



CLiX ENRICH for Clinical Trials

Service Definition

Clinithink Limited

Date: 30/03/2017
Pages: 12
Version: 1.0
Author: Clinithink



THIS DOCUMENT IS CONFIDENTIAL and intended solely for the use of the individual to whom it is addressed or any other recipient expressly authorised by Clinithink limited, in writing or otherwise, to receive the same. If you are not the addressee or authorised recipient of this document, any disclosure, reproduction, copying, distribution, or other dissemination or use of this communication is strictly prohibited.

This document remains the property of Clinithink Limited.
All contents are © copyright 2017.

Contents

Overview	2
Solution Overview	4
Business Requirement	4
Unstructured clinical data is rich content if it can be accessed and interrogated	4
Solution – CLiX ENRICH for Clinical Trials.....	5
Solution use cases and benefits	10
Clinical trial protocol optimization.....	10
Benefit.....	10
Feasibility studies.....	10
Benefit.....	11
Accelerating subject recruitment	11
Benefit.....	11
Delivering new data services and tools	12
Benefit.....	12

Overview

Founded in 2009, Clinithink successfully tackled the challenge of transforming unstructured clinical narrative into structured, actionable clinical information. Today, Clinithink develops industry-leading business solutions for the healthcare and life sciences markets, powered by its CLiX clinical natural language processing (CNLP) technology. Spanning the US and the UK, Clinithink is backed by venture funding and has offices in London and Bridgend in the UK and Atlanta and New York in the US.

CLiX ENRICH, Clinithink's uniquely powerful clinical data platform, provides mission-critical clinical data insights for some of the world's leading academic healthcare organizations, healthcare solution technology providers and research organizations, revolutionizing clinical trials and pharmaceutical development.

- CLiX ENRICH enables:
 - Quick access to answers from the interrogation of vast amounts of data for more insightful population health management;
 - Easy integration and publishing within existing business intelligence (BI) analytics platforms for more effective predictive modelling;
 - Richer, stronger predictive models for increased accuracy in risk assessment.
- CLiX ENRICH *for Clinical Trials* can:
 - Increase predictability during bid phase feasibility assessment;
 - Increase the yield of patients eligible for enrollment from the existing population;
 - Dramatically accelerate the identification and enrollment of patients for clinical trials.

CLiX ENRICH for Clinical Trials is deployed for use typically by secondary healthcare provider organizations. Such organizations deliver healthcare and operate major healthcare enterprise information systems to support patient care processes, including Electronic Medical Record and other clinical documentation systems. Consequently, customers have access to large volumes of clinical data that can be leveraged to support research. This data is typically both structured and unstructured. Normally, existing business intelligence and reporting tools and infrastructure support the analysis of structured data in this environment.

However, clinical trials, post market surveillance and population health analytics are all aggregate data research activities undertaken and/or supported by secondary healthcare

providers that would benefit greatly from the enrichment which unstructured narrative offers. This is because there are circumstances in each of those activities where narrative contains important data signals and patterns which are not reflected in available structured data.

Clinithink's CLIX ENRICH *for Clinical Trials* solution can overcome this challenge, making a large amount of previously inaccessible clinical data available for analysis alongside existing structured data sources, hence supporting all the business activities listed above. The remainder of this document describes the solution in more detail.

Solution Overview

Clinithink's CLiX ENRICH *for Clinical Trials* solution is designed to support the aggregate analysis of unstructured clinical data, delivering value to a number of different business contexts within the life sciences field.

Business Requirement

Unstructured clinical data is rich content if it can be accessed and interrogated

Investment in technology to support analytics in healthcare is increasing rapidly. Inputs to such solutions are typically structured data obtained from Electronic Medical Record (EMR) and other similar transaction processing systems used by health systems. Valuable actionable insights can be gleaned from this approach, especially in relation to operational efficiency and public health improvement. However, a number of business contexts require analysis of data relating to detailed clinical criteria to define and retrieve specific patient cohorts to drive benefit and value. These contexts include:

Identification of potential candidates for clinical trials

Typically, inclusion and exclusion criteria for clinical trials are complex clinical statements which define patients of interest for the trial. Gathering patients for trials is a costly, time-consuming and predominantly manual process which greatly impacts time to market once products have reached Phase III clinical trials. Leveraging unstructured clinical narrative to identify candidates can dramatically reduce the time taken to achieve enrolment targets for provider and research organizations, enabling the completion of an effective trial on time and to budget. This therefore increases speed to market for pharmaceutical companies, the contract research organizations (CROs) or academic research organizations (AROs) running the trials on their behalf, creating value for each.

