Citation: CPT Pharmacometrics Syst. Pharmacol. (2015) 4, 316–319; doi:10.1002/psp4.57

PERSPECTIVE

Pharmacometrics Markup Language (PharmML): Opening New Perspectives for Model Exchange in Drug Development

MJ Swat[1]*, S Moodie[2], SM Wimalaratne[1], NR Kristensen[3], M Lavielle[4], A Mari[5], P Magni[6], MK Smith[7], R Bizzotto[8], L Pasotti[6], E Mezzalana[6], E Comets[8], C Sarr[9], N Terranova[10], E Blaudez[11], P Chan[7], J Chard[12], K Chatel[11], M Chenel[13], D Edwards[14], C Franklin[15], T Giorgino[5], M Glont[1], P Girard[10], P Grenon[16], K Harling[17], AC Hooker[17], R Kaye[12], R Keizer[17], C Kloft[18], JN Kok[19], N Kokash[19], C Laibe[1], C Laveille[13], G Lestini[8], F Mentré[8], A Munafo[10], R Nordgren[17], HB Nyberg[12,17], ZP Parra-Guillen[18], E Plan[17], B Ribba[20], G Smith[21], IF Trocóniz[22], F Yvon[1], PA Milligan[7], L Harnisch[7], M Karlsson[17], H Hermjakob[1] and N Le Novère[1,23]

The lack of a common exchange format for mathematical models in pharmacometrics has been a long-standing problem. Such a format has the potential to increase productivity and analysis quality, simplify the handling of complex workflows, ensure reproducibility of research, and facilitate the reuse of existing model resources. Pharmacometrics Markup Language (PharmML), currently under development by the Drug Disease Model Resources (DDMoRe) consortium, is intended to become an exchange standard in pharmacometrics by providing means to encode models, trial designs, and modeling steps.
CPT Pharmacometrics Syst. Pharmacol. (2015) **4**, 316–319; doi:10.1002/psp4.57; published online on 15 June 2015.

The field of pharmacometrics has progressed at an impressive pace in the past decade. New software tools and methods have become available for estimation tasks as well as for clinical trial simulation and optimal design.[1] The development of new tools, however, brings challenges for the users in terms of integrating them into existing workflows. Typically, this requires manual translation of the underlying pharmacometric model for each tool being used, not only due to differences in model formulation/language but also due to tool capabilities, software-specific methods, and algorithms. Such translation along with the often-needed conversion of associated datasets may introduce errors and takes unnecessary time, as no converters exist. A common exchange format within pharmacometrics, which would reduce the efforts needed to exchange models, is clearly needed.

This was recognized by the partners within the NonLinear Mixed Effects (NLME) Consortium several years ago, and an initial specification for such a format was drafted. Unfortunately, development did not continue beyond the first version. These initial results were not lost altogether and were the starting point when the idea was picked up again in 2011, with the initiation of the Drug Disease Model Resources (DDMoRe) project (http://ddmore.eu)[2] under the European Innovative Medicines Initiative (IMI). DDMoRe aims to "build and maintain a universally applicable, open source, model-based framework, intended as the gold standard for future collaborative drug and disease Modeling & Simulation." An integral part of this framework is the Pharmacometrics Markup Language (PharmML) exchange format. Using PharmML, the DDMoRe framework specifically aims to integrate existing tools (**Figure 1**) such as NONMEM, Monolix, win(open)BUGS, PFIM, PopED, PsN, Xpose, SIMCYP Simulator, MatLab, R, as well as new tools developed within DDMoRe, e.g., Simulx and infix2pharmml (**Table 1** in **Supplemental Material**). Another major part of DDMoRe, the Model Repository (http://repository.ddmore.eu), is the place where PharmML-coded models can be stored, retrieved, and shared with the community. In this article we discuss the scope and structure of PharmML v. 0.6, which was released publicly in January 2015 as the second public release (see related websites http://ddmore.eu/pharmml and http://www.pharmml.org for more information).

