

**DATABASE**

SwissPK^{cdw} – A clinical data warehouse for the optimization of pediatric dosing regimens

Roland Goers¹ | Diana Coman Schmid² | Vera F. Jäggi³ | Paolo Paioni³ |
 Michal J. Okoniewski² | Althea Parker² | Beat Bangerter³ | Sofia Georgakopoulou⁴ |
 Thierry Sengstag^{4,5} | Julia Bielicki⁶ | Romy Tilen^{1,3} | Swen Vermeul² |
 Stefanie D. Krämer⁷ | Christoph Berger³ | Bernd Rinn^{2,5} |
 Henriette E. Meyer zu Schwabedissen¹

¹Biopharmacy, Department of Pharmaceutical Sciences, University of Basel, Basel, Switzerland²Scientific IT Services, ETH Zürich, Zürich, Switzerland³Division of Infectious Diseases and Hospital Epidemiology, University Children's Hospital Zürich, Zürich, Switzerland⁴Center for Scientific Computing, University of Basel, Basel, Switzerland⁵Swiss Institute of Bioinformatics, Lausanne, Switzerland⁶Pediatric Pharmacology and Pharmacometrics Research, University Children's Hospital Basel, Basel, Switzerland⁷Biopharmacy, Institute of Pharmaceutical Sciences, Department of Chemistry and Applied Biosciences, ETH Zürich, Zürich, Switzerland**Correspondence**

Stefanie D. Krämer, Vladimir-Prelog-Weg 1-5/10, 8093 Zürich, Switzerland.
 Email: stefanie.kraemer@pharma.ethz.ch

Christoph Berger, Steinwiesstrasse 75, 8032 Zürich, Switzerland.
 Email: christoph.berger@kispi.uzh.ch

Bernd Rinn, Mattenstrasse 26, 4058 Basel, Switzerland.
 Email: brinn@ethz.ch

Henriette E. Meyer zu Schwabedissen, Klingelbergstrasse 50, 4056 Basel, Switzerland.
 Email: h.meyerschwabedissen@unibas.ch

Funding information

This work was funded by the Swiss Personalized Health Network (SPHN). The authors are grateful for their support. Moreover, the publication was supported by the Open Access Publication Fund of the University of Basel.

Abstract

Clinical trials have been performed mainly in adults and accordingly the necessary information is lacking for pediatric patients, especially regarding dosage recommendation for approved drugs. This gap in information could be filled with results from pharmacokinetic (PK) modeling, based on data collected in daily clinical routine. In order to make this data accessible and usable for research, the Swiss Pharmacokinetics Clinical Data Warehouse (SwissPK^{cdw}) project has been set up, including a clinical data warehouse (CDW) and the regulatory framework for data transfer and use within. Embedded into the secure BioMedIT network, the CDW can connect to various data providers and researchers in order to collaborate on the data securely. Due to its modularity, partially containerized deployment and open-source software, each of the components can be extended, modified, and re-used for similar projects that require integrated data management, data analysis, and web tools in a secure scientific data and information technology (IT) environment. Here, we describe a collaborative and interprofessional effort to implement the aforementioned infrastructure between several partners from medical health care and academia. Furthermore, we describe a

Roland Goers and Diana Coman Schmid contributed equally.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *CPT: Pharmacometrics & Systems Pharmacology* published by Wiley Periodicals LLC on behalf of American Society for Clinical Pharmacology and Therapeutics.

real-world use case where blood samples from pediatric patients were analyzed for the presence of genetic polymorphisms and the results were aggregated and further analyzed together with the health-related patient data in the SwissPK^{cdw}.

INTRODUCTION

During development of a new molecular entity (NME) preclinical and clinical studies are conducted to ensure the drug's efficacy and safety. Included in this assessment are pharmacokinetic parameters describing absorption, distribution, metabolism, and excretion of the NME. Most clinical trials were limited to adults, for practical and economic reasons and, consequentially, the differences in requirements for pediatric and adult patients are not visible.¹⁻³ For example, some drug-metabolizing enzymes reach adult levels shortly after birth, whereas others remain at lower activity for months to years.⁴ In consequence, one drug may be metabolized at a similar rate as in adults, whereas another remains unchanged much longer in pediatric patients and accumulates with multiple doses. To give another example, renal excretion is disproportionately low until 2 years of age, affecting the half-life of drugs which are mainly renally excreted in adults.^{4,5}

Pediatric drug development and drug use have significantly changed over the last 50 years from "therapeutic orphans,"⁶ to a situation where a robust regulatory framework has been created to ensure that new drugs are reviewed for potential use in pediatric patients.⁷ But there is still a gap in information on drug use in pediatric patients and the information on pediatric use in the drug labels is still scarce, especially for drugs already approved, which leads to immense off-label usage in this patient population.⁸ Unfortunately, pediatricians do not have another choice and, thus, in Switzerland, about 50% of the drugs prescribed to infants below 1 year and neonates are off-label.⁹⁻¹⁴

One possibility to enrich the information on pediatric drug use, is the systematic collection and transparent provision of clinical experience and literature data from pediatrics, as conducted in the national program SwissPedDose.¹⁵ It was launched and supported by the eight largest Children's Hospitals in Switzerland and the Swiss Federal Office of Public Health. In detail, currently applied dose recommendations are compiled, which are then supplemented with information from literature and evaluated by experts in the field to define consensus dose recommendations for Switzerland.¹⁶ The dose finding and certainly dose optimization can be supported by pharmacokinetic models, based on available data. Compartmental methods utilize kinetic models to approximate the behavior of the drug concentration in plasma over time.

