German Medical Data Sciences 2023 — Science. Close to People.
R. Röhrig et al. (Eds.)

doi:10.3233/SHTI230695

Modelling Adverse Events with the TOP Phenotyping Framework

Christoph BEGER[a,1], Anna Maria BOEHMER[b], Beate MUSSAWY[c], Louisa REDEKER[d], Franz MATTHIES[a], Ralph SCHÄFERMEIER[a], Annette HÄRDTLEIN[e], Tobias DREISCHULTE[e], Daniel NEUMANN[a,2], and Alexandr UCITELI[a,2]

[a] *Institute for Medical Informatics, Statistics and Epidemiology, Leipzig University*
[b] *Institute of Pharmacy, Department of Clinical Pharmacy, University of Bonn*
[c] *Hospital Pharmacy, University Medical Center Hamburg-Eppendorf*
[d] *Department of Clinical Pharmacology, School of Medicine, Faculty of Health, Witten/Herdecke University*
[e] *Institute of General Practice and Family Medicine, University Hospital, LMU Munich*

Abstract. The detection and prevention of medication-related health risks, such as medication-associated adverse events (AEs), is a major challenge in patient care. A systematic review on the incidence and nature of in-hospital AEs found that 9.2% of hospitalised patients suffer an AE, and approximately 43% of these AEs are considered to be preventable. Adverse events can be identified using algorithms that operate on electronic medical records (EMRs) and research databases. Such algorithms normally consist of structured filter criteria and rules to identify individuals with certain phenotypic traits, thus are referred to as phenotype algorithms. Many attempts have been made to create tools that support the development of algorithms and their application to EMRs. However, there are still gaps in terms of functionalities of such tools, such as standardised representation of algorithms and complex Boolean and temporal logic. In this work, we focus on the AE delirium, an acute brain disorder affecting mental status and attention, thus not trivial to operationalise in EMR data. We use this AE as an example to demonstrate the modelling process in our ontology-based framework (TOP Framework) for modelling and executing phenotype algorithms. The resulting semantically modelled delirium phenotype algorithm is independent of data structure, query languages and other technical aspects, and can be run on a variety of source systems in different institutions.

Keywords. algorithms, adverse events, electronic health records, computable phenotypes

1. Introduction

The Medical Informatics Initiative (MII) aims to support medical research and improve patient care through IT solutions with real world data. This can support tailored and personalised diagnostic and treatment decisions, create new insights for effective and sustainable disease control, and contribute to the continuous improvement of healthcare.

[1] Corresponding Author: Christoph Beger, E-mail: christoph.beger@imise.uni-leipzig.de.
[2] Shared last authorship: DN for POLAR and AU for TOP.

Within the MII, clinical use cases shall bridge the gap between patient care, different MII strategies and clinical and bioinformatics research. One of such use cases is POLAR ('POLypharmacy, drug interActions, Risks'), which uses the methods and processes of the MII to contribute to the detection of medication-related health risks, such as medication-associated adverse events (AEs) [1]. As described by de Vries et al., "an AE is usually defined as an unintended injury or complication that results in prolonged hospitalisation, disability at discharge or death, and is caused by healthcare management rather than the patient's underlying disease process" [2]. According to a systematic review on the incidence and nature of in-hospital AEs, 9.2% of hospitalised patients suffer an AE, and approximately 43% of these AEs are considered to be preventable [2].

AEs are documented in electronic medical records (EMRs) or can be identified using algorithms that operate on EMR. Such algorithms normally consist of structured filter criteria and rules to "identify individuals who exhibit certain phenotypic traits, such as the same diseases, characteristics, or set of comorbidities" [3], thus are referred to as phenotype algorithms (PAs). Mo et al. [4] and Chapman et al. [3] compiled desiderata for the development of EMR-based PAs that outline many important aspects to facilitate development and reusability.

PAs must be represented in a standardised way to enable sharing and reuse in other institutions. Typically, models are shared in very heterogeneous formats and in many cases only in textual form, making the application to other institutions a very challenging task [5,6]. Other formats include specific programming or query languages such as R, Python and, in particular, the Clinical Quality Language, which has a high expressive power and supports the representation of complex phenotype definitions [7].

