Comparing semi-automatic systems for recruitment of patients to clinical trials

Marc Cuggia¹, MD PhD, Paolo Besana¹, PhD, David Glasspool², PhD

¹Unité Inserm U936, IFR 140, Faculté de Médecine, University of Rennes 1, France
²School of Informatics, University of Edinburgh, 10 Crichton Street EH8 9AB, Edinburgh, United Kingdom

Address Correspondence to:
Marc CUGGIA
Unité Inserm U936, IFR 140,
Faculté de Médecine,
2 Avenue du Professeur Léon Bernard
35043 Rennes cedex
Tel: +33 299284215
marc.cuggia@univ-rennes1.fr
Objectives: (i) To review contributions and limitations of decision support systems for automatic recruitment of patients to clinical trials (Clinical Trial Recruitment Support Systems, CTRSS). (ii) To characterise the important features of this domain, the main classes of approach that have been used, and their advantages and disadvantages. (iii) To assess the effectiveness and potential of such systems in improving trial recruitment rates.

Data sources: A systematic MESH keyword-based search of Pubmed, Embase, and Scholar Google for relevant CTRSS publications from January 1st 1998 to August 31st 2009 yielded 73 references, from which 33 relevant papers describing 28 distinct studies were chosen for review, based on their report of a novel decision support system for trial recruitment which reused already available patient data.

Method: The reviewed papers were classified using a modified version of an existing taxonomy for clinical decision support systems, using 10 axes relevant to the trial recruitment domain.

Results: It proved possible and useful to characterize CTRSS on a relatively small number of dimensions and a number of clear trends emerge from the study. Only nine papers reported a useful evaluation of the effectiveness of the system in terms of trial pre-inclusion or enrolment rate. While all the systems reviewed re-use structured and coded patient data none attempts the more difficult task of using unstructured patient notes to pre-screen for trial inclusion. Few studies address acceptance of systems by clinicians, or integration into clinical workflow, and there is little evidence of use of interoperability standards.

Conclusions: System design, scope, and assessment methodology vary significantly between papers, making it difficult to establish the
impact of different approaches on recruitment rate. It is clear, however, that the pre-screening phase of trial recruitment is the most effective part of the process to address with CTRSS, that clinical workflow integration and clinician acceptance are critical for this class of decision support, and that the current trends in this field are towards generalisation and scalability.

**Keywords:** Clinical Trial; Patient Selection; Clinical Decision Support Systems; Systematic Review.

**Introduction**

Clinical trials (CTs) are the gold standard for testing therapies or new diagnosis techniques that may improve clinical care. Because they rely on adequate sample sizes it is common for CTs to fail in their objectives because of the difficulty of meeting the necessary recruitment targets in an effective time and at reasonable cost. With the development of “omics” sciences in medicine and more focused therapies, adequate patient enrolment has become even more challenging as eligibility criteria become more complex.

Moreover, most CTs are run at multiple centres and involve multidisciplinary actors. New strategies are necessary to efficiently recruit patients from a population, at a large scale, in a cost effective way.

The recruitment process is therefore a particular point of weakness for clinical trials. The development of IT in medicine, and particularly in hospitals, offers a good opportunity to support and improve the recruitment process. Digital patient data, either acquired during the clinical care process or contained in Electronic Health Records (EHR), might be reused to automatically apply eligibility criteria. Indeed, much work
has been carried out on this approach using different methods and technologies. However while computer-based Clinical Decision Support Systems (CDSS) have been widely surveyed in the literature, none of these reviews has focused specifically on computer decision support for CT recruitment. This is the purpose of this paper.

1. Objectives

In this paper, a literature review of automatic Clinical Trial Recruitment Support Systems (CTRSS) is presented. This review is relevant for two reasons. First, there is no similar work in the literature. More general reviews of CDSS face a difficult task summarizing CDSS features or evaluating evidence of their efficiency due to the heterogeneity of the domain, and focusing on a sub domain of CDSS will allow better characterization of its features. Second, the recruitment process is a well-known bottleneck in clinical trial procedure and huge efforts are currently under way to optimize this process. We expect a focused review of enrollment decision support systems to be particularly useful in guiding this work.

The main aim of this review is to inform researchers involved in the clinical trial process of the potential benefits of CTRSS, and to review the contributions and limitations of previous works carried out in this field. In the next section we review the main reasons for low recruitment rates in trials. After presenting the methodology of our review, each reviewed system is then analyzed according to a set of criteria based on a taxonomy of decision support systems defined by Berlin [1]. Finally we discuss the issues identified and implications for CTRSS.
2. **Background**

In this section, we describe the mains steps of CT recruitment process and the main barriers to enrollment identified in the literature. We will use this analysis to classify existing works or systems that aim to optimize the recruitment process.

1.1. **The recruitment process of clinical trials**

The two most important parameters that must be decided early in the design of a CT are the population of interest, which determines the eligibility criteria of the trial, and the sample size required to give sufficient statistical power for analysis. These two factors generally interact strongly: A well-focused, sharply defined population of interest may well be too small to easily yield enough willing participants to achieve the desired sample size. Reduced sample size reduces the power of the study, but relaxing eligibility criteria to allow a larger population of interest (and hence a larger pool from which to recruit) introduces a confounding bias where factors that are not the prime focus of the study cannot be excluded. Defining the population of interest simultaneously determines both the eligibility criteria and the sample size [2].

