e-Health – For Continuity of Care
C. Lovis et al. (Eds.)

doi:10.3233/978-1-61499-432-9-53

Explicit Evidence for Prognostic Bayesian Network Models

Barbaros YET [a,1] Zane PERKINS [b] Nigel TAI [c] and William MARSH [a]
[a] *School of EECS, Queen Mary University of London, UK*
[b] *Centre for Trauma Science, Queen Mary University of London, UK*
[c] *The Royal London Hospital, London, UK*

Abstract. Many prognostic models are not adopted in clinical practice regardless of their reported accuracy. Doubts about the basis of the model is considered to be a major reason for this as the evidence behind clinical models is often not clear to anyone other than their developers. We propose a framework for representing the evidence behind Bayesian networks (BN) developed for prognostic decision support. The aim of this evidence framework is to be able to present all the evidence alongside the BN itself. We illustrate this framework by a BN developed with clinical evidence to predict coagulation disorders in trauma care.

Keywords. Bayesian Networks, Clinical Evidence, Knowledge Engineering, Prognostic Models, Clinical Decision Support

Introduction

A large number of prognostic models are developed to support clinical decision making but many of them are not adopted into clinical practice [1]. The predominant reason for this are concerns regarding model accuracy. Accuracy alone, however, does not ensure use of a model: clinicians will be reluctant to use some prognostic models regardless of their accuracy [2] while some models with mediocre performance are widely used [3]. Clinicians tend to reject a prognostic model if they are not convinced that the model's performance, for their patients, will be similar to its published performance in validation studies [2,3]. Understanding the clinical evidence supporting the model is an important factor for clinicians to evaluate its prospective performance in their practice [2].

Bayesian networks (BN) offer a powerful framework for building evidence-based models as their graphical structure is well suited for representing causal relations. However, the current representation of BNs is not descriptive enough to show the details of clinical evidence behind a BN. In many clinical BNs, the names of variables are often short and ambiguous, the relations modelled by edges are not explained, and the corresponding evidence from the clinical literature and data is not presented. Consequently, domain knowledge and evidence behind these BNs are clear only to their developers. In this paper, we present an evidence framework that organises and presents the evidence behind clinical BN models. We illustrate the application of the evidence framework by a BN that has been developed to provide decision support in trauma care by predicting acute traumatic coagulopathy (ATC) [4]. In the remainder of this paper,

[1] Corresponding Author. E-mail: b.yet@qmul.ac.uk

Section 1 discusses the challenges of describing evidence behind BNs and presents our evidence framework, and Section 2 presents the conclusions.

1. Evidence Framework

The challenges of describing evidence behind a prognostic BN can be summarised in two points. The first challenge is to organise the evidence in a structured way. The evidence about a BN must be detailed enough to prevent ambiguities about its relevance, source and type. The evidence can be relevant to different parts of the BN: some can be relevant to a particular variable or relation whereas others can be relevant to an entire group of variables. Moreover, there may be different items of evidence. A part of the BN may have evidence from the data, and another part may be based on clinical publications. Conflicting evidence must also be taken into account in order to have a comprehensive description of the evidence. The evidence framework must contain all of the relevant evidence even if it belongs to something not modelled in the BN. The meaning of the modelled variables and relations must also be clear.

The second challenge is to present evidence in a user-friendly environment without losing any important information. The users do not necessarily need to see the technical details of the evidence structure when browsing the evidence. In order to overcome these challenges we propose an evidence framework composed of two elements: a structure for organising the evidence data (Section 1.1) and a web page generated from this structure for browsing the evidence (Section 1.2).

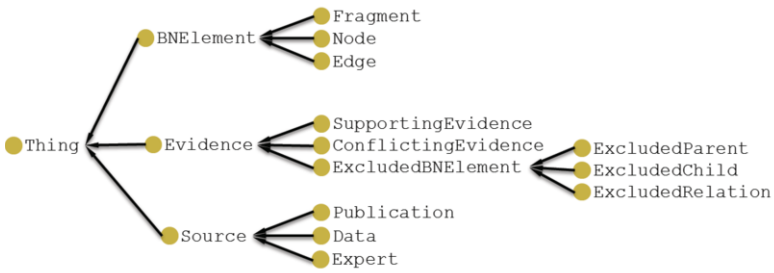

Figure 1. Class hierarchy for the evidence ontology

1.1. Evidence Ontology

We use the web ontology language (OWL) framework in the Protégé software (version 4.3.0) for organising the evidence data. Our primary reason for using ontologies is their flexibility in building and modifying a data structure. In the remainder of this paper, the classes, individuals, object and data properties of ontologies are represented by yellow circles, purple diamonds, blue and green rectangles respectively.

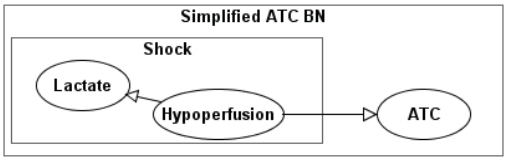

Figure 2. Simplified ATC BN

Our evidence structure is based on three main classes: BN element, evidence and source (Figure 1). We illustrate how data is entered to the evidence ontology by using a simplified version of the ATC BN that has 1 fragment, 3 variables and 2 edges (Figure 2). Evidence for the complete ATC BN can be found at [5].

