Creating a single instance / multi tenant eClinical Platform

Paper

October 2016

CLINPAL

Revision History		
Version	**Description of Change**	**Date Released**
V0.1	Internal Review release	1st Oct 2016
V1.0	Public Release	9th Oct 2016

Table of Contents

CLINPAL

eClinical System Architectures

Information technology systems to support clinical trials have existed since the 1980's. Their evolution has not followed an entirely equivalent path to other industry segments. A number of factors come into play in how systems architectures evolve, we will start by looking back at how technology platforms supporting clinical research have evolve to today. We will then define how single instance / multi-tenant platforms fueled by a number of disruptive catalystic changes will adjust the way IT does not just support the existing clinical trial, but change the way in which trials run.

The migration from CDM Systems

Clinical Data Management systems ruled the roost. All medium to large scale firms typically using systems such as DLB Recorder (acquired by → ERT → Omnicomm), Domain Pharma's Clintrial (acquired by → PhaseForward → Oracle) and Oracle Clinical. They were all Oracle based and placed a focus towards supporting the heads down data entry, paper CRF management, and structured data. The architecture of these systems was defined by their use case and physical environment. They were relational database tools – SQL based – that operating as database table or view builders. The term 'Database build' often referenced in CRO Bid grids today dates back to these days when the act of setting up the system was in effect 'building a database' to support the clinical trial data.

In the early 2000's, EDC systems emerged driven by the internet and a wish to move away from the shifting of paper CRF's. The early systems supplemented CDM, but did not wholly replace them. They typically had limited support for heads down data entry, double data entry and critically, they generally had poor structured data support.

Unlike CDM, most EDC systems operated on a study by study basis with non-enforced data structure standardization. Their focus instead was presenting good looking forms to sites and for the corresponding monitoring activity.

CLINPAL

A move towards better underlying data, and better support for patients has driven the take up of ePRO – electronic Patient Reported Outcome systems. However, these have tended to be bolt on solutions either from different vendors, or, from the same vendor as a separate product development.

Sponsor or CRO companies that work with multiple vendors encounter compounded problems. Many studies on many diversified technologies creates a significant headache.

The resulting architectural outcomes has been an increasingly prevalence of data aggregation and data warehouse tools that pulled together data diversified data. In certain respected, these downstream solutions are sledgehammer fixes for an upstream problem where disconnected solutions define and describe things independently.

The evolution of eClinical technologies

It takes many years for an eClinical software company to grow. Sponsor companies that run individual clinical trials will not be prepared to commit volumes of studies to an unproven (by themselves or others) vendor. As a result, the vendor marketplace is dominated by a combination of heavily invested companies that had to survive beyond the initial years of investment versus the niche players that were able to grow beyond the initial early adoption phase by addressing a niche area.

A number of companies have taken this slow growth of a large number of 'silo' niche vendors to create a virtual suite built up from grouping a set of technologies to create an apparent end to end platform. From a marketing perspective, this can be convincing – products can look and feel similar – but from an architectural perspective, the systems rarely effectively share metadata or even data effectively. The implementation of Service Bus architectures under these legacy solutions may appear to address the lack of integration, but in reality they remain held down by their difficulty to consolidation and standardizing metadata from products that were not built with this in mind.

'Circumstance led' architectures will result in overall solutions where sponsor companies are left 'scratching their heads' trying to determine why costs and complexities are high, and integrations are difficult.

Best of breed shopping list

Clinical trial budget holders are left with a dilemma. Go to the marketplace looking for a set of different 'best of breed' technologies that support their trial and hope that the lack of effective integration will

not compromise their study, OR, go with potentially a lesser number of systems that provide a more integrated approach, but accept the more limiting functionality. It is not uncommon for a clinical trial at sites to rely on 5 or more separate systems to support the operations.

Historical 'all-in-one' systems

Sponsor companies have often been reticent about using a single integrated platform. This can be attributed to 3 primary reasons.

1). The risk of putting 'all their eggs in one basket' and

2). The historical functional limitations of the integrated systems on the market and

3). A prior commitment from areas of their business to already standardize on single system within their own silo.

A number of factors emerging today are changing things.

Domain Specific (Language) Platforms & PaaS

In the 1980's and 90's, it was popular to use 4th Generation languages (4GL's) and application development platforms to rapidly develop business applications. Tools such as Progress, Seachange and Powerbuilder allowed companies to rapidly produce software that would take many times longer to develop in traditional lower level languages. 4GL's went out of fashion initially with the emergence of windows Application Development Environments (ADE's) such as Delphi and Visual Studio. Their demise continued with the evolution of the internet where a rich development and user experience was difficult to achieve.

Clinsoft with Clintrial made a valiant attempt to move into a rapid development world with the creation of ClinTrial v4. However, the constraints that an off-the-shelf 4GL placed on the application meant that this was limited.

Platforms As a Service[1] is a category of cloud computing services that provides a platform allowing customers to develop, run, and manage applications without the complexity of building and maintaining the infrastructure typically associated with developing and launching an app.

