

COMPUTERVERARBEITBARE REPRÄSENTATION VON TEILNAHMEKRITERIEN KLINISCHER STUDIEN MITTELS ASPIRE UND ERGO: WAS IST MÖGLICH?

Schwarzthans F¹, Haider J¹, Hübner-Bloder G², Duftschmid G¹

Kurzfassung

Diese Arbeit analysiert die Ansätze ASPIRE und ERGO zur formalen Repräsentation von Teilnahme Kriterien klinischer Studien hinsichtlich ihres Potentials und Anwenderfreundlichkeit bei der Umsetzung von 689 als Freitext formulierten Kriterien aus 50 Diabetes-spezifischen Studien. Während ERGO einen breiteren Anwendungsbereich und grundsätzlich höheren Abdeckungsgrad aufweist, ist ASPIRE wesentlich intuitiver und weniger zeitaufwändig in der Anwendung.

Abstract

In this paper we analyze the potential and usability of the representation languages ASPIRE and ERGO for the formal representation of 689 textually represented eligibility criteria of 50 diabetes-specific clinical trials. Whereas ERGO enables broader applicability and in principle a higher rate of coverage, ASPIRE is significantly easier and less time-consuming to use.

Keywords – ERGO, ASPIRE, Clinical Trials, SnomedCT

1. Einleitung

Klinische Studien stellen heutzutage eine zentrale Erkenntnisquelle für den medizinischen Fortschritt dar [1]. Die Rekrutierung der Studienpatienten bildet einen häufig langwierigen und problembehafteten Teilprozess klinischer Studien, vor allem wenn es sich hinsichtlich des benötigten Patientenkollektivs um sehr spezifische Studien handelt, für die meist nur wenige Patienten verfügbar sind. Gemäß Sung et al. nehmen in der Regel nur etwa 7% jener Patienten, die die Teilnahme Kriterien einer Studie erfüllen, an dieser auch teil [2]. Im Bereich onkologischer Studien fällt dieser Anteil sogar auf nur 3% [3]. Generell schließen klinische Studien in 86% der Fälle die Rekrutierungsphase im Vergleich zum Zeitplan des Studienprotokolls verspätet ab [4].

Die Probleme bei der Rekrutierung resultieren unter anderem aus dem aufwändigen und fehleranfälligen manuellen Vergleich von Patientendaten mit den als Freitext formulierten Teilnahme Kriterien. Eine Optimierung dieses Vorgangs durch einen automatisierten Vergleich wäre insofern denkbar, als zumindest ein Teil der in Studienteilnahme Kriterien geprüften Patientendaten oft bereits zu Studienbeginn als Resultat vorangegangener Behandlungsprozesse vorliegt und heutzutage in der Regel über elektronische Gesundheitsakten (EHRs) der Patienten verfügbar sind.

¹Institut für Medizinisches Informationsmanagement und Bildverarbeitung, Medizinische Universität Wien

²Institut für Medizinische Informatik, UMIT – The Health & Life Sciences University

Aus diesem Grund wurde in den letzten Jahren begonnen, Ansätze zur computerverarbeitbaren Repräsentation von Studienteilnahmekriterien (z.B. ASPIRE, ERGO, OaSIS, OncoLink, caMatch CRFQ, EON, SAGE) zu entwickeln, welche durch Abfrage der in EHRs vorliegenden Daten unter anderem eine Vorselektion von potentiellen Studienteilnehmern ermöglichen könnten [5].

Das Ziel der vorliegenden Arbeit besteht darin, zwei bedeutende Vertreter der oben genannten Ansätze – ASPIRE [6] und ERGO [7] – dahingehend zu analysieren, welcher Abdeckungsgrad bei der computerbasierten Umsetzung von mittels Freitext formulierten Teilnahmekriterien existierender klinischer Studien jeweils erreichbar ist. Weiters sollen Erfahrungswerte hinsichtlich der Anwendbarkeit der beiden Ansätze gesammelt werden, d.h. wie einfach und intuitiv die Formulierung der Kriterien jeweils möglich ist. Die zwei Repräsentationssprachen ASPIRE und ERGO wurden unter anderem deshalb für die Analyse ausgewählt, weil sie konträre Modellierungsansätze verfolgen. Während ASPIRE auf einem vordefinierten Set von fixen Datenelementen aufsetzt, basiert ERGO auf einem generischen Datenmodell, mittels dessen beliebige Kriterien darstellbar sein sollen.

