MID3 Thoughtflow

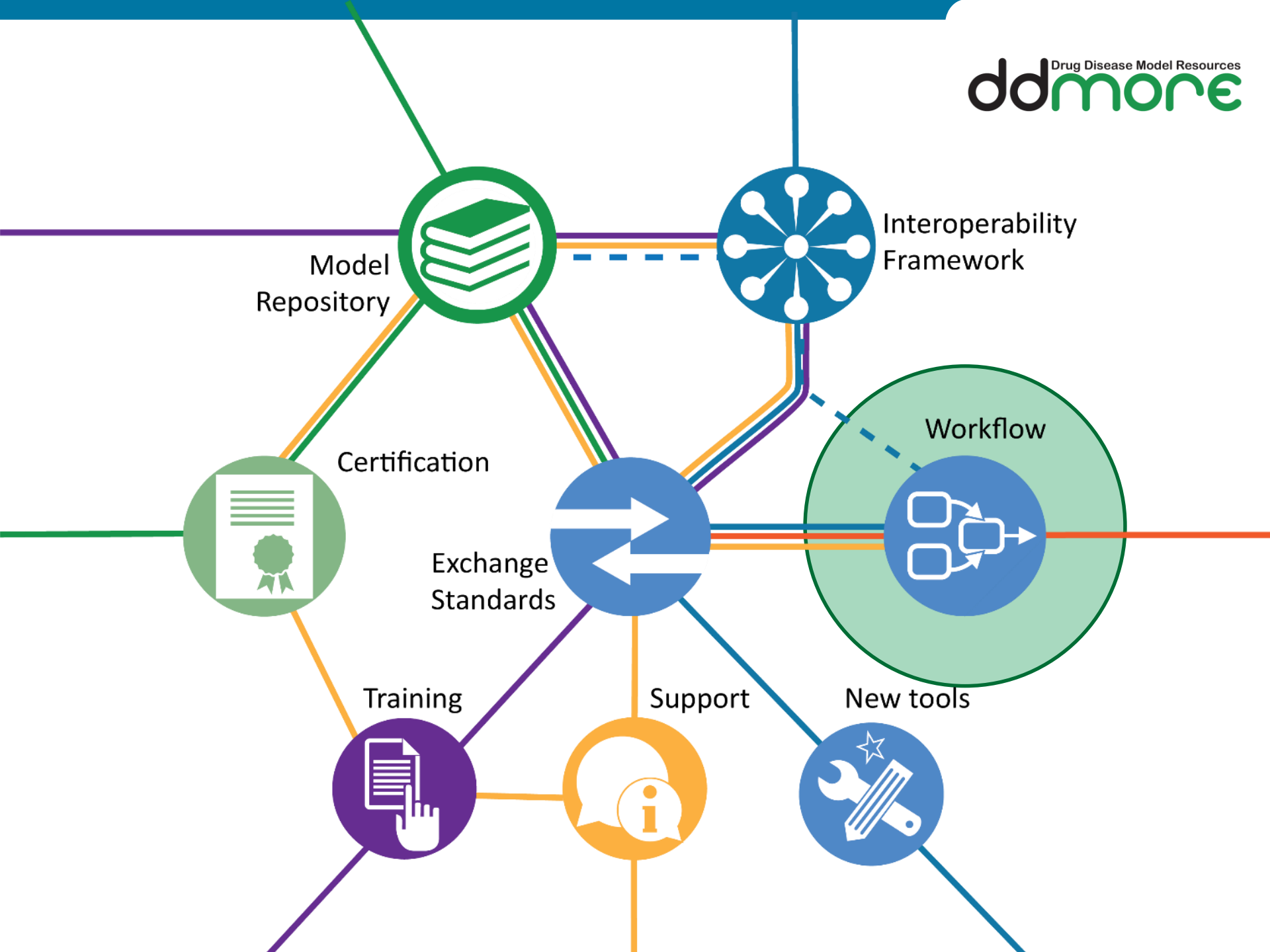
Standards and Tools

Workshop PAGE 2016

Presented by:

Justin Wilkins

On behalf of the DDMoRe consortium


Drug Disease Model Resources
ddmore
Model Repository
Interoperability Framework
Workflow
Certification
Exchange Standards
Training
Support
New tools

Overview

- **Background on Model Informed Drug Discovery and Development (MID3)**
- **Motivation for the Thoughtflow standards and tools**
  - Existing tools and missing components
- **MID3 Thoughtflow standards and tools**
  - Potential benefits
  - As standards and as tools
  - A PROV-O primer
- **A brief technology demonstration**

Model-Informed Drug Discovery and Development (MID3)

Citation: CPT Pharmacometrics Syst. Pharmacol. (2016) 5, 93–122; doi:10.1002/psp4.12049

WHITE PAPER

Good Practices in Model-Informed Drug Discovery and Development: Practice, Application, and Documentation