---

[1] https://en.wikipedia.org/wiki/Platform_as_a_service

Catalysts for change

1) the 'responsive' patient

There is a design approach to websites called Responsive Web Design (RWD). This is where a single webpage (or web application) can be produced in such a fashion as to operate effectively on different dimensions of devices. PDA, Tablet or Computer browsers. In 2005 – this was a nice to have – in 2016, this is a mandate for internet users. Patients are internet users. What this means to a patient that has a direct involvement in a clinical trial is that they have an expectation that they will be able to use any internet device to support their involvement in a trial. The platform that supports them must be 'Responsive'. A lack of support will hamper or limit their engagement, and therefore impact data quality and/or compliance.

2) Need for change[1]

The way in which clinical trials function today MUST change. Costs are increasing, recruitment is dropping and timelines to successful development are extending. The methods used to date must change if PharmaBio and Devices companies are going to continue to operate profitably.

3). Data standards compliance

The FDA have committed to the standardization of data they receive for data submitted as part of an NDA[2]. Throwing time and money at the problem has been the traditional approach. EDC has negatively impacted the industries ability to meet these submission standards with standards such as CDISC CDASH only partly addressing the issue. Standard eCRF forms do not guarantee standard data – they suggest and imply it, but they leave a situation that demands costly trial by trial programming and data aggregation. The 2 year implementation target of Dec 2016 is looming [3]for when the FDA were aiming to move from a recommendation to a mandate (waiver aside) for the submission of data in CDISC standard formats.

CLINPAL

4). Site burden

Some physicians may be reluctant to refer patients to clinical trials because they believe that their involvement in a clinical trial will be an excessive administrative or financial burden to their practice (EDICT, 2008). In addition, This is not good for the site or the sponsor. It is not sustainable. Measures must be taken to revise the technologies that are offered to sites to not hinder by to assist in their clinical trials. Forcing transposition of source data into an EDC system is often nothing more than a data typist operation. Providing an environment that can support eSource in form and function will help sites as well as offering other aids that make the clinical trial experience less onerous.

- Single Sign On
- Integrated eLearning
- Cross function task and query management

Rich web development and user experience environment

With the prevalence of HTML 5, we have for the first time since desktop applications in the mid 90's, a means to create a rich user experience via browser. Software tools such as Lucidcharts, Smartsheet and Confluence show that it is now possible to not just create applications, but to create tools that operate through a web browser.

SaaS Single Instance / Multi-Tenant eClinical platform

Definition

A single instance / multi-tenant [4] eClinical platform provides a single system that supports 'n' – sponsors, 'n' clinical trials, 'n' - sites and 'n' – users on a single system where 'n' is only limited by the overall capacity of the cloud implementation (theoretically unlimited) provided by the vendor. Unlike CDM and legacy EDC tools, rather than setting up multiple installations and databases for each sponsor / trial, a SI/MT system is always available, always on, and hosts all studies. But why? Why not simply create multiple instances of a system? Does this not make security and deployment easier?

Enforcing Data Standards

How many definitions of Date of Birth do we need to have in the world? Answer – One.

If you go to a typical pharmaceutical company and ask the question how many definitions they have for Date of Birth across the different systems that are used to support a clinical trial, it could be as high as ;

(Number of Trials * Number of systems)

That might be a broad brush statement – some of these systems are not metadata based and therefore have a single definition of date of across all trials. Also, some systems may implement CDISC standards and use BRTHDTC. However, where is the enforcement when many different systems are involved? Computer software does not handle 'should' or 'generally' very well. Either the date of birth is called a standard name or it is not. Non integrated systems fall back to the lowest common denominator OR the case of data standards the poorest solution at enforcing the standardization of data elements.

Single instance does not mean a single system

A single instance / multi-tenant solution can be better described as a hub. A hub for the storage of metadata and data. How the data is captured and cleaned may vary based on specific use cases. What is important is that the definition of the data and the aggregation of the data are linked. Data quality is tightly linked with metadata standardization.

The principle of the 'hub' approach is increasingly attractive with the emergence of micro-services and single-sign-on standards. The approach is being seen across industries other than life sciences. Today, I can use my corporate google app's account login to login to our wiki, create a swim-lane diagram using an entirely separate SaaS application and cross reference changes in our own SaaS change management system. As a user, it comes across as a single solution. Technical, these are 3 entirely separate applications linked through a SSO solution and interfaced through micro-services.

Applying this same approach to eClinical with the Clinpal system – I might login using my OpenAuth 2 ID, I must be presented with a task to code a field – I go to the grid that presents the field in context, and the coding appears as an option I can apply to this field. I do not login again. A single user interface is provided, but 2 systems are involved and interfacing. In the example, the metadata and data are retained inside Clinpal, but the act of coding is carried out aided by OpenAuth 2.0 in a separate application.