Die vorliegende Arbeit entstand als Vorstudie eines Projekts, das sich mit der automatisierten Identifikation von Patienten für klinische Studien im Kontext von Diabetes Mellitus (DM) beschäftigt, und fokussiert daher ebenfalls auf DM-spezifische Studien. Die im Rahmen der Arbeit gewonnenen Ergebnisse basieren auf der Umsetzung von insgesamt 689 in Freitext formulierten Teilnahmekriterien von 50 DM-spezifischen Studien mittels ASPIRE und ERGO.

1. 1. ASPIRE

ASPIRE (Akronym für „Agreement of Standardised Protocol Inclusion Requirements of Eligibility“) definiert zwei Typen von Studienteilnahmekriterien – einen Satz krankheitsübergreifender Kriterien (z.B. Alter und Geschlecht) und pro Erkrankung einen spezifischen Satz von Kriterien [6]. ASPIRE ist also nur für Studien jener Krankheitsdomänen anwendbar, für die bereits ein Satz spezifischer Kriterien spezifiziert wurde. Wie in [5] beschrieben, werden derzeit nur die beiden Domänen Diabetes Mellitus und Brustkrebs unterstützt.

Die vordefinierten ASPIRE-Kriterien basieren auf sogenannten „Common Data Elements“ (CDEs), die in Zusammenarbeit mit CDISC [8] entwickelt werden [5]. Bei CDEs handelt es sich um Dateneinheiten, die sowohl einen Namen als auch eine Definition besitzen und für welche zusätzlich eine Liste erlaubter Werte angegeben werden kann [9]. Für DM existieren z.B. CDEs, die sich auf den Diabetes Typ, die Dauer der Erkrankung, den HbA1c-Wert, etc. beziehen. Für den Diabetes Typ sind z.B. die Werte „Typ1“, „Typ2“ oder „beide Typen“ erlaubt.

1. 2. ERGO

ERGO (Akronym für „The Eligibility Rule Grammar and Ontology“) repräsentiert Teilnahmekriterien mittels eines objektorientierten Informationsmodells [7] und wurde als Protégé 3.4 Ontologie umgesetzt [10]. ERGO ist sehr generisch, im Gegensatz zu ASPIRE gibt es keine vordefinierten Kriterien. In Kriterien referenzierte medizinische Variablen, wie z.B. HbA1c-Wert, Diabetes Typ, oder Schwangerschaft werden mittels Codes standardisierter Terminologien, wobei speziell SNOMED-CT [11] empfohlen wird, als Objekte repräsentiert. Diese Objekte werden dann mittels logischer Operatoren zu den benötigten Teilnahmekriterien verknüpft. Dieses Konzept erlaubt es, allgemeine sowie spezifische Kriterien mit einem einzigen Modell umzusetzen.

2. Methoden

Im ersten Schritt wurde zum Stichtag 22.02.2012 im Studienverzeichnis ClinicalTrials.gov [1] eine Abfrage nach DM-spezifischen Studien (Suchbegriff „diabetes“) durchgeführt, wobei auf offene Studien eingeschränkt wurde. Von der daraus resultierenden Liste wurden die ersten 50 Studien gewählt, da das Ergebnis der Abfrage in ClinicalTrials.gov in absteigender Relevanz bezüglich der inhaltlichen Nähe zu den tatsächlichen Suchbegriffen gereiht wird.