Depending on the observed kinetics, a model with one, two, or more compartments is chosen to approximate the plasma concentration-time curve.¹⁷⁻¹⁹ Physiologically-based pharmacokinetic modeling combines compartment models with physiological parameters, such as organ plasma flow or renal filtration rate, and with in vitro determined parameters, if available, such as permeability coefficients or intrinsic clearance.²⁰ Population pharmacokinetics allow to investigate the influence of specified features, such as body weight or genetic polymorphisms in drug-metabolizing enzymes on the drug concentrations found in a target population. If a factor has a measurable impact on the drug's dose-concentration relationship, the dose recommendation can be adjusted accordingly.²¹

Within the Swiss Pharmacokinetics Clinical Data Warehouse (SwissPK^{cdw}) project we developed and implemented a secure by design modular data and IT application encompassing several infrastructure and software application layers (see Figure 1). It provides tools for active data management (module DM), statistical data analysis (module DA), and domain interpretation (module DI; here: exploration and recommendation [module ER]), embedded in a secure environment (Leonhard Med,²² the data and IT platform of the BioMedIT Node Zürich). The project is a collaborative effort of the children's hospitals in Zürich (Kispi) and Basel (UKBB), the academic Biopharmacy groups in Zürich and Basel, and the BioMedIT node Zürich, represented by the Scientific IT Services, ETH Zürich.

The patient data used in this project were so far provided from Kispi and UKBB and were transferred securely within the IT infrastructure hosting the project space, using the process recently established in the Swiss Personalized Health Network²³ (SPHN; <https://sphn.ch/en/home/>). This included the approval from the local ethics committee, the signing of a data transfer and usage agreement (DTUA) between the involved parties and the establishment of secure data transfer systems. The protocols and agreements included the transfer of biological material between the children's hospitals and the University of Basel for genetic analysis.

CONSTRUCTION AND CONTENT

In 2017, the SPHN initiative was launched, aiming to drive the development of secure scientific IT

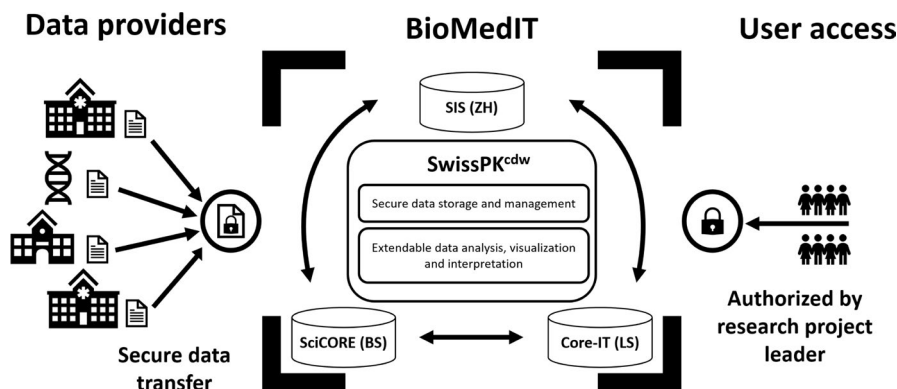


FIGURE 1 Schematic representation of the BioMedIT infrastructure, which hosts the Swiss Pharmacokinetics Clinical Data Warehouse (SwissPK^{cdw}). BioMedIT's computing network is provided by the three partners Scientific IT Services (SIS, ETH Zürich), sciCORE (University of Basel), and Core-IT (Swiss Institute of Bioinformatics and University of Lausanne) (central box). SwissPK^{cdw} is built up based on the standard BioMedIT services that include a secure project space and secure data transfer process. The added value of this project is the development of three modules: data management (DM), data analysis (DA) and visualization exploration and recommendation (ER), which extend the project space functionality with user friendly applications for DM, DA, and visualization ER. Confidential patient data are transferred securely from data providers to the project space in the BioMedIT network.²³ The access to the SwissPK^{cdw} project space and resources in Leonhard Med is restricted to authorized users designated by the project leader

infrastructures for interoperability and accessibility of healthcare data in Switzerland.²⁴ Part of this effort is the provision of an IT infrastructure, which fulfills the strict data protection and security requirements for sensitive patient data. These goals are addressed within the framework of the BioMedIT network,²³ and have been applied to the herein described SwissPK^{cdw}, which started in summer 2020 as an infrastructure project supported by the SPHN.

Patient consent and data security

Patient data are considered sensitive personal information. The herein described SwissPK^{cdw} is hosted in Switzerland, and thus Swiss law applies. In detail, the data collection and use in research is governed by the Federal Act on Data Protection (FADP) and the Human Research Act (HRA). This legislation regulates the use of patient data in treatment and research. The FADP classifies health data as sensitive personal data (FADP, art. 3c) which means that, to ensure data protection, certain requirements must be met prior working with the data. In order to comply with these legal requirements, the patient's consent is necessary. Within SwissPK^{cdw} projects, we are collecting data from patients that gave a general consent (GC) with entry in the respective hospital or retrospectively.

The GC allows the use of patient data and biological material for future research and allows the patient to impose certain limitations (e.g., type of project). This consent

can be withdrawn by the patient at any time. Within the project we asked for and were granted ethical approval for the use of data from patients (BASEC-2019-01004). The de-identified (pseudonymized) patient data that are used in the SwissPK^{cdw} classify as confidential (SPHN IT Security Policy²⁴). The sensitive nature of this clinical data implies that the computational infrastructure needs to comply with the highest data security standards at all stages of data handling (including data transfer, management and analysis) and with the research policies on data FAIRness (SPHN Ethical Framework²⁴).

Application for authorization

Any researcher or clinician can apply for authorization to access the SwissPK^{cdw} infrastructure. To be granted authorization, the user must first pass the SPHN Data Privacy and IT security test offered by the SPHN Data Coordination Center (DCC). The applying researcher has to read, understand, and sign the SwissPK^{cdw} policy and the Acceptable Use Policy (AUP) of Leonhard Med, which is sent to the applicant upon request and explicit authorization by the SwissPK^{cdw} project leader. Finally, prior to being granted access the applicant has to provide a valid ethical approval by the corresponding Ethical Committee for the intended analysis. Authorization requests can be sent per mail to SwissPKcdw@kispi.uzh.ch. In the future, we additionally plan to grant open access to our web application containing only nonsensitive data.

