Article published online: 2024-02-07

Creating a Medication Therapy Observational Research Database from an Electronic Medical Record: Challenges and Data Curation

Wolfgang Rödle[1] Hans-Ulrich Prokosch[1] Eva Neumann[3] Irmgard Toni[2] Julia Haering-Zahn[2] Antje Neubert[2] Sonja Eberl[2]

[1] Chair of Medical Informatics, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany
[2] Departmant of Paediatrics and Adolescent Medicine, Universitätsklinikum Erlangen, Erlangen, Germany
[3] Dr Margarete Fischer Bosch Institute of Clinical Pharmacology, Stuttgart, Germany

Address for correspondence Wolfgang Rödle, Chair of Medical Informatics, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen 91058, Germany (e-mail: wolfgang.roedle@fau.de).

Appl Clin Inform 2024;15:111–118.

Abstract

Background Observational research has shown its potential to complement experimental research and clinical trials by secondary use of treatment data from hospital care processes. It can also be applied to better understand pediatric drug utilization for establishing safer drug therapy. Clinical documentation processes often limit data quality in pediatric medical records requiring data curation steps, which are mostly underestimated.

Objectives The objectives of this study were to transform and curate data from a departmental electronic medical record into an observational research database. We particularly aim at identifying data quality problems, illustrating reasons for such problems and describing the systematic data curation process established to create high-quality data for observational research.

Methods Data were extracted from an electronic medical record used by four wards of a German university children's hospital from April 2012 to June 2020. A four-step data preparation, mapping, and curation process was established. Data quality of the generated dataset was firstly assessed following an established 3 × 3 Data Quality Assessment guideline and secondly by comparing a sample subset of the database with an existing gold standard.

Results The generated dataset consists of 770,158 medication dispensations associated with 89,955 different drug exposures from 21,285 clinical encounters. A total of 6,840 different narrative drug therapy descriptions were mapped to 1,139 standard terms for drug exposures. Regarding the quality criterion correctness, the database was consistent and had overall a high agreement with our gold standard.

Conclusion Despite large amounts of freetext descriptions and contextual knowledge implicitly included in the electronic medical record, we were able to identify relevant data quality issues and to establish a semi-automated data curation process leading to a high-quality observational research database. Because of inconsistent dosage information in the original documentation this database is limited to a drug utilization database without detailed dosage information.

Keywords
- pediatrics
- data curation
- Data Quality Assessment
- medication data
- electronic medical records

received
April 26, 2023
accepted after revision
August 28, 2023

DOI https://doi.org/10.1055/s-0043-1777741.
ISSN 1869-0327.

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Background and Significance

Evidence-based medicine needs to be enhanced by large-scale pharmacoepidemiological studies. The promise of real-world evidence (RWE) is especially relevant to pediatrics, where medicines prescribed for children are often used without evidence derived from randomized clinical trials.[1–7]

Observational studies, e.g., performed by the Observational Health Data Sciences and Informatics (OHDSI) collaboration[8] have shown their potential to lead to a better understanding of drug prescription patterns. In line with the recommendations of the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP),[9] they can therefore excellently complement traditional clinical trials and help to establish safer drug therapy, provided that quality standards are met.

A systematic review pursued by Lasky et al[7] has, however, pointed out that the use of RWE is yet not fully developed in pediatrics and has suggested to further leverage administrative and electronic health record databases to study medication safety and effectiveness in children. Given this background, the aim of this study was to establish a pediatric drug therapy observational research database based on the Children's and Adolescent's Hospital of Erlangen University Hospital (UKER) departmental electronic medical record (EMR).

When creating an observational research database, however, one needs to understand the limitations and data quality problems of data originally documented to support patient care without any research aims in mind. Thus, the second objective of our research was to investigate the inherent data quality issues of the original data, to establish a data curation process, and to finally assess data quality of the established observational research database.

