

Generating Medical Logic Modules for Clinical Trial Eligibility Criteria.

Craig G. Parker, MD^{1,2}, David W. Embley, PhD¹

¹Computer Science Department, Brigham Young University, Provo, UT

²Intermountain Health Care, Salt Lake City, UT

ABSTRACT

Clinical trials are an important part of modern medical research, however the effort required to find candidates for participation in such trial is significant. With the increasing prevalence of electronic medical records, automated or semi-automated solutions become feasible. We present an semi-automated approach for determining clinical trial eligibility based on information available in an electronic medical record.

INTRODUCTION

Clinical trials form the experimental basis for much of modern medicine. As with other experimental sciences, a sufficiently large target population must be sampled to gain the statistical power needed to have confidence in the results of the study. For clinical trials this means we must find a significant number patients who meet the eligibility criteria of the trial. A trade-off exists between the number patients who meet the eligibility criteria and the strictness of the criteria. The quality of the trial results can suffer from either too few enrollees or too broad eligibility criteria. Systems that help identify eligible patients^{1,2} can thus improve the quality of clinical trial results.

METHODS

The system we propose aims to assist healthcare organizations involved with clinical trials in identifying eligible candidates. The input to our system will be clinical eligibility criteria in the form of first order predicate logic. Each of the individual criteria will be represented as a simple predicate. These predicates will be related through a logical expression in conjunctive normal form. While automatic generation of these predicates and expressions from natural language trial documents is outside the scope of this project, an effort to this end is in progress at our institution.

Each of the predicates will be classified by structure to assist in the process of creating a mapping to the target database. For example, predicates made up of a single noun phrase (e.g. pregnant) are likely to be observations or diagnoses. Predicates that contain a noun phrase and a numeric value (e.g. gestational age > 26.0 wks.) are likely to represent things such as

time values or laboratory measurements. Predicates consisting of two noun phrases are likely to be name-value pairs.

Based on these classifications the system will attempt to map the terms in the predicate to concepts in the target database³. The mapping process will make use of the data dictionary of the target database, and the schema of the target database, as well as other medical vocabularies and ontologies. We will use the results of the mapping process and the logical expressions to generate medical logic modules in Arden Syntax.

We recognize that a complete mapping of all terms will not be possible for many reasons: not all necessary data will be in the target database, not all predicates will be precise enough to be mapped, and some predicates require logic that is higher than first-order. To handle this situation we will first determine eligibility to the extent possible based on the information that is mappable. If eligibility is still possible after this evaluation we will then present the user with a questionnaire to complete the determination of eligibility.

EVALUATION

To evaluate our system we will calculate the precision and recall of both our individual term mappings and entire predicate mappings. We will also evaluate the correctness of the Arden syntax modules that we generate.

REFERENCES

1. Ohno-Machado L, Wang SJ, Mar P, Boxwala AA. Decision Support for Clinical Trial Eligibility Determination in Breast Cancer. Proc AMIA Symp; 1999. p. 340-4.
2. Butte AJ, Weinstein DA, Kohane IS. Enrolling Patients Into Clinical Trials Faster Using Real-Time Recruiting. Proc AMIA Symp; 2000. p. 111-5.
3. Embley DW. Multifaceted Exploitation of Metadata for Attribute Match Discovery in Information Integration. In: Proceedings of the International Workshop on Information Integration on the Web (WIIW'01); 2001 Apr; Rio de Janeiro, Brazil; p. 110-7.