EFPIA MID3 Workgroup: SF Marshall[1]*, R Burghaus[2], V Cosson[3], SYA Cheung[4], M Chenel[5], O DellaPasqua[6], N Frey[3], B Hamrén[7], L Harnisch[1], F Ivanow[8], T Kerbusch[9], J Lippert[2], PA Milligan[1], S Rohou[10], A Staab[11], JL Steimer[12], C Tornøe[13] and SAG Visser[14]

This document was developed to enable greater consistency in the practice, application, and documentation of Model-Informed Drug Discovery and Development (MID3) across the pharmaceutical industry. A collection of "good practice" recommendations are assembled here in order to minimize the heterogeneity in both the quality and content of MID3 implementation and documentation. The three major objectives of this white paper are to: i) inform company decision makers how the strategic integration of MID3 can benefit R&D efficiency; ii) provide MID3 analysts with sufficient material to enhance the planning, rigor, and consistency of the application of MID3; and iii) provide regulatory authorities with substrate to develop MID3 related and/or MID3 enabled guidelines.

CPT Pharmacometrics Syst. Pharmacol. (2016) **5**, 93–122; doi:10.1002/psp4.12049; published online 14 March 2016.

Acknowledgements:
Efthymios Manolis (EMA/MSWG)
Terry Shepard (MHRA/MSWG))
Ine Skottheim-Rusten (NMA/MSWG/PDCO)

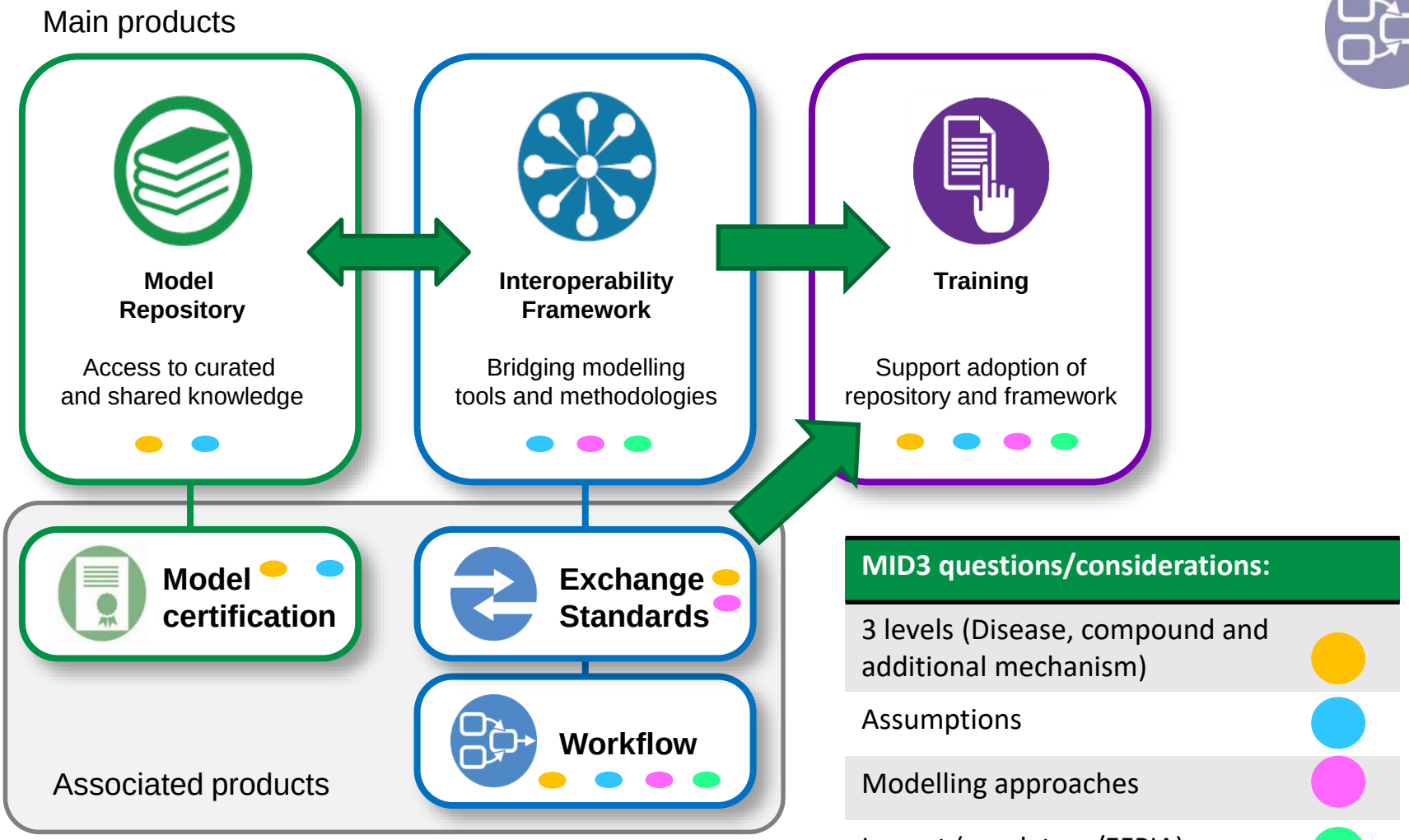
You can listen to the podcast on:
http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)2163-8306/homepage/podcasts.htm