Methods

The Setting: The VMobil Electronic Medical Record

In four wards of the pediatric clinic of UKER a departmental EMR named VMobil was used for drug therapy documentation from April 2012 to June 2020.[10] It was interfaced with the hospital-wide electronic patient record (EPR) (which during those years did not support the medication process yet) and turned off in mid-2020 when comprehensive medication documentation functionalities were added in the UKER EPR. In order to leverage the treasure of real-world data from those 8 years of computerized drug therapy documentation a project was initiated to prepare this historic medication data for future use in retrospective observational studies and additionally invest in a comprehensive data curation and quality assurance process.

Data Processing and Data Curation

Initially all VMobil data items to be transferred to the observational research database were defined. Secondly the format of the target database was defined by the pharmacists, with the results of a first manual data quality analysis in mind. It was decided to provide the target database information in one flat file associated with a single drug dispensation per row. Each drug dispensation row comprises

- Dispensation data: medication ID, dispensation ID, active ingredient, route of application, dose and unit, ATC code.
- Encounter-related data: encounter ID, encounter start date and duration, hospital ward.
- Patient-related data: patient ID, age in days, most recent body weight and height along with the corresponding date of measurement.

In order to achieve standardized terms for e.g., drug substances (active ingredient), route of administration, and pharmaceutical dose forms it was decided to map respective items to (1) the official terminology and codes from the World Health Organization-maintained ATC classification system (https://www.whocc.no/) for active drug ingredients and (2) standard vocabulary terms from the European Directorate for the Quality of Medicines and Healthcare (EDQM).[11]

The full data export, transformation, and curation process consisted of four steps: (1) automated creation of a mapping table in which medication freetext entries were mapped to a standardized vocabulary; (2) manual data curation process by a pharmacist to enhance the quality of this mapping table; (3) an independent data validation process by the pharmacists team on a 10% subset of the mapping table; and (4) an automated transfer of VMobil data into the target observational database using the quality assured mapping table.

Step 1: The freetext field was supposed to be documented according to a clinic internal specification in the following syntax:

> "drug substance incl. drug strength" followed by
> "administration route" enclosed in brackets "[…]"
> followed by
> "pharmaceutical dose form" enclosed in parentheses "(…)"
> e.g., Mesalazin 500 mg [oral] (tablet)

However, the pharmacist team inspection showed that this specification was not consistently followed in the patient care documentation. Thus, data quality issues were categorized and data curation steps were defined for each of the issues detected.

The first challenge was to create a mapping table with standardized terms for the field "active ingredient" from the VMobil medication name freetext field. One pharmacist's manual inspection of the content of the medication freetext field and the coded terms for ATC code, route of application, and dosage yielded that the freetext included the most valid truth and that entries for coded fields were often empty or even wrong. Thus, first the freetext drug name field was parsed and every single word in this field was compared with the ATC classification ingredient names. Whenever a match with an official ingredient name was found the respective ingredient name and ATC code were used as the curated substitutes for the original freetext in the new mapping table. With a second parsing step all terms in the freetext field enclosed within square brackets (administration route) were provided as a list for inspection by the pharmacist. The

result of this inspection, however, showed that the route of application was not always documented correct in the brackets. Thus, it was decided not to transfer this information automatically to the target database. Instead, a list of all documented information in brackets was created, which the pharmacists corrected manually in order to use this in the next step. The same parsing was pursued with the dosage form enclosed in round brackets.

Step 2: With both information together the quality of the VMobile data was high enough, so that the pharmacist could manually curate the mapping table created in Step 1 and add the mappings between the original narrative text and the respective standard EDQM terms for the route of administration.

Step 3: The third data curation step then consisted of a further independent manual quality control process pursued by members of the pharmacist team on a randomly created subset of about 10% of the generated mapping table entries.

Step 4: In the final extraction-transformation-loading (ETL) process every single dispensation entry was automatically transferred, based on the generated mapping table, into the target database. Additionally, the patients age (in days) was calculated by date of birth and day of dispensation. The date and time of administration were also included in the target database. Demographic patient/encounter data items were transferred to the target database without any transformation.

Quality Assessment of the Target Database

Our first quality assessment of the target database applied the 3×3 Data Quality Assessment (DQA) Framework developed by Weiskopf et al.[12] We examined primarily the quality dimensions of validity "completeness" and "correctness" and secondarily the third validity dimension "currency." Thereby, we considered the two data dimensions "patients" and "variables." Owing to the anticipated use of the database, the evaluation of the third data dimension, "time," was waived.