MOTIVATION

The current situation in pharmacometrics, specifically the lack of a common exchange format, resembles that in other

[1]EMBL-European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, UK; [2]Eight Pillars Ltd, Edinburgh, UK; [3]Novo Nordisk A/S, Bagsværd, Denmark; [4]Inria Saclay, Paris, France; [5]National Research Council, Institute of Biomedical Engineering, Padova, Italy; [6]Dipartimento di Ingegneria Industriale e dell'Informazione, Università degli Studi di Pavia, Pavia, Italy; [7]Global Clinical Pharmacology, Pfizer, Sandwich, UK; [8]INSERM, IAME, UMR 1137, Paris, France, University Paris Diderot, IAME, UMR 1137, Sorbonne Paris Cité, Paris, France; [9]Advanced Quantitative Sciences (AQS), Novartis, Basel, Switzerland; [10]Merck Institute for Pharmacometrics, Merck Serono, Lausanne, Switzerland; [11]Lixoft, Orsay, France; [12]Mango Solutions, Chippenham, Wiltshire, UK; [13]SGS Exprimo NV, Mechelen, Belgium, Clinical Pharmacokinetics and Pharmacometrics, Institut de Recherches Internationales Servier, Suresnes, France; [14]Simcyp (a Certara company), Sheffield, UK; [15]CPMS Technology and Development, Southall, UK; [16]CHIME, University College London, London, UK; [17]Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden; [18]Freie Universtitaet Berlin, Germany, Institute of Pharmacy, Department of Clinical Pharmacy and Biochemistry, Berlin, Germany; [19]Leiden Institute of Advanced Computer Science (LIACS), Leiden University, Leiden, The Netherlands; [20]Inria Grenoble - Rhône-Alpes, Grenoble, France; [21]Scientific Computing Group, Cyprotex Discovery Limited, Macclesfield Crewe, UK; [22]Department of Pharmacy and Pharmaceutical Technology, University of Navarra, Pamplona, Spain; [23]Babraham Institute, Babraham Research Campus, Cambridge, UK. *Correspondence: MJ Swat (mjswat@ebi.ac.uk)
Received 17 November 2014; accepted 6 May 2015; published online on 15 June 2015. doi:10.1002/psp4.57

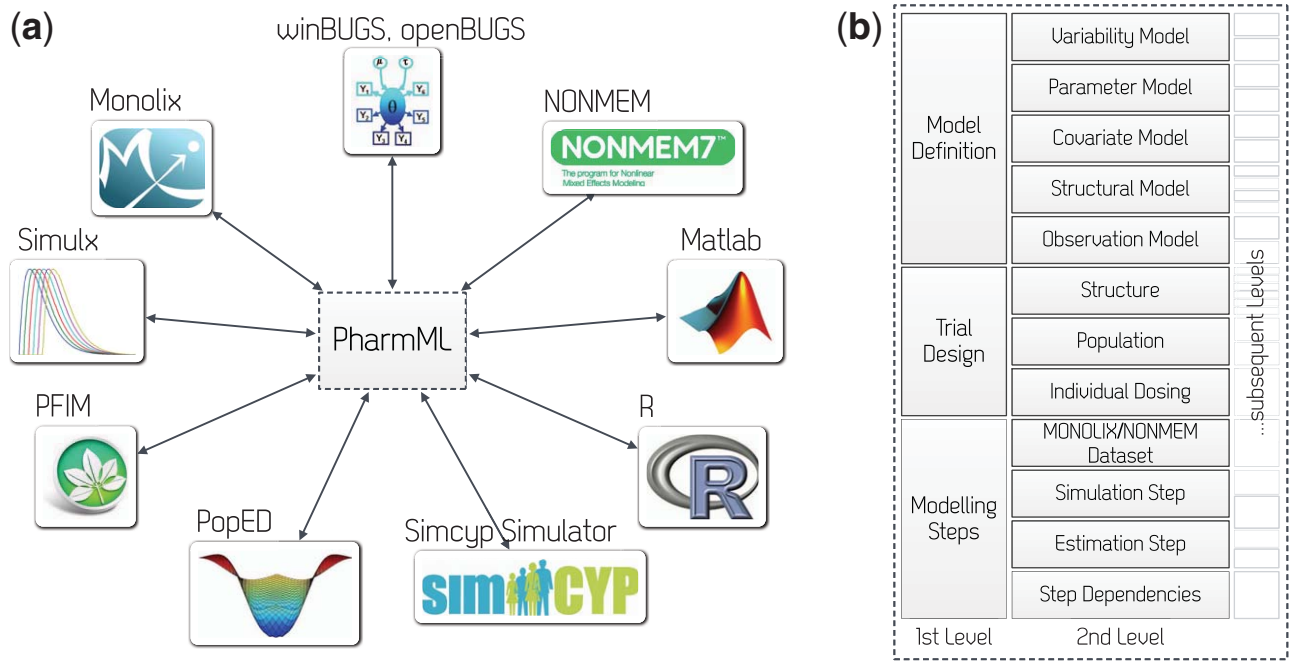

Figure 1 (a) PharmML as lingua franca for the DDMoRe platform and its target tools. **(b)** The basic structure of PharmML with the first two layers shown. The first one consists of Model Definition, Trial Design, and Modeling Steps; the second has a finer-grained structure with submodels or other specialized elements. For more details see text and **Figure 3** (**Supplemental Material**).