MID3: Assumptions

Important Assumptions	Justification	New/ Established	Testable/ Not-Testable	Test/Approach to assess impact	Evaluation
Pharmacological assumptions					
Physiological assumptions					
Disease assumptions					
Data assumptions					
Mathematical and statistical assumptions					

Source: MID3 workgroup communication slides

Linking MID3 and DDMoRe



MID3 questions/considerations:
3 levels (Disease, compound and additional mechanism)
Assumptions
Modelling approaches
Impact (regulatory/EFPIA)

Source: MID3 workgroup communication slides

Introducing Thoughtflow

Why Thoughtflow?

Pharmacometric data analysis is a complex and multifactorial task…

Keeping track of all the steps in analyses in a project and coherently documenting them afterwards is challenging, tedious, error-prone, and time-consuming.

Existing tools don't quite measure up

Numerous "workflow"/information management tools already exist,
e.g. Taverna, Knime, Activiti, Kepler, Navigator, Improve, Pipeline Pilot, etc

These tools mostly track input-output for tasks (addresses reproducibility), but do not capture more complex relationships such as assumptions, decisions, external influencers. They also lack version control.

Accurately tracking model development requires capturing relationships between entities and activities in a more useful and comprehensive way than traditional workflow tools.

MID3 Thoughtflow

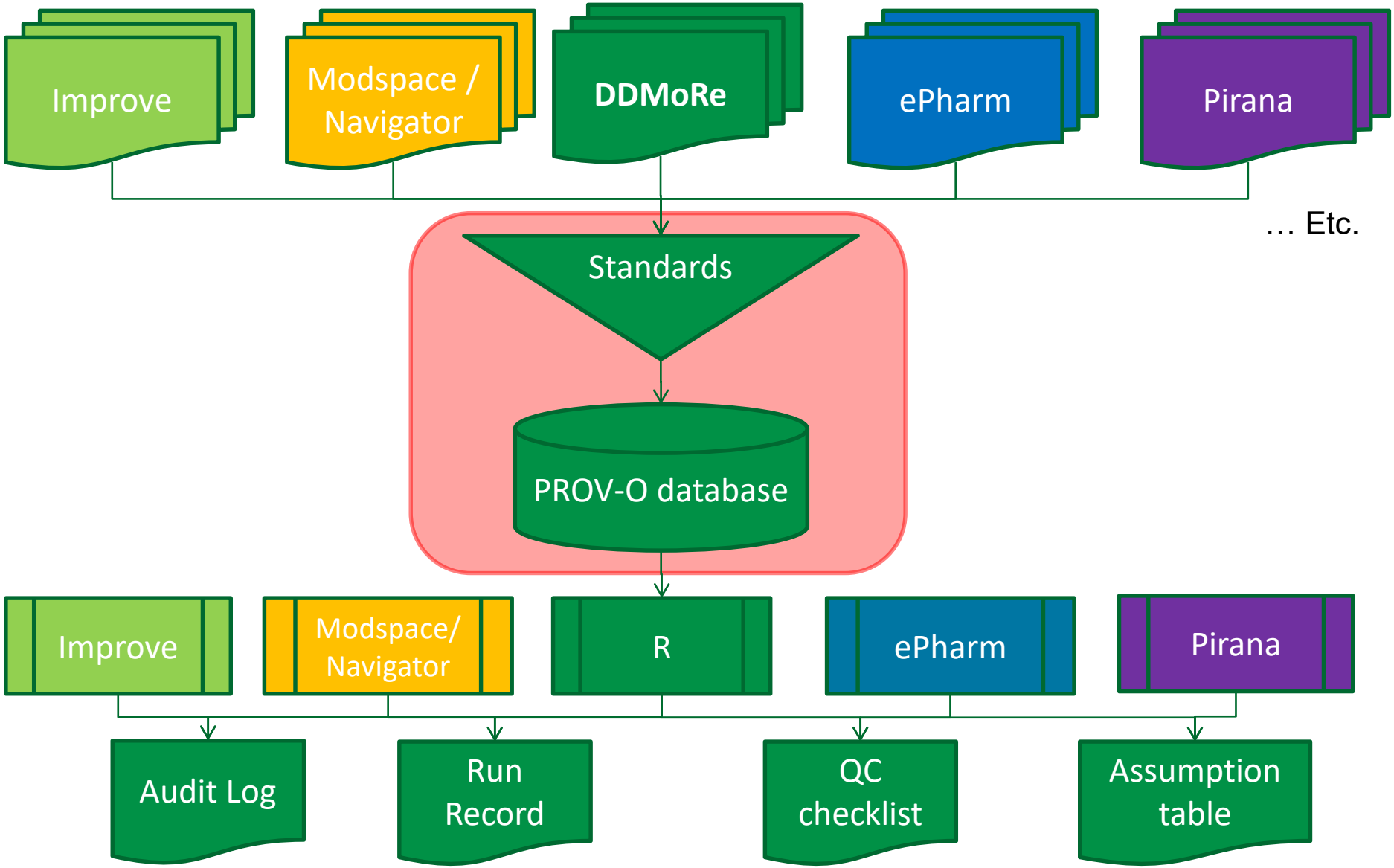
Drug Disease Model Resources

ddmore

What will it be able to do?