In addition to this intra-database DQA we secondly evaluated the quality of the target database by comparing it with a gold standard database created within the AVOID II study.[13,14] This study investigated drug-related problems in cohorts before and after the introduction of the VMobil departmental EMR. The quality of the VMobil drug therapy documentation for this study was assured by a pharmacist team which conducted full chart reviews to establish a gold standard documentation for the respective time period (May–September 2014).

Statistical Analyses

In checking the conformity of the dataset to the external gold standard, we calculated Cohen's kappa (κ) for categorical data. This statistic is commonly used to measure interrater reliability, whereby we followed McHugh's advice to interpret the degree of agreement ("almost perfect" [$\kappa > 0.9$], "strong" [$0.9 > \kappa > 0.8$], "moderate" [$0.8 > \kappa > 0.6$], "weak" [$0.6 > \kappa > 0.4$], "minimal" [$0.4 > \kappa > 0.2$], and "none" [$0.2 > \kappa > 0.0$]).[15] To assess agreement for continuous variables, we used Bland–Altman plots, i.e., plots of the difference between the two measurements (y-axis) against the average of the two measurements (x-axis).[16,17] For the other analyses, proportions were provided. Where appropriate, we report a 95% confidence interval. For the statistical analyses we used Microsoft Excel and IBM SPSS 28.

Results

Dataset Overview

The VMobil EMR comprised of 19,583 patient records with 31,060 different inpatient encounters. Medication therapy was documented for 21,285 encounters (respectively, 13,934 patients). Thereby, a total of 770,158 different individual medication dispensations were recorded. These correspond to 89,955 different drug exposures.

While in the VMobil source database the different freetext medication entries comprised 6,840 different texts, those could be mapped to only 1,139 different standardized terms in the mapping table created in Step 1 of the data curation process.

The entries from the EMR database could thus be reliably assigned to the newly defined standard terms via the developed mapping table. For the different drug names the overall proportion of vocabulary with no plausible assignment to the standard terms was very low. In 1.1% (77 out of 6,840) of the entries, no ATC code could be assigned and in 0.4% (29 out of 6,840) neither an ATC code nor an active substance term could be assigned. No reliable indication of the route of administration category was given by the vocabulary in 2.7% (183 out of 6,840) of entries and 3.0% (205 out of 6,840) of the names were affected by at least one missing aspect.

The corresponding target database comprises one flat file associated with a single drug dispensation per row, associated with more than 21,000 inpatient encounters with almost 90,000 different drug exposures.

Categories of Data Quality Issues

►Fig. 1 illustrates the VMobil drug therapy order entry screen. Medication is a freetext field which should be used to describe a drug dispensation according to the syntax specification described in chapter 2.2 above. Annotation is a secondary freetext field, which is sometimes used to further specify the medication freetext.

Problems identified in the curation process related to the entry screen:

(a) Proposed default values (e.g., unit) were not corrected Since from the context it was clear for all members of the care team that in the above example not 1 mg, but 1 tablet with 800 mg was given at a single drug dispensation, physicians did not change the proposed unit default "milligram" to "tablets" (in order to speed up the ordering process). Thus, drug dosage/units combinations needed a careful manual inspection to be corrected in the curation process. It was, however, finally decided by the pharmacists team to not transfer such information into the target database because of too many inconsistent documentations, which could not fully be clarified retrospectively.

Fig. 1 Schematic illustration of the VMobil drug therapy order entry screen (translated into English). Note: The date provided in the figure is a fictional example and has no relation with any patient.

(b) Clinic internal syntax specification for the medication name freetext field

Despite the clear syntax specified numerous freetext entries not following this specification were found, e.g.,:

- NaCl 0,9% [rectal] Microlax enema dilution (clyster) instead of the drug substance name, the tradename was used and it was not written at the beginning of the text
- Mesalazin [oral] 500 mg tablets
  pharmaceutical dose form "tablets" is not enclosed in "(…)"
- Pankreatin [oral] capsules (enteric-coated)
  pharmaceutical dose form "capsules" is not enclosed in "(…);" instead an additional specification to it is enclosed in the parentheses

(c) Tradename instead of drug substance was documented

In some of those cases the drug substance name was given in the second freetext field (annotation).

