MedUse: A Visual Analysis Tool for Medication Use Data in the ABCD Study

H. Bartsch[1,2] and L. Garrison[5] and S. Bruckner[5] and A. (Szu-Yung) Wang[4] and S. F. Tapert[3] and R. Grüner[1]

[1]Mohn Medical Imaging and Visualization Centre, Haukeland University Hospital, Bergen, Norway
[2]Center for Multimodal Imaging and Genetics, University of California San Diego, La Jolla, United States
[3]Department of Psychiatry, University of California San Diego, La Jolla, California, United States
[4]Laboratory of Neuroimaging, National Institute on Alcohol Abuse and Alcoholism, Bethesda, Maryland, United States
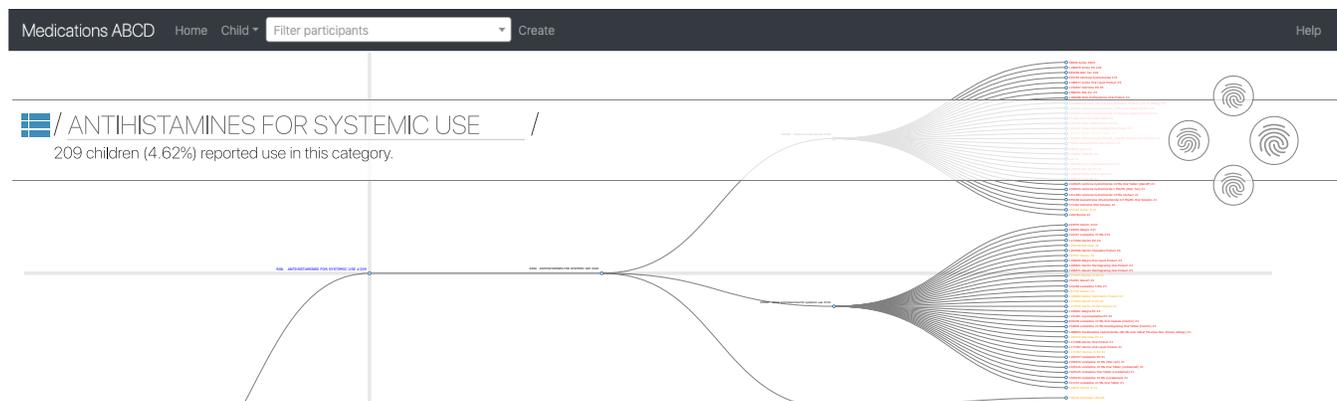[5]Department of Informatics, University of Bergen, Norway

Figure 1: *MedUse: Medication use analysis for the largest long-term study of brain development and child health in the United States.*

Abstract
The RxNorm vocabulary is a yearly-published biomedical resource providing normalized names for medications. It is used to capture medication use in the Adolescent Brain Cognitive Development (ABCD) study, an active and publicly available longitudinal research study following 11,800 children over 10 years. In this work, we present medUse, a visual tool allowing researchers to explore and analyze the relationship of drug category to cognitive or imaging derived measures using ABCD study data. Our tool provides position-based context for tree traversal and selection granularity of both study participants and drug category. Developed as part of the Data Exploration and Analysis Portal (DEAP), medUse is available to more than 600 ABCD researchers world-wide. By integrating medUse into an actively used research product we are able to reach a wide audience and increase the practical relevance of visualization for the biomedical field.

CCS Concepts
• Human-centered computing → Information visualization; Activity centered design;

1. Introduction

The evolution of data sharing platforms into data analysis systems is driven by the complexity of the analysis required to generate publishable results, the complexity of the data shared, and the costs involved in downloading and post-processing data. Using conventional methods to access the data generated with observational studies such as the Adolescent Brain Cognitive Development (ABCD) study [CCC*18], a researcher needs to download,