Databases

Many approaches have been taken in creating an effective data platform for an eClinical environment. Focusing on EDC – databases are often defined 'per sponsor' with a set of static tables storing application data who's structure does not change combined with either dynamic tables managed by the application or more commonly 'hyper-skinny[5]' form. Hyper-skinny tables are where a single value is stored in a single record. This has the advantage with EDC in ensuring the schema remains the same, but creates complexities and difficulties in reporting and data extracts.

New types of databases – NoSQL – in both Document form and Object Oriented – have existed for many years. It is only in recent years where their cost effectiveness in comparison to commercial SQL tools, AND their capacity to handle very large volumes of data across multiple hosts and data centers means they are not just competing in certain fields, but are over taking.

With the implementation of the Clinpal system we moved from a well known SQL database to a Document database. We did this for a number of reasons;

- Scalability – the ability to scale a virtual database to terabytes and beyond
- Reliability – the ability to run the database across many different servers and data centers

- Schema-less – the ability to be closer to 'the code' and not be constrained by a fixed relational tabular data structure
- Break the VPN fallback – many EDC vendors when faced with the challenges of providing reporting or data to their sponsors fall back to opening up their databases to a VPN connection.

A document database provides the level of variance required when offering support for multi-application metadata driven solutions.

The purpose of a clinical trial is to capture data. Not to capture forms. Forms are simply manifestations of data. With the evolution of mobile devices, the eCRF as the only factor for the representation of entered clinical trial data is waning.

When a form is opened on a tablet – how many questions should I see? If the same form is opened on a mobile device?

The answer is it depends – it is variable. The tie to a fixed number of fields per form is not relevant in todays multi-form factor world.

One platform, many uses

A single platform with single sign-on, and a generic environment for Forms, Workflow and Messaging opens the door to supporting not just 'EDC' or 'ePRO' but many many variants of the application space required to support the successful execution of a clinical trial. This vast simplification in the architectural complexity drives enormous benefits with significant reductions in the demand for metadata and data synchronization across platforms and products.

Open API's versus Open Source

For many organizations that either had limited budget or required a solution that demanded extensive integrations, open source eClinical solutions have been the preferred route. These systems bring the benefit of a development community and the means to self manage / self host / self change a product. However, this has come with drawbacks. Creating and managing an effective validated hosting

environment together with a validated software development lifecycle process does not come cheap. In addition, although it may be possible to leverage new releases of the open source software, implementations often become small silo's in themselves with the availability of the platform for a wider site and patient audience limited.

Open API solutions are a new breed of software as a service solution that provide an always on secure, but always available API. With the means to push and pull configuration metadata as well as data through these API's, other integrating systems have significant power in determining how the Open solution operates. The API's are general persistent over many many years, even when the underlying system, architecture and capabilities change. This approach enables multiple systems to co-exist without the inherent ties that might occur in custom data integrations with an open source solution.

Highly configuration based systems combined with an Open API bring the means to fully adapt a systems setup without recourse to 'programming'. This can address the validation concerns as well as avoid the custom bespoke trap.

Integrated Trial Development Environment

In a 4GL environment popular in the early 90's, it was possible to use an IDE to quickly create tables, create forms, define logical data triggers, menus, reports…. A full stock control system could be developed in 3 days… This was great for high level business programmers.

We have a similar situation in the world of clinical trials. We don't need the full flexibility of a 3GL, with all its complexities, but study designers do wish to create data structures, forms and logic, and deploy quickly. Creating tools to support this has not been possible until recent developments with cloud computing and HTML5.

The Data structures are SDTM, or a flavor favored by the sponsor.

The Forms are CDASH style (although we don't tie ourselves to a single medium)

The database triggers – or edit check logic & workflow – define how the forms interact and how data is validated.

Conclusion

A combination of factors including the pervasiveness of Cloud computing, cost control demands, out of the box inter-operability needs and simplified patient engagement will cause the emergence of Single Instance / Multi Tenant platform solutions such as Clinpal™ that will not only improve the efficiency of clinical trials, but will actually change the way in which clinical trials are executed.

Disclosure

So in 2012, eClinicalHealth undertook the development of a platform that met the aforementioned criteria and to date have deployed it globally in studies ranging from a single 'virtual' site remote studies to large scale 10,000+ patient global Observational studies. The platform continues to evolve with end to end patient engagement an integrated feature.

CLINPAL

[1] **Transforming Clinical Research in the United States: Challenges and Opportunities: Workshop Summary.**

Institute of Medicine (US) Forum on Drug Discovery, Development, and Translation.

Washington (DC): National Academies Press (US); 2010.

[2] **Study Data Standards for Submission to CDER**

http://www.fda.gov/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/ucm248635.htm

[3] http://www.fda.gov/downloads/Drugs/UCM433492.pdf

[4] Wikipedia – generic definition of Multi-tenant solutions

[5] Practical guide to clinical data management by Susanne Prokscha, ASIN: B008V04MQ4 , pages 34-35,