On Selecting a Clinical Trial Management System for Large Scale, Multi-Centre, Multi-Modal Clinical Research Study
Hugo Leroux, Simon McBride and Simon Gibson
The Australian eHealth Research Centre
{firstname.lastname}@csiro.au

Abstract. Clinical research studies offer many challenges for their supporting information systems. AIBL assembled 1112 participants who volunteered crucial information for a comprehensive study on neurodegenerative diseases. This paper discusses the shortcomings of the clinical trial management system chosen to record the results of the study. A set of guidelines was devised and a critique of five systems ensued. OpenClinica was selected as the most appropriate option. The main contribution of this paper is: (i) proposing a set of guidelines to determine the appropriateness of Clinical Trial Management Systems (CTMS) solution; (ii) providing a brief critique of existing commercial and open-sourced CTMS; and (iii) alluding to some data migration issues and providing cues on how to address them. We conclude that open-source CTMS are viable alternatives to the more expensive commercial systems to conduct, record and manage clinical studies.

Keywords. Clinical trial management systems, clinical study, case report form

1. Introduction

Clinical research studies offer many challenges for their supporting information systems. The distribution of data collection points, the heterogeneity and complexity of that data and long term nature of the Australian Imaging, Biomarker and Lifestyle Study of Ageing (AIBL) [1] is a particularly challenging case study. AIBL instigated a longitudinal study into Alzheimer’s disease involving a cohort of over a thousand participants residing in Perth and Melbourne. Each participant volunteered a range of information including demographic, medical history, diet and lifestyle. They completed neuropsychological and blood tests and a subset underwent brain imaging tests [1].

An 18-month collection was conducted and a number of limitations were identified in the proprietary system providing data entry. These issues stem from the problem of receiving and processing data from disparate systems and were categorised as: data quality and integrity, usability, flexibility, lack of audit trail and defects introduction.

In addition to the technical limitations of the system, contractual issues with the vendor became apparent as the maintenance contract expired and renewal was priced beyond budget constraints. The combination of issues led to the decision to ponder an alternative clinical trial management system (CTMS). This paper describes the evaluation of several CTMS alternatives and the data migration process to the selected system. Section 2 addresses the criteria for a suitable system, a review of prospective CTMS and introduces OpenClinica as the preferred CTMS. Section 3 provides a discussion and presents our results. Section 4 summarises and concludes this paper.
2. Methodology

To remedy the data consistency and integrity issue we introduced some rigorous software engineering approaches to the process:

1. Addition of source control to provide audit trail to the process
2. Introduction of a bug-tracking software Jira [2] to document all irregularities within the data.
3. Addition of validation scripts, in a continuous integration loop, to ensure consistency of data items from several parts of the system.

The result is a process that is customisable and auditable but most importantly all data modifications are accountable. However, this is purely an intermediate and temporary measure. Below is our long-term solution to the data integrity issue.

2.1. Survey of the requirements for CTMS

The US Food and Drug Administration (FDA) have outlined a guidance document addressing computerised systems used in clinical investigations (CSCI), known as 21 CFR Part 11[3]. In Australia, the corresponding body, the Therapeutic Goods Administration (TGA) have also developed a set of notes on Good Clinical Practice (GCP) [4]. These guidelines serve as reference point for designing CTMS. Their main objective is to act as a standard to facilitate acceptance and auditing of clinical trials by the regulatory authorities. We have drafted a subset of these guidelines that are desirable functionalities for the prospective system.

1. Implement security measures and protocols that prohibit unauthorised access to the study and data.
2. Provide adequate audit trail to ensure that all changes pertaining to the conduct of the trial are well documented.
3. Incorporate features to encourage the consistent use of clinical terminology and to alert users that data is out of range.
4. Provide suitable safeguards to isolate identifiable information from the study and ensure that retrieved data regarding each subject is only attributable to that subject.
5. Provide satisfactory backup and recovery protocols to guard against data loss.
6. Provide support for several types of fields (such as dates, text, numerical values) and in various formats (such as files, x-ray images).
7. Facilitate data extraction and the ability to swiftly generate reports.
8. Uphold the cost effectiveness of the system.
9. Endorse minimal development efforts
10. Advocate an advantageous type of licensing.
11. Promote adherence to industry standards, such as the Clinical Data Interchange Standards Consortium (CDISC) [5].