Many attempts have been made to create tools that support development of algorithms and application to electronic health records [8–10]. Yet, there are still gaps in terms of functionalities of such tools, such as standardised representation of algorithms and complex Boolean and temporal logic [9]. Especially complex logic is in fact of high importance for many PAs [5].

In this sense, the MII junior research group 'Terminology and Ontology-based Phenotyping (TOP)' (part of the MII SMITH consortium [11]) aims to develop an easily applicable ontology-based framework (TOP Framework) for modelling and executing PAs. The most important advantage of our approach is a clear separation between the modelling of the domain knowledge (by medical staff, biometricians, etc., i.e., non-IT experts) and the implementation. The semantically modelled PAs are thus independent of the structure of the data, query languages as well as other technical aspects and can be executed on different source systems. In this work, we focus on the modelling of PAs, i.e. domain knowledge (exemplified by the detection of adverse events) by domain experts. The ontological and technical aspects of the framework are outlined in the chapter 'Methods' and explained in more detail in separate publications.

2. Methods

2.1. Example Adverse Event: Delirium

Delirium is an acute brain disorder that affects mental status and attention, triggered by an acute event (e.g., infection, surgery) [12]. The incidence of delirium is remarkably high, ranging from 4% to 54% depending on the patient population, screening method and study design [13]. Detection and prevention of delirium is a major challenge in

hospitals, as this condition is associated with a wide range of serious adverse outcomes, including mortality, increased length of stay, long-term cognitive impairment, and risk of institutionalisation (e.g., nursing home placement, hospital readmission) [14,15]. Despite its high prevalence and serious consequences, delirium is significantly underdiagnosed in clinical practice [16]. An algorithm that diagnoses delirium due to any cause based on factors automatically extracted from EMRs can provide rapid and continuous screening for prevalent delirium, as well as improved retrospective assessment. Such an algorithm should include more than the International Classification of Diseases, 10th Revision (ICD-10) diagnosis codes, as the incidence and prevalence of delirium is underestimated when determined by ICD-10-coded diagnoses alone [17–20].

Halladay et al. constructed an electronic prediction rule for prevalent delirium based on the NICE meta-analysis of risk factors for delirium, consisting of ICD-10 diagnosis codes for cognitive impairment as well as age, infection, and sodium level [21]. On the basis of this work as well as the evaluation of the performance of ICD-10 codes and antipsychotic medication use in the identification of delirium by Kim et al., we developed an algorithm for the automatic detection of delirium due to any cause in EMR data as part of the POLAR project [17,21]. The resulting algorithm was developed in the TOP Framework and is described in Section 3.

2.2. Ontological Data Model of Phenotypes

The data model used in the TOP Framework (TOP API) [22] is based on the Core Ontology of Phenotypes [23]. Here, we consider phenotypes as individual characteristics, such as the weight of a person, but also complex (composite or derived) properties such as the body mass index (BMI) or the sequential organ failure assessment score (SOFA score) of a person. Abstract entities instantiated by phenotypes are called phenotype classes. We distinguish between single (e.g., age, weight, height) and composite phenotypes (e.g., BMI, SOFA score), which are made up of other phenotypes.

Single phenotype classes are specified by giving the title, data type, unit of measurement, description, medical terminology codes, etc. The terminology associations play the most important role. Established terminologies (such as Logical Observation Identifiers Names and Codes (LOINC), ICD-10, etc.) are used to semantically describe the phenotypes and enable comparability of the data, but also to map to the data sources and to generate and execute the necessary queries.

The composite phenotype classes are specified by an evaluable expression representing either a phenotype class, a constant or a function with any number of arguments, allowing nesting. For example, the expression of the phenotype class BMI, which is calculated by dividing body weight by height to the square, can be represented as "quotient(weight, power(height, 2))". Not only mathematical, but also logical and ontological functions are supported, and the set of supported functions is extensible.

