Data modelling methods in clinical trials: Experiences from the CTMND project (ctmnd.org)

Athanasios Anastasiou, Emmanuel Ifeachor, John Zajicek & the CTMND Consortium

University of Plymouth

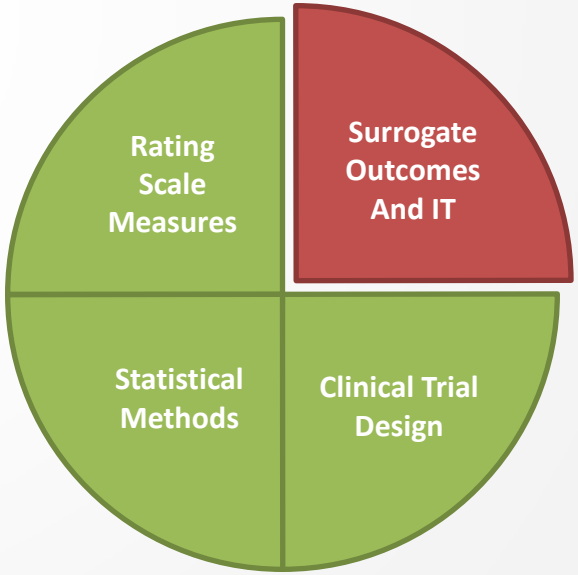
Peninsula College of Medicine & Dentistry

Outline Of The Talk

- The CTMND Project
- Requirements
- State of the art data models
- The dual model approach and openEHR
- Implementation Details
- Q&A

Clinical Trial Methods In Neurodegenerative Diseases

"Improving Treatment Evaluations And Management For People With Neurodegenerative Diseases"



Funded by the UK - N.I.H.R

CTMND
Surrogate Outcomes And IT

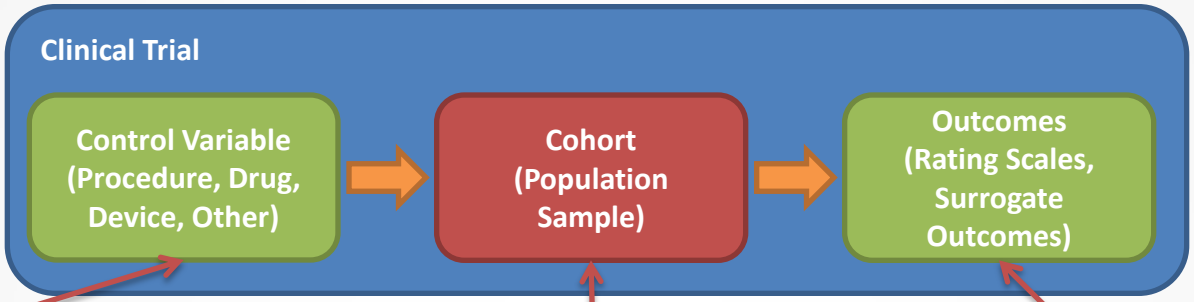
- What Are We Designing For?

"An Online Data Collection And Analysis System To Facilitate Clinical Trials And Relevant Research Taking Into Account, Wherever Possible, Routinely Collected NHS Data."

Requirements

Requirement	Description	Typical Use Case
Patient Registry	A record of subjects interested in participating to clinical trials.	Study Feasibility, Recruitment to clinical trials, local subject population statistics.
Clinical Data Management	Manage the data generated by (potentially) multicentre clinical trials .	Clinical Trial Data Collection
Novel Outcome Description	Describe and manage clinical trial outcomes commonly employed in neurodegenerative disease studies, including surrogate data if required as well as novel outcomes.	Clinical Trial Data Collection
Communicate With External Data Sources	Communicate with external data sources. Perform record linkage (wherever this is possible)	Epidemiological research, subject access to healthcare services for service planning applications and mining routinely collected NHS data (e.g. pharmacy data) for subjects with specific patterns associated with Neurodegenerative Diseases (ND)

Seeking A Future Proof & Flexible Data Model



Which Control Variable?
Which Subject Received it?
Who administered it? (Research Staff)
When was it administered?
How much of it was administered? (Quantity)
How was it administered? (Protocol)
What was the unit?
Any Additional Notes?