- The DDMoRe MID3 Thoughtflow standards and tools are intended to provide cross-platform, robust support for tracking, reporting and replicating elements of a modelling and simulation PKPD project (across all phases of development), either in part or as an entire analysis
- The model development process is summarized in the form of audit logs, run records, QC checklists, and assumption tables
- Such a tool can also support regulatory submission by recording metadata and other source (input) information; modelling and simulation outputs, assumptions, extrapolation, interpretation, decisions and documentation
- Although integrated with the DDMoRe environment and infrastructure, it can exist independently

Diagrammatic view of Thoughtflow



Potential benefits

- ***Analysts***
  - Rapidly generate documentation for analyses in a reproducible and structured manner
    - Audit logs, run records, QC checklists, analysis reports
  - Facilitate knowledge management by providing rich metadata: plans, assumptions, rationale, annotations, decisions and QC/QA pass/fail
    - Makes it easier to pick up/transfer/communicate work from others to review and contribute
  - Avoids the repetition or duplication of prior work, and helps assure quality, traceability, and reproducibility
- ***Managers***
  - Much easier to track analyses, impact on decisions, dependencies, assumptions
    - What work is impactful? What aspects of analyses typically drive decisions? What assumptions underpin inferences? How long do we spend on QC? What are the rate limiting steps?
  - Easily track the progression of the analysis with respect to project timelines, and can thus properly allocate resources to meet potential deadlines
- ***Reviewers/regulators***
  - Assumptions are clearly documented and tied to inputs (data sources), models (which model aspects are underpinned by assumptions) and inferences (what is the consequence of the assumption). Improve transparency and thus enhance clear communication between sponsors and regulators

Both as standards and tools

As standards

MID3 [1] has emerged as a driver for better transparency around provenance, which defines the relationships between entities and activities.

We have translated the concepts, terms and processes set out in the MID3 paper and develop a standard that permits the capture, storage and analysis of those concepts and allows traceability from root concepts such as a problem definition, or question to be answered, through to a decision, recommendation or conclusion.

We have also adapted and expanded the World Wide Web Consortium's PROV-O [2] standard for pharmacometric applications.

ex: Entities (models, datasets, analysis scripts), activities (model estimation), assumptions, decisions
Elements are linked to each other (ex: covariate model derived from base model, decisions made based on model results/diagnostics)

Usefulness of the standards

The MID3 Thoughtflow standards can allow for incorporation into third-party software tools, driving further development for capturing pharmacometric analyses.

Demonstrations of the technology will soon be incorporated into other third-party software tools such as Mango Development Solutions' Navigator and scinteco's Improve

[1] Marshall et al (2016). Good Practices in Model-Informed Drug Discovery and Development: Practice, Application, and Documentation. *CPT:PSP* 5(3):93-122.
[2] http://www.w3.org/TR/prov-o/

PROV-O: The W3C Provenance Ontology

Drug Disease Model Resources

Provenance is information about entities, activities, and people involved in producing a piece of data or thing.

- Can be used to form assessments about its quality, reliability or trustworthiness
- The W3C PROV Family of Documents defines a model, corresponding serializations and other supporting definitions to enable the inter-operable interchange of provenance information in heterogeneous environments such as the Web

https://www.w3.org/TR/prov-o/

Capturing Pharmacometrics Workflow Concepts with PROV-O

PROV-O defines a range of terms that are used to capture all the information necessary to define the provenance of items:

- **Entity:** An entity is a physical, digital, conceptual, or other kind of thing with some fixed aspects; entities may be real or imaginary (e.g. a model, a dataset, an output file, a script, a decision, an assumption)
- **Activity:** An activity is something that occurs over a period of time and acts upon or with entities; it may include consuming, processing, transforming, modifying, relocating, using, or generating entities (e.g. a model execution, a visual predictive check)
- **Agent:** An agent is something that bears some form of responsibility for an activity taking place, for the existence of an entity, or for another agent's activity. Can be an Organisation (e.g. Leiden, Pfizer), a Person (e.g. a User) or a Software Agent (e.g. R, Monolix).