store, merge, and filter the available project's data. All of these data management steps incur costs for the data sharing platform and the individual researcher, which can hinder biomedical research. Data sharing platforms such as the NIMH Data Archive [Nata] therefore add features for online data analysis such as the Data Analysis and Exploration Portal (DEAP) [DEA]. This provides a unique opportunity for visualization to impact ongoing research in the fields of developmental science and neuropsychology. The DEAP system provides access to ABCD study data to individual researchers
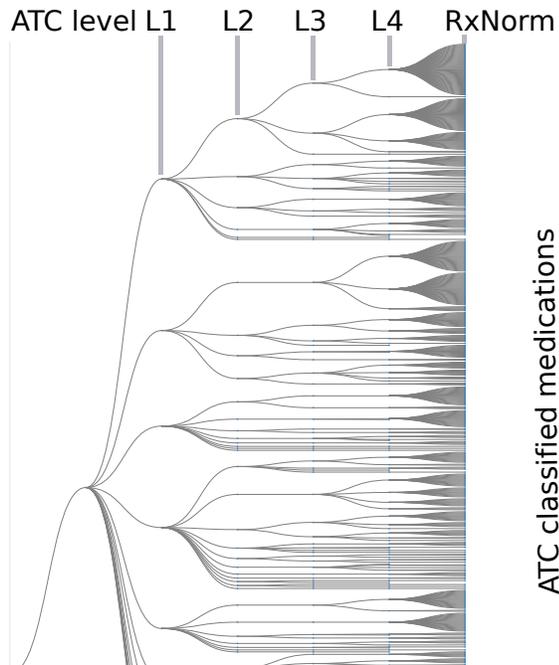
Figure 2: *ABCD medication use data organized by medUse. The data is separated into medications recognized by ATC. Individual ATC levels are marked as L1 through L4. The label* **RxNorm** *indicates the position of the captured medication names in RxNorm nomenclature.*

and performs direct hypothesis tests. Its modular design and large user base provides a unique opportunity for new visual analysis tools to be tested by the scientific community. In this work we have extended DEAP, creating the first visualization-based application focusing on the exploration of medication use data collected by the ABCD study, a long-term study of brain development and child health in the United States, collecting longitudinal data from 11,878 children. This data resource includes current medication use and medication use by the mother during pregnancy captured in parent questionnaires as structured medication names from the RxNorm vocabulary [U.S11]. This data is shared annually as a spreadsheet with collected medication identifiers and names as columns per participant [ABC].

The objective of medUse is to a) enable the user of the ABCD study to interactively browse the RxNorm coded medication use data using a meaningful level of abstraction and to b) allow the user to create novel scores that sum over the individual medication entries and integrate into the DEAP system as first-class scores.

2. Related Work

The literature on tree visualization is extensive and mainly focused on the effective use of space [KHL*07]. Separated into space filling and node-linked visualization methods these algorithms use a wide range of visual organizational principles such as 2D or 3D embedding [DE01], distortions [LRP95], rectangular [SCF11], radial, or

spiral arrangements [Sch11, Cui] with the goal to produce visually pleasing graphical representations.

Space filling visualizations such as treemaps [Shn92] are more space conserving compared to many node-linked methods. They are well suited for displaying information located at the leaf nodes of the hierarchy [GBP02], but they are more difficult to use for the task of selecting non-leaf nodes in large trees. Selection and navigation between neighboring nodes in the tree can result in focus movements in both horizontal and vertical planes with distance as the conserved organizational principle. Treemaps can be adjusted to allow for parent labels to be displayed [LMS*18, KH14] but examples presented in those implementations suggest that they are more tailored for short labels. Support for a focused display of labels at different scales has been suggested [KvK*19]. Using camera distance and occlusion to selectively show and position hierarchical labels of multi-instance objects for volumetric displays.

Focus and context approaches have been suggested to help users to show particular tree nodes while providing context for the placement of the node in the hierarchy [LRP95, SCF11]. The 2D distortion introduced to highlight a focus point in the graph can be severe and require global changes in the visual representation during navigation. Focus and context approaches can also be applied to more traditional approaches for displaying large hierarchical structures such as indented lists [SQX*06]. The large available horizontal space is ideal for longer node names but is more difficult for the user to navigate and highlight the current location without extensive mouse or keyboard interactions.