How was the population selected?
Where from?
Where are the subjects?
Where will each subject go?
Which other studies does this subject participates in

Which Outcome(s)? (Exact term)
From Which Subject?
Who collected it?
When was it collected?
Who collected it?
How was it collected? (Protocol)
Where is the original dataset?
What procedure (if any) was used to derive the marker? (Computational)
What are its units?

How will these data be organised in the CTMND system?

State Of The Art Data Models

- CDISC – ODM
  - A single model for archival and interchange
  - Expressed in a simple schema
  - Limited provision for semantics in describing and re-using novel parameters / forms

- HL7 and BRIDG
  - CDISC + HL7 = BRIDG.
    - CDISC's BRIDG model expressed through HL7
  - A model for the interchange of "messages" (transactions between systems)
  - Expressed in a multitude of schemas
  - An extensive and complex model requiring close collaboration with committees to define new structures.

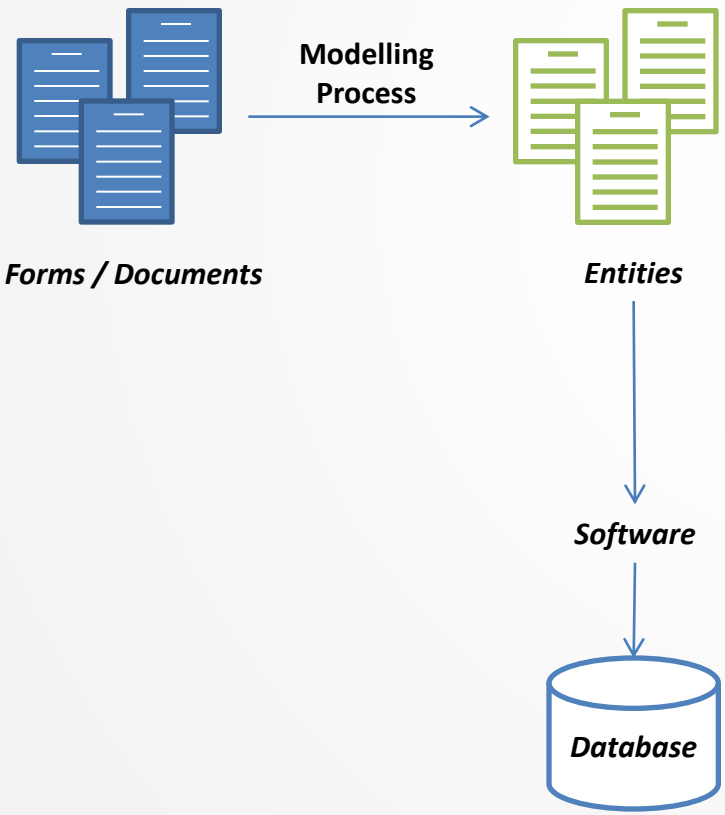
- openEHR
  - A dual model that is part of a health computing platform
  - Expressed in a small set of schemas
  - A very flexible approach to describe novel data requirements and re-use them.
  - Open participation, freely available standard specifications and tools