areas of life sciences 10–15 years ago, most notably within systems biology and neurosciences. There, standards such as CellML, SBML, and NeuroML have been developed to encode models of biochemical, cellular, and multicellular processes (**Table 2** in **Supplemental Material**). These standards have transformed the corresponding areas of science. Many new tools have appeared after their establishment and model exchange has become much easier.[3]

Developing a similar, powerful exchange format for pharmacometrics is challenging, because the tools available in this area use a variety of approaches for model encoding. For example, the most popular tool in the field, NONMEM with the NMTRAN language for model specification,[4] allows users to encode virtually any conceivable modeling scenario in an assignment-based style, which gives great flexibility but also makes standardization difficult. On the other hand, Monolix with the MLXTRAN language for model specification[5] uses a declarative style with a clearly defined vocabulary, grammar, and clear language boundaries. Ensuring that models formulated using both approaches can be implemented in an exchange format is very demanding and requires an adaptable structure. The different data formats or data file layouts being used add to the complexity of the problem.

The challenge the field is facing is also a consequence of the complexity and scope of pharmacometrics, with models at different scales (from models of intracellular pathways to whole body models) being applied for different purposes (from descriptive models to clinical trial simulation models) to address a wide range of problems in drug development. Tool support for PharmML is a demanding engineering task because it requires good understanding of both the pharmacometrics and computational science.

Despite the complexity of the task, we firmly believe that we are able to make the daily work for pharmacometricians easier by alleviating the burden of translating models and converting datasets. Most important, as recoding is a potential source of errors, a model should only have to be encoded once, regardless of how many different tools use it within a given workflow.

To summarize, a common exchange format is expected to facilitate:

- Smooth and error-free transmission of models between tools.
- Use of complex workflows via standardized model and output definitions (**Figure 2**).
- Reproducibility of research.
- Easier reporting and bug tracking.
- Improved interaction with regulatory agencies regarding modeling and simulation.
- Reuse of existing model resources, e.g., BioModels database.[6]
- Development of new tools and methods.
- Expanding the community developing/applying pharmacometric models.

The creation of a tool-independent format for unambiguous model formulation is the key step for the successful achievement of these goals.

PharmML BASICS

PharmML is based on XML (http://www.w3.org/XML), a standard developed by the World Wide Web Consortium.