2.3. Terminology- and Ontology-based Phenotyping Framework

The TOP Framework consists of three main components (see Figure 1), namely a user interface (frontend), a REST server for storing phenotype definitions (backend), and various services that use phenotype models to perform specific tasks. In this section frontend, backend, and the service component for the execution of phenotypic queries

are outlined briefly. A detailed description of the whole framework is planned for future work, but for the time being documentation is available in our public GitHub repository[3].

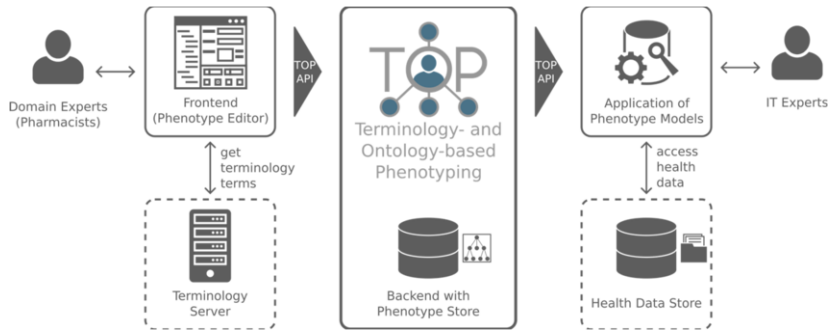

Figure 1. TOP Framework Overview. Domain experts create and use phenotype models via the Phenotype Editor, while IT experts are responsible for implementing interfaces to clinical and trial data in Health Data Stores. The backend conforms to the TOP API specification for all inbound and outbound data flows.

The TOP frontend is a JavaScript web application that allows domain experts to collaboratively manage organisations and repositories containing phenotype definitions. In the TOP Framework, organisations act as a simple permission system. Users only have access to public repositories or repositories of organisations to which they belong to. The frontend has an intuitive graphical user interface (Phenotype Editor), where phenotypes are displayed in a tree structure and metadata of a phenotype can be defined via input fields. Submissions are sent to the TOP backend via the TOP API.

The most important component of the TOP Framework is the backend, where all phenotype models are stored in a relational database management system. The backend is a Spring Boot REST server that can be integrated in a local setup with an OAuth2 authentication server. This allows institutions to reuse authentication mechanisms already in place and improve security. A plugin mechanism is available to connect phenotype services to the backend and extend its functionality.

PAs can be defined by specifying (single or composite) phenotype classes as inclusion/exclusion criteria. The phenotypic query service generates queries for all criteria, translates them into the corresponding query language of the source system, and executes them using an adapter [24,25]. The query results are used to evaluate expressions of the composite phenotypes. The result of a PA is a set of individuals that match all criteria, i.e. possess the corresponding phenotypes.

3. Results

In this section, the modelling process of the TOP Framework as well as the resulting delirium algorithm are described in detail.

When developing an algorithm in TOP, it is generally recommended to start with the basic data material needed to evaluate the algorithm. In the case of delirium, the data materials are properties of individuals that are directly related to entries in the EMR, namely the subject's current age, sodium level in blood, diagnostic conditions (e.g.,

[3] Public repository on GitHub: https://github.com/Onto-Med/top-deployment

cognitive impairment), medication administrations, and performed operations. All of these properties are mapped to unrestricted single phenotype classes with a corresponding data type (see Table 1). They also need semantic descriptions by terms of standard terminologies (e.g., LOINC), which can easily be added in the TOP Framework by selecting a terminology and providing the codes of the terms.