The Dual Model Approach & openEHR

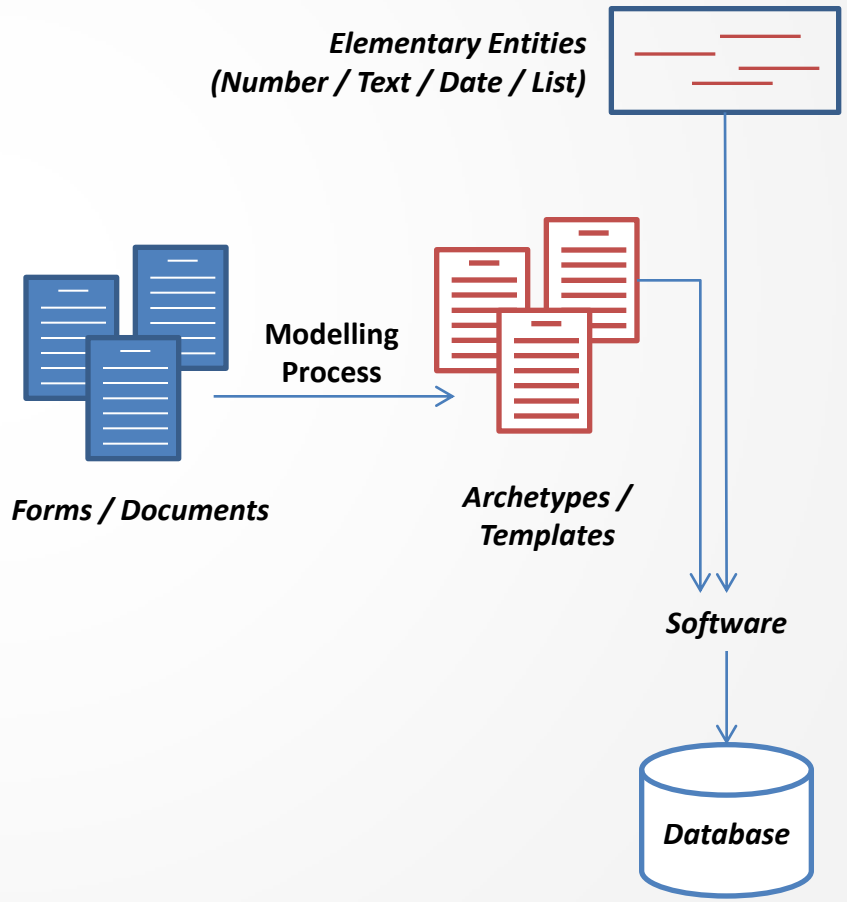
- Motivation
  - Similarities between clinical practice and clinical trials
    - Both have a requirement for an electronic health record.
    - A time stamped collection of data
    - But also differences…
      - Observation / Intervention Encodings (!)
  - Clinical Trials
    - Accurate documentation of all aspects of data collection
      - To avoid systematic errors / miscommunication errors / garbage input
    - Dynamic Environment
  - CTMND Project
    - Description of novel outcomes
      - Still in debate…

The Dual Model Approach & openEHR

- Single Model



- Dual Model



The Dual Model Approach & openEHR

- Archetypes
  - Are used to define new data entries
    - E.g. Blood Pressure
  - Complete Description
    - Terminology (with provision for different languages)
    - Encodings
    - Unit / Nominal Range and other data

- Templates
  - Are used to define compositions of Archetypes
    - Correspond to ‘Forms’

The Dual Model Approach & openEHR

- Archetypes & Templates
  - Can have different versions
  - Can be extended
    - E.g. Substance Use
      - Alcohol Consumption
      - Caffeine Consumption
      - Tobacco Use
  - Can be re-used
    - Within a system
    - Across systems
  - Are openly debated and structured

- An extensive collection of Archetypes is already available from the openEHR Clinical Knowledge Manager.

The Dual Model Approach & openEHR

- Tools & Resources
  - Archetype Editing
    - ADL Workbench
    - LinkEHR Archetype Editor
    - openEHR Clinical Knowledge Manager
  - openEHR Infrastructure
    - openEHR Reference Implementation
      - Java, Python, Eiffel, .Net
    - Opereffa Project
    - LinkEHR Normalisation Platform
    - oceanEHR (Commercial)

The Dual Model Approach and openEHR

- Complete definition of data structures for interventions and outcomes and re-use at similar settings
  - Parameters, Forms (Collections of parameters)
  - Specific terminology and encodings
- Isolation Of The Domain Specialists / Software Developers
  - Minimal Maintenance
- Data Exchange Capability
  - Detailed descriptions of the instruments that were used in a specific clinical trial in an unambiguous way.
  - Other researchers can replicate specific aspects of a trial exactly
- Enables the integration of clinical trial data with a subject's electronic health record
- The definitions of those data structures will form part of the documentation of novel outcomes