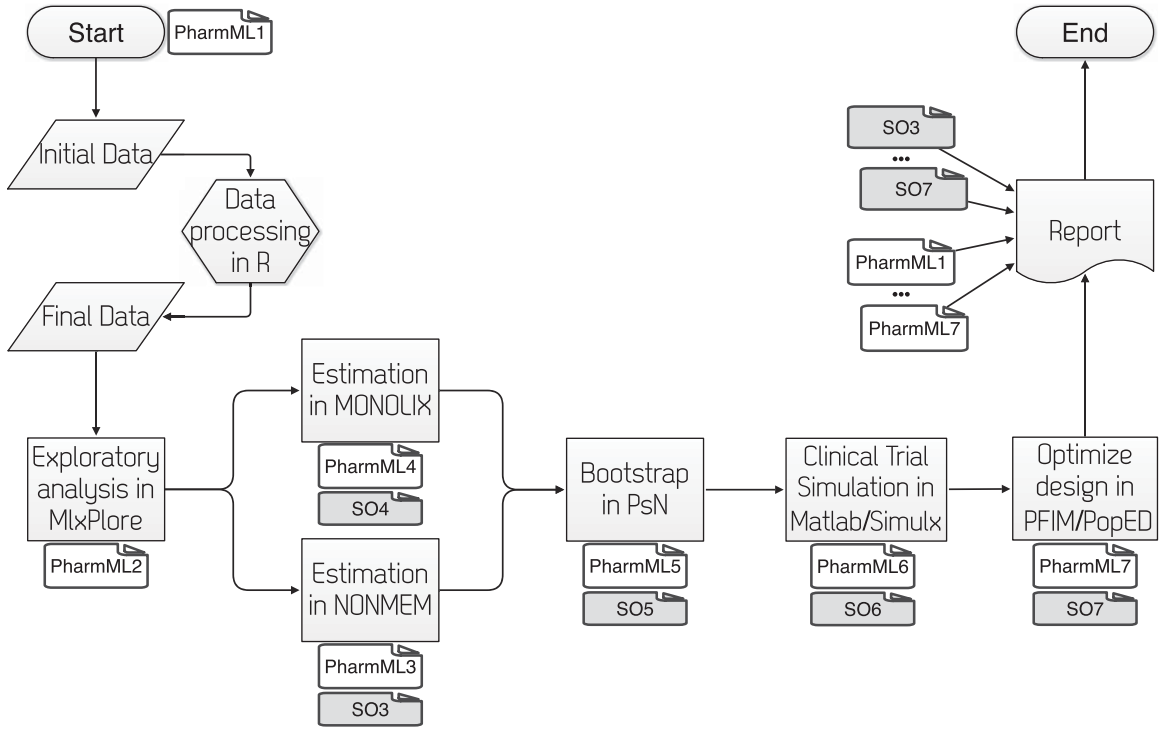

Figure 2 PharmML and Standardized Output (SO) supporting a typical workflow in pharmacometrics featuring major target tools of the DDMoRe platform. Here, it starts with data processing in R, which can consist of data formatting, merging, and/or missing data imputation. After that an explanatory analysis is carried out in MlxPlore, followed by estimation using either Monolix or NONMEM. Subsequent steps are bootstrapping using PsN, clinical trial simulation in MatLab/Simulx, and finally Optimal Design in either PFIM or PopED. At every step of the workflow, the PharmML model can be stored and the results following each step can be recorded in the corresponding SO file. Documenting workflows in such a detailed way can potentially simplify reporting and ensures reproducibility.

XML has a number of distinct features making it very useful for machine-to-machine exchange languages such as PharmML. It is extendable and allows named, meaningful tags to be defined and annotated. It is ideal for interacting with and reusing other XML-based formats. Moreover, there are many tools available to support XML processing and XML schema development. PharmML consists of three elements to encode 1) models, 2) trial designs, and 3) modeling steps (**Figures 1, 3** in **Supplemental Material**).

The **Model Definition**, the core section of PharmML containing five submodels, was developed based on the mathematical formalism of NLME models.[7,8] The *Variability Model*, formulated as a nested hierarchy, describes the parameter and residual error-related variability structures. Any number of variability levels is allowed, each of them fully defined by a covariance matrix. The *Parameter Model* comes with a flexible structure to support a range of possible formulations. The default one is the Gaussian model, which assumes the parameters to be normally distributed up to a transformation and which can include either a linear or a nonlinear covariate model. Alternatively, the parameters can be described using an arbitrary expression. Additionally, the correlation structure for the random effects can be defined pairwise or using different matrix types. The *Covariate Model* describes information about covariate transformation, e.g., allometric scaling, continuous, or discrete distribution and interpolation. The *Structural Model* supports algebraic equations, ordinary differential equations (ODEs) with initial conditions, and delay differential equations (DDEs) with history definition. Pharmacokinetic (PK) models, as the most frequently used models, can also be encoded using *PK macros*, a concept borrowed from MLXTRAN, which allows an equation-free encoding of a vast number of compartment models using predefined macros. The *Observation Model* supports continuous data models with a flexible residual error model as well as different types of discrete data models, e.g., categorical, count, and time-to-event. Here, both declarative and assignment-based encoding styles are also supported.

The **Trial Design** section is based on a CDISC standard[9] and plays a central role in encoding of simulation and optimal design tasks, but it can also be used for estimation tasks. In contrast to the traditional approach, where the trial design is implemented within the dataset, this element permits formulating a study design in a dataset-independent manner. Using only a few basic elements, it is possible to encode complex designs, e.g., crossover trials with multiple arms, epochs, occasions, treatment types, and/or washout events. Within the **Trial Design** element, PharmML also has distinct placeholders for covariates, dosing records, and observations, and, depending on the task, only the relevant records have to be provided.