1.2. **Eligibility criteria**

The eligibility criteria correspond to a set of characteristics shared by all participants, which accurately define the features of the population of interest. These characteristics determine the rules to be applied for building the sample of subjects. They may include age, gender, medical history and current health status. Eligibility criteria for treatment studies often require that patients have a particular type and stage of their disease
Enrolling participants with similar characteristics helps to ensure that the results of the trial will be due to is the feature under study and not other factors (confusion bias). In this way, eligibility criteria help researchers achieve accurate and meaningful results. A second function of eligibility criteria is to exclude patients who are likely to be put at excessive risk by the study.

Eligibility criteria comprise in most cases a set of inclusion criteria defining the characteristics mandatory in the population and a set of exclusion criteria defining the characteristics to be avoided. Usually the negation of an exclusion criterion becomes an inclusion criterion and vice versa. The number of criteria is very different from one study to another. The more numerous and/or specific they are, the more difficult it is to determine the eligibility of a subject (either because few patients match a useful proportion of the criteria, or because sufficiently detailed patient data are not available to determine the match).

1.3. Sample size

While the sample size has an important influence on the accuracy of the results, the power of the tests and the scope of the conclusions, this number is often determined ad-hoc. A sample size that is too small may turn a potentially important study into one that is indecisive or even a complete failure [3].
1.4. **Why is recruitment so hard?**

Several lines of evidence show that many CTs fail to achieve their recruitment objectives [4-6], which in turn can jeopardize the conclusion of the studies.

1.4.1. **Current situation**

The enrolment rate for a CT is the percentage of eligible patients that are finally enrolled on the trail. At an international level there is insufficient information to accurately evaluate CT enrollment rates.

Most published papers focus on the oncology field. With an enrollment rate of 35% to 46% depending on the country, oncology represents the greatest part of all CT. An important study carried out in 2004 [7] and based on data provided by the NCI Surveillance, Epidemiology and End Results program has shown that trial participants represented approximately 1.7% of the total number of incident cancer cases (for breast, colorectal, lung and prostate cancers) diagnosed during the 2000 through 2002 study period. In the United States, the National Cancer Institute reports that approximately 3 to 5% of adult cancer patients are enrolled in clinical trials at any given time. With an enrollment rate of 14% of Britain’s annual cancer incidence, the United Kingdom seems to have the highest rate of cancer clinical trials participation of any country in the world [8]. Whatever the rates, a recent study [9] shows that the recruitment phase takes up an increasing amount of the CT timeline and sometimes leads investigators to stop trials because recruitment targets cannot be achieved.

Recent studies [10] carried out by 19 pharmaceutical companies across all medical domains, compared the enrollment rates of different countries worldwide in 2004, 2006 and 2008. This report shows a consistent increase in numbers of CTs and a significant
change in enrollment rate for each country during the last six years. A strong increase in
CT enrollment rate is observed in Eastern Europe and Asia, while a decrease is seen in
the US. Europe remains stable, the UK being the most active country.

1.4.2. Factors identified

Barriers to CT recruitment have been broadly studied in the literature [7,11-26].
Although all medical specialties face problems in recruiting patients, the oncology
domain seems to the most affected [7,11-16,17-19,21,23,26]. The main factors cited in
the literature can be summarized into three categories.

System or organizational related barriers:

These mainly concern a lack of staff and research nurses, difficulties communicating at
different levels (intra-organizational or national level) and difficulties determining the
impact of clinical trial design on internal planned resources. For example, it has been
shown that physicians whose practices were based in university settings or had help
from a cooperative group had higher rates of accrual onto studies [11]. Organizational
barriers also result from a lack of access to health centers offering clinical trials. A 2003
Harris poll of 6,000 adult cancer patients [27] indicates that in the US, 85% of cancer
patients were either unaware or unsure that participation in clinical trials was an option,
though 75% of these patients said they would have been willing to enroll had they
known it was possible. The system or organizational related barriers likely depend on
the country and healthcare organization.

Individual barriers:

Individual barriers appear both at patient and the physician levels. The factors
commonly reported by patients deciding not to enroll in a CT are feeling uncomfortable with experimentation, feelings of uncertainty, or the desire to not lose control of decision-making. Patients dislike the idea of possible adverse effects or the feeling that a trial might not offer the best treatment option.

For the physician, the most often cited factor is the time and effort required to recruit patients. Physicians participating in a busy practice may feel that they do not have the time for efficiently identifying eligible study subjects, or they may simply forget to offer and enroll patients in open trials [28]. Completing these tasks also requires a level of access to clinical trial information. Important efforts have been made to make CT information available online at an international level [29,30] such as clinicaltrials.gov or at a national level [31] such as www.inca.fr in France. These tools list the open clinical trials by diagnosis, phase, modality, sponsorship, drug, and geographic location. But they are still rarely used “at the bedside” or during multidisciplinary team meetings in most local practice settings [32]. Information concerning the CTs in which a hospital is involved is still often available only on paper documents. And an important drawback of these tools is that they don't take into account patient data in order to automatically check CT criteria

**CT design related barriers**

The most commonly identified design factors are the complexity of some CTs and the difficulty of explaining them to the patient. The number of available clinical trials is quite high (e.g. 35,496 recruiting trials in October 2010 referenced on clinicaltrials.gov), and it is impossible for physicians or other medical staff to be aware of all of them.

In addition, overly strict or non-adapted criteria are often cited as a barrier to finding
eligible patients. For example, some trials require assessment of performance of patients using different methods according to their age or their condition. Conversely it has been shown [33] that a too low a number of exclusion criteria leads to reluctance by the physician to enroll older patients even if they only have a mild functional impairment. Criteria are still written manually by experienced protocol authors or committees of clinical experts, and the wording of criteria may not be clear to the physician trying to determine the eligibility of a patient. Moreover, it is possible for some criteria to be redundant or for criteria to conflict within the same protocol[34]. Patients dislike the idea of randomization, or of any eventual inconvenience or constraints in their everyday life caused by CT protocol design. Several other reasons for low recruitment based on protocol design have also been identified, related to extreme ages [18,26,33], social condition [21,23] or even ethnic or racial factors [7,35,36].