Quantifying unmet need

Pharmaceutical companies collaborate with academic organizations to assess the potential market sizes of patients with unmet need that could be addressed with new products. The characteristics of such patient groups often include detailed clinical criteria which are best assessed against narrative clinical data.

Solution – CLiX ENRICH for Clinical Trials

To address the business requirements described above and similar ones in related contexts, Clinithink has developed a solution that enables organizations to import narrative patient data, process it with the company's unique CLiX CNLP platform and then interrogate the output using Clinithink's powerful query capability. Users can also write and run their own clinical queries to support the specific information requirements of the organization. The output from these activities is stored in a relational database that can be interrogated by 3rd party data warehouse and business intelligence solutions alongside existing structured data.

Figure 1 sets out the solution architecture with the paragraphs following providing a description of the components of the solution.

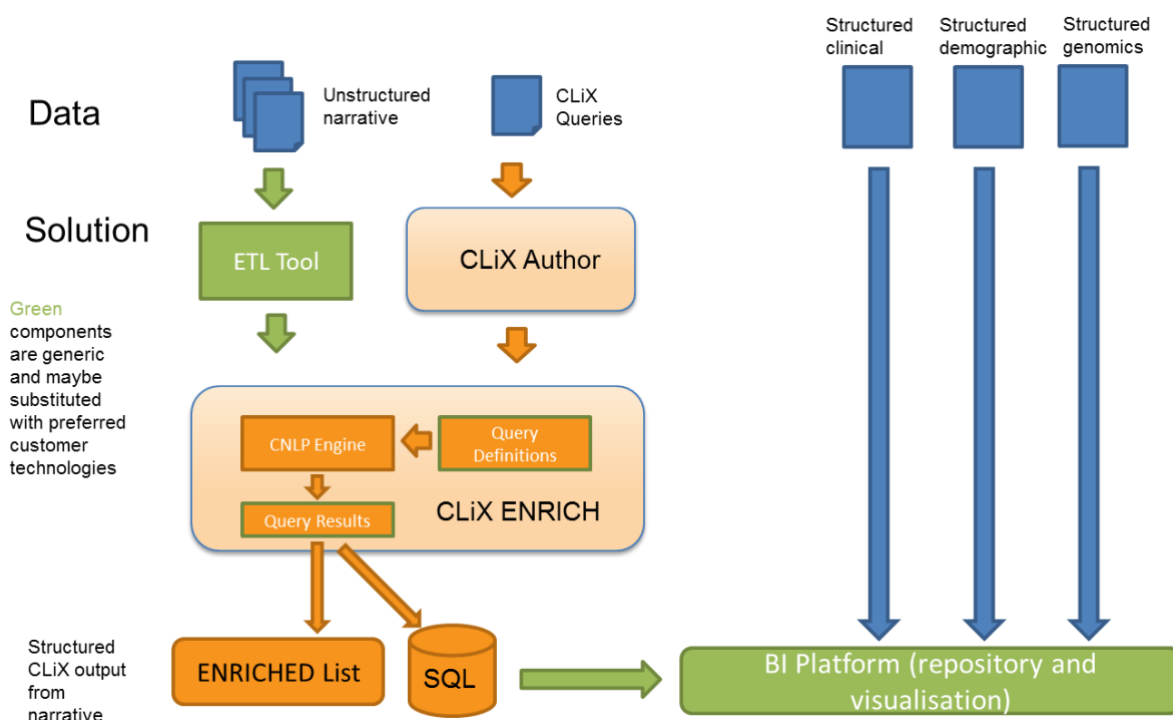


Figure 1. CLiX ENRICH for Clinical Trials solution – high level architecture.