SwissPK^{cdw} infrastructure design

The herein presented clinical data warehouse is embedded in the IT infrastructure of the BioMedIT network,²³ which is a nationwide secure infrastructure for data and compute resources. It consists of three regional nodes in Zürich, Basel, and Lausanne operated by the Scientific IT Services (SIS; ETH Zürich), sciCORE (University of Basel), and Core-IT (Swiss Institute of Bioinformatics/DCSR; University of Lausanne), respectively. BioMedIT provides a scientific computing environment specifically designed for the secure transfer, storage, and processing of sensitive biomedical patient data. As such, the Leonhard Med platform of the node in Zürich grants a secure environment for bringing together various patient data types (e.g., routinely collected clinical data and large scale omics data) from distinct sources, for actively managing research data, and for data analysis using versatile compute and software resources. Following the BioMedIT process for secure data transfer, data providers (e.g., hospitals and universities) are connected to the network and transfer data securely into dedicated secure project spaces hosted at the BioMedIT network nodes (see Figure 1). The data transfer is coordinated by the DCC of BioMedIT, access to SwissPK^{cdw} in Leonhard Med is restricted to authorized users, and security controls are imposed at infrastructure and policy level.^{23,25}

The SwissPK^{cdw} is designed as a modular system, which allows the DM, DA, and DI/ER modules to be used separately (e.g., by data managers and/or data scientists) or combined in the web application (e.g., by clinicians).

The DM is built around openBIS,^{26,27} a versatile tool for active data management that has features for secure handling of confidential data (e.g., user-based role access and data audit trail), can capture diverse data modalities (e.g., patient clinical, genetic data, and experimental data) and can be integrated with upstream (e.g., data validation) and downstream processes (e.g., statistical data analyses and web applications).

The DA module provides the tools for extensive data analysis. This includes the R²⁸ stack with specialized libraries for pharmacokinetics,²⁹ Python,³⁰ and Jupyter Notebooks,³¹ as well as the possibility to deploy commercial software like MATLAB (MathWorks, Natick, MA, USA).

Finally, the SwissPK^{cdw} web application provides graphical user access to predefined statistical analyses (from module DA) that are applied to data managed with openBIS (module DM) and currently offers exploration functionalities (module ER) for user selected parameters (e.g., patient cohort, medication, clinical variables, etc.). As such, SwissPK^{cdw} actively contributed to the development of the secure nationwide infrastructure within the

framework of SPHN as a complete, scalable, and reusable proof-of-concept pipeline for pediatric biomedical research: from secure data transfer, to active data management,^{26,27} data analysis,²⁸ and interpretation.

Data management module

The pseudonymized patient data that is used in the SwissPK^{cdw} classify as confidential and requires security protection over the full data life cycle. The DM module within SwissPK^{cdw} uses openBIS^{26,27} as the active research data management system, which fulfills the security requirements and is deployed on the secure platform of Leonhard Med. This DM is pivotal in its role in the whole infrastructure. It provides the tools for ensuring data quality checks (i.e., semi-supervised data validation), advanced query functionality (e.g., search all patients with age in specified range and with specified gender), data audit trails (e.g., track any newly added or modified data for a patient), user role-based access (e.g., “observer” with read-only permission, “user” with editing permission, etc.), and patient consent tracking (e.g., consent type and status defined in the data model). The DM module is based on a customizable and flexible data model that links observations (e.g., laboratory values, gene variants, biological samples, etc.) to patients and uses standard ontologies defined in the SPHN core dataset.³² Furthermore, openBIS has additional functionalities (i.e., support for integration with archiving and publishing of datasets and metadata in research repositories). As such, openBIS as resource data management system for the SwissPK^{cdw} project streamlines and enforces the proper data management on research and patient data.

OpenBIS is deployed as a client-server application and provides access to the stored data in a relational database via application programming interfaces and web-based user interfaces. In order to ensure good scalability, the metadata is stored in a separate database as the corresponding data. The data organization within openBIS is described and defined by a data model designed in a customized way for the project.^{26,27}

The openBIS model uses controlled vocabularies for metadata and captures the SwissPK^{cdw} project structure and data modalities. As such, it comprises spaces for drug-related patient data (e.g., gentamicin space and cyclosporine space) and spaces that can be used for research-related scripts or literature-related data. Access to openBIS spaces may be restricted using the user-based role access functionality of openBIS. Each space contains collections of metadata: patient consent (e.g., status as accepted or refused), routinely collected patient clinical information (e.g., age, gender, weight, etc.), measurements

(e.g., height, height unit, date, and time of measurement), laboratory related information (e.g., laboratory coding system; i.e., L4CHLAB, data and time of data collection, kit name, and instrument name), diagnosis (e.g., diagnosis code and coding system; i.e., ICD10GM version 2018), and genetic results (e.g., laboratory identifier, gene name, single-nucleotide polymorphism [SNP], and method used for detection). In openBIS, each individual patient and the aforementioned related data is associated with a unique research identifier or pseudonym.

The project's data manager and other authorized project members operate on the DM module via the graphical user interface of openBIS. Data can be provided at source (e.g., hospital and research laboratory) in any format from SQL database dumps, comma separated value (csv) files, tab delimited text files (txt), or RDF files. The data transfer is done via the Secure Encryption and Transfer Tool³³ (SETT) of BioMedIT. SETT is an application to package, cryptographically encrypt, and sign, as well as transfer data via secure channels within the BioMedIT nodes. The files are encrypted with the public key of the data manager and signed by the data provider. SETT transfers the encrypted data from data providers, via the secure landing zones of the BioMedIT nodes and routes it to the SwissPK^{cdw} secure project space at Leonhard Med, where the data manager of the project performs the decryption.

Next, confidential patient data on the SwissPK^{cdw} secure project space is registered into the project's openBIS instance. In brief, a customized tool reads the raw SQL database dumps (csv or txt files), standardizes values (e.g., applies correct UCUM [Unified Code for Units of Measure] codes for units, and makes timestamps compatible with openBIS), validates the data (e.g., checks that the weight of a patient is within an expected range) and reports the possibly invalid values, thus allowing the data provider to review and correct the erroneous values before they are imported into openBIS. The new data is then compared with already imported data, creating new entries or updating old entries accordingly. To achieve this comparison quickly, an intermediate SQLite database is used. Finally, the validated, normalized, and compared data are then imported and assigned to the correct spaces and projects in the openBIS instance. Successfully imported data are also written to the SQLite database for data comparison with future data imports.

To connect the data management module with the data analysis module, a custom Python script was developed, which queries for all data in a given openBIS space and exports the data to formats, such as csv, so the data can be used by other processes and tools (e.g., R scripts for statistical data analysis and web applications for data exploration and visualization).

