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

doi:10.3233/978-1-61499-432-9-970

PAT: An Intelligent Authoring Tool for Facilitating Clinical Trial Design

Anastasios TAGARIS[a,1], Vassiliki ANDRONIKOU[a], Efstathios KARANASTASIS[a], Efthymios CHONDROGIANNIS[a], Charalambos TSIRMPAS[a], Theodora VARVARIGOU[a] and Dimitris KOUTSOURIS[a,1]
[a]*Institute of Communication and Computer Systems (ICCS), National Technical University of Athens, (NTUA), Greece*

Abstract. Great investments are made by both private and public funds and a wealth of research findings is published, the research and development pipeline phases quite low productivity and tremendous delays. In this paper, we present a novel authoring tool which has been designed and developed for facilitating study design. Its underlying models are based on a thorough analysis of existing clinical trial protocols (CTPs) and eligibility criteria (EC) published in clinicaltrials.gov by domain experts. Moreover, its integration with intelligent decision support services and mechanisms linking the study design process with healthcare patient data as well as its direct access to literature designate it as a powerful tool offering great support to researchers during clinical trial design.

Keywords. Clinical trial, patient recruitment, semantics, decision support systems

Introduction

Clinical research aims, among others, at revealing the therapeutic potential of substances, methodologies and devices in order for them to be developed into real world therapies. A recent study indicates that over $139 billion are spent annually on health research in the USA [12] with the estimated average cost per drug candidate being *more than € 1 billion*, while recently reported figures show that the overall required investment may reach even € 8 billion [1]. The prolonged therapy investigation and development process (being 11.3 years [2] on average) and the low productivity that clinical research faces (1 out of 5,000 compounds under research enter the market [3]), which is highly affected by bad study design (such as missing, poorly scientifically justified or inconsistent parameters and vague and/or rigid eligibility criteria) leaves millions of patients remaining untreated or following treatments of low/medium efficacy or side effects which affect their health. Weng et al. [4] stress out that key issues, which need to be addressed in clinical trials include incompleteness, ambiguity and inconsistency with most errors being identified within the protocol writing process. A major determinant for the success of a trial, is the set of EC specified for the study. EC are quite often poorly justified, vague, ambiguous or overly rigid. In fact, an analysis of inclusion and exclusion criteria across several trials [5] showed that more than 36% of the medication related criteria, 64% of the medical

[1] Corresponding Author.

comorbidity related ones and 78% of the age related ones are poorly justified. In fact, vague and/or ambiguous criteria may increase patient risks and weaken the ability to utilize the criteria to establish appropriate labeling. Moreover, overly rigid criteria may limit the generalizability potential of the study results, lead to a study population failing to represent the real world patient population, limit patient accrual and, hence, study costs, duration and feasibility.

Within this context, this paper focuses on presenting a series of novel mechanisms integrated and visualized through an authoring tool, which aims at improving three important processes in clinical research; specification of study parameters, determination of the eligible patient population and patient recruitment. It comprises part of the work performed within the PONTE project, which has been has been partially funded by the European Commission's 7th Framework Programme under contract number 247945 and worked towards the intelligent design of clinical trials.

1. Methods

The PONTE Authoring Tool (PAT) is a GUI specially designed to expose the functionality of the underlying PONTE platform's components to the end user. It facilitates the *CTP management* and the *design,* the specification of *study parameters* (e.g., study duration, number of patients), the *description of the eligible study population* and the determination *of healthcare entities* which will constitute the *pool for the patient recruitment* process. PAT is also used *as a collaborative environment* (essential for multicenter trials) as users can collaboratively work on CTP authoring.

The following methodological steps have been followed in order to define the technical and functional requirements of the PAT: (i) A **round of interviews with clinical research experts** took place in order to sketch out the CTP structure, the dependencies among its parameters and their semantic relations. The initial basis for these interviews was a series of CTP templates provided by the PONTE project clinical experts. These templates were analysed in terms of structure and information organisation. The analysis also involved identification of the informational elements which could benefit from decision support and specification of the dependencies (both direct and indirect ones) among the parameters in the CTP. Each expert (8 in total) focused on a specific set of parameters and indicated a series of questions that they often need to have answers for, prior to the specification of the parameter. (ii) **Thorough analysis of the information available at the clinicaltrials.gov [11] site**. The xml description of the studies was downloaded and the variety of values for study parameters (such as study type, blinding methods, control types and endpoints, among others) was analysed. A particular part of the analysis involved the EC, in terms of structure, categorisation, level of detail and individual parameters specified. The analysis was based on 1000 criteria selected across all studies at the portal. In order to ensure that different types of criteria are fed in the analysis, the respective studies were chosen so that a wide spectrum of disorders and active substances is covered. The analysis of those data (in terms of structure, terminology and types used) was taken into consideration to build the EC Ontology [6]. Hence, we followed an ontology based approach for capturing the semantics and allowing for semantic inference, in order to enable researchers to apply EC questions to Electronic Health Records connected to PONTE. (iii) The "access from anywhere", "collaboration among experts" requirements, as posed by the PONTE clinical experts, together with the "diversity in