CLiX ENRICH for Clinical Trials enables the combination of pre-existing structured data sources (e.g. clinical, demographic or genomic data) already being managed and consumed by healthcare provider organizations with data from unstructured narrative by exposing the unstructured data in a structured normalized form. The solution is composed of a number of components providing differing capabilities fully integrated to resolve a business problem such as those set out above.

An extract, transform and load (ETL) tool is used to format the narrative into the standard input format supported by CLiX ENRICH *for Clinical Trials* (HL7 2.3 or CSV). This provides both the narrative content to be processed by the CLiX Engine as well as the patient identifier necessary to link back to the structured data.

Once loaded, the narrative can be processed and analyzed and the relevant workflows managed from a simple and easy to use dashboard shown in Figure 2 below.

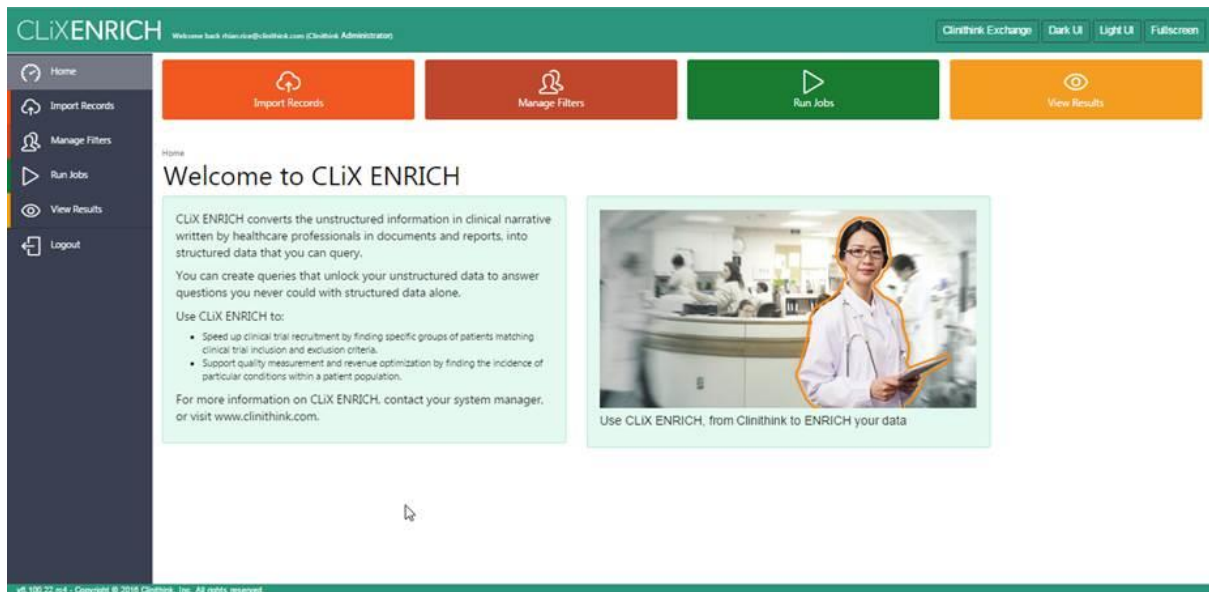


Figure 2. CLiX ENRICH Dashboard

Figure 3 below shows the CLiX Notes tool, demonstrating encoding of narrative data into structured information.