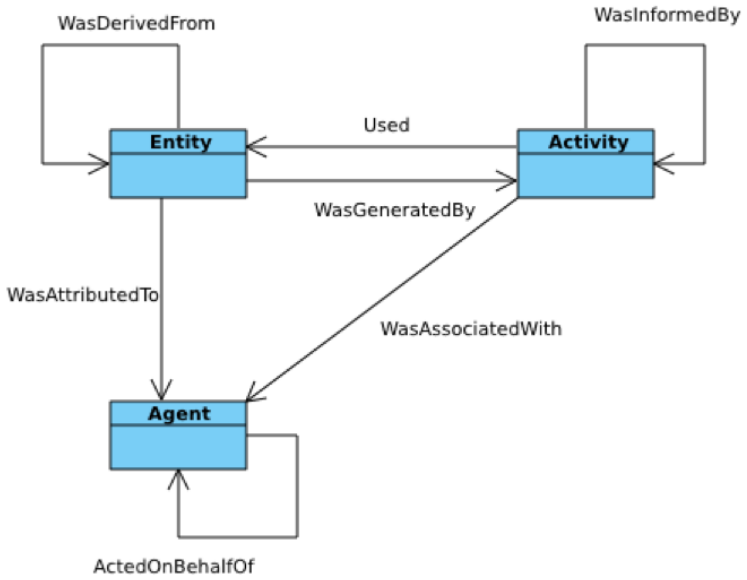
https://www.w3.org/TR/prov-o/

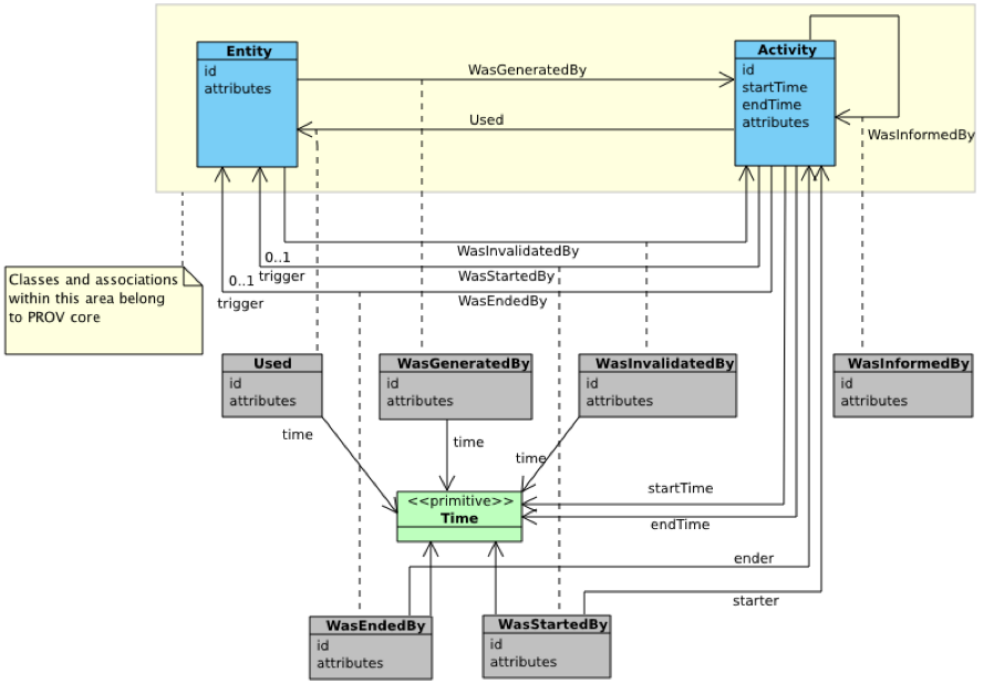
Capturing Pharmacometrics Workflow Concepts with PROV-O

Drug Disease Model Resources

PROV-O defines a range of terms that are used to capture all the information necessary to define the provenance of items:

- **Relationships:** PROV-O defines a set of relationships that describe the interactions between entities, activities and agents.