- e.g., medication: Antra 10 mg (capsules)/annotation: Omeprazole 10 mg

(d) Abbreviation variety

Various abbreviations were used, e.g., for substance names, administration routes, and pharmaceutical dose forms, requiring interpretation by pharmacists from the complete order context and manual mapping to standard terminologies (ATC, EDQM).

Despite the heterogeneity of those freetext medication descriptions this freetext documentation still contained the most valid contents. Thus, the content of this field was used as the main source of information by the pharmacist creating the required mapping table.

Results of the Quality Assessment

►Table 1 shows the main results of the intra-database assessment according to the DQA framework described by Weiskopf et al[18] and the additional extra-database assessment by comparison with the independent AVOID II study documentation. Regarding the quality criterion correctness, the variables tested showed plausible distributions compared with population data and the database was internally consistent. Additionally, good agreement was found between the VMobil data and the AVOID II data (>90% for demographic data and ≥98% for medication data). Currency in terms of the time interval between weight or height measurements and medication was also good, as weight and height were recorded close to the time of medication for >96% of all drug exposures.

The intra-database assessment did not indicate any incompleteness (completeness was at least 95% for all variables). The comparison of the VMobil data with the AVOID II data allowed a detailed analysis of the validity dimension "completeness" with regard to drug exposure. Initially, we found almost 20% (214/1,131) missing drug exposures in VMobil, with over 40% (137/320) of patients affected by at least one incompleteness. However, further evaluation revealed that mainly certain drug (groups) are systematically affected by this incompleteness with 42.6% (91/214) of the missing drug exposures being due to insulins, midazolam, anesthetics, and laxative rectal solutions, and another 11.7% (25/214) to ibuprofen, diazepam, and metamizole. The remaining 98 missing exposures concern another 66 different drugs without systematic clusters. Because of our experience with the ward procedures we know that the seven main affected drug groups were subject to a different documentation workflow (e.g., initial separate documentation outside the patient curve) and thus were susceptible to systematic incompleteness. Therefore, incompleteness is not a general issue for all drugs. Additionally, the number of missing exposures is still low compared with the number of documented exposures for ibuprofen, diazepam, and

Table 1 Quality assessment by intra-database assessment according to the data quality assessment framework described by Weiskopf et al and the additional extra-database assessment by comparison with the independent AVOID II study documentation: main results of validity dimensions completeness, correctness, and currency in combination with the appropriate data dimensions

	Intra-database assessment	Extra-database assessment
Completeness—dimension of data: patients		
Number of admissions	Number of admissions as expected	100% completeness (of AVOID II admissions included in the VMobil database)
Demographic data	>95% completeness (age, sex, weight, height, length of hospital stay)	100% completeness (age, sex, weight, height, length of hospital stay)
Medication data	>99% completeness (drug name/ATC codes >96% completeness (route of administration)	57% completeness[a] (drug names/ATC codes/ route of administration)
Completeness—dimension of data: variables		
Demographic data	≥99% completeness (age, sex, length of hospital stay, weight >95% completeness (height)	100% completeness (age, sex, weight, height, length of hospital stay)
Medication data	≥99% completeness (drug names, ATC codes, route of administration)	>80% completeness[a] (drug names/ATC codes/ route of administration)
Correctness—dimension of data: patients		
Demographic data	Age and sex distribution as expected	N/A
Correctness—dimension of data: variables		
Demographic data	>98% plausible age-appropriate weight data	>90% conformity of age, sex, weight, height
Medication data	N/A	≥98% conformity of medication data
Currency—dimension of data: variables		
Demographic/ medication data	>94% weight and height recorded close to time of medication	N/A

Abbreviation: N/A, not applicable.
[a]The analysis of which drugs were affected by missing exposure data showed that >54% of these are attributable to seven drug /groups (insulins, midazolam, anaesthetics, laxative rectal solutions, ibuprofen, diazepam, and metamizole). These medications were susceptible to systematic incompleteness because of documentation workflows. For ibuprofen, diazepam, metamizole, however, the number of missing exposures is still very low compared with the number of documented exposures; thus, the effect of missing exposures in drug utilization studies is likely to be small.

metamizole (≤10%). Thus for these, the effect of missing exposures in drug utilization studies is likely to be small.