Im nächsten Schritt wurden die insgesamt 689 Teilnahmekriterien der 50 Studien extrahiert und in einem Protokoll einzeln dokumentiert. Dabei wurden drei Arten von Kriterien markiert, die für die weitere Umsetzung in den zwei Repräsentationssprachen ASPIRE und ERGO nicht berücksichtigt wurden. Dies waren Kriterien, die aus Sicht der Autoren (a) inhaltlich unklar formuliert waren (z.B. „self-report of diagnosis of diabetes by a doctor“), (b) unselektiv waren und damit nichts zur Abgrenzung geeigneter Studienpatienten beitrugen (z.B. „all are welcome“), bzw. (c) nicht automatisiert mit Patientendaten verglichen werden können, da die betreffenden Daten mit hoher Wahrscheinlichkeit nicht in einem EHR vorhanden wären (z.B. „Willing to complete 24 months of study follow up“). Die Kategorien (a) und (b) stellen Ungenauigkeiten dar, wie sie in per Freitext repräsentierten Studienkriterien häufig beobachtet werden [5].

Studie	ASPIRE	ERGO	NB
Study29 http://ClinicalTrials.gov/show/NCT01333176			
Gender Eligible for Study: Both	Gender(Both) NU(GESUND)	OR(male, female) NU(SNOMEDCT)	
Accepts Healthy Volunteers: Yes			
Members of Seabird Island or surrounding communities	NU(ORT)	NU(SNOMEDCT)	
incl self-identified as First Nation1			Y
14 years old and older	MinAge(14) NU(SONSTIGE)	Age>=14 NOT(type 1 Diabetes, type 2 Diabetes)	
Known diabetes			
age less than fourteen years old			Z
pregnancy or less than six weeks post-partum medications known to interfere with glucose metabolism	Pregnancy(not Pregnant) NU(SONSTIGE)	NOT(Pregnancy) (<= 6 Weeks before now)	
acute illness,			Y
excl life-expectancy of less than six months			X

Abbildung 1: Beispielhafte Umsetzung der Kriterien für eine Studie. Nicht umsetzbare Kriterien sind mit „NU“ und der Kategorie des Kriteriums markiert (GESUND: Akzeptieren gesunder Freiwilliger; ORT: Geographische Verortung des Patienten; SNOMEDCT: kein passender SNOMED-CT Code gefunden). Spalte „NB“ markiert die nicht berücksichtigten Kriterien (X: Daten nicht in EHR; Y: unklar formuliert; Z: unselektiv).

Für die restlichen Kriterien wurde dann einzeln geprüft, ob und wenn ja mittels welchen Sprachkonstrukts diese jeweils in ASPIRE bzw. ERGO repräsentiert werden könnten. Bei ASPIRE wurde zu diesem Zweck nach einem jeweils passenden CDE mit geeigneter Parametrierung gesucht. Bei ERGO wurden zunächst für die im Kriterium referenzierten medizinischen Variablen entsprechende SNOMED-CT Codes identifiziert und daraus Teilkriterien gebildet. Zur Recherche der Codes wurde der NCI Term Browser¹ verwendet. Danach wurde das Kriterium, wie in *Abbildung 1* zu sehen ist, durch logische Verknüpfung der Teilkriterien zusammengesetzt.

Jene Kriterien, die nicht repräsentiert werden konnten, wurden markiert und inhaltlich in Kategorien eingeteilt. Bei ASPIRE wurden dabei die inhaltlichen Kategorien „Akzeptieren gesunder Freiwilliger“, „Aktive Teilnahme an anderen Studien“, „Blutglukose“, „Psychiatrische Vorerkrankungen“, „Geographische Verortung des Patienten“, sowie „Medikation“ gebildet.

¹URL: <http://nciterms.nci.nih.gov>

Kriterien, die weniger oft als fünf Mal auftraten, wurden der Kategorie „Sonstige“ zugeordnet. Hier ist allerdings anzumerken, dass diese Einteilung aufgrund der relativ geringen Studienzahl nur für die verwendeten 50 Studien anwendbar ist und keinen Anspruch auf Vollständigkeit stellt. Weiters ist zu beachten, dass ASPIRE zwar ein Feld für die Diagnosegruppe mit der Ausprägung "Healthy Subjects" vorsieht - da manche Studien aber als Einschlusskriterium sowohl DM verlangen als auch Gesunde akzeptieren, könnte dies zu Widersprüchen führen, weshalb das Kriterium „Akzeptieren gesunder Freiwilliger“ zur Auswertung hinzugefügt wurde. Bei ERGO ergaben sich drei Kategorien von nicht repräsentierbaren Kriterien – solche, zu denen kein passender SNOMED-CT Code gefunden werden konnte, solche, die bestimmte Arten von Zeiteinschränkungen beinhalteten, die derzeit in ERGO nicht umgesetzt sind [12] und Kriterien der Kategorie „Sonstige“, die aus jeweils unterschiedlichen Gründen (z.B. bedingte Präferenzen für bestimmte Teilnehmer, oder Kriterien die nichts mit dem Patienten zu tun haben) nicht modelliert werden konnten.

