of clinical complications preventive actions should be taken. The screening for VTE is mandatory in three high-risk groups, namely women who are prescribed oral estrogen preparations, pregnancy morbidity and patients undergoing major orthopedic surgery (Wu et al., 2006). Thus, more studies are necessary to reach a correct identification of factors associated with these diseases in order to assess the individual risk of thrombosis, and promote more targeted prophylactic and therapeutic alternatives.

On the one hand, in this study the complex pathophysiologic features of these hypercoagulable states (i.e., genetic, environmental and acquired risk factors) are discussed, being the problem tackled with Artificial Intelligence (AI) based methodologies and techniques for problem-solving (Chandrasekaran, 1983). Indeed, this work will focus on the development of an AI grounded Decision Support System aiming at the early diagnosis of thrombophilia, and also to signalizing of patients with hypercoagulable states. On the other hand, this work reports the founding of a computational framework that is centred on a Logic Programming (LP) based approach to knowledge representation and reasoning (Neves, 1984; Neves, Machado, Analide, Abelha, & Brito, 2007), complemented with a computational framework based on Artificial Neural Networks (ANNs) (Cortez, Rocha, & Neves, 2004). One of its main features relies on the unvarying way the incomplete information is fingered, i.e., how it is treated, either quantitative or qualitative, as it will be shown below.

**BACKGROUND**

**Knowledge Representation and Reasoning**

Logic Programming (LP) has been used for knowledge representation and reasoning, representing a point of convergence in the disciplines of Logic, Mechanical Theorem Proving, and Computer Science. It may be given in terms of elements of Model Theory (Gelfon & Lifschitz, 1988; Kakas, Kowalski, & Toni, 1998; Pereira & Anh, 2009), or Proof Theory (Neves, 1984; Neves et al., 2007). In the present work, the Proof-Theoretical approach is followed as an extension to LP. Indeed, an Extended Logic Program is a finite set of clauses in the form:
Comparison of Step Length Estimators from Wearable Accelerometer Devices
www.igi-global.com/chapter/comparison-step-length-estimators-wearable/12948?camid=4v1a