Further details of the quality assessment results can be found in the ►**Supplementary Materials 1** and **2 (available in the online version only).**

Discussion

In their 2020 systematic review of observational studies published in 2016 that used real-world data to assess medication safety or effectiveness in children. Lasky et al identified 29 relevant publications.[7] In almost half of the studies, however, medication exposure was reported only at the class level, without details about the specific drug entities used. They further noted that fewer than a third used administrative claims or electronic health records databases, whereas the majority used elaborate and time-consuming manual chart review and primary data collected in single institutions. They concluded that their systematic review documents that the high level of activity in RWE in general has had less of an impact on pediatrics[7] and illustrated the need for establishing pediatric real-world drug utilization databases. In 2021 Horton et al published a commentary about RWE for assessing treatment effectiveness and safety in pediatric populations, illustrating opportunities, and also challenges for generating such RWE. They also concluded that rises in pediatric medication use, overall and off-label, on one side and an increasing availability of real-world data on the other, could provide valuable and novel opportunities to harness RWE and improve children's health.[18]

In our literature search, we found only few observational studies of pediatric inpatients that applied real-world databases. These included investigations that employed the U.S. databases, "Premier Perspective Database," or "Pediatric Health Information System," such as drug utilization studies.[19–24] Outside the United States, there was even less evidence of research that involved EMR databases. Two examples are Chinese studies that used single hospital databases of EMR data to explore signals of drug-induced renal injury[25] or to assess medication safety in pediatrics by analyzing clinical conditions.[26] However, both studies did not mention any assessments or measures to ensure data quality while using EMR data. Choi et al reported the establishment of a neonatal database for observational studies in Ruanda to monitor the quality of care for neonates in the academic setting exemplary for a low-income country. This database did not have a direct focus on medication data. However, similar to ours, it was found that there were

significant challenges in optimizing this database during the process.[27] Reinecke et al[28] describe three algorithms that extract medication data from the electronic health record, assess them for their structuredness, and improve them. However, their study focuses on the algorithms' description and not on the use or usability of the data. Kapsner et al[29] have recently implemented a comprehensive real-world data-based data quality framework, but no results from the application of this framework have yet been published.

Thus, our work may be seen as one of very few efforts to establish a real-world drug utilization database and further make this database openly available upon reasonable requests. As described above, we could successfully achieve all aims of our investigation: firstly, we were able to investigate and even categorize the data quality issues of the original data; secondly, we established a data curation process; thirdly, the planned pediatric drug therapy observational research database was established; and finally, we assessed data quality of the established observational research database.

The most important lesson learned in our project was that optimizing electronic medication ordering software is a multifaceted complex challenge. On one side one would aim at fully structured data entry, strictly based on standardized drug substance catalogues and standardized medication ordering terminologies to enhance data quality for secondary use. On the other side, however, drug order data entry as freetext may still be preferred and appreciated by clinicians for routine care documentation in stressful daily practice. Therefore, the design of an electronic medication ordering tool, which today should typically provide some kind of clinical decision support and thus needs to be approved and certified according to the medical device regulation, always requires a careful and risk-adapted consideration of usability, patient safety aspects, clinical processes, and secondary use requirements.