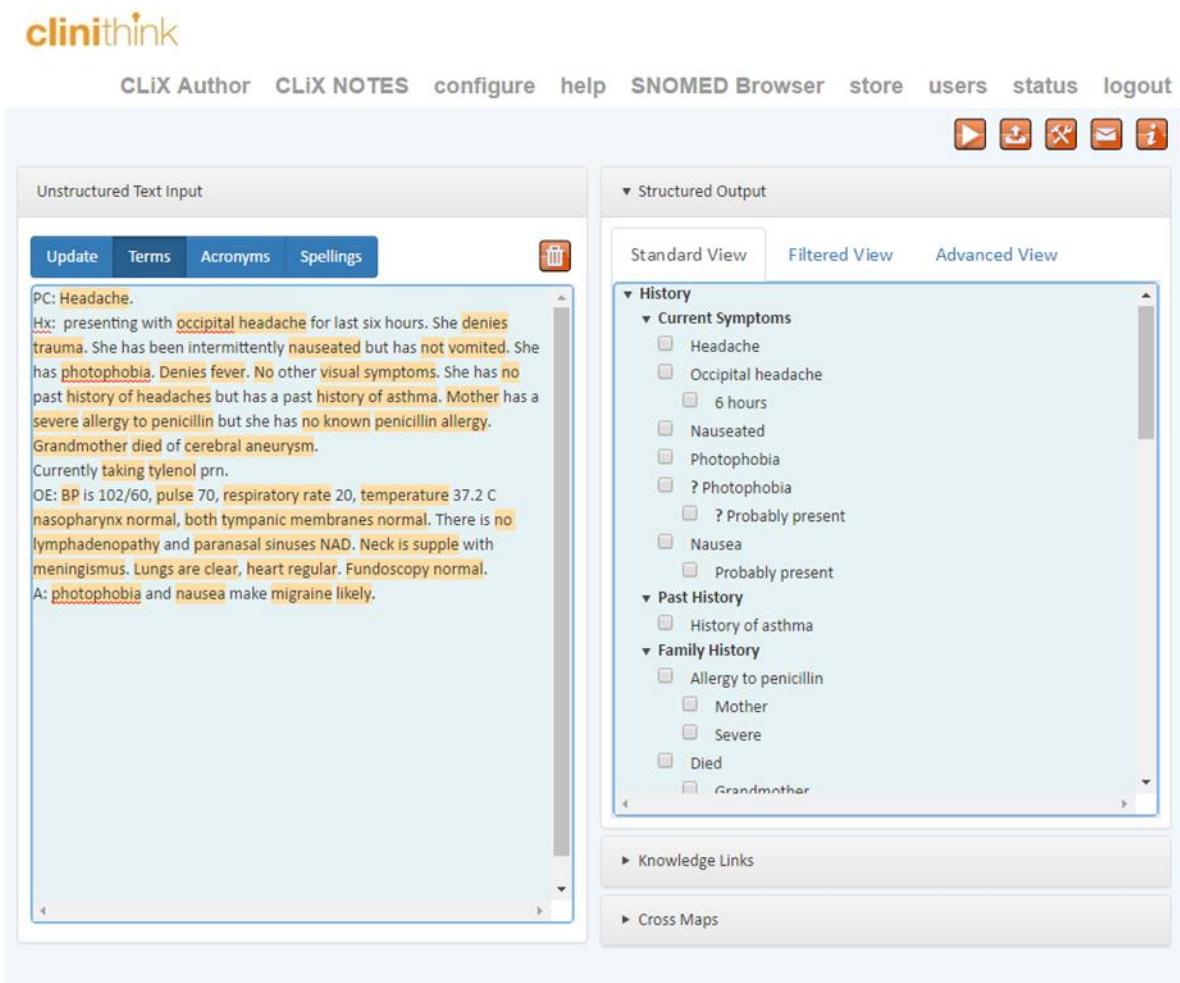


Figure 3. CLiX Notes showing encoding of narrative.

The CLiX Engine produces structured output from the narrative that is of a standard format based on SNOMED CT. CLiX Author is used to write CLiX Queries which are queries to interrogate the structured SNOMED output. CLiX Queries can be authored using both an intuitive graphical user interface for business users as well as a script-based interface for power/technical users. In the clinical trials context, CLiX Author, shown in Figure 4 below, is used to create CLiX Queries that correspond to protocol inclusion and exclusion eligibility criteria.

Home > Author > procedure > Queries

procedure NEW

Author Test

Filter results

Asthma Q

Asthma Compound C

Search Text

☐ Hide descendants

Focus Concept INCISION OF APPENDIX

Advanced

Temporal context Current or past BROWSE

Subject relationship context Subject of record BROWSE

Procedure context Done BROWSE

Name Incision of appendix

Comments

SUBMIT

Figure 4. CLiX Author

Customers that have extensive 3rd party BI tooling can export the output from CLiX ENRICH for further analysis. However, CLiX ENRICH offers powerful concept frequency outputs that are helpful in quickly exploring and understanding a new narrative data set as shown in Figure 5 below.