1.5. Could Hospital Information Systems (HIS) help to improve the recruitment process?

There is considerable evidence that health information systems, and in particular Hospital Information System (HIS), could improve the recruitment process. Dugas et al. [4] have estimated the accrual rate in clinical trials by manually assessing routine data from HIS. Their study shows that complete, high quality and accurate data can significantly help the recruitment process, at a very preliminary stage, by defining the expected rate of recruitment.

However, most relevant patient information still remains non-structured (e.g. clinical notes, clinical assessments). Moreover, relevant information is often coded for a specific
aim (e.g. billing, DRG system) and so may not correspond to the requirement of the CT (e.g. because of lack of accuracy, or the risk of “up-coding” [37] the diagnosis in order to classify the patient in a more expensive group).

After having identified the main barriers for the recruitment process and having identified the potential of reusing data coming from the E.H.R, then a new question arises: What significant work has been carried out towards automating patient recruitment? This is the main goal of this review - to provide an overview for researchers involved in the clinical trial process of the potential benefits of CTRSS, and to examine the contributions and limitations of previous work carried out in this field.

3. Methods

1.6. Literature search:

The publication selection and review process is described in the figure 1. A systematic search of Pubmed and Scholar Google was undertaken using a combination of the following MeSH terms or keywords: “Clinical trial”; and ”patient selection” or “recruitment” or ”enrollement” with “decision support system” or ”medical records systems, computerized”. For instance, for Pubmed, we used the following query: ("Clinical Trials as Topic"[Mesh] AND ( "Patient Selection"[Mesh] OR "recruitment" OR "enrollement") ) AND ("Decision Support Systems, Clinical"[Mesh] OR "Medical Records Systems), Computerized"[Mesh]) . We searched CTRSS publications from 1998 to October 2009.

Figure 1: Flowchart of the publications selection and review process.
As this study is focused on CTRSS, we broadly defined as inclusion criteria, any computer system or applied computer method that helps health professional in the recruitment process by reusing patient data derived from the E.H.R. We excluded, for example, systems or methods easing CT accessibility for patients and/or physicians (such as the ClinicalTrial.gov website), but that do not clearly reuse patient data.

In order to consistently address a CDSS subset, we decided not to exclude papers that did not report real patient outcomes. Furthermore, as will be discussed, many of the most relevant recent papers report systems which are still at a prototype stage. Despite the lack of full clinical results we chose to include these recent works in order to provide the broadest possible review.

1.7. Analysis axes

A systematic review was carried out of both research and industrial CTRSS systems along with a brief review of the decision support methods and interoperability standards employed in the field. Once the publications had been identified, they were classified according to a simplified version of the Clinical Decision Support Systems Taxonomy (CDSS Taxonomy) proposed by Berlin et al [1], which addresses the technical, workflow, and contextual features of any kind of CDSS.

This taxonomy was originally used for a general and complete review of CDSS. It contains 26 axes in five categories (Context, Knowledge and Data Source, Decision Support, Information Delivery, and Workflow). For example, the category “context” contains five axes (clinical setting, clinical task, unit of optimization, relation to point of care, external behavior modification programs and potential barrier). A description is provided for each axis (e.g. Potential barriers: Potential barriers to
completion of the action recommended by CDSS (socioeconomic barriers, conflicting reinforcement).

Because we focus on CTRSS, some Berlin taxonomic axes appear to us to be either non adapted for the domain (e.g.: the “External behavior modification program” axis) or irrelevant (e.g. the “clinical urgency” axis). These were removed, and we added an “evidence of efficiency” axis which addresses the direct or indirect proof of efficiency of the CDSS corresponding to an increase of the enrollment rate in clinical trials. Our final taxonomy includes 10 axes (see table I).

Table I: CDSSc taxonomy (the axes are gathered in five broad categories).

Each paper was analyzed against this taxonomy by two authors. In case of disagreement between authors, a consensus was sought. We have structured and summarized the results according to the five categories (see table II).

4. Results

From Pubmed and Scholar Google, we found 73 references. Our review of these identified 28 distinct studies of CTRSS spread over 33 publications. Sometimes the same team publishes several papers focused on one system, dealing with its improvement or for instance its assessment. In this case, we have summarized the contribution of these papers. Conversely, if a team publishes papers indicating a break with the previously described system, then we consider these papers as a new original work. This is the case for instance, for Machado’s papers published in 1993 then in 1999.

The level of detail provided in the papers varies greatly, in particular concerning the methods and protocols used for evaluating the systems (evaluation methodology or
efficiency indicators).
1.8. **Context:**

1.8.1. **Who are the actors and what are the related potential barriers addressed?**

Most of the systems (20/28) address “individual” barriers, at the level of the physician during the critical prescreening step. During this phase, physicians are the main actors involved in the pre-inclusion decision. Five works concern actors involved especially in the CT process (e.g. research assistants). Three works [32,38,39] identify nurses as potential users of the systems. Only one paper [38] describes patients as potential users. Beside these works dealing with individual barriers, four works are more focused on recruitment barriers in CT design [40-42]. These papers describe methods and tools to help CT promoters design eligibility criteria. These works propose that adopting a formalism for representing the criteria could dramatically improve their quality, allow indexing or annotation of the CT, and open the way to automatic recruitment of patients.