3. Resultate

Von den insgesamt 689 Kriterien wurden 124 (18,0%) Kriterien nicht für die Umsetzung mittels ASPIRE bzw. ERGO berücksichtigt: 20 (2,9%) Kriterien wurden als „inhaltlich unklar formuliert“ eingestuft, weitere 27 (3,9%) als „unselektiv“. Der Kategorie „mit hoher Wahrscheinlichkeit nicht in einem EHR vorhanden“ wurden 77 (11,2%) Kriterien zugeordnet.

Von den berücksichtigten 565 Kriterien waren 308 (54,5%) direkt mittels vorhandenen ASPIRE-Kriterien repräsentierbar. Von den nicht repräsentierbaren Kriterien bezogen sich 50 (41,0%) auf das „Akzeptieren gesunder Freiwilliger“, 24 (19,7%) auf „Psychiatrische Vorerkrankungen“, 20 (16,4%) auf den „Blutglukosewert“, 12 (9,8%) auf die „Medikation“, 8 (6,6%) auf die „Aktive Teilnahme an anderen Studien“, sowie 8 (6,6%) auf die „Geographische Verortung des Patienten“. 135 (23,9%) der nicht repräsentierbaren Kriterien wurden der Kategorie „Sonstige“ zugeordnet.

Von den berücksichtigten 565 Kriterien waren 359 (63,5%) direkt mittels ERGO repräsentierbar. Grundsätzlich hat ERGO das Potential weit mehr abzubilden, jedoch konnten mit den im Rahmen der vorliegenden Arbeit zu Verfügung stehenden Ressourcen teilweise keine offensichtlich geeigneten SNOMED-CT Codes gefunden werden. Geht man davon aus, dass bei höherem Ressourceneinsatz passende Codes für die betroffenen 184 Kriterien eruiert wären, könnte mit ERGO ein Abdeckungsgrad von 96,1% der untersuchten Kriterien erreicht werden. 3 (0,5%) der Kriterien sind deshalb mit ERGO nicht repräsentierbar, weil sie spezielle, in ERGO nicht umgesetzte Zeiteinschränkungen beinhalten. 19 (3,4%) Kriterien konnten aus jeweils unterschiedlichen Gründen nicht mit ERGO modelliert werden.

Hinsichtlich der Anwendbarkeit bestätigte sich die Feststellung von Weng et al., dass ASPIRE sehr einfach verständlich und vor allem sehr intuitiv anzuwenden ist [5]. Ein Großteil der allgemeinen Kriterien zum Patienten, sowie grundlegende, krankheitsspezifische Kriterien können schnell und problemlos durch Selektion und Parametrierung eines der vorgegebenen CDEs umgesetzt werden.

ERGO ist hinsichtlich seiner Anwendbarkeit bei weitem nicht so intuitiv wie ASPIRE. Einerseits erfordert die Modellierung einzelner Kriterien teilweise komplizierte Verschachtelungen und andererseits stellt die Recherche passender SNOMED-CT Codes, wie bereits in [13] festgestellt, eine komplexe und zeitaufwändige Aufgabe dar. Letzteres Problem wird dadurch weiter verkompliziert, dass inhaltlich gleiche Kriterien in verschiedenen Studien oft uneinheitlich benannt werden und tlw. Synonyme verwendet werden, die im NCI Term Browser nicht auffindbar sind.