Data analysis and domain interpretation modules

In addition to the means to aggregate data from different providers and store them securely, SwissPK^{cdw} offers several ways for researchers to work on the data. Pharmacokinetic modeling is often conducted with the help of commercial or license-bound tools, such as NONMEM (ICON, Dublin, Ireland) and Monolix (Lixoft, Paris, France), which have become the de facto gold standard within industry and academia.²⁹ Within this project, we wanted to provide means to conduct the same workflows but relying on open-source software which is more in the spirit of open access research and sharing of data and results. Thus, the goal was to implement modeling capabilities of the SwissPK^{cdw} platform by primarily using R. To this end, we relied on the R libraries *nlmixr*²⁹ and *RxODE*,³⁴ whereas *nlmixr* is specifically designed for fitting pharmacokinetic and pharmacokinetic/pharmacodynamic models with individual and population data (see Figure 2). Its nomenclature and syntax are similar to the aforementioned commercial/license-bound software and thus allows a smooth transition for already experienced researchers. Due to the openness of the software, features and capabilities can be implemented by collaborating researchers and the platform extended to new needs.

User authorized to work in the SwissPK^{cdw} secure project space at Leonhard Med may use a broad range of scientific software tools in addition to R and RStudio (e.g., Jupyter Notebooks currently via the openBIS plugin, Python stack, or pending permission, use of commercial software such as MATLAB is possible).

UTILITY

Besides the project's main goal to drive the development and implementation of versatile, scalable, and reusable components in the secure infrastructure of BioMedIT (i.e., the modular SwissPK^{cdw} platform), using this resource for research (i.e., feeding real-world patient data into the system) was the next step. The aforementioned infrastructure of the SwissPK^{cdw} was used to extract patient information from the source systems of the two hospitals (Kispi and UKBB), and to transfer this information securely into the SwissPK^{cdw} (see Figure 3). During the project, the process to add laboratory data gathered in a research laboratory was also set up and tested. Here, we collected leftover EDTA whole blood samples obtained from the selected patients, pseudonymized them, and transferred them from the hospitals to the laboratories of the Biopharmacy group at the University of Basel in order to analyze them for the presence of specific genetic SNPs which can impact

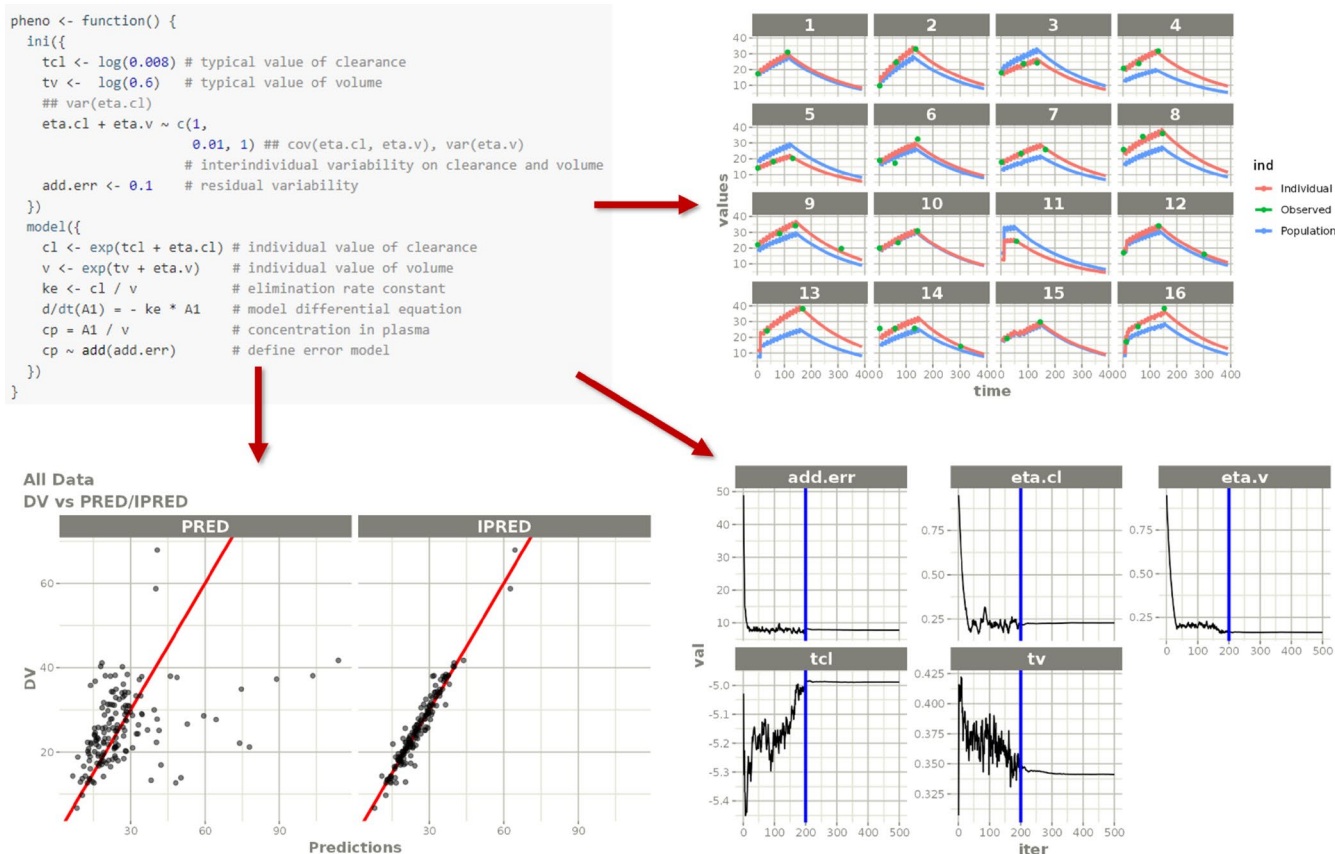


FIGURE 2 The nlmixr provides an easy syntax to define pharmacokinetic models and fit them to the data (top left). The visualization includes the derived model kinetics (top right) as well as graphical assessment of the quality of the fit (bottom left) and traces of the parameter estimations (bottom right). Adapted from <https://nlmixrdevelopment.github.io/nlmixr/articles/addingCovariances.html>³⁸

the pharmacokinetics of selected drugs. The data obtained in the analysis of SNPs combined in panels defined by a certain drug were encrypted and securely transferred. The results of the genotyping can then be combined with the extracted patient information for modeling of the pharmacokinetics (see Figure 3).