Finally, one paper tries to figure out the “potential patient barriers”. Rollman et al. compare two strategies of recruitment for mental health conditions, one using a reminder system embedded in the EHR which warns the physician of a potential recruitment versus a traditional case finding method carried out by a study recruiter stationed in a practice waiting room. This study shows that the rate of pre-inclusion of minority (here non-white) subjects was fivefold higher for the DSS strategy versus waiting-room recruitment strategy (the pre-enrollment rate by DSS being slightly lower than the waiting-room recruitment strategy for the rest of the subjects).
The authors argue that the strategy using the E.H.R during the interview could give physicians all the necessary and useful cues to recruit the patients. This approach would foster the “patient-physician” relationship, in comparison to a “systematic” interview carried out in the waiting room by a recruiter.

1.8.2. Application or experimentation domains

Out of the 28 papers reviewed, nine publications relate to oncology, three to AIDS, two to endocrinology and cardiology, and one each to rheumatology, emergency department and psychiatry.

Six papers do not specify the domain of application. For example, in Pakhomov et al [42], the presented system aimed to automatically build a cohort of patients relevant to epidemiology research; however the authors consider their system to be also applicable to the recruitment process for clinical trials. Most of these papers chose an application domain for testing their systems but do not exclude their use in other domains.

1.8.3. Workflow

An effective clinical Decision Support System needs to fit into the user’s workflow [43]. Success with alerts, guidelines, and algorithms depends substantially on integrating suggestions with practice. For example, it may be a requirement that use of the system should require minimal additional time and effort and should not constitute an interruption to the normal flow of work for the users. In our context, such systems should focus on the very critical phases of the recruitment process, such as the prescreening steps.
Understanding clinician workflow, particularly when designing applications for the outpatient setting, is critical. However most of the papers reviewed do not explicitly mention the issue of workflow and focus on describing an implicit application of the system, either at bedside or at specific times during the care process (at admission, at discharge or during multidisciplinary consensus meetings).

A few papers [44-46,48] describe the interactions between the different actors involved in the recruitment process. Embi et al [44-46] give a precise description of the workflow involving physicians, main investigators and clinical research assistants. They stress the necessity to solicit as little as possible from the physicians due to their workload. Breitfel et al [47] carried out a preliminary study of workflow showing that 25% of physician and nurse time was spent walking from one place to another. This waste of time justifies, according to the authors, access to the decision support system at the point of care. In their work, they propose to use a small computer that easily fits in a coat pocket.

CDSS systems can be divided into two types: systems that are automatically triggered by an event during the clinical process, and systems that require a voluntary action from the user. The “automatic” operating mode has an effect on the recruitment workflow as such systems automatically generate alerts which someone has to deal with - recording them in the patient medical record or informing the appropriate actors – which will in general interrupt their workflow. The appropriate threshold to trigger such alerts is, according to several authors, difficult to determine [44-48]. If the system is too sensitive, physicians may be overloaded by inappropriate alerts. Inversely, triggering on very specific rules implies a risk of missing some eligible subjects.

The appropriateness of the alerts depends on human factors as well. In particular,
automatically sending a message presupposes that someone is able to read and deal with it. For instance, the system proposed by Afrin et al [45] sends alerts, asynchronously readable, through email to main investigators 24 hours a day and messages to pagers only during working hours of the physicians.

Integrating the recruitment and healthcare processes within the same workflow may bring difficulties and lead to a new arrangement of some activities or to setting up triggering rules. For instance, the system described by Butte et al can identify in real time patients whose glycaemia is abnormal in order to enroll them in a clinical trial. In this experiment, authors had to deal with the timing of laboratory tests. Occasionally, a specimen of blood for glucose testing might be sent before patient registration (e.g. in case of extreme emergency). Or, conversely, sometimes a delay may occur before the result is available in the EMR (because it requires a validation by a senior biologist). In these two cases, the recruitment system failed to detect eligible subjects because of a desynchronization between the recruitment process and the health care process. Specific timing rules had to be implemented to solve this issue.

1.9. Data sources and data integration

1.9.1. Data sources:

All systems require the input of both patient data and eligibility criteria. Patient data may be provided in two different ways. Some systems (14 out of 28) require the user to manually enter the patient’s data. The user may have to interpret the patient’s data in order to adapt them to the decision support system (for instance by changing a unit, by calculating a score).
Other systems automatically integrate the patient’s data. Patient data is retrieved from
the HIS of one or more hospitals. It can either come from the patient electronic record,
be captured on the fly at the moment of production or be pooled extracted from a data
warehouse.

The eligibility criteria come either from local clinical trials or from the national clinical
trial registry (generally clinicaltrials.gov). The eligibility criteria are always presented in
a non-structured textual format that must be in turn coded to be usable by the decision
support system. The coding may be manual when an operator codes criteria as rules or
requests, or automatic when automatic processes (generally NLP methods) are applied
to the original textual version of the eligibility criteria. Whether manual or automatic,
the coding of the criteria remains difficult for most of the authors, due to heterogeneity,
complexity or inconsistency of criteria.

In order to overcome these difficulties, several solutions have been proposed:

- Olasov et al [49] use UMLS for disambiguating the criteria.

- Brietfield et al select criteria according to their capability to maximize the pre-
  screening process. They show that most such criteria are, in fact, simple and that
  it is not worth encoding the complex ones.

- Ash et al [50], propose that the authors of CT protocols do the encoding of the
criteria themselves, if possible using libraries of encoded criteria when available.
  This approach would lead to a more accurate and consistent encoding of the
  eligibility criteria.

- Gennari et al [51] developed an ontology of eligibility criteria based on the
  Protégé tool associated with an authoring tool. The goal here is to provide the
terminology and the logic needed to express eligibility criteria in the most
formal way.