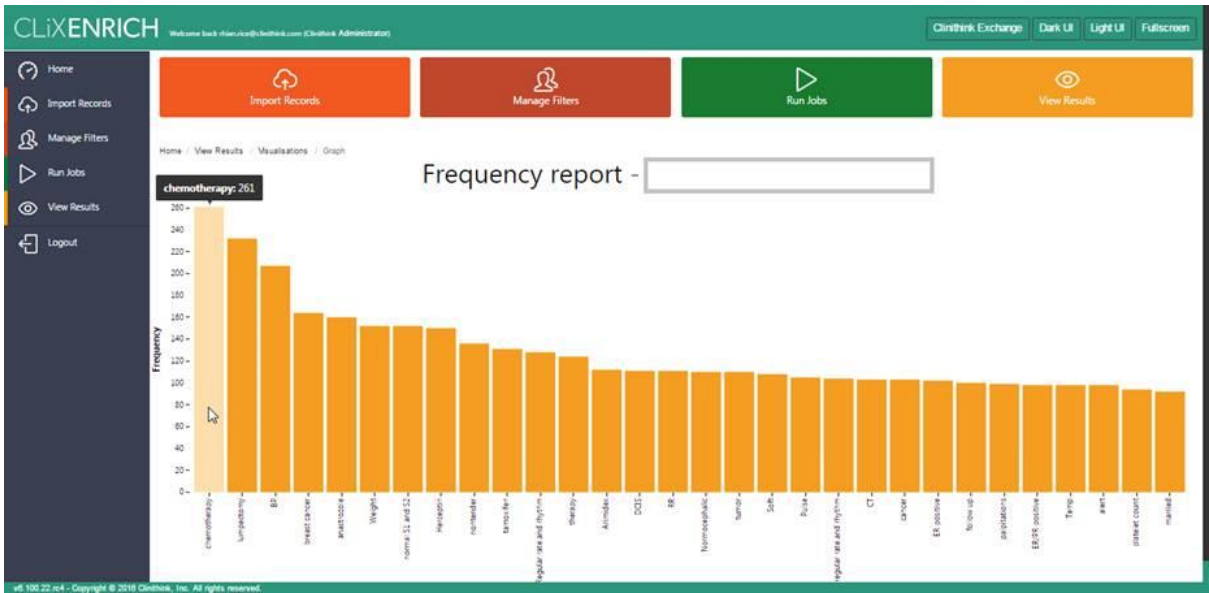


Figure 5. CLiX ENRICH concept frequency output

In addition to exporting the output from the solution, CLiX ENRICH for Clinical Trials enables users to work directly with ENRICHed lists – lists of patients generated locally for review by trial investigators and ranked in order of likely eligibility, as shown in Figure 6 below.