The RxNav browser is a web application for exploring the RxNorm and ATC classification schemes [ZBKN06]. The software supports a *class view* visualization that supports zoom and pan after searching for a specific medication name. Only a single parent chain from a single medication is displayed. Navigation to neighboring nodes is cumbersome as it involves repeated folding operation using drop-down menus. The tool does not allow the user to visualize medication data for a study cohort. An interesting service that allows computations on study cohort data is RxMix [PMNB13]. RxMix provides a graphical user interface to compose medication data specific processing pipelines. Applied to user-provided lists of medications such a pipeline can, for example, translate between different medical nomenclatures.

3. Tasks

A typical analysis using the ABCD data resource would look at medication use behavior or psychological and cognitive side effects of classes of medications. **Assigning the class labels to each participant** in the cohort the researcher **compares multi-dimensional class assignments statistically**. For example, if there is a correlation between the reported use of ADHD medications and a cognitive outcome measure that correlates with attention or short-term memory. The large number of similar prescriptions and brand and generic versions of the same drug requires significant effort and expertise for the researcher to **derive medication classes** for such a system level analysis. For ADHD related medications, for example, our tool reports that there are 53 individual medications selected by the parents. As the study includes healthy kids, a large number

Table 1: *ACT classification for the medication Adderall by organ system, therapeutic use, pharmacological subgroup, and chemical subgroup. An additional ATC level 5 for chemical substances is not used to simplify the overall tree layout.*

level	principle use	class id	node name
	root		Medications
	drug in ATC		Anatomical Therapeutic Chemical (ATC1-4)
1	organ system	N	Nervous System
2	therapeutic subgroup	N06	Psychoanaleptics
3	pharmacological subgroup	N06B	Psychostimulants, Agents used for ADHD and Nootropics
4	chemical subgroup	N06BA	Centrally acting sympathomimetics
RxNorm	ABCD captured medication	84815	Adderall

of research topics would benefit from the inclusion of medication data. The effort required to generate, document and update these manual classifications prompted the development of our medUse application.

4. Approach

To support research in the field of developmental science and neuropsychology we selected the Anatomical Therapeutic Chemical Classification System (ATC) [Wor], an ontology that provides information to group individual medications hierarchically by organ system, the mechanism of action or therapeutic use, and by pharmacological and chemical subgroup (see Table 1).

The ATC classification system provides a fixed depth of five hierarchy levels. We introduce one additional level to include the more than 800 observed RxNorm medication types in Figure 2. Using a data-driven generation of the medication class hierarchy we limit the ATC classifications to entries that have at least one reported medication use in the selected group of study participants. For medication items that belong to multiple ATC classes the specific therapeutic use is unclear, so we use multiple assignments to allow for a correct statistical coverage of the multiple choices. This approach leads to a combined tree with more than 1,000 nodes and an average branching factor of 24 (see Figure 2).

For visualization-naive users, the main target group of medUse, we developed a four-point navigation paradigm to guide navigation. This allows for more control by separating movements in the horizontal direction to indicate class specificity from movement vertically in the tree to select alternative drug classes at the same specificity (Figure 3). To support class comparison across different branches we changed the default navigation strategy, allowing users to move between branches using the Up and Down arrow navigation (see example in Figure 3). Each branching node also tracks the last navigation choice as the user traverses the tree (right to left). If the navigation direction is reversed, the previously recorded choice is selected during tree traversal. Both of these changes increase saliency of user choice and can prevent feelings of being lost after navigation steps. Overall, this four-point navigation system supports the user in recovering already visited nodes in the tree. Limiting the mental load required to perform the task medUse attempts to improved the overall user satisfaction. For experienced users we also included a full text search as a drop-down option. This allows them to navigate quickly to a specific location, providing a

method to ignore hierarchical information and instead search for keywords found in node descriptions. Our application supports all combinations of the available navigation modes i) mouse/touch, ii) keyboard/cursor, and iii) text searches and updates all components of the user interface accordingly for maximum view flexibility.

Based on its intended use, one medication item is always displayed in the center of the screen. From there it can be exported as a binary classification vector coding 1 for study participants at a given visit that had at least one medication use reported for the class and 0 otherwise. With the integration of medUse into the DEAP application this new data vector becomes part of the ABCD study data and is available in the statistical modules of DEAP.

An active class level removes the need for frequent mouse clicks to select a new level for inspection and provides