The VMobil user interface for example includes many input fields for the necessary information like (date of) prescription, medication, route of application, unit, time period (for the prescription), pharmaceutical dosage form, and dose for individual time point dispensations. Even though most of those data elements were captured in a structured manner based on predefined value lists, two of them (medication and annotation, see ►Fig. 1) were freetext fields. Thus, in daily clinical practice, due to time constraints, often only the minimal information necessary for clinical care and legal requirements was documented. Therefore, the entries were not standardized as in a research database and consequently toward this end, the VMobil data displayed some inconsistencies. This did not cause problems in daily clinical practice, as medical staff can nonetheless understand and interpret the data correctly, based on their context knowledge. Without this context knowledge, interpretation and curation of the VMobil database contents required a partially manual analysis and interpretation of the database, involving an interdisciplinary team of database specialists and researchers with pharmaceutical background. Within this process we could detect different categories of data quality issues in the VMobil database and successfully establish data curation strategies to at least partially cope with those issues. As a result, we were able to create a database to be used for observational drug utilization research comprising more than 21,000 inpatient encounters with almost 90,000 different drug exposures. The database has a completeness of >95% for linked clinical information such as age, sex, weight, height, and length of hospital stay. In our quality assessment, we were able to verify the reliability of the database (►Supplementary Materials 1 and 2 [available in the online version only]). Hence, it was already valuable in a first application to analyze the safety profile of metamizole use in children.[30]

Limitations

As reported above we could not achieve to transfer all drug dispensation data from the VMobil system to our anonymized real-world database. Similar to the experience reported by Cars et al[31] in our project the quality of drug dispensation dosage information (extracted from the unstructured freetext field in the VMobil database) would have been too low, so it was decided to omit this information in our target database. Further, within the extra-database quality assessment an incompleteness concerning drug exposures was detected, which, however, was restricted to only few drug (classes). Since this could be explained with a different documentation workflow on the four pediatric wards particularly for those limited drug classes, it was an isolated effect that did not occur for all other drug exposures.

The categories of quality issues detected and described above relate to our unique setting with one particular commercial EMR system. Due to different workflows and software, problems may vary in other settings. However, the problem of extensive freetext documentation is still an issue in many electronic health records products. Thus, we are convinced that similar problems may arise when data from other commercial drug ordering systems shall be used for medical research. Therefore, the description of our quality issue categories may be helpful in similar projects at other locations. Furthermore, the data quality of our harmonized observational database can be assessed before it is used by an existing freely available framework.[32]

Conclusion

Our project has illustrated that using medication data from clinical documentation for medical-/pharmaceutical research and creation of a real-world drug utilization database is possible. It has however also pointed out that careful analysis of the original clinical database and reflecting its contents with background knowledge about the original data entry user interface, as well as clinicians' preferences in the resulting documentation processes, are important steps to identify the need for data curation within the ETL processes. Even though manual data curation steps may be required, this is a worthwhile effort to be invested in order to achieve high data quality.

Currently, the value of our drug utilization database is still limited, since the drug administration data are not yet linked

with comprehensive additional clinical data. Such clinical data (e.g., laboratory results, patients' diagnosis, and procedures performed), however, have been provided in the FHIR-based research data repository of our university hospital within the German Medical Informatics Initiative (MII)[33] as a result of the MIRACUM project.[34] Therefore, we have already initiated projects to (1) transfer the contents of our drug utilization database into the fast healthcare interoperability resource (FHIR) format (according to the implementation guidelines of the MII) and thus integrate it with the comprehensive set of clinical data from many other clinical systems of our university hospital and (2) to map it into the observational medical outcomes partnership (OMOP) common data model, which we have established at Erlangen University Hospital to be able to participate in international data sharing projects of the OHDSI consortium.[35]

Clinical Relevance Statement

The work presented here has created a database that can be used for drug utilization research and illustrated the need for data curation processes.

Multiple-Choice Questions

1. Which is the hardest challenge in creating a research database out of clinical documentation?
   a. Time lost in documentation
   b. Data quality in documentation
   c. Too many different patients
   d. Interoperability of the systems

   Correct Answer: The correct answer is option b. The gaps and inconsistencies in documentation make it necessary to transform and create some entries manually. It also needs medical and pharmaceutical as well as internal knowledge about the procedures of the clinic and wards.

2. According to which main drug ordering information were the data prepared?
   a. Patients
   b. Encounters
   c. Prescriptions
   d. Dispensations

   Correct Answer: The correct answer is option d. The entries of the created research database have the dispensation ID as primary key and are therefore structured accordingly.