2.2. Survey of prospective CTMS and fit to our criteria

We have reviewed several of the Electronic Data Capture (EDC) systems outlined in [6]. They include Oracle® Clinical, Rave, InForm, DADOS Prospective and OpenClinica. These systems, along with the existing system have subsequently been applied to the list of criteria to determine their suitability.
Oracle® Clinical [7] is a commercial EDC system that facilitates the design and management of clinical studies. It integrates several functionalities relating to study design and management into a single application architecture.

Rave [8] is another leading commercial EDC system that provides software-as-a-service based clinical data management solution in a single platform. It is very flexible, allows the sharing of data across several partner technologies and offers wide support for industry standards.

Inform [9] is yet another commercial EDC system developed in a modular service-oriented-architecture platform that accommodates the full spectrum of study scenarios. Its modular design offers flexibility and scalability and seamless integration of components.

DADOS Prospective [10] is a web-based open-source EDC system for data management activities. It supports easy creation of Case Report Forms (CRFs) and electronic signatures, consolidates raw research data, source documentation and regulatory files and audit trails in a single environment. However, it does not yet adhere to any of the industry standards, such as CDISC.

OpenClinica [11] is an open source platform for collecting and managing clinical data trials. It supports the creation of customisable studies and the design of user-defined eCRFs, supports industry standards and adheres to regulatory guidelines.

Table 1. Fit of existing EDC systems to our criteria

<table>
<thead>
<tr>
<th>EDC Criteria</th>
<th>Oracle® Clinical</th>
<th>InForm</th>
<th>Rave</th>
<th>DADOS Prospective</th>
<th>OpenClinica</th>
<th>Current System</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>5</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>6</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>7</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>8</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>9</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>10</td>
<td>Proprietary</td>
<td>Proprietary</td>
<td>Proprietary</td>
<td>GPL</td>
<td>LGPL</td>
<td>Proprietary</td>
</tr>
<tr>
<td>11</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 1 provides a comparative fit of the six systems to our set of requirements. Despite faring well in terms of functionalities, Clinical, Rave and Inform were not considered as candidates because: (i) we could not justify the prohibitive costs associated with their acquisition; and (ii) issues surrounding licensing and not being open-sourced would force us into expensive maintenance contracts with their respective providers. The two remaining systems offered practically identical benefits in terms of functionality. OpenClinica was ultimately chosen because it addressed all of our needs.

The study has been restructured in OpenClinica and we have modelled the study protocol around the new framework illustrated by Figure 1. Each of the branches of the
study has been redesigned as an event defined by an electronic CRF (eCRF). Doing so, we facilitate the organisation of the data, ensure concurrent execution of the tasks, provide better security and ensure that each data item is attributable to only one subject.

One of the strengths of OpenClinica is in the richness of eCRF creation. It facilitates the organisation of the data at a coarse- or fine-grained level by effortlessly classifying related data items into groups or sections. Being a strongly-typed system, it imposes a data type on each data value and response field defined within the eCRF. This has enabled us to remodel the data from a loose data system comprising strings to a strongly-typed system encompassing a rich data model. Associated with each of these fields is an efficient validation mechanism, implemented using regular expressions and pre-defined functions, which is executed upon data entry.

Furthermore, while our current system had limited support for data integrity, OpenClinica provides two of its most important features. First is the generation of ‘Discrepancy Notes’ for data entered that do not adhere to the stringent validation criteria described previously. And second, it provides an audit log of all electronic activities performed on a particular record, comprising a timestamp of the change, the user, the reason for the change and the old and new values of the record.

We regard the generation of reports as another crucial functionality for this study. Reports correspond to what is sent to the collaborators to facilitate their research. OpenClinica supports the CDISC [5] standard and the data can be exported in either the CDISC Operational Data Model (ODM) 1.2 or 1.3 formats, the OpenClinica ODM 1.2 or 1.3 formats or as a tab-separated text file.