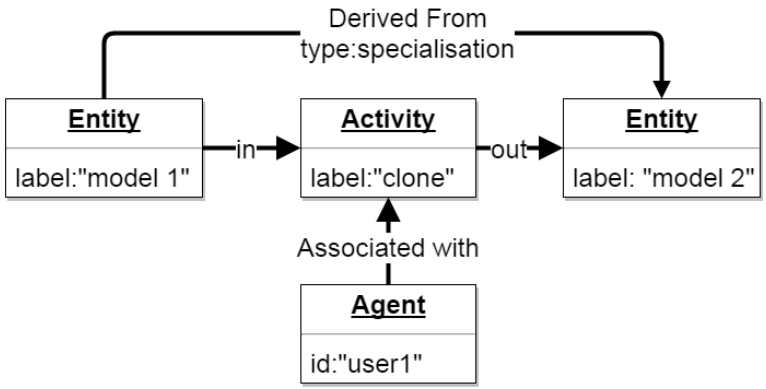




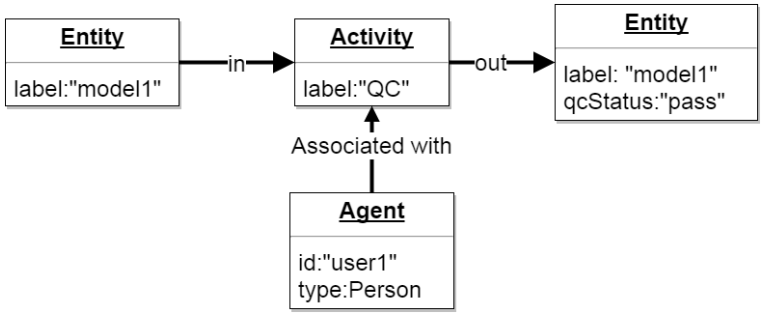
Capturing Pharmacometrics Workflow Concepts with PROV-O