Table 1. Mapping of delirium items to TOP phenotype class types (excerpt). Categories were added to group similar phenotype classes. Other available data types: 'string', 'date/time'. Abbreviations: International Classification of Diseases, 10th Revision (ICD-10), Anatomical Therapeutic Chemical (ATC) classification system, Operation and Procedure Classification System (OPS), Logical Observation Identifiers Names and Codes (LOINC)

	Delirium item	**FHIR Resource (Coding System)**	**TOP phenotype class type**	**Data type**
Step 0	Parameters	-	Category	-
Step 1	Coded diagnoses	Condition (ICD-10)	Single Phenotype	Boolean
	Age	Patient & Encounter	Single Phenotype	numeric
	Antipsychotic medication	Medication (ATC)	Single Phenotype	Boolean
	Operation	Procedures (OPS)	Single Phenotype	Boolean
	Sodium level	Observations (LOINC)	Single Phenotype	numeric
Step 2	Age above 80 years	-	Restricted Single Phenotype	Boolean
Step 3	2/3 factors	-	Composite Phenotype	Boolean
	Extended algorithm	-	Composite Phenotype	Boolean

The next step is to check whether the algorithm requires some properties of an individual to match certain criteria. These cases are modelled as restricted single phenotype classes (subclasses of the corresponding unrestricted single phenotype classes). For instance, the delirium algorithm contains a check for the subject's current age to be higher than 80 years, which means that a restricted single phenotype class must be added as a subclass of the phenotype class 'age'. This new restricted phenotype class has a value range restriction as shown in Figure 2.

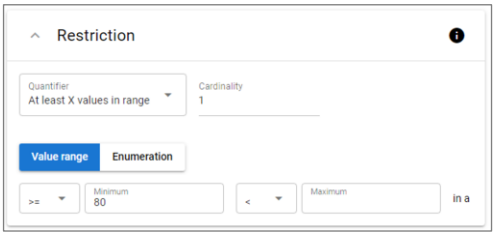

Figure 2. Input form to define value range restriction of the restricted single phenotype class for subject's age above 80 years. One can also provide a set of valid values (enumeration). Quantifier and cardinality are used to restrict how many values of the phenotype should match the restriction. Other quantifier options are 'all values in range', 'exactly x values in range', and 'not more than x values in range'.

In the final step of the modelling process composite phenotype classes are developed, which are made up of unrestricted or restricted single phenotype or other composite phenotype classes (as described in Section 2.2). Composite phenotype classes have an expression consisting of a function and a set of arguments, which can also be

expressions. To simplify the construction of expressions, the TOP frontend has a special input form (see Figure 3, where users can select a function from a list of available functions. In the next step, they are prompted for function arguments. An argument is defined by searching for a phenotype in the current repository, selecting a constant from a predefined list, or entering a numeric, date, or string value. The final modelling step is repeated until all algorithm computations are mapped to the TOP Framework. Typically, this results in a composite phenotype class that directly or transitively references all other phenotype classes in the algorithm. In the case of delirium, the phenotype 'Extended algorithm' was defined that combines all other phenotypes in a logical expression. If the expression evaluates to true, the subject is considered to have a delirium.

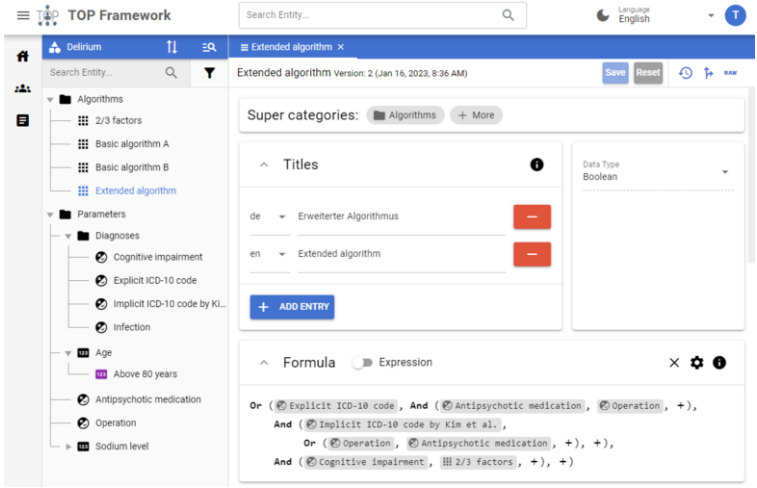

Figure 3. TOP Framework Screenshot. Shown is the delirium repository that contains all phenotype classes (left side) relevant to the phenotype algorithm for detection of delirium. On the right side, the definition of the phenotype 'Extended algorithm' is displayed, with a formula that is composed from logical functions and other phenotype classes as arguments. A high-resolution version is available in the attachments.