4. Diskussion und Schlussfolgerung

Ross et al. analysierten 1000 zufällig ausgewählte Eignungskriterien von ClinicalTrials.gov hinsichtlich ihrer „Komplexität, der semantischen Bedeutung, der klinischen Inhalte und der Datenquellen“ [14]. Dabei wurde eine „signifikante klinische und semantische Variabilität“ zwischen den Kriterien festgestellt. Etwa 7% aller Kriterien wurden als „nicht in einem EHR vorhanden“ eingestuft, während 3,2% als unklar formuliert bzw. 3,1% als unselektiv klassifiziert wurden. Diese Anteile sind im Wesentlichen mit den in der vorliegenden Studie ermittelten Werten (11,2% nicht in EHR vorhanden, 2,9% unklar formuliert, 3,9% unselektiv) vergleichbar. Um die Quote der unverständlich formulierten Kriterien zu reduzieren, wäre die Entwicklung von Standards bzw. best practice patterns hinsichtlich der qualitätsgesicherten und einheitlichen Spezifikation von Studien-Teilnahmekriterien wünschenswert [14] [15].

Ein wesentlicher Unterschied zu den Resultaten von Ross et al. ist in den jeweils beobachteten Kriterien, welche Zeiteinschränkungen beinhalten, zu erkennen. Während dieser Anteil bei Ross et al. bei 40% liegt, lag er in der vorliegenden Studie bei nur 4,6%. Eine Erklärung für diese Diskrepanz könnte sein, dass wir auf DM-spezifische Studien fokussierten, während Ross et al. Studien beliebiger Domänen analysierten.

Eine Limitation der vorliegenden Studie besteht darin, dass sie ausschließlich auf DM-spezifische Studien fokussiert und somit nicht automatisch auf Studien anderer Domänen verallgemeinerbar ist. Auch die Wahl des Registers ClinicalTrials.gov als Datenquelle könnte einen Selection Bias darstellen in Richtung Studien mit mehr quantitativen als qualitativen Kriterien und Studien mit höherer Studiendesignqualität als nicht registrierte Studien [14]. Weiters wurde die medizinisch korrekte Umsetzung der Kriterien zwar von einer medizinischen Expertin qualitätsgesichert, die Codierung medizinischer Sachverhalte mittels vorgegebener Terminologien wie SNOMED-CT ist jedoch selbst für Experten eine ausgesprochen komplexe Aufgabenstellung, für die es häufig keine eindeutige Lösung gibt.

Hinsichtlich der Bewertung von ASPIRE ist festzuhalten, dass dieser Ansatz nur für Studien jener Krankheitsdomänen sinnvoll einsetzbar ist, für die bereits spezifische Kriterien definiert wurden. Mit dem allgemeinen Satz der krankheitsübergreifenden Kriterien ist nur eine geringe Abdeckung (in der vorliegenden Studie waren dies 32,4% der berücksichtigten Kriterien) erreichbar. Bewegt man sich jedoch innerhalb einer von ASPIRE unterstützten Krankheitsdomäne, so wie in unserem Fall von DM-spezifischen Studien, so sind zumindest in etwa die Hälfte der Kriterien einer Studie abbildbar. Die im Rahmen unserer Studie identifizierten Kategorien nicht repräsentierbarer Kriterien bieten Anhaltspunkte für künftige Erweiterungen der ASPIRE-Kategorien. So wäre z.B. bei Aufnahme der 3 häufigsten Kategorien „Akzeptieren gesunder Freiwilliger“, „Psychiatrische Vorerkrankungen“, und „Blutglukosewert“ in unserer Studie eine Abdeckungsquote von 71,2% erreichbar. Eine offensichtliche Stärke weist ASPIRE im Bereich Anwendbarkeit auf, im Gegensatz zu ERGO ist der Umsetzungsprozess hier sehr intuitiv und mit geringem Zeitaufwand möglich.

ERGO's Stärken liegen in seiner unbeschränkten Einsatzbarkeit für Studien beliebiger Krankheitsdomänen und seinem hohen erreichbaren Abdeckungsgrad existierender Kriterien. Beide Stärken sind jedoch abhängig von der Ausdrucksstärke der genutzten Terminologie und nur mit hohem Zeitaufwand und entsprechender Expertise hinsichtlich der Anwendung der Terminologie erzielbar. So konnte im Rahmen dieser Arbeit beispielsweise auch nur eine knapp höhere Abdeckungsquote als bei ASPIRE erzielt werden. Der Umstand, dass jeder Bezug zu einer medizinischen Variable mittels einer standardisierten Terminologie repräsentiert werden muss, wirkt sich im Vergleich zu ASPIRE negativ auf die Anwenderfreundlichkeit aus.