Table 1. Defining the lactate variable in the evidence ontology

Lactate	
Type: BNElement Node	Object Properties: within Shock hasParent Hypoperfusion hasIncommingEdge HypoToLactate hasExcludedParent LactateExParEv1 Data Properties: description "The amount of lactate measured by…"

The BN element class has 3 subclasses: fragment, node and edge. A BN is composed of nodes representing variables and edges representing relations between those variables. It is necessary to clarify the meaning of these elements in order to describe knowledge behind the BN model. For example, a node named 'Heart Rate' may be sufficient to show that it represents a measurement of the patient's heart rate; but this name may not be descriptive enough if the time and location of this measurement is important for the use of the BN. Table 1 shows how the Lactate node in the ATC BN is defined in the evidence ontology. Associated fragments, edges, evidence and neighbor nodes are defined by object properties, and the node's description is recorded by a data property.

Table 2. Defining the 'Hypoperfusion → Lactate' edge in the evidence ontology

HypoToLactate	
Type: BNElement Edge	Object Properties: comesFrom Hypoperfusion pointsTo Lactate hasSupportingEvidence LactateSuppEv1 hasSupportingEvidence LactateSuppEv2

In order to include an edge in a BN model, the BN developers must have evidence that the relation represented by the edge exists. Therefore, all edges in a BN must have supporting evidence in an evidence-based BN. The meaning of a modelled relation is recorded in its associated evidence items. Conflicting evidence should also be taken into account. For example, one publication may claim that two variables are independent, whereas another publication may find correlation between those variables. Table 2 shows the data entered for the Hypoperfusion → Lactate edge in the evidence ontology. Its object properties define the variables connected by this edge, and relevant items of supporting evidence.

Table 3. Defining the shock fragment in the evidence ontology

Shock	
Type: BNElement Fragment MeasurementIdiom	Object Properties: contains Hypoperfusion contains Lactate Data Properties: description "This part of the model measure…"

A BN substructure represents a part of the BN that describes an important concept in its domain. Our ontology defines BN substructures in a general way as 'BN fragments'. The fragment class organises the information about description, evidence and member variables of a BN fragment. If the BN fragment has a specific type, such as idioms [6], this can be defined as a subclass of the fragment class. Table 3 shows the data about the Shock fragment recorded in the evidence ontology. This fragment contains 2 variables and its type is a measurement idiom.

Table 4. Supporting evidence item for the 'Hypoperfusion → Lactate' edge

LactateSuppEv1	
Type: Evidence SupportingEvidence	Object Properties: hasSource RixenEtAl2005 hasSource VanDrommeEtAl2010 Data Properties: statement "Lactate is produced in anaerobic…"

The evidence class organises the data about the type and statement of evidence related to the BN elements. There are three types of evidence in our evidence ontology: supporting evidence, conflicting evidence and evidence about relevant BN elements that are not included in the model. A BN element can have multiple items of evidence, and an item of evidence can have multiple sources. Table 4 shows one of the supporting evidence items for the Hypoperfusion → Lactate edge (Table 2). The evidence statement is recorded by a data property and publications related to this evidence are linked by the 'hasSource' object property. The source class describes the source of evidence; which can be a dataset, a domain expert or a scientific publication. Table 5 shows an example of a publication recorded in the source class. The evidence framework offers a convenient way to audit completeness of evidence by using SPARQL queries in OWL. For example, a user can review the edges without evidence by using the query below:

```
SELECT DISTINCT ?x
WHERE {?x a :Edge. MINUS {?x :hasSupportingEvidence ?evidence}}
```

1.2. Browsing Evidence

The evidence ontology is not a convenient tool to browse evidence especially if the user is not proficient with OWL. Therefore, our evidence framework generates a web page for browsing evidence after the data is entered to the evidence ontology using Protégé. The web page generated for the complete ATC BN can be found at [5].

Table 5. Details of a published evidence item stored in the source class

RixenEtAl2005	
Type:	Data Properties:
Source	hasPMID 16277731
Publication	hasRefDetails "D. Rixen and J. H. Siegel (2005), Critical Care, vol. 9, no. 5, p. 441…"

2. Conclusion

This paper proposed an evidence framework that complements clinical BNs by representing relevant clinical evidence. The ATC BN is used as a case-study to illustrate the evidence framework [4,5]. Our evidence framework is able to organise and present the evidence in more detail than it is possible with the existing BN representation. The evidence framework can store various types of evidence, including evidence supporting or conflicting with the BN, and offers a convenient way to query completeness of evidence. It can store evidence related to different levels of BN representation including BN fragments, nodes and edges. Although the evidence is stored within a complicated ontology structure, users can browse the evidence in a web page that is automatically generated from the ontology, without dealing with any of the underlying technical details. The evidence framework is not specific to the ATC BN; it could be applied to other BN models in different clinical domains. The main challenge of applying it to other domains is to identify the relevant evidence from the clinical literature. This step requires significant amount of effort from domain experts but it is essential for building evidence-based models. Our next step is to run a qualitative evaluation session with a group of clinicians. The probability values from the BN could also be included in the evidence framework to allow it to be used as a standardised format for representing BN models, and the data could be matched with medical terminologies such as SNOMED to have universally consistent descriptions.

References

[1] Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. *BMJ*. 2009;338(7708):1432–5.
[2] Wyatt JC, Altman DG. Commentary: Prognostic models: clinically useful or quickly forgotten? *BMJ*. 1995;311(7019):1539–41.
[3] Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ.* 2009;338(7709):1487–90.
[4] Yet B, Perkins ZB, Fenton NE, Tai NR, Marsh DWR. Not just data: A method for improving prediction with knowledge. *J Biomed Inform.* 2014;48:28–37.
[5] Acute Traumatic Coagulopathy Bayesian Network Evidence Browser [Internet]. [updated 2014 Feb 22; cited 2014 Apr 14]. Available from: http://www.traumamodels.com/atcbn/ATC_EBase
[6] Neil M, Fenton NE, Nielsen L. Building large-scale Bayesian networks. *Knowl Eng Rev.* 2000;15(3):257–84.