4. Discussion

Several approaches have been used to develop EMR-based PAs. For instance, the Quality Data Model (QDM) has been reported as a promising format for use in clinical research [26–28]. The QDM allows PAs to be represented in a structured, machine-interpretable form using terminologies for clinical entities and features. Unfortunately, QDM has no support for sharing logic between algorithms, which could lead to reimplementation of subcomponents and make portability difficult [26].

Recent developments in machine learning and especially deep neural networks led to new models for patient classification based on EMR data [29,30]. At first glance, this development contradicts our approach, as we use rules-based algorithms based on expert knowledge. However, we believe that expert knowledge is the most important aspect for the development of PAs (rules-based or machine learning [31]). Furthermore, rules-based algorithms are explainable, can be developed by domain experts, and are less error-prone to the data quality of EMRs. PAs created with the TOP Framework are exchangeable in a standardised format (TOP API specification) and can even serve as

the basis for machine learning models, because the knowledge of the domain expert is already included in the model.

An important limitation of the TOP Framework is its restriction to structured data. It is not possible to classify (or to calculate) phenotypes described in unstructured text. In order to combine our approach with unstructured data, natural language processing must be used to extract and to structure information that should be used by TOP PAs. The framework allows the creation of expressions of any complexity and even includes the specification of temporal logic. This allows a wide range of scenarios to be covered, but also negatively correlates with the match-rate of PAs [32].

Our next step is to extend the TOP Framework with new features. We plan to add more composite phenotype expression functions that are commonly used in PAs, and to support terminology versioning. Another important aspect is the provision of rich metadata and the adoption of existing standards to enhance the findability and interoperability of phenotype definitions. We will use the Resource Description Framework and established terms to provide appropriate metadata. Finally, we will model and publish further PAs from POLAR and other MII use cases utilising the TOP Framework. For example, the presented algorithm detects the presence of delirium due to any cause. In a subsequent step, this algorithm will be combined with potentially delirium-causing drugs to detect drug-induced delirium.

5. Conclusion

The TOP Framework is a platform that is designed for domain experts to define and manage phenotype specifications. In this work, we present an exemplary application of the framework to the modelling of an algorithm for the detection of the adverse event delirium. We show that the whole algorithm can be built with the framework, and we provide the resulting algorithm as a structured phenotype model.

Declarations

Conflict of Interest: The authors declare that there is no conflict of interest.

Author Contributions: CB, AU: Conceptualization, Methodology, Writing - Original Draft, Review and Editing, Validation; AB, BM, LR, AH, TD: Methodology, Writing - Original Draft, Review and Editing; FM, RS, DN: Writing - Original Draft; DN, AU: Supervision All authors have approved the manuscript as submitted and accept responsibility for the scientific integrity of the work.

Acknowledgement

The work on ‘Modelling Adverse Events with the TOP Phenotyping Framework’ within the MII junior research group 'Terminology and Ontology-based Phenotyping’, the overarching use case of the German Medical Informatics Initiative ‘POLAR_MI - POLypharmacy, Drug interActions, Risks’ and the MII SMITH consortium is supported

by the German Federal Ministry of Education and Research (BMBF; 01ZZ2018, 01ZZ1910[A-Z], and 01ZZ1803A)

References

[1] Scherag A, Andrikyan W, Dreischulte T, Dürr P, Fromm MF, Gewehr J, et al. POLAR – „POLypharmazie, Arzneimittelwechselwirkungen und Risiken“ – wie können Daten aus der stationären Krankenversorgung zur Beurteilung beitragen? Prävent Gesundheitsförderung 2022. https://doi.org/10.1007/s11553-022-00976-8.

[2] de Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA. The incidence and nature of in-hospital adverse events: a systematic review. Qual Saf Health Care 2008;17:216–23. https://doi.org/10.1136/qshc.2007.023622.