Protection of Human and Animal Subjects

A positive ethics vote exists for the study (Application No 561_20 BC, ethics commission of the Friedrich-Alexander-Universität chaired by Prof. Dr. med. Renke Maas).

Supplementary Materials

The ►**Supplementary Material 1** (available in online version only) is a detailed description of the thorough intra-database quality assessment. The ►**Supplementary Material 2** (available in online version only) is a detailed description of the thorough extra-database quality assessment.

Note

The data that support the findings of this study are available from the senior author S.E. (Sonja.eberl@uk-erlangen.de), upon reasonable request.

Conflict of Interest

None declared.

Acknowledgments

The present work was performed in (partial) fulfilment of the requirements for obtaining the Degree "Dr. rer. biol. hum." from the Friedrich-Alexander-Universität Erlangen-Nürnberg (W.R.).

References

1 Conroy S, Choonara I, Impicciatore P, et al; European Network for Drug Investigation in Children. Survey of unlicensed and off label drug use in paediatric wards in European countries. BMJ 2000; 320(7227):79–82
2 Frattarelli DA, Galinkin JL, Green TP, et al; American Academy of Pediatrics Committee on Drugs. Off-label use of drugs in children. Pediatrics 2014;133(03):563–567
3 Gore R, Chugh PK, Tripathi CD, Lhamo Y, Gautam S. Pediatric off-label and unlicensed drug use and its implications. Curr Clin Pharmacol 2017;12(01):18–25
4 Knopf H, Wolf IK, Sarganas G, Zhuang W, Rascher W, Neubert A. Off-label medicine use in children and adolescents: results of a population-based study in Germany. BMC Public Health 2013; 13:631
5 Mühlbauer B, Janhsen K, Pichler J, Schoettler P. Off-label use of prescription drugs in childhood and adolescence: an analysis of prescription patterns in Germany. Dtsch Arztebl Int 2009;106 (03):25–31
6 Kimland E, Odlind V. Off-label drug use in pediatric patients. Clin Pharmacol Ther 2012;91(05):796–801
7 Lasky T, Carleton B, Horton DB, et al. Real-world evidence to assess medication safety or effectiveness in children: systematic review. Drugs Real World Outcomes 2020;7(02):97–107
8 Lu Y, Van Zandt M, Liu Y, et al. Analysis of dual combination therapies used in treatment of hypertension in a multinational cohort. JAMA Netw Open 2022;5(03):e223877
9 The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology. Published June 30, 2022. Accessed August 22, 2023 at: https://www.encepp.eu/standards_and_guidances/methodologicalGuide.shtml
10 Bleisinger B, Hartmann P, Schels T. Elektronische Patientenkurve: Schluss mit Notizzetteln. Deutsches Ärzteblatt 2013. Accessed February 11, 2022 at: https://www.aerzteblatt.de/archiv/136891/Elektronische-Patientenkurve-Schluss-mit-Notizzetteln
11 Council of Europe EDQM Standard Terms - Internal Controlled Vocabularies for Pharmaceutical Dose Forms
12 Weiskopf NG, Bakken S, Hripcsak G, Weng C. A data quality assessment guideline for electronic health record data reuse. EGEMS (Wash DC) 2017;5(01):14
13 Toni I, Wimmer S, Rascher W, Neubert A. AVOID - Kann eine elektronische Patientenkurve mit Dosierungsrechner die Arzneimitteltherapiesicherheit in der Pädiatrie erhöhen? 5. Deutscher Kongress für Patientensicherheit bei medikamentöser Therapie - Arzneiverordnung in der Praxis. Accessed August 22, 2023 at: https://www.akdae.de/arzneimitteltherapie/arzneiverordnung-in-der-praxis/

ausgaben-archiv/ausgaben-ab-2015/ausgabe/artikel/2018/2018-04/avoid-kann-eine-elektronische-patientenkurve-mit-dosierungsrechner-die-arzneimitteltherapiesicherheit-in-der-padiatrie-erhohen2019