1.9.2. Data integration:

Data integration refers to the methods used to acquire the data from the sources discussed above and make it available for the decision support system to use. Three different integration methods for patient data are used: Half (14/28) of the recruitment systems are clearly separated from the patient data sources and the data are integrated manually (from the patient medical records or the data warehouse). The other systems are designed with a connection to the patient data sources. Either systems analyze the dataflow exchanged with the HIS (this is the case for triggering systems) or systems send queries to data warehouses.

In most of the published works, data integration is based on proprietary formats or methods. Using interoperability standards remains the exception rather than the rule. Therefore systems integration relies on a strong linkage between subsystems.

There is some evidence that the use of standardized and structured information or knowledge models can improve data coding both in terms of consistency and coverage. For example, starting from a corpus of free text CT criteria, Ash et al [50] improved Lucida Machado’s work [52,53] by adopting an elaborated data model and standard vocabularies (through UMLS) that allowed the number of criteria that could be encoded to be increased from 50% (the rate reached in Machado’s work) to 98.9%. Moreover, they succeeded in encoding more complex criteria than Machado.

The Common Data Elements (CDE) developed by the National Cancer Institute, for the design, execution and analysis of cancer clinical trials was used as a data model in three works [38,54,51,52].
Lonsdale et al propose to use the Arden Syntax, an artificial intelligence frame-based grammar for representing and processing medical conditions and recommendations. The Arden syntax is relevant here as it defines the mechanism for accessing data through a Virtual Medical Record (VMR). Each value-paired attribute defined in a specific EHR is mapped to the VMR. This abstraction of medical records ensures that any number of healthcare organizations can write, maintain and share clinical decision logic no matter what the structure of their own repositories.

Apart from the works cited above, no publication uses interoperability standards like HL7 or OpenEHR, or other initiatives for data interoperability in clinical research like CDISC (Clinical Data Exchange Standards Consortium) [55] or CaBIG [56]. Nonetheless, some authors argue that interoperability standards could be a good answer to the problem of system scalability and reusability [38,51-52,54,57-59-61].

Current efforts focus on using a standardized terminology or ontology to map either patient data or criteria on to a common referential for coding both patient data and the CT criteria. In this respect, the three most cited (and actually used) terminologies are UMLS (as a terminology system) [49,54,62-64,], SNOMED CT [57,64-68] and ICD (International classification of diseases) [44,62-63,69,70].

Decision support methods

Essentially three methods have been adopted for inference in recruitment decisions: rule-based methods, query-based methods and Bayesian methods (although the Bayesian method is always combined with a rule based method).

1.9.3. Rule-based methods

In these systems, inference rules are used to define patient eligibility. An inference rule
is a statement that has two parts, an “if” clause and a “then” clause. Rules may be generated either manually or automatically. These rules might be processed by an inference engine, or interpreted by an algorithm.

The representation formalism for rules varies from one system to the other. It may be specific to a system. Afrin et al [45] implement rules for triggering alerts using a PERL script that regularly checks if the exams of hospitalized patients are abnormal. If so, the system automatically triggers the pre-screening workflow. Two publications [38,52] represent the eligibility criteria with the Arden syntax [71] which can be directly treated as logical rules.

Patel et al [67-68,72] represent inclusion criteria with description logic (DL) and implement automatic reasoners (Pellet et Racer [73-74]) in order to check the consistency of the criteria and make inferences about patient eligibility. The power of Semantic Web tools such as these is due to the performance and scalability of their automatic processing.

1.9.4. Query-based methods

This category of systems uses databases and query languages to define patient eligibility. Patient data is searched for patients meeting the eligibility criteria using database queries. For example, all patients having a specific condition (e.g. diabetes mellitus) might be found by querying the E.H.R database, and the subset of those having an abnormal blood test result (e.g renal clearance) determined by querying the Lab database.

Query-based systems are mainly used on relatively large-scale patient data warehouses [40,54,62,70,78]. Calinescu et al for instance present a system for querying the data
warehouses used in a health network of hospitals specialized in oncology. SQL is the most commonly cited query language.

Large scale data warehouses are also queried by systems using Semantic Web methods. Patel et al [67,68] propose to use the RDF and SPARQL languages to define queries with a high level of semantic abstraction. This approach falls within the scope of the Semantic Web paradigm and makes the query of heterogeneous and distributed data warehouses possible.

1.9.5. Bayesian methods

A Bayesian network is a probabilistic model that represents a set of variables and their conditional dependencies via a directed acyclic graph. Three reviewed publications use Bayesian approaches [50,52,75,76]. Pakhomov et al [63] compare Naïve Bayesian networks to the perceptron method for distinguishing cases of acute and chronic forms of congestive heart failure. The authors concluded that both of the methods were efficient, the perceptron method being more accurate and less sensitive.

Ash [50] uses a Bayesian network to infer missing data based on population-based probabilities of patient characteristics. The probabilities were obtained from the medical literature and known statistical databases.

1.9.6. Other methods

Seroussi et al [61] use a specific formalism, close to a guideline representation formalism, in order to represent criteria for a whole set of CTs. These are represented as a decision tree, in which nodes describe patient states and leaves propose the best corresponding CT. The decision tree is coded manually. In this work, the system asks
for the most common criteria, shared by the maximum set of CTs, and step after step, focuses on more specific criteria, eventually yielding a set of eligible trials.