[3] Chapman M, Mumtaz S, Rasmussen LV, Karwath A, Gkoutos GV, Gao C, et al. Desiderata for the development of next-generation electronic health record phenotype libraries. GigaScience 2021;10:giab059. https://doi.org/10.1093/gigascience/giab059.

[4] Mo H, Thompson WK, Rasmussen LV, Pacheco JA, Jiang G, Kiefer R, et al. Desiderata for computable representations of electronic health records-driven phenotype algorithms. J Am Med Inform Assoc 2015;22:1220–30. https://doi.org/10.1093/jamia/ocv112.

[5] Conway M, Berg RL, Carrell D, Denny JC, Kho AN, Kullo IJ, et al. Analyzing the heterogeneity and complexity of Electronic Health Record oriented phenotyping algorithms. AMIA Annu Symp Proc AMIA Symp 2011;2011:274–83.

[6] Kirby JC, Speltz P, Rasmussen LV, Basford M, Gottesman O, Peissig PL, et al. PheKB: a catalog and workflow for creating electronic phenotype algorithms for transportability. J Am Med Inform Assoc 2016;23:1046–52. https://doi.org/10.1093/jamia/ocv202.

[7] Brandt PS, Kiefer RC, Pacheco JA, Adekkanattu P, Sholle ET, Ahmad FS, et al. Toward CROSS-PLATFORM electronic health record -DRIVEN phenotyping using Clinical Quality Language. Learn Health Syst 2020;4. https://doi.org/10.1002/lrh2.10233.

[8] Rasmussen LV, Kiefer RC, Mo H, Speltz P, Thompson WK, Jiang G, et al. A Modular Architecture for Electronic Health Record-Driven Phenotyping. AMIA Jt Summits Transl Sci Proc AMIA Jt Summits Transl Sci 2015;2015:147–51.

[9] Xu J, Rasmussen LV, Shaw PL, Jiang G, Kiefer RC, Mo H, et al. Review and evaluation of electronic health records-driven phenotype algorithm authoring tools for clinical and translational research. J Am Med Inform Assoc 2015;22:1251–60. https://doi.org/10.1093/jamia/ocv070.

[10] Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. Stud Health Technol Inform 2015;216:574–8.

[11] Winter A, Stäubert S, Ammon D, Aiche S, Beyan O, Bischoff V, et al. Smart Medical Information Technology for Healthcare (SMITH): Data Integration based on Interoperability Standards. Methods Inf Med 2018;57:e92–105. https://doi.org/10.3414/ME18-02-0004.

[12] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Fifth Edition. American Psychiatric Association; 2013. https://doi.org/10.1176/appi.books.9780890425596.

[13] Gibb K, Seeley A, Quinn T, Siddiqi N, Shenkin S, Rockwood K, et al. The consistent burden in published estimates of delirium occurrence in medical inpatients over four decades: a systematic review and meta-analysis study. Age Ageing 2020;49:352–60. https://doi.org/10.1093/ageing/afaa040.

[14] Siddiqi N, House AO, Holmes JD. Occurrence and outcome of delirium in medical in-patients: a systematic literature review. Age Ageing 2006;35:350–64. https://doi.org/10.1093/ageing/afl005.

[15] Pandharipande PP, Girard TD, Jackson JC, Morandi A, Thompson JL, Pun BT, et al. Long-Term Cognitive Impairment after Critical Illness. N Engl J Med 2013;369:1306–16. https://doi.org/10.1056/NEJMoa1301372.

[16] Geriatric Medicine Research Collaborative. Delirium is prevalent in older hospital inpatients and associated with adverse outcomes: results of a prospective multi-centre study on World Delirium Awareness Day. BMC Med 2019;17:229. https://doi.org/10.1186/s12916-019-1458-7.

[17] Kim DH, Lee J, Kim CA, Huybrechts KF, Bateman BT, Patorno E, et al. Evaluation of algorithms to identify delirium in administrative claims and drug utilization database: Delirium Identification in Claims Data. Pharmacoepidemiol Drug Saf 2017;26:945–53. https://doi.org/10.1002/pds.4226.