A data dictionary is a very useful component that holds the definition of the data item in the study and is sent to collaborators to assist them in their research. It is derived from the eCRF files, which ensures that it is always consistent with the eCRF.

The migration of the study data to OpenClinica has already started and the data integrity issues in the current system has hindered its smooth operation. To hasten the process, we have introduced migration and validation processes as depicted in Figure 2. The migration process retrieves the data from the current study and transforms it into a format required by OpenClinica. The migrated data is then analysed to see whether the value of each data set at the target matches the value at the source. Subsequently, a validate process has been introduced to ensure data integrity. Validation entails first analysing each data item and first ensuring that it conforms to the correct data type. Second, in the case of a numerical data item, a list or range of acceptable values is obtained and mapped against the data item. If any discrepancy is encountered, then not only is the entire process rolled-back but this data item is flagged to indicate that it is
invalid. This latter step not only ensures that the data that is copied into OpenClinica is consistent with the current system but that it also passes data quality and integrity measures.

3. Discussion and Results

Determining the most suitable CTMS for a study entails a number of factors:
1. **Open-source or Commercial:** This is determined by the size and scope of the study. Commercial systems would appeal to industry-funded clinical trials or suit larger research groups working on multi-modal trials, whereas open-source systems are generally free, although some may charge a minimal cost for support and would suit smaller groups.
2. **Need for customisation:** Commercial systems do not offer access to their source code, preferring instead to charge for customisation requests. Open-sourced systems by their very nature provide more flexibility for personalisation and could be the only option available to study administrators who need a tailored approach to their study.
3. **Licensing:** Ownership constraints and licence have a significant impact on studies’ operating costs. Commercial CTMS users have the option of an upfront permanent licence or a renewable one. Open-source systems operate either on a General Public Licence (GPL) or a Limited General Public Licence (LGPL).
4. **Adherence to industry standards:** Most CTMS systems adhere to industry standards, although some of the emerging open-sourced systems may not. Non-compliance to industry standards may restrict the extent of collaboration between various study groups due to incompatibilities and could also hinder the publication of study results.
5. **Cost:** Cost can be the ultimate determining factor as discussed in [6].

The current trend to sharing and persisting research data [12] reinforces the need for strict guidelines for creating, storing and managing clinical trials data. Walport and Brest [12] make three suggestions to improving access to research data but more importantly, they intimate to the creation of incentives for researchers and a shift in culture to one that rewards data creators and curators equally.

The migration process has demonstrated that it is possible to migrate from a rigid proprietary system to an open-sourced system. However, owing to the lack of adequate safeguards for data entry and validation in the current system, we have had to make slight adjustments to migrate the data successfully. Furthermore, remodelling the tasks into eCRFs to include customised lists, groups and sections have necessitated

![Diagram](image-url)
additional steps to translate the data. The next step is to demonstrate the OpenClinica system to the end-users. Following a trial period, we intend to go live with the OpenClinica system and slowly phase out the current system.

4. Conclusion

A comprehensive study into the effects of neurodegenerative diseases was instigated by the AIBL research group. They assembled a cohort of over a thousand participants who underwent a battery of tests. These tests were collected, recorded and managed in a commercial clinical trial management system (CTMS), which had shortcomings. A framework was introduced to remedy the situation but led to data integrity and security issues.

This paper has discussed the steps undertaken to remedy these shortcomings. Doing so, it has outlined the challenges facing health informaticians when choosing, designing and implementing a CTMS. A set of eleven desirable guidelines was proposed. This was followed by a critique of five CTMS and how they fit the eleven criteria. The OpenClinica system was chosen as the most suitable option and a brief description of some of its key functionality ensued. Data from the current study has been partially migrated to the OpenClinica platform with relative success so far.

The main contribution of this paper is threefold. (1) It has proposed a set of guidelines to determine whether a commercial or open-sourced solution is more appropriate for clinical trials and under what conditions; (2) it has provided a brief critique of existing CTMS providing study directors with a starting point in determining the most appropriate one for their clinical trial; (3) it has alluded to data migration issues and provided cues on how to address them.

We conclude that open source CTMS are viable alternatives to the more expensive commercial systems to conduct, record and manage clinical studies.

References