1.10. Assessment and Efficiency

The objectives and evaluation methodologies vary a lot from one publication to the next. We chose to group the publications in three categories (see table III):

Table III: Criteria for evaluation quality

Because there are only nine papers in category A we will briefly discuss each one individually here. The assessment methodology in these studies depends strongly on the system type. Different indicators which reflect the systems' effectiveness more or less directly are evaluated.

The measure on which we would really like to compare the reviewed papers is the effect of CDSS on CT recruitment rate. For this we need to know recruitment rates both with and without CDSS, all other factors remaining stable. This is however very difficult to achieve in practice, and even in the case of the category A papers this specific measure is often not directly evaluated.

Butte et al. [48] have developed a real time messaging system to recruit patients with a hypoglycaemia condition in endocrinology CTs. This trigger-based system checks every two minutes if a blood sugar test was lower than a defined value. During the four months of assessment, 13 cases were detected by the system, and 11 patients actually met the other inclusion criterion for the study (which was having no previous episode of hypoglycaemia). Eventually, 10 patients consented to be included in the trial. The paper does not mention how many patients were actually tested by the system, but as all the blood sugar tests were systematically checked, the authors claim a sensitivity of 100 %
Afrin et al. [45] developed a similar system for rheumatology. They evaluated their system in clinical use for 10 months. Eligible patients were required to have certain suggestive diagnostic laboratory test results (e.g., significant proteinuria and a positive anti-nuclear antigen antibody or a positive anti-double stranded-DNA antibody). The system detected 1768 trigger events from the HIS corresponding to abnormal proteinuria. After applying additional rules (immunological criteria), 70 patients were automatically pre-screened. Then, for each patient, a notification was sent by email and paging to the ordering physician (or the attending physician) to complete the screening process. Eventually, it turned out that only three patients were enrolled in the trial. According to the authors, this low rate of inclusion is due to diverse problems: Some notifications were not sent due to missing email addressees, and communication issues arose between the ordering or attending physicians and the principal investigator who was in charge of determining full eligibility.

Ash et al. [50] used a system combining a rules based approach with Bayesian networks for recruiting patients in oncology CTs. For evaluation they encoded 10 CTs containing a mean of 27.2 criteria per trial. They succeeded in encoding 98.9% of the criteria with their formalism. They tested their system on data from a 20 patient sample (each patient could meet complete or partial criteria of one or several CTs). In this study, it turned out that none of the 20 patients was considered as completely eligible, that is, meeting all criteria. All the criteria checked by the system were compared to an expert physician’s decisions. The agreement was good (between-rater reliability kappa test = 0.86 p<0.05).

Seroussi et al. [61] present a recruitment system for breast cancer CT, based on a decision tree formalism. Their system was tested during multidisciplinary team
meetings. Amongst 127 cases analysed with the system, 23 patients were potentially eligible for the available CT. 12 patient were actually recruited. The relative accrual rate was thus 52% (12/23) and the absolute accrual rate 9% (12/127). Of the 11 potentially eligible patients that were not actually recruited, 64% (7/11) were not recruited due to failure to verify eligibility conditions, none refused to give consent, and 36% (4/11) were not recruited for other reasons.

Fink et al. [77] used a rule-based based system for oncology CT. A transverse study was carried out on 261 patients. System efficiency was compared to manual recruitment. The system and humans agreed on 113 matches. The system provided 173 new matches and missed 43 matches because of a lack of data. These outcomes indicated in this case that clinicians may miss up to 60% of matches. The authors argued that their system could increase the potential number of enrolled patients by a factor of 2.5.

Embi et al [44,47,69] present a real time ruled based recruitment system. The system detects eligible patients during clinical encounters, then a physician (who carried out the clinical exam) can refer them to the trial coordinator by sending a warning message. The system was tested in clinical use for 4 months, with 48 physician participants. The system automatically triggered 4780 alerts and 238 were referred by the physicians to the investigators. In terms of participation, the number of physicians referring patients after system activation increased more than eightfold from five to 42 and the number of enrolments generated doubled from five before to 11 after. In terms of pre-inclusion rate, 238 patients were referred by the physicians to the protocol coordinators. After having checked all the CT criteria, only 121 patients remained potentially eligible. Ultimately 24 patients were actually enrolled.

Grundweier and al. [79] have developed a ruled based system for paediatric CT. In their
system, if one or more patients are found to be eligible, either a reminder is directly displayed in the E.H.R, or a list of eligible patients is sent to research assistants. The authors compared the efficiency of these both methods. During the two year period of assessment, 11 trials were tested with the reminder method and three with the eligibility list method. The range of potentially eligible patients was between 17 and 1162 depending on the trial, and the number of patients actually enrolled ranged from 3% to 25%. Overall, across both methods, 2/3 of the tested trials reached their goals in term of recruitment.

Rollman and al [39] used a similar approach based on an EHR and prompting system. In this case, they compare this strategy to a systematic interview of patients carried out by a research assistant in the waiting room. The system was evaluated over 22 months. 8095 patients were interviewed and 193 (2.4%) met all the criteria and agreed to enrol. During the same period, thanks to the prompting system, physicians referred 794 patients. Amongst them, 176 (22%) met all the criteria and agreed to enrol. The accrual rate without the system is not reported.

5. Discussion

1.11. Limits of the study

This study focuses on a sub type of decision support system, CTRSS. The number of relevant papers related to CTRSS is relatively low (considering the same period of publication, that is from 1989 to 2009, and excluding review papers, the rate of CTRSS papers represents around 1% of all papers about decision support systems). Moreover, many of them either carry out no assessment of effect on recruitment to CTs, or are not
directly focused on recruitment (for example, works on transforming criteria to a machine-readable form) although they were still considered relevant for this review.