[18] Bui LN, Pham VP, Shirkey BA, Swan JT. Effect of delirium motoric subtypes on administrative documentation of delirium in the surgical intensive care unit. J Clin Monit Comput 2017;31:631–40. https://doi.org/10.1007/s10877-016-9873-1.

[19] McCoy TH, Snapper L, Stern TA, Perlis RH. Underreporting of Delirium in Statewide Claims Data: Implications for Clinical Care and Predictive Modeling. Psychosomatics 2016;57:480–8. https://doi.org/10.1016/j.psym.2016.06.001.
[20] Casey P, Cross W, Mart MW-S, Baldwin C, Riddell K, Dārziņš P. Hospital discharge data under-reports delirium occurrence: results from a point prevalence survey of delirium in a major Australian health service. Intern Med J 2019;49:338–44. https://doi.org/10.1111/imj.14066.
[21] Halladay CW, Sillner AY, Rudolph JL. Performance of Electronic Prediction Rules for Prevalent Delirium at Hospital Admission. JAMA Netw Open 2018;1:e181405. https://doi.org/10.1001/jamanetworkopen.2018.1405.
[22] Beger C, Matthies F. Onto-Med/top-api: v0.7.2 2023. https://doi.org/10.5281/ZENODO.7788494.
[23] Uciteli A, Beger C, Kirsten T, Meineke FA, Herre H. Ontological representation, classification and data-driven computing of phenotypes. J Biomed Semant 2020;11:15. https://doi.org/10.1186/s13326-020-00230-0.
[24] Uciteli A, Beger C, Wagner J, Kirsten T, Meineke FA, Stäubert S, et al. Ontological modelling and FHIR Search based representation of basic eligibility criteria. GMS Med Inform 2021:Biometrie und Epidemiologie; 17(2):Doc05. https://doi.org/10.3205/MIBE000219.
[25] Beger C, Matthies F, Schäfermeier R, Kirsten T, Herre H, Uciteli A. Towards an Ontology-Based Phenotypic Query Model. Appl Sci 2022;12:5214. https://doi.org/10.3390/app12105214.
[26] Thompson WK, Rasmussen LV, Pacheco JA, Peissig PL, Denny JC, Kho AN, et al. An evaluation of the NQF Quality Data Model for representing Electronic Health Record driven phenotyping algorithms. AMIA Annu Symp Proc AMIA Symp 2012;2012:911–20.
[27] Mo H, Jiang G, Pacheco JA, Kiefer R, Rasmussen LV, Pathak J, et al. A Decompositional Approach to Executing Quality Data Model Algorithms on the i2b2 Platform. AMIA Jt Summits Transl Sci Proc AMIA Jt Summits Transl Sci 2016;2016:167–75.
[28] Pacheco JA, Rasmussen LV, Kiefer RC, Campion TR, Speltz P, Carroll RJ, et al. A case study evaluating the portability of an executable computable phenotype algorithm across multiple institutions and electronic health record environments. J Am Med Inform Assoc 2018;25:1540–6. https://doi.org/10.1093/jamia/ocy101.
[29] Nogues I-E, Wen J, Lin Y, Liu M, Tedeschi SK, Geva A, et al. Weakly Semi-supervised phenotyping using Electronic Health records. J Biomed Inform 2022;134:104175. https://doi.org/10.1016/j.jbi.2022.104175.
[30] Chiu P-H, Hripcsak G. EHR-based phenotyping: Bulk learning and evaluation. J Biomed Inform 2017;70:35–51. https://doi.org/10.1016/j.jbi.2017.04.009.
[31] Wilcox AB. The Role of Domain Knowledge in Automating Medical Text Report Classification. J Am Med Inform Assoc 2003;10:330–8. https://doi.org/10.1197/jamia.M1157.
[32] Hamidi B, Flume PA, Simpson KN, Alekseyenko AV. Not all phenotypes are created equal: covariates of success in e-phenotype specification. J Am Med Inform Assoc 2023;30:213–21. https://doi.org/10.1093/jamia/ocac157.