Substances where medication levels in blood are measured in daily clinical routine are cyclosporine, tacrolimus, and gentamicin. The application of the antibiotic gentamicin requires the monitoring of its plasma levels in order to avoid adverse drug effects which are linked to high plasma levels over long time periods.³⁵ Gentamicin is metabolically inert and is excreted renally. In contrast, cyclosporine and tacrolimus are immunosuppressants also known for their narrow therapeutic index, but their metabolism and excretion involve a complex system of various drug metabolizing enzymes (CYP3A4 and CYP3A5) and drug transporters (ABCB1), for which genetic variants are known (for details please refer to <https://www.pharmgkb.org/pathway/PA165986114>).

Within the laboratory of the Biopharmacy group in Basel, genetic SNP panel testing was conducted using commercially available chemistry (Thermo Fisher Scientific, TaqMan Genotyping assays, and TaqMan Master Mix).

The panel of genetic polymorphisms selected for cyclosporine was following the consensus report published by Brunet et al.,³⁶ even though the recommendations were gathered for adult patients (see Table 1).

At the time of writing, SwissPK^{cdw} contains three datasets for the aforementioned drugs, namely cyclosporine, tacrolimus, and gentamicin (see Table 2). Thirteen patients are registered within the cyclosporine data set, having an average age of 9 years. The timespan in which the patients were treated with the drug and data has been logged is, on average, 200 days, whereas the longest registered timespan is 1414 days. The gentamicin routine dataset is the largest in terms of number of patients (266)³⁹ whereas they were, on average, 2 years old at the time of treatment. The data set of a prospective pharmacokinetic study comprises 126 patients which were treated with three doses on average for a duration of 7 days in total. Additionally, each dataset contains further information on treatment intervals, applied doses, comedication, and searchable free text notes. At the time of writing, the stored data are integrated into the data warehouse and can be used by researchers for pharmacokinetic modeling and analysis. The detailed analysis of the pharmacokinetics is beyond the scope of this paper.

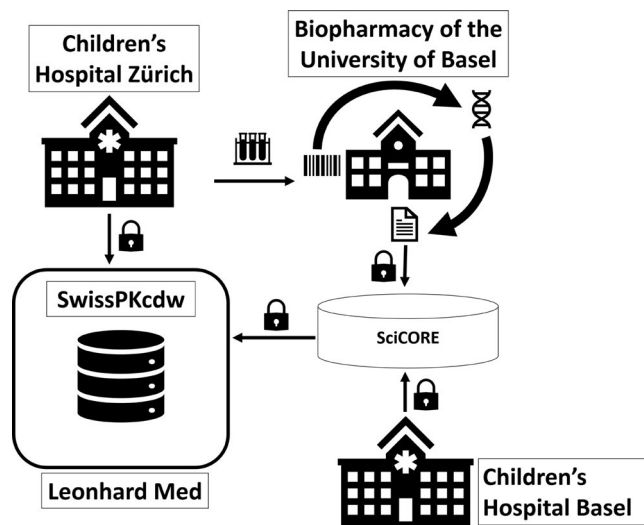


FIGURE 3 Schematic representation of the collaborative workflow between the Children's Hospital Zürich and the Biopharmacy group of the University of Basel. The pooled information from the clinical and laboratory information systems of the hospital is stored in the data lake and extracted from the latter. Upon extraction, the data is securely transferred into the dedicated project space in Leonhard Med, following the BioMedIT process.²³ Similarly, patient samples are analyzed for the presence of genetic polymorphisms and the results are transferred encrypted via the sciCORE BioMedIT node in Basel. After transfer, data are decrypted by designated project personnel (i.e., data manager), validated and registered in the openBIS data management system hosted within the secure project space. Similar to the Children's Hospital Zürich, the Children's Hospital Basel extracts the patient information from their systems and transfers the data via the sciCORE node

Web application

In order to provide easy and user-friendly ways to explore SwissPK^{cdw} and generate dose recommendations in the future, we have designed and developed a web application (see Figure 4). It builds upon the DM and DA modules and is accessible by authorized users via the secure remote desktop login to the project space at Leonhard Med.

In detail, the SwissPK^{cdw} web application is implemented in JavaScript and uses openCPU,³⁷ which provides easy integration of R and JavaScript via RPC calls and an HTTP API. The web application deployment in Leonhard Med is based on container technology (i.e., Docker) and consists of two main components: (1) the exploration component, which generates graphical and tabular summaries of the data upon user input, and (2) the dosage recommendation component, which will be capable to calculate a recommended dose based on the derived models. The exploration module has been implemented (see Figure 4) and allows the user to select a drug from a list of available drugs (e.g., gentamicin) and to vary the ranges of

TABLE 1 Summary of genetic variants tested for in the cyclosporine panel indicating the respective polymorphism by rs number, the exp. MAF (European population, revision 20201027095038), and the obs. MAF as obtained by analyzing 25 patients

rs number	Polymorphism	exp. MAF	obs. MAF
rs1045642	ABCB1 c.1236 C>T	0.481	0.460
rs2229109	ABCB1 c.1199 C>T	0.042	0.060
rs35599367	CYP3A4g.20493 G>A	0.046	0.100
rs776746	CYP3A5g.12083 G>A	0.930	0.980
rs10264272	CYP3A5g.19787 G>A	0.001	n.d.
rs1057868	POR g.75587 C>T	0.286	0.120
rs4253728	PPAR α g.68637 G>A	0.270	0.400
rs4823613	PPAR α g. 56877A>G	0.282	0.420

Abbreviations: exp. MAF, expected minor allele frequency; n.d., Minor allele not detected in the cohort; obs. MAF, observed MAF.

parameters, such as the patients' age or gender. Based on this selection, a report containing graphical and tabular summaries is generated and displayed. Parts of the report are downloadable in various formats.

The future recommendation module will be based on pharmacokinetic models developed by researchers and clinicians within the collaboration. As soon as a model has been derived which describes the data accurately, it can be integrated in the SwissPK^{cdw} web application and be used to predict a dose recommendation for a patient. The patient's characteristics, such as age, weight, or polymorphisms, can be specified by the user as covariates in the underlying models.