14 Wimmer S, Toni I, Botzenhardt S, Trollmann R, Rascher W, Neubert A. Impact of a computerized physician order entry system on medication safety in pediatrics-the AVOID study. Pharmacol Res Perspect 2023;11(03):e01092
15 McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(03):276–282
16 Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: measures of agreement. Perspect Clin Res 2017;8(04):187–191
17 Zaki R, Bulgiba A, Ismail R, Ismail NA. Statistical methods used to test for agreement of medical instruments measuring continuous variables in method comparison studies: a systematic review. PLoS One 2012;7(05):e37908
18 Horton DB, Blum MD, Burcu M. Real-world evidence for assessing treatment effectiveness and safety in pediatric populations. J Pediatr 2021;238:312–316
19 Lasky T, Ernst FR, Greenspan J, Wang S, Gonzalez L. Estimating pediatric inpatient medication use in the United States. Pharmacoepidemiol Drug Saf 2011;20(01):76–82
20 Lasky T, Greenspan J, Ernst FR, Gonzalez L. Morphine use in hospitalized children in the United States: a descriptive analysis of data from pediatric hospitalizations in 2008. Clin Ther 2012;34 (03):720–727
21 Feudtner C, Dai D, Faerber J, Metjian TA, Luan X. Pragmatic estimates of the proportion of pediatric inpatients exposed to specific medications in the USA. Pharmacoepidemiol Drug Saf 2013;22(08):890–898
22 Griffith HG, Dantuluri K, Thurm C, et al. Considerable variability in antibiotic use among US children's hospitals in 2017-2018. Infect Control Hosp Epidemiol 2020;41(05):571–578
23 Dai D, Feinstein JA, Morrison W, Zuppa AF, Feudtner C. Epidemiology of polypharmacy and potential drug-drug interactions among pediatric patients in ICUs of U.S. children's hospitals. Pediatr Crit Care Med 2016;17(05):e218–e228
24 Getz KD, Miller TP, Seif AE, et al. Opioid utilization among pediatric patients treated for newly diagnosed acute myeloid leukemia. PLoS One 2018;13(02):e0192529
25 Yu Y, Nie X, Zhao Y, et al. Detection of pediatric drug-induced kidney injury signals using a hospital electronic medical record database. Front Pharmacol 2022;13:957980
26 Yu G, Zeng X, Ni S, et al. A computational method to quantitatively measure pediatric drug safety using electronic medical records. BMC Med Res Methodol 2020;20(01):9
27 Choi J, Urubuto F, Dusabimana R, et al. Establishing a neonatal database in a tertiary hospital in Rwanda - an observational study. Paediatr Int Child Health 2019;39(04):265–274
28 Reinecke I, Siebel J, Fuhrmann S, et al. Assessment and improvement of drug data structuredness from electronic health records: algorithm development and validation. JMIR Med Inform 2023;11:e40312
29 Kapsner LA, Mang JM, Mate S, et al. Linking a consortium-wide data quality assessment tool with the MIRACUM metadata repository. Appl Clin Inform 2021;12(04):826–835
30 Zahn J, Eberl S, Rödle W, Rascher W, Neubert A, Toni I. Metamizole use in children: analysis of drug utilisation and adverse drug reactions at a German university hospital between 2015 and 2020. Paediatr Drugs 2022;24(01):45–56
31 Cars T, Wettermark B, Malmström RE, et al. Extraction of electronic health record data in a hospital setting: comparison of automatic and semi-automatic methods using anti-TNF therapy as model. Basic Clin Pharmacol Toxicol 2013;112(06):392–400
32 Schmidt CO, Struckmann S, Enzenbach C, et al. Facilitating harmonized data quality assessments. A data quality framework for observational health research data collections with software implementations in R. BMC Med Res Methodol 2021;21(01):63
33 Semler SC, Wissing F, Heyder R. German medical informatics initiative. Methods Inf Med 2018;57(S 01):e50–e56
34 Prokosch HU, Acker T, Bernarding J, et al. MIRACUM: medical informatics in research and care in university medicine. Methods Inf Med 2018;57(S 01):e82–e91
35 Maier C, Lang L, Storf H, et al. Towards implementation of OMOP in a German university hospital consortium. Appl Clin Inform 2018; 9(01):54–61