Drug Disease Model Resources

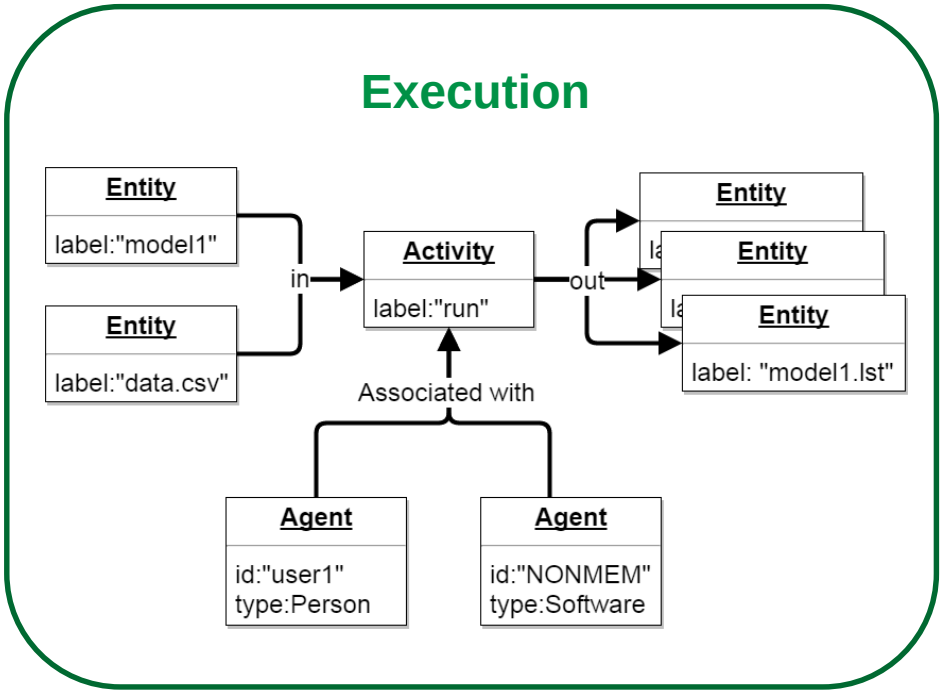
Model derivation



Model characterisation



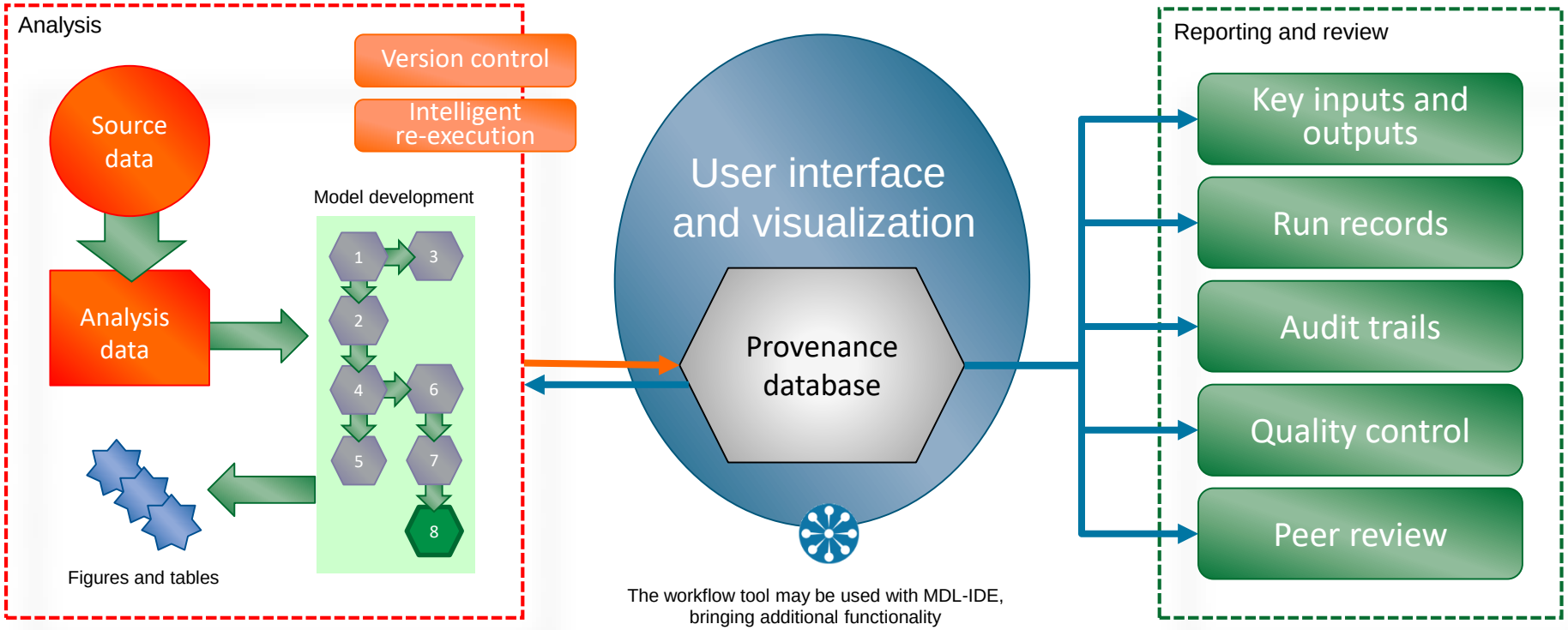
Execution



As a tool

We intend to launch a free, open source software implementation at the end of August 2016.

Although not all features will be available in the first release, we believe it will provide a convincing demonstration of concept and solid foundation for further development.

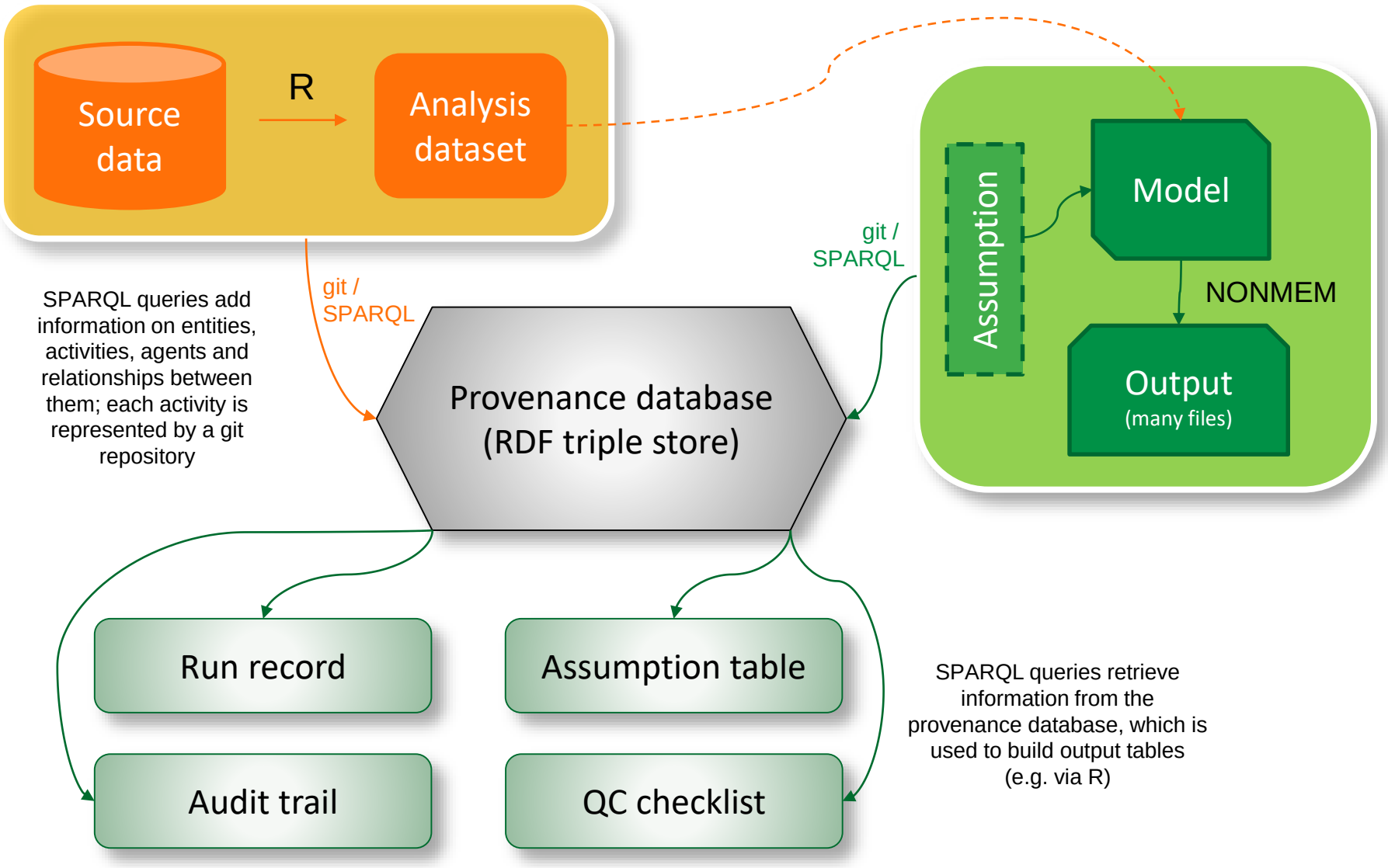


Inputs, scripts, models, assumptions, outputs and decisions are versioned and their provenance is tracked