DISCUSSION

We are reporting details on the establishment of a sustainable infrastructure which can be applied in pharmacokinetic modeling research using data generated within clinical routine. This data extracted from the record system of healthcare providers can be supplemented with data generated within laboratories in an academic setting. The design of the infrastructure is modular and grants the possibility to aggregate and use data in a secure way. Use of clinical data obliges to a very high standard in data security. Because SwissPK^{cdw} contains confidential patient data, the project required a legal framework for transferring and using the data (i.e., DTUA). Furthermore, each researcher needs to apply for authorization from the SwissPK^{cdw} project leader to access the project's data. These restrictions can be seen as limitations by scientists, which are used to open access of data and methods. However, as a clinical project

	Cyclosporind/ tacrolimus	Gentamicin (routine)	Gentamicin (study)
Number of patients	13	266	126
Avg. age (yr)	9 (± 5.2)	2 (± 3.8)	0 (± 1.7)
Avg. number of doses	167 (± 48.7)	1.3 (± 0.53)	2.5 (± 2.89)
Avg. timespan (days)	200 (± 406)	21 (± 100)	7 (± 15)

TABLE 2 Current number of patients stored in SwissPK^{cdw}, grouped by their treatment

The average age highlights the focus on pediatric patients.

Abbreviations: Avg., average; SwissPK^{cdw}, Swiss Pharmacokinetics Clinical Data Warehouse.

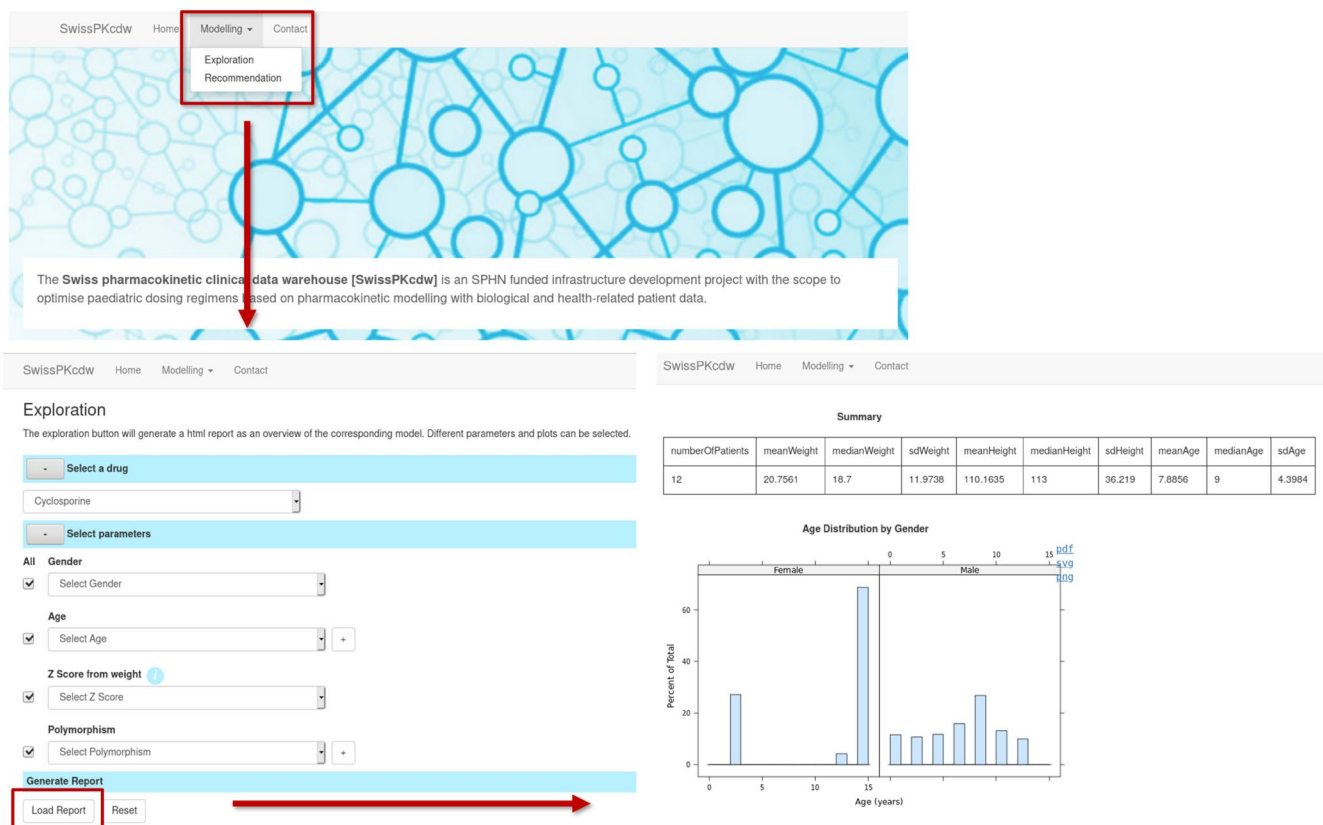


FIGURE 4 Overview of the web application functionality and workflow. The exploration module as currently implemented is shown in the panels. The user can select which drug dataset should be used and narrow down the patients' parameters, if necessary. Based on this selection, a report containing tabular and graphical summaries is generated

containing confidential patient data, SwissPK^{cdw} must comply to the regulations.

From a technical perspective, we were able to successfully use existing BioMedIT solutions (e.g., secure project space and secure data transfer, DTUA). SwissPK^{cdw} has added further value by developing and testing a proof-of-concept modular platform for data management, analysis, exploration, and visualization within the secure environment of BioMedIT. Although this project has specific components (e.g., use of openBIS for data management and web app integrating openBIS with R-based analytical workflows), it aims to be as maintainable and sustainable as possible by using exclusively open-source software alongside with containerized web app deployment. Thus,

it has the potential to establish generic blueprints for similar future projects.