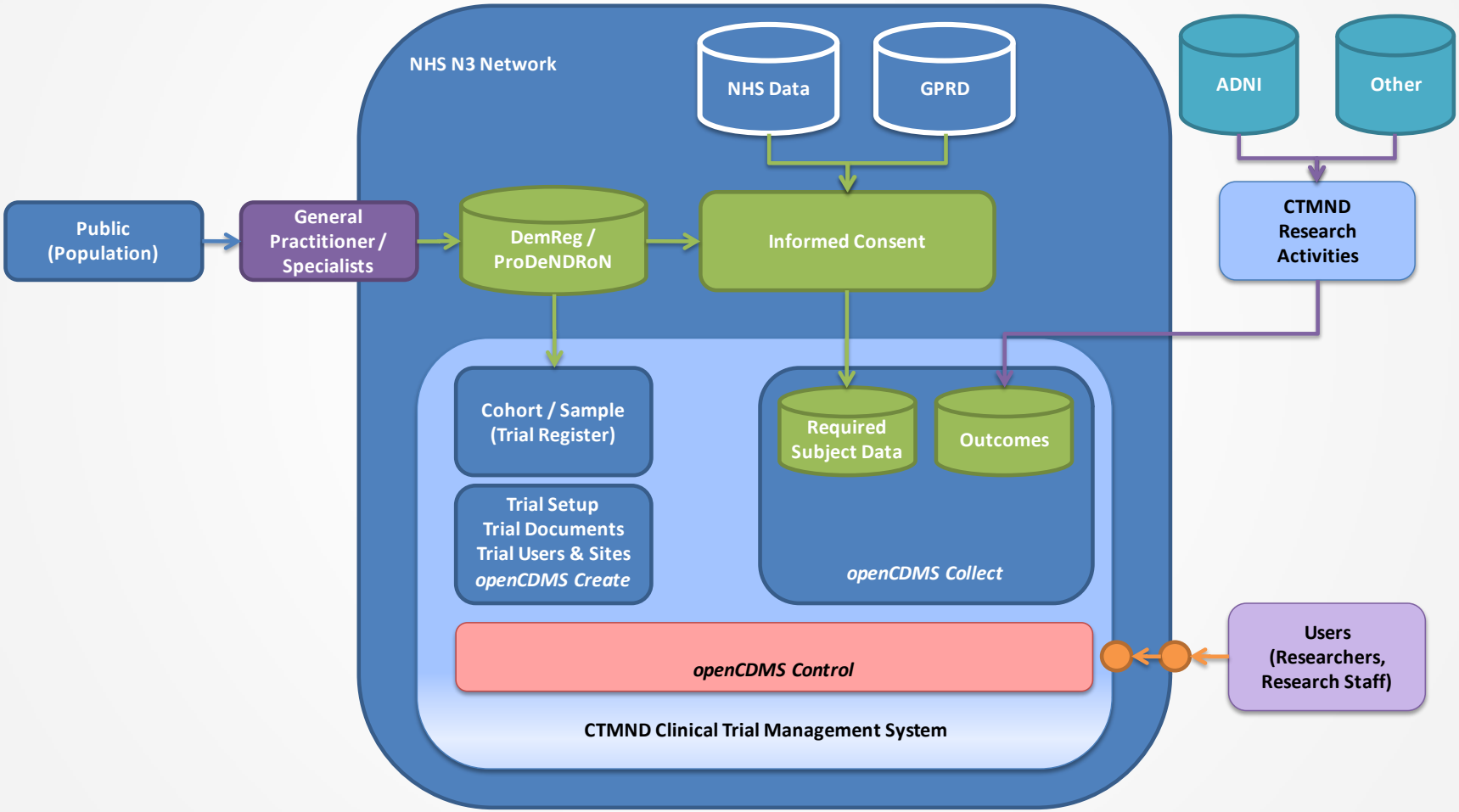
Previous Work

- Christian Knoll, Sebastian Garde, Peter Knaup, *Facilitating Secondary Use of Medical Data by Using openEHR Archetypes*, Proc of MIE2008
- Use openEHR as the data model for clinical trial data, re-use existing medical data for clinical trials
- openSDMS prototype
- Defined a set of administration Archetypes, re-used clinical data Archetypes
  - Covering: Complete Clinical Trial Definition, Data Collection Templates

Current Work

- Review and model existing neurodegenerative diseases scales and outcomes (ongoing)
- Model the novel outcomes currently in production by the project
  - As openEHR Archetypes
- Produce a clinical trial data management system employing these concepts.
  - Currently to be based on openCDMS

Implementation Details



Concluding Remarks

- The CTMND project's research in novel scales and outcomes required a flexible solution to collect and process them electronically.
- The dual modelling approach provides certain attractive benefits to the clinical trials domain
- openEHR offers a flexible electronic health records platform that is also suitable for the clinical trials domain
- Within the CTMND project, work is currently underway to implement a clinical trials management system based on openEHR

Q & A