It was also difficult to categorize the different systems. For example, although we distinguish rule-based systems from query-based systems, these are clearly strongly related as a query can be considered the expression of a rule (e.g. a query on a data warehouse for all patients having a sugar blood rate > 5.55 mmol.l-1). We consider a system “rule-based” rather than “query-based” when rules are managed separately from the data source, and where several heuristics could be checked at the same time by an inference engine.

Finally, we found that comparing the impact of different approaches on recruitment rate was difficult. System designs, scope, and assessment methodologies vary significantly between papers. Despite the lack of clear data at this point in the development of the CTRSS field, it is nonetheless possible to see a number of clear tendencies emerging from the present study. Below we summarise the main tendencies we find in our analysis.

The best-assessed systems use simple methods or technologies.

The systems we reviewed which take the simplest technical approach are those which scan for the occurrence of a particular event during the care process, and trigger an alert to one or more actors involved in the recruitment process. Although it is difficult to directly compare assessment methods from one system to another, this kind of system in general seems to be effective, since none of them failed in their goals.

This does not of course mean that systems that have not been assessed in real situations (categories B and C) are not potentially powerful. It is difficult to move from an “in silico” experiment to full clinical evaluation of a decision support system, and it may be
that the more complex systems which have started to appear more recently are more effective than the simpler systems of earlier papers, but that there is simply not sufficient comparable data on their impact to be able to make such a general comparison.

Moreover, there is often a gap between technologies used in HIS, and technologies used in research work. For instance, using SPARQL [72,67] (which is a semantic-web query language) in HIS may be premature considering that currently, most of the HIS don’t use RDF databases or more generally semantic web technologies.

Few studies focus on systems acceptance by physicians.

We have seen little work assessing the acceptance or perception of the system by users. However, this aspect seems crucial, because recruitment systems are by definition linked to the care process, and must be able to provide meaningful assistance without damaging patient care.

The importance of workflow integration

CTRSS plays its role at the junction of two distinct processes: the care process and the research process. A comprehensive understanding of organizations and processes is therefore essential during system design. Few of the reviewed works stressed this aspect, for example by giving a detailed analysis of use cases or an impact assessment of the CTRSS implementation.

For instance, it is clear that some actors involved exclusively in the care process (such as physicians and nurses) understandably resist changes to their work patterns to accommodate CTRSS. It is therefore important to make the wider context of such systems clear in order to counterbalance possible negative perceptions (e.g. by recognizing staff participation, taking into account their own time constraints, and
giving feedback on the effect of the system on inclusion rates at the end of the process). Perhaps more than with other kinds of CDSS systems, it is crucial to consider workflow integration in CTRSS. The complexity of the CT workflow means that even minor problems in organization or communication can carry a system to failure even if it performs well in itself.

1.12. **Pre-screening phase is the most important process to address.**

If we exclude purely theoretical works, CTRSSs all tackle a sub part of the inclusion process, the pre-inclusion phase. All these systems aim at spotting candidates amongst a population of subjects (generally patients) who meet a restricted set of broad eligibility criteria, for instance demographic characteristics (e.g. sex, age) or medical characteristics (e.g. tumour stage, surgical history). The inclusion phase of the trial starts only once this set of likely candidates has been identified, and more precise criteria are now tested. These “inclusion” criteria are generally quite strict and require specific investigations, for instance lab tests (e.g. a Prostate Specific Antigen dosage) or disease staging based on imaging.

In order to get an homogenous sample (particularly in the case of multicentre trials), these stricter criteria are generally systematically retested, using standardized techniques, even if the information was previously available in the E.H.R. The final screening process actually focuses on pre-identified individuals, and generally, it is carried out by actors who are fully involved in the research clinical field (e.g. research assistants, investigators, etc).

The pre-screening phase is a critical step in the recruitment process because it
allows the individuals likely to enter a clinical trial to be spotted in a given population. It is based on a set of coarse criteria, and on information that is likely to be available in the patient record.

It should be mentioned that pre-screening criteria are not currently clearly identified in the documents describing clinical trials. The decision to consider criteria as pre-screening or screening criteria is made by the investigator or the clinical assistant. Moreover, once a likely candidate individual has been spotted all the pre-screening criteria are meticulously re-checked during the screening phase.

Clarifying the category of criteria, during the design phase of CTs, would clearly be very useful in optimising CTRSS efficiency.

1.13. All CTRSS reuse structured and coded patient data.

If we exclude the 15 CTRSS requiring manual input of patient data, all the other systems (14) reuse patient data that were previously structured and coded. These data are stored in the HIS, and usually in the EHR.

Patient data are already coded with a local terminology (e.g. MED) or often with a reference terminology (e.g: Snomed CT). Thus, none of the reviewed systems uses NLP for processing patient data. It should however be mentioned that unstructured and non-coded clinical data is still the common rule in HER. Structured and coded data come more often from the biology and imaging domains or the billing process. And as we have noted, this kind of information does not correspond well with pre-screening criteria. Concerning data recorded for the billing process (e.g. diagnosis or clinical procedures carried out), because these data have not been collected for a clinical purpose they may lack precision or could be biased [80].
Fortunately, there are some medical fields where clinical data tend to be coded due to demands such as epidemiologic studies, or for forensic purposes. This is particularly the case in oncology, where a great effort towards standardization of processes and data has been made for many years. In this kind of domain CTRSS should be more appropriate and efficient than in other fields where clinical data is still broadly stored in free text (e.g. psychiatry or internal medicine).