SwissPK^{cdw} is certainly still far away from providing data for hundreds of drugs based on thousands of patients, but it seems worth mentioning that an important aspect of scientific evaluation of clinical data is the controllership of the data. Especially if setting up a clinical data warehouse, where data from multiple providers are transferred into, defining data ownership becomes rather complex. Certainly, personal data is owned by the person providing it (the patient), however, without giving any solutions for that topic, we want to mention that there are multiple stakeholders that could claim "ownership" or "controllership" at least in the context of the scientific conduct on

handling intellectual property of a data set. It is part of the SPHN initiative to find a national and perhaps legal solution for that important question. The DTUAs used by this and other SPHN projects offer a first solution to this complex question.

A drawback of the focus on pediatric patients is that the number of available patients is much lower compared to adult patients. In SwissPK^{CDW}, the number of entries is in the lower hundreds range. This could limit the number of possible analysis technologies for the time being. However, the available data will increase over time and, thus, the relevance of this point will diminish. We decided to go for genetic panel testing, which certainly means that we are limited to the drug-gene associations known at the moment of selection. The project, as approved by the ethical committee, foresees no storage of samples at the research laboratory and, consequently, all samples have to be destroyed after assessment, quality control, and data transfer. Accordingly, there is no possibility yet to add additional variants that may be identified to be predictive for drug efficacy and safety in the future.

We are aware that in order to be applied for clinical decision making, the dosage recommendation system must be approved as a digital medical device by the authorities. Thus, we are not providing a medical decision support system at the moment and we are just reporting on the possible implementation of a web app.

To our knowledge at the time of writing, SwissPK^{CDW} as a clinical data warehouse, specializing in pediatric patients and providing the necessary infrastructure to aggregate, process, and analyze the data, is unique among the projects in personalized health care. However, as mentioned above in the limitations, the available data from patients is currently low. Access to clinical data from a high number of patients and from a wide variety of drugs would not only result in optimized dosing regimens, but would in addition allow to investigate the physiological processes and their age-dependence involved in pharmacokinetics. Due to the modular design of the infrastructure, extending the available data and even the functionality of SwissPK^{CDW} can be done conveniently.

SUMMARY

The SwissPK^{CDW} is the first database in Switzerland which focuses on the aggregation of routine clinical pharmacokinetic data of pediatric patients. It provides all necessary computational resources to analyze the data in a law-conforming way. We believe that the availability of such a data warehouse will encourage additional researchers to join our efforts and build better pharmacokinetic models

as well as motivate further hospitals to provide their data in order to extend the warehouse toward various drugs. Our project is committed to data FAIRness and computational workflows interoperability, while using data protection measures to ensure confidentiality of sensitive patient data. The CDW will be part of a platform providing online pharmacokinetic modeling applications intended to analyze the data and optimize dosage regimens.

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

ORCID


Paolo Paioni  <https://orcid.org/0000-0002-3904-1606>

Romy Tilen  <https://orcid.org/0000-0002-1627-1513>

Swen Vermeul  <https://orcid.org/0000-0002-1307-3052>

Stefanie D. Krämer  <https://orcid.org/0000-0002-0426-4340>

Christoph Berger  <https://orcid.org/0000-0002-2373-8804>

Henriette E. Meyer zu Schwabedissen  <https://orcid.org/0000-0003-0458-4579>

REFERENCES

1. Di Pietro ML, Cutrera R, Teleman AA, Barbaccia ML. Placebo-controlled trials in pediatrics and the child's best interest. *Ital J Pediatr.* 2015;41:11.
2. Davidson AJ, O'Brien M. Ethics and medical research in children. *Pediatric Anesthesia.* 2009;19:994-1004.
3. Kauffman RE, Kearns GL. Pharmacokinetic studies in paediatric patients: clinical and ethical considerations. *Clin Pharmacokinet.* 1992;23:10-29.
4. Krämer SD, Testa B. The biochemistry of drug metabolism – An introduction. *Chem Biodivers.* 2009;6:1477-1660. <https://doi.org/10.1002/cbdv.200900233>
5. Cella M, Knibbe C, Danhof M, Della Pasqua O. What is the right dose for children? *Br J Clin Pharmacol.* 2010;70:597-603. <https://doi.org/10.1111/j.1365-2125.2009.03591.x>
6. Shirkey H. Therapeutic orphans. *J Pediatr.* 1968;72:119-120. [https://doi.org/10.1016/s0022-3476\(68\)80414-7](https://doi.org/10.1016/s0022-3476(68)80414-7)
7. Burckart GJ, Kim C. The revolution in pediatric drug development and drug use: therapeutic orphans no more. *J Pediatr Pharmacol Ther.* 2020;25:565-573. <https://doi.org/10.5863/1551-6776-25.7.565>
8. Wimmer S, Rascher W, Neubert A. How often do SmPCs contain contraindications and special warnings that are specific for the paediatric population. *Klin Padiatr.* 2019;231:191-198. <https://doi.org/10.1055/a-0898-4076>
9. Di Paolo ER, Stoetter H, Cotting J, et al. Unlicensed and off-label drug use in a Swiss paediatric university hospital. *Swiss Med Wkly.* 2006;136:218-222.
10. de Souza AS, dos Santos DB, Rey LC, et al. Off-label use and harmful potential of drugs in a NICU in Brazil: a descriptive study. *BMC Pediatr.* 2016;16:13. <https://doi.org/10.1186/s12887-016-0551-8>
11. Yackey K, Stukus K, Cohen D, Kline D, Zhao S, Stanley R. Off-label medication prescribing patterns in pediatrics: an update. *Hosp Pediatr.* 2019;9:186-193. <https://doi.org/10.1542/hpeds.2018-0168>