A lightweight but powerful GUI brings visualization and control

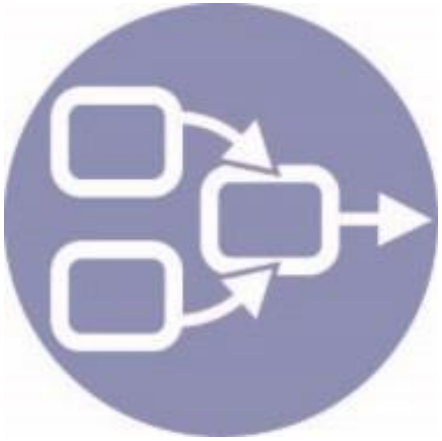
The database can be queried to streamline reporting and review

How it all works – some examples



Overview

- **Background on Model Informed Drug Discovery and Development (MID3)**
- **Motivation for the Thoughtflow standards and tools**
  - Existing tools and missing components
- **MID3 Thoughtflow standards and tools**
  - Potential benefits to analysts, managers, reviewers/regulators
  - As standards and as tools
  - A PROV-O primer
- **A brief technology demonstration**



Thank you!

Phylinda Chan[1], Jonathan Chard[2], S Y Amy Cheung[3],
Andrew Dunn[2], Chris Franklin[4], Richard Kaye[2], Enrique Larios[5],
Stuart Moodie[1], Maria Luisa Sardu[6], Gareth Smith[7],
Mike K Smith[1], Evan Wang[8], Justin Wilkins[1,8]

Pfizer[1], Mango[2], AstraZeneca[3], GSK[4],
Leiden University[5], Merck KGaA[6], Cyprotex[7], Lilly[8]

Discussion

- Do you think you'd be able to use Thoughtflow concepts and an implementation in your daily work? Do you think working this way would improve your life?
- Which features and/or functionality you like the most? Which do you like the least?

Visit our poster!

II-42 Jonathan Chard

Pharmacometrics workflow: standards for provenance capture and workflow definition

Wednesday 15:10-16:30

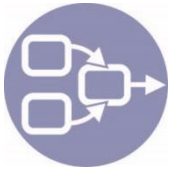
Backup

Drug Disease Model Resources

Core values
MID3 Workflow environment

Value	How DDMoRe delivers this
Traceability	Workflow tool records all individual steps throughout analysis with time stamps and dependencies
Clarity	On-the-fly and post-hoc visualisation of the steps of the analysis and their interdependencies (model “Tree” view)
Audit trail	Workflow tool may be used to generate a complete audit trail for an analysis
Decision tracking	Decisions may be documented at any point in model development, linked to any entity (model, data, output); visualization of decision tree
Facilitated review process	Activities and entities can be marked as Qced – reviewer can see dependent activities and entities and whether these are Qced. Assumptions are linked to data, models / model components are linked to assumptions, assumptions can be invalidated. Easy to see impact of assumptions on inferences (MID3).
One-click re-execution and reproducibility	Upon changes made to input entities (such as datasets), all dependencies may be re-generated using a single click; entire analysis may be repeated in this way
"Lab Book"-like documentation	Workflow tool automatically collects information that can be queried in a reporting, scientific, QC or audit context
Stand-alone installation	Runs independently of other DDMoRe software components, although delivers benefits if these are present

Core features
MID3 Workflow environment



Drug Disease Model Resources
ddmore

Feature	Description
Tracks provenance	Recording of relationships between components of an analysis (scripts, models, outputs, assumptions, decisions, …), minimizing burden on users
Artifact version control	Detects when a component has changed, records the nature of the change, and determines its impact on other components
Updates analysis components to reflect upstream changes ("re-running")	When changes to an analysis component (such as a data file) have been made, all dependencies may be updated at the click of a mouse
Exports run records, QC checklists and audit trails	Run records, QC checklists and audit trails may be assembled and exported in a convenient format for reporting, minimizing analyst need to do this manually
Assembles report components	Key graphs and tables can be prepared and assembled in a convenient location and format for reporting and QC
GUI facilitates project and provenance visualization	A streamlined, portable GUI centralizes all functionality without getting in the way
Independent but synergistic	DDMoRe tools are not required for core functionality, but their availability will bring benefits
Built on standards	Underlying technology is built on widely-adopted standards for defining provenance (PROV-O) and programming tools (Java)
A solid foundation	The tool will be designed to support further development and expansion