the structure of data to handle", made apparent the third step regarding the technology used to **implement the PAT as a web tool following the Model View Controller (MVC) Design Pattern approach** [7]. MVC is an architecture that separates the representation of information from the user's interaction with it. The central idea behind MVC is code reusability and separation of concerns. The PAT Web Client uses JavaScript technology in order to provide a rich user-friendly browser based interface. This is done through the PAT Web Client Library which utilises the jQuery library[2] and Ajax[3] programming in order to create a more interactive and dynamic interface. Its server implementation (PAT Web Server Application) is developed using JAVA EE and runs on Apache Tomcat Application Server.

2. Results

Figure 1 depicts the PAT architecture and its 3 main modules: (i) The **PAT Web Client** is the part of the application running on the web browser. It contains a JavaScript library component (PAT Web Client Library) implementing the *presentation logic* of the application, providing rich user interaction capabilities and communicating with the server-side components. (ii) The **PAT Web Server Application** running on a web server host implements the core functionality of PAT and communicates with the rest of the PONTE platform. It encapsulates the *user interface rendering logic* and the orchestration of user actions, and business workflows. It contains the Request Handler which receives requests from the browser (PAT Web Client), creates the corresponding user interface elements from the UI Controls Library and renders the response back to the browser. The Request Handler is also responsible for communicating with the Authentication /Authorisation module in order to establish a security context for the execution of the request handling. The PAT Business Logic is implementing the core application logic and is responsible for interacting with the other PONTE components such as the Decision Support or the Ontology Based Search Engine components. The Business Logic also uses the PAT Local Storage to temporarily store the CTP parameters filled in by the users and other application state information. (iii) The **PAT Local Storage module** contains a local database which provides *persistence services* for the needs of the PAT Web Server Application. All CTP parameters are persisted on the local storage throughout the authoring lifecycle.

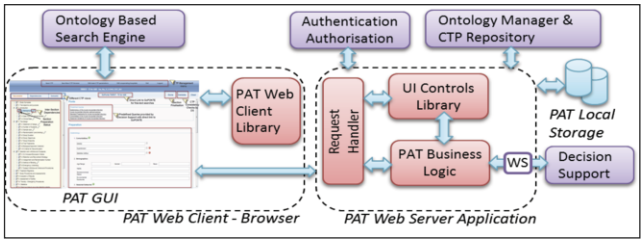

Figure 1. Internal PAT Architecture

As mentioned above, PAT facilitates the user in filling in each section of the CTP by providing useful information wherever possible, in order to ensure *CTP parameters*

[2] http://jquery.com/
[3] http://en.wikipedia.org/wiki/Ajax_(programming)

consistency, study efficacy and patient safety. In brief, PAT allows the PIs to manage their CTPs and the healthcare entities they cooperate with for patient recruitment purposes. It presents the user with three different CTP views: a hierarchical one (based on the CTP structure), a dependency one (based on dependencies between parameters) and a semantic one (built around three key concepts: the investigational active substance, the study disorder and the drug target). Hence, the PIs are offered a very flexible interface to navigate across the study parameters and fill in their data. Validation checks are provided whenever PIs save their work at a section or the CTP as a whole. For each section, PAT presents to the PI a set of automatically generated questions, the answer to which, may facilitate their (research) work. These questions are directly linked with the semantic search engine GoPONTE [9], to automatically retrieve the results of the selected ones.

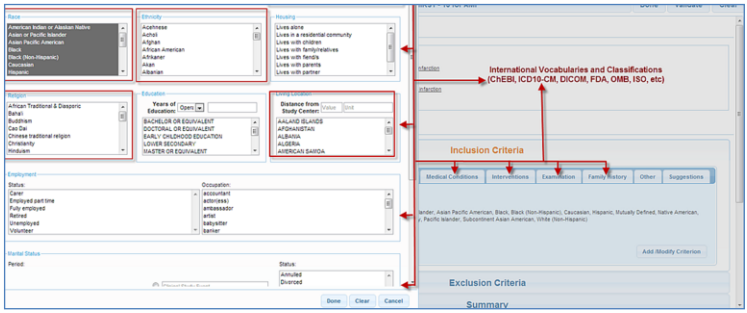

Figure 2. The EC categorization and links to vocabularies in PAT

PAT allows the specification of the study EC (Figure 2) based on the aforementioned EC-Ontology [6] and using international vocabularies and classifications for each EC category. It also communicates with the decision support mechanisms in order to present the user with suggestions regarding the EC. At any point during CTP design, the PI may request for a size estimation of the eligible patient population at each collaborating healthcare entities, based on the current EC specified. When the study design is completed, the PI can finalise the CTP and submit it to the Ethics committee. Upon approval of the CTP, patient recruitment may be initiated.