1.14. **More recent work tends to increase CTRSS scalability.**

Unlike earlier work which proposed ad-hoc CTRSS focused within a single domain, or which were connected to a specific data source, more recent work implements new methods allowing automatic recruitment at large scale to be considered. There is a trend towards generalisation and scalability. This is made possible by the use, for instance, of semantic web technologies [72,67] or by new architectures like federated data warehouses [81][45] where eligible patients can be spotted on a large scale. In the same spirit, a recent trend has been to define machine-readable formalisms to encode CT criteria. Two main approaches are emerging. On one hand, methods using NLP techniques to transform full text written criteria into a formal representation. This approach faces the common pitfalls of NLP techniques (disambiguation, context representation, etc.) [82].

On the other hand, the second approach is to support the CT promoters, during the CT design phase, in order to define the criteria in a machine-readable formalism. Several works focus on authoring tools providing, for instance, quality control of criteria (to avoid ambiguities), the syntax and the semantic framework to encode the criteria, or to study the discriminatory effect of a criterion by comparing it in former trials.
The authoring tool approach appears to us very relevant. If it seems to be, at first sight, more laborious to encode the criteria “at the source”, it opens up interesting prospects for automatically processing criteria downstream in CTRSS. PAS CLAIR

1.15. A lack of interoperability standards implementation.

Most of the CTRSS we have analysed are based on a tight coupling between data sources (e.g. EHR or patient data warehouse) and the CTRSS itself. This approach makes connection between the CTRSS and other heterogeneous patient data sources difficult. Interoperability standards seem to be a relevant way to address this problem [83,84,58], but these standards have almost never been implemented in CTRSS. However, many recent published works describe how to use standards for instance for DSS interoperability. A generic specification of HL7, described by Kawamoto et al [85], defines a Service Oriented Architecture for Clinical Decision Support system integration at large scale.

In the oncology domain the CaBIG initiative should be mentioned, in which interoperability methods and standards are defined for integrating clinical health information systems and research information systems. More specifically, within the CaBIG initiative, the BRIDG project [86] defines an interoperability framework, filling the gap between standards belonging to the health care domain (e.g. the HL7 standard) and the CT domain (CDISC standards which are for instance used for structuring data elements of Case Report Forms). These initiatives make it possible to “plug” a CTRSS into different sources of patient data and to cope with different system implementations, and eventually to foresee automatic recruitment at a large scale.

As we have seen, coded patient data are used by CTRSS as input. But today, most
clinical data are still in semi-structured form and are rarely coded in current electronic health information systems. Fortunately a huge effort is currently under way to define standards for structuring and coding the information in electronic healthcare systems. For instance, OpenEHR or HL7 are defining respectively archetypes and templates [87] [88], in order to define the structuring and the semantics of data elements for catching and coding patient data. All these efforts at standardization will benefit the patient recruitment process.

Conclusion: Toward a translational information system?

At the end of this analysis, we consider that the automatic recruitment issue is still open. It is difficult to make any strong statements about how effective automatic recruitment is, or about what makes a good decision support system for clinical trial recruitment, because so few have been evaluated for their overall impact on recruitment. Much work is still needed in terms of syntactic and semantic definition of both patient data and eligibility criteria. Greater scalability of CTRSS is essential for larger scale patient recruitment; the adoption of distributed architectures and interoperability standards may address this problem. The workflow integration issue is in our opinion often underestimated, either because systems have not been tested in the real world, or because workflow impact is not sufficiently taken into account during system design. Finally, we have highlighted the difficulty of transforming eligibility criteria from text to a machine-readable form. Authoring tools, used from the very first step of the CT design, could help work around the difficulties of natural language processing.

For us, it appears that CTRSS is clearly part of the new paradigm of translational medicine. The AMIA has defined translational bioinformatics as [89] "... the development of storage, analytic, and interpretive methods to optimize the
transformation of increasingly voluminous biomedical data into proactive, predictive, preventative, and participatory health."

Reusing patient data from the EHR for clinical research purposes illustrates this paradigm very well. In this context, several consequences have to be considered: The need for sharing of data by clinical and research processes is going to increase dramatically in the near future. It makes sense to take this data reusability into account by collecting from the source, that is during the healthcare process, some of the more relevant information for recruitment, for instance corresponding to the most common pre-screening criteria. Still within this paradigm, new perspectives arise through recent initiatives such as ELIXIR (European Sciences Infrastructure for biological information) [90] or IMI (Innovative Medical Initiative) [91]. These projects aim at building international e-health infrastructures to gather clinical patient data from heterogeneous sources (for instance from different hospital electronic health records) and that are managed by “trusted third parties” to broker data between healthcare, health administration and medical research in a safe, secure and ethical manner. On these infrastructures it will be possible to provide different services for carrying out population studies, utilising a unified approach across different research centres. One example of such a service is to help promoters figure out where a CT would have the best chance to recruit, by spotting a population of interest before launching recruitment itself.

Finally, with the development of “omics” sciences, genomics eligibility criteria will increasingly have a place in clinical trial recruitment. CTRSS will take advantage of this situation, because genomic data are more easily machine-processable than unstructured clinical data. Conversely, targeting very specific populations with particular genomic
features will make recruitment harder without computer support.

6. Acknowledgements

The authors would like to thank Delphine Rossille for her help.

David Glasspool was supported by EPSRC grant EP/F057326/1 and by a programme grant from Cancer Research UK.

Conflict of interest statement

Summary table

References


Annual Symposium, 2009.


[55] “Clinical Data Interchange Standards Consortium.”


[88] “AMIA Strategic” [https://www.amia.org/inside/stratplan] access date May 2010


Figure 1: Flowchart of the publications selection and review process.
Table I: CDSSc taxonomy (the axes are gathered in five broad categories).

Table II – Comparison of the different works

Table III: Criteria for the evaluation quality