12. Committee on Drugs. American Academy of Pediatrics Uses of drugs not described in the package insert (off-label uses). *Pediatrics*. 2002;110:181-183.
13. Conroy S, Choonara I, Impicciatore P, et al. Survey of unlicensed and off label drug use in paediatric wards in European countries. European Network for Drug Investigation in Children. *BMJ*. 2000;320:79-82. <https://doi.org/10.1136/bmj.320.7227.79>
14. Schrier L, Hadjipanayis A, Stiris T, et al. Off-label use of medicines in neonates, infants, children, and adolescents: a joint policy statement by the European Academy of Paediatrics and the European society for Developmental Perinatal and Pediatric Pharmacology. *Eur J Pediatr*. 2020;179:839-847. <https://doi.org/10.1007/s00431-019-03556-9>
15. SwissPedDose. SwissPedDose – Nationale Datenbank zur Dosierung von Arzneimitteln bei Kindern. Accessed March 31, 2021. <https://swisspeddose.ch/>.
16. Tilen R, Panis D & Aeschbacher S, et al. Development of the Swiss Database for Dosing Medicinal Products in Paediatrics. *Eur J Pediatr*. in press.
17. Mould DR, Upton RN. Basic concepts in population modeling, simulation, and model-based drug development. *CPT Pharmacometrics Syst Pharmacol*. 2012;1:6. <https://doi.org/10.1038/psp.2012.4>
18. Mould DR, Upton RN. Basic concepts in population modeling, simulation, and model-based drug development-part 2: introduction to pharmacokinetic modeling methods. *CPT Pharmacometrics Syst Pharmacol*. 2013;2:38. <https://doi.org/10.1038/psp.2013.14>
19. Upton RN, Mould DR. Basic concepts in population modeling, simulation, and model-based drug development: part 3-introduction to pharmacodynamic modeling methods. *CPT Pharmacometrics Syst Pharmacol*. 2014;3:88. <https://doi.org/10.1038/psp.2013.71>
20. Villiger A, Stillhart C, Parrott N, Kuentz M. Using physiologically based pharmacokinetic (PBPK) modelling to gain insights into the effect of physiological factors on oral absorption in paediatric populations. *AAPS J*. 2016;18:933-947. <https://doi.org/10.1208/s12248-016-9896-z>
21. Jullien V, Urien S, Chappuy H, et al. Abacavir pharmacokinetics in human immunodeficiency virus-infected children ranging in age from 1 month to 16 years: a population analysis. *J Clin Pharmacol*. 2005;45:257-264. <https://doi.org/10.1177/0091270004272215>
22. ETH Zürich Scientific IT Services. Leonhard Med: Secure Scientific Platform for Confidential Data. Accessed March 31, 2021. <https://sis.id.ethz.ch/services/confidentialresearchdata/>.
23. Coman Schmid D, Cramer K, Oesterle S, et al. SPHN – The BioMedIT Network: a secure IT platform for research with sensitive human data. In: Pape-Haugaard LB, Lovis P-H, Cort Madsen I, Weber P, Hostrup Nielsen P, Scott P, eds. *Digital Personalized Health and Medicine*. Studies in Health Technology and Informatics. Vol 270. IOS Press; 2020:1170-1174.
24. Swiss Personalized Health Network (SPHN). Accessed March 31, 2021. <https://sphn.ch/>.
25. ETH Zürich Scientific IT Services. Confidential Research Data Support. Accessed April 1, 2021. <https://sis.id.ethz.ch/services/confidentialresearchdata/>.
26. Bauch A, Adamczyk I, Buczek P, et al. openBIS: a flexible framework for managing and analyzing complex data in biology research. *BMC Bioinform*. 2011;12:468. <https://doi.org/10.1186/1471-2105-12-468>
27. Barillari C, Ottoz DSM, Fuentes-Serna JM, Ramakrishnan C, Rinn B, Rudolf F. openBIS ELN-LIMS: an open-source database for academic laboratories. *Bioinformatics*. 2016;32:638-640. <https://doi.org/10.1093/bioinformatics/btv606>
28. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2020. Accessed April 2, 2021. <https://www.R-project.org/>.
29. Fidler M, Wilkins JJ, Hooijmaijers R, et al. Nonlinear mixed-effects model development and simulation using nlmixr and related R open-source packages. *CPT Pharmacometrics Syst Pharmacol*. 2019;8:621-633. <https://doi.org/10.1002/psp4.12445>
30. Van Rossum G, Drake FL. *Python 3 Reference Manual*. CreateSpace; 2009.
31. Kluyver T, Ragan-Kelley B, Pérez F, et al. Jupyter Notebooks: A publishing format for reproducible computational workflows. In: Loizides F, Schmidt B, eds. *Positioning and Power in Academic Publishing: Players, Agents and Agendas*. IOS Press; 2016:87-90.
32. SPHN core dataset, version 2020_1. *SPHN*. Accessed April 23, 2021. <https://sphn.ch/document/sphn-dataset/>.
33. BioMedIT-SPHN sett – BioMedIT’s Secure Encryption and Transfer Tool. Accessed March 31, 2021. <https://sett.readthedocs.io/en/dev-r/index.html>.
34. Wang W, Hallow KM, James DA. A tutorial on RxODE: simulating differential equation pharmacometric models in R. *CPT Pharmacometrics Syst Pharmacol*. 2016;5:3-10. <https://doi.org/10.1002/psp4.12052>
35. Touw DJ, Westerman EM, Sprij AJ. Therapeutic drug monitoring of aminoglycosides in neonates. *Clin Pharmacokinet*. 2009;48:71-88. <https://doi.org/10.2165/00003088-200948020-00001>
36. Brunet M, van Gelder T, Åsberg A, et al. Therapeutic drug monitoring of tacrolimus-personalized therapy: second consensus report. *Ther Drug Monit*. 2019;41:261-307. <https://doi.org/10.1097/ftd.0000000000000640>
37. Ooms J. *The OpenCPU System: Towards a Universal Interface for Scientific Computing through Separation of Concerns*. *arXiv:1406.4806 [stat.CO]*; 2014. Accessed April 2, 2021. <https://arxiv.org/abs/1406.4806>.
38. nlmixr. *Example Model: Phenobarbitol with correlations*. Accessed April 1, 2021. <https://nlmixrdevelopment.github.io/nlmixr/articles/addingCovariances.html>.
39. Paioni P, Jäggi V, Tilen R, et al. Gentamicin population pharmacokinetics in pediatric patients—A prospective study with data analysis using the saemix Package in R. *Pharmaceutics*. 2021;13:1596. <https://doi.org/10.3390/pharmaceutics13101596>

How to cite this article: Goers R, Coman Schmid D, Jäggi VF, et al. SwissPK^{cdw} – A clinical data warehouse for the optimization of pediatric dosing regimens. *CPT Pharmacometrics Syst Pharmacol*. 2021;00:1–10. doi:[10.1002/psp4.12723](https://doi.org/10.1002/psp4.12723)