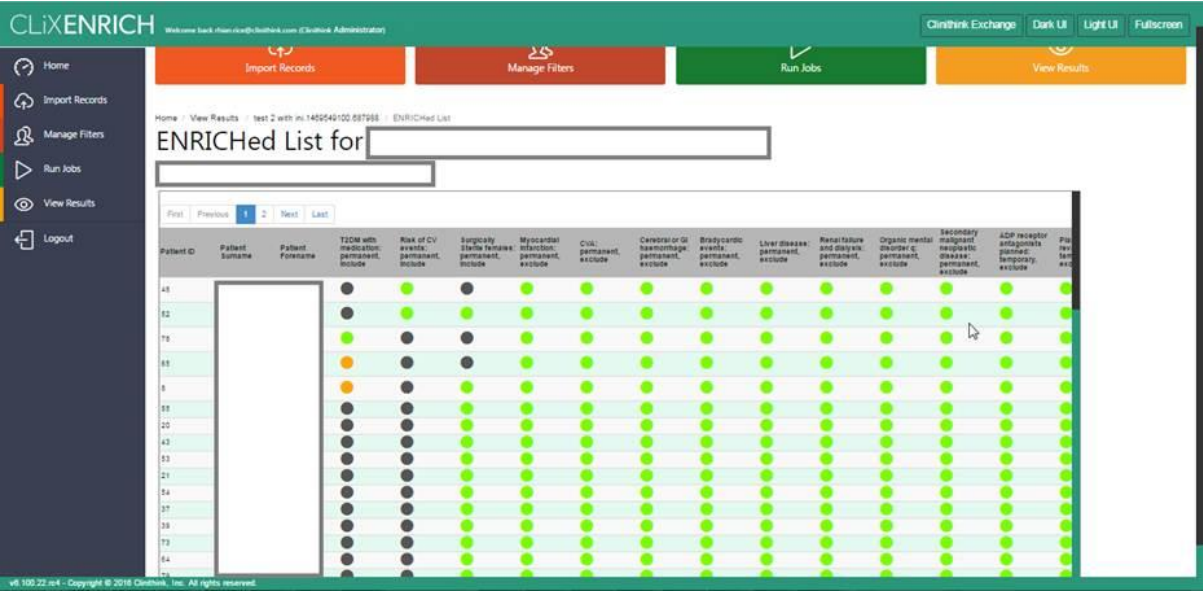


Figure 6. CLiX ENRICH output: ENRICHed List

Solution use cases and benefits

A number of use cases are enabled by the CLiX ENRICH *for Clinical Trials* solution.

Clinical trial protocol optimization

At the clinical trial planning stage, feasibility work is undertaken to understand which clinical criteria should be used to define subjects for the trial. If the criteria define a cohort that is too specific, the power of the study to detect an effect of the drug/intervention is likely to be high but the number of recruitable patients too low. Conversely, if the trial criteria are insufficiently specific, the recruitment target will likely be achievable but the study power will be limited, reducing the chances of detecting an effect. The ability to model the impact that altering trial criteria has on the pool of recruitable patients before the trial protocol is frozen is currently limited by the fact a proportion of those criteria rely on data features that are only described in narrative.

Benefit

Incorporating output from CLiX ENRICH *for Clinical Trials* to add information from the unstructured data enables more granular evaluation of potential eligibility and improves the effectiveness of protocol optimization significantly. More effective optimization in turn reduces the risk of

- a) protocol re-design mid-trial, a very expensive and time consuming impact
- b) failing to detect an effect due to insufficient study power.

Feasibility studies

Traditionally, when assessing how many patients are likely eligible subjects against a trial protocol, ad hoc and anecdotal data is used, rendering feasibility estimates highly inaccurate. If the estimates were optimistic, cost and time is consumed fruitlessly in attempting to meet an unachievable target committed to the trial sponsor.

Benefit

Using CLiX ENRICH *for Clinical Trials*, this risk is removed, a much more precise estimate of likely eligible candidates can be produced very quickly. This enables healthcare provider organizations to become a more competitive and efficient partner for sponsors, while avoiding the risk of bidding for trials for which adequate numbers of suitable candidates do not actually exist.

Accelerating subject recruitment

Once a trial protocol has been frozen with defined inclusion and exclusion criteria and the likely number of eligible candidates assessed, enrollment begins with pre-screening – the identification of patients who are broadly eligible against the criteria. Some criteria (such as mental competence to consent) must be assessed face to face at screening. Most of the criteria assessed at pre-screening are extremely granular and not described in available structured data sources. Instead, they must be assessed through manual chart review. This is an extensive, time-consuming, labor-intensive process, resulting in a decision to invite a candidate to a screening clinic or process. Even then, many clinically eligible patients may decline the opportunity to become subjects in the trial.

Benefit

Using CLiX ENRICH *for Clinical Trials* enables research organizations to save time and money in rapidly finding eligible candidates for subject recruitment. Typically, the solution will deliver much higher yields of eligible patients than is possible using current manual techniques and will do so in a fraction of the time needed to undertake manual chart review. The relevant ENRICHed List is output by the solution in rank order of eligibility, enabling investigators to focus on more eligible candidates first. Since this activity may be conducted on behalf of an external sponsor, there may be revenue and/or performance goals for customers that are more likely to be achieved as a result of leveraging CLiX ENRICH *for Clinical Trials*.

Delivering new data services and tools

Healthcare organizations typically use a range of tools to support both internal and external analytics activities. Currently these relate predominantly to structured data.

Benefit

Leveraging CLiX ENRICH *for Clinical Trials'* output enables the incorporation of unstructured data management and analysis capability into existing and new tools and services. This creates additional value and utility for users, internal or external.