The third section, **Modeling Steps**, is used to define basic tasks to be performed with the model. Currently, two are supported: estimation and simulation. For estimation, the default option is that all information about the underlying trial design is given implicitly by the dataset (the alternative, **Trial Design,** is in such cases not needed) with appropriate mapping to relevant parts of the **Model Definition**. In addition, initial estimates with or without boundaries have to be provided along with basic settings for the particular purpose, e.g., estimation of individual parameters, estimation of population parameters, or calculation of the Fisher information matrix. For simulation, information about the underlying trial design can either be specified in the **Trial Design** section or sourced from a dataset. Parameter values and basic task settings again have to be provided. Finally, because one PharmML file can define multiple tasks, their dependencies can be encoded.

INTEROPERABILITY AND FUTURE PLANS

PharmML has been designed for the exchange of models between tools. Users will be able to write models using a human readable language also developed within DDMoRe, the Modeling Description Language (MDL) (http://ddmore.eu/mdl). To facilitate its use, an Integrated Development Environment tool (MDL-IDE) is available, within which the model is automatically translated to PharmML and can be passed to PharmML-compatible tools. Development of the MDL and the MDL-IDE is still ongoing, but initial results are very promising. Alternatively, in cases when only the *Structural Model* is required, modelers can use the web-editor infix2pharmml (http://infix2pharmml.sourceforge.net/) to draft ready-to-use PharmML models.

Another key element of the DDMoRe framework is libPharmML (https://sourceforge.net/projects/libpharmml.ddmore.p), an Application Programming Interface (API), the development of which follows the updates of PharmML. LibPharmML provides basic programmatic functionality for working with PharmML-coded models, such as methods for reading, writing, and validating PharmML. DDMoRe aims to integrate a number of target tools and libPharmML is essential for achieving this goal.

Work is ongoing on a number of new PharmML elements, including a Standardized Output (SO), support for Optimal Experimental Design (OED) and Bayesian estimation. The SO element is designed to be a tool-independent storage format for results typically produced in pharmacometrics. The OED element builds on the **Trial Design** with an additional "design space" for domain-specific optimization settings. Support for SBML-coded[10] structural models, within the **Model Definition**, is under development.

CONCLUSION

PharmML is an open-source exchange format for models, intended to facilitate smooth, error-free interoperability between the software tools required in pharmacometrics today. Using a standardized model and output definition, PharmML has the potential to streamline complex workflows, increase the reproducibility of research, ease reporting and bug tracking, and improve the reuse of existing models. It is anticipated that the adoption of PharmML within the field will act as a catalyst for development of novel software and that PharmML will become a widely used standard.

Acknowledgments. We thank Wendy Aartsen, Landry Cochard, Elisabetta Cargnello, Geraldine Dupin, and Angel Rafael for their outstanding support. The research leading to these results received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement 115156, resources of which are composed of financial contributions from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. The DDMoRe project is also supported by financial contribution from Academic and SME partners.

Conflict of Interest. The authors declared no conflict of interest.

1. Mould, D.R. & Upton, R.N. Basic concepts in population modeling, simulation, and model-based drug development. *CPT Pharmacometrics Syst. Pharmacol.* **1**, e6 (2012).
2. Harnisch, L., Matthews, I., Chard, J. & Karlsson, M.O. Drug and disease model resources: a consortium to create standards and tools to enhance model-based drug development. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e34 (2013).
3. Sauro, H.M. & Bergmann F. Standards and ontologies in computational systems biology. *Essays Biochem.* **45**, 211–222 (2008).
4. Beal, S.L., Sheiner, L.B., Boeckmann, A.J. & Bauer, R.J. NONMEM User's Guides. Technical report, Icon Development Solutions, Ellicott City, MD, USA (2009).
5. Lixoft, Monolix 4.3. <http://lixoft.com>.
6. Li, C. *et al.* BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Syst. Biol.* **4**, 92 (2010).
7. Lavielle, M. *Mixed Effects Models for the Population Approach Models, Tasks, Methods & Tools* (Chapman & Hall/CRC Biostatistics Series 2014).
8. Bonate, P. *Pharmacokinetic-Pharmacodynamic Modeling and Simulation* 2nd ed. (Springer, 2011).
9. CDISC SDM-XML Technical Committee. CDISC Study Design Model in XML (SDM-XML), Version 1.0. Technical report, Clinical Data Interchange Standards Consortium. <www.cdisc.org> (2011).
10. Hucka, M. *et al.* The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* **19**, 524–531 (2003).

Supplementary information accompanies this paper on the *CPT: Pharmacometrics & Systems Pharmacology* website (http://www.wileyonlinelibrary.com/psp4)