3. Discussion

Given the complexity of the study design, the variety of EC and the greatly heterogeneous data required for analysis and processing, the development of services and solutions to facilitate researchers in CTP design comprises a great challenge. PAT encapsulates intelligent, semantically-assisted mechanisms for study parameter specification and EC determination. The thorough analysis of CTP templates performed by experienced clinical experts, allowed for the development of a model, based on which PAT organizes CTPs around 3 different views: the hierarchy of its parameters, dependencies among them and their semantic linking with the CTP main concepts (investigational active substance, study disorder and drug target). These views were considered as “very helpful” during trial design by the experts, as they allowed for “easy and fast detection of inconsistencies” and “less time consuming parameter specification” although the dependencies should be further enriched and the 3 main concepts could be expanded to also include “clinical trials” and “patient” concepts. The

CTP parameters are directly linked with literature through templates of questions [10] that researchers would be interested in finding answers for, during design. According to the clinical experts, this functionality boost the scientific validity of the parameters and allows for faster referencing of the CTP. Nevertheless it was noted that this functionality could be enhanced through the introduction of more questions and the linking with data sources beyond literature, such as drug data sheets.

Analysis of the EC across studies published at clinicaltrials.gov, together with reviewing of the extracted patterns by clinical experts, led to a series of EC templates for describing the study population. PAT links with mechanisms applying these criteria onto patient data at healthcare entities (i.e., hospitals, clinics, etc.) in order to provide an estimation of the eligible population size. This functionality was considered of great value by the clinical experts (especially PIs) as it provides an indication about the *feasibility of the study* as well as the *potential market share*. Furthermore, it offers automatic retrieval of eligible patients who could potentially participate in the trial. According to the clinical experts, this is expected to significantly reduce the patient recruitment resources, both in terms of time and effort required, but also, will allow for fast and effective site recruitment. The EC model and resulting templates were evaluated as well structured and capable to cover frequently used criteria with a modest degree of complexity. However, further analysis is required in order to capture even more complicated EC but also to enrich the underlying vocabularies.

4. References

[1] Herper, M., 2012. The Truly Staggering Cost Of Inventing New Drugs, s.l.: Forbes

[2] J. D. Masi, R. Hansen, and H. Grabowski. The Price of Innovation: New Estimates of Drug Development Costs. Journal of Health Economics, 22:151{185, 2003.

[3] PhRMADrug Discovery and Development: Understanding the R&D process Pharmaceutical Research and Manufacturers of America. 2007. 2007.

[4] C Weng, DW. McDonald, J.H. Gennari, Participatory Design of A Collaborative Clinical Trial Protocol Writing System, IT in Health Care: Socio-technical Approaches 2nd International Conference, 13-14 September 2004, Portland, Oregon

[5] Van Spall, Harriette (21 Mar 2007). "Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review.".The Journal of American Medical Association 297 (11): 1233–40

[6] Chondrogiannis, E., Andronikou, V., Mourtzoukos, K., Tagaris, T., & Varvarigou, T. (2012). *A novel Query Rewriting Mechanism for Semantically interlinking Clinical Research with Electronic Health Records*. Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics. Craiova: ACM. doi:10.1145/2254129.2254189

[7] MSDN Model-View-Controller, "*Microsoft Developer Network: Model-View-Controller,*" 2014. MSDN Model-View-Controller, "Microsoft Developer Network: Model-View-Controller," 2014. [Online]. Available: msdn.microsoft.com/en-us/library/ff649643.aspx.

[8] Tagaris, A., Chondrogiannis, E., Andronikou, V., Tsatsaronis, G., Mourtzoukos, K., Roumier, Varvarigou, T. (2012). *Semantic Interoperability between Clinical Research and Healthcare: the PONTE approach.* Semantic Interoperability in Medical Informatics (SIMI2012). Workshop co-located with the ESWC2012.

[9] Tsatsaronis, G., Mourtzoukos, K., Andronikou, V., Tagaris, T., Varlamis, I., Schroeder, M., Matskanis, N. (2012). *PONTE: A Context-Aware Approach for Automated Clinical Trial Protocol Design.* 38th International Conference on Very Large Databases.

[10] Roumier, J., Liebman, M., Tsatsaronis, G., Matskanis, N., Tagaris, A., Mourtzoukos, K., .Schroeder, M. (2012). *Semantically-assisted Hypothesis Validation in Clinical Research.* Lisbon: eChallenges 2012.

[11] ClinicalTrials.gov A service of the U.S. National Institutes of Health (US) [cited 2014 May 4]. Available from: http://clinicaltrials.gov

[12] Research America, ARRA Saves Research from Decline, Available from: http://www.researchamerica.org/uploads/healthdollar09.pdf