Zusammenfassend ist festzuhalten, dass ERGO insgesamt ein höheres Potential zuzuschreiben ist, sofern das erforderliche hohe terminologische Know-How vorhanden ist. Ist dies nicht gegeben und bewegt man sich in einer von ASPIRE abgedeckter Krankheitsdomäne, so überwiegt wohl die hohe Anwenderfreundlichkeit letzteren Ansatzes.

5. Referenzen

- [1] National Institute of Health, N.I.H. [Online]. [cited 2013 1 17]. Available from: <http://clinicaltrials.gov/>
- [2] Sung N, Crowley WJ, Genel M, Salber P, Sandy L, Sherwood L, et al. Central challenges facing the national clinical research enterprise. JAMA. 2003 Mar 12; 289(10): p. 1278-1287.
- [3] Group CTPR. Report of the National Cancer Institute Trials Program Review Group. [Online].; 2007 [cited 17.1.2013]. Available from: http://deainfo.nci.nih.gov/advisory/bsa/bsa_program/bsactprgmin.pdf
- [4] Lamberti M. Enrollment delays are getting worse. In: Lamberti MJ, ed. An Industry in Evolution. 4th ed. Boston, Mass: Thomson CenterWatch; 2003:79. .
- [5] Weng C, Tu SW, Sim I, Richesson R. Formal representation of eligibility criteria: a literature review. J Biomed Inform. 2010 Jun; 43(3): p. 451-467.
- [6] Niland J. ASPIRE: agreement on standardized protocol inclusion requirements for eligibility. [Online].; 2007 [cited 2013 1 17]. Available from: <http://hssp-cohort.wikispaces.com/file/detail/ASPIRE+CDISC+Intrachange+July+10+2007+Final.ppt>
- [7] Tu S, Peleg M, Carini S, Bobak M, Rubin D, Sim I. The Eligibility Rule Grammar and Ontology (ERGO). [Online]. [cited 2013 1 17]. Available from: <http://rctbank.ucsf.edu/home/ergo>
- [8] CDISC. [Online]. [cited 2013 1 17]. Available from: <http://cdisc.org/mission-and-principles>
- [9] Common data elements - harmonizing information, streamlining research. [Online]. [cited 2013 1 17]. Available from: http://www.ninds.nih.gov/research/clinical_research/toolkit/NINDS_CDE_brochure.pdf
- [10] protégé. [Online]. Available from: <http://protege.stanford.edu/> " <http://protege.stanford.edu/> .
- [11] NCI Term Browser. [Online]. [cited 2013 1 17]. Available from: http://ncit.nci.nih.gov/ncitbrowser/pages/multiple_search.jsf
- [12] Tu S, Peleg M, Carini S, Rubin D, Sim I. ERGO: A Template-Based Expression Language for Encoding Eligibility Criteria. [Online].; 2008 [cited 2013 1 17]. Available from: http://ontology-of-clinical-research.googlecode.com/files/ERGO_Technical_Documentation.pdf.
- [13] Meizoso Garcia M, Iglesias Allones J, Martinez Hernandez D, Taboada Iglesias M. Semantic similarity-based alignment between clinical archetypes and SNOMED CT: an application to observations. Int J Med Inform. 2012; 81(8): p. 566-578.
- [14] Ross J, Tu SW, Carini S, Sim I. Analysis of Eligibility Criteria Complexity in Clinical Trials. AMIA Summits Transl Sci Proc. 2010; 2010: p. 46-50.
- [15] Tu SW, Peleg M, Carini S, Bobak M, Ross J, Rubin D, et al. A Practical Method for Transforming Free-Text Eligibility Criteria into Computable Criteria. J Biomed Inform. 2011 April; 44(2): p. 239-250.

Corresponding Author

Florian Schwarzthans

Institut für Medizinisches Informationsmanagement und Bildverarbeitung

Medizinische Universität Wien

Spitalgasse 23, 1090 Wien, Österreich

n0909548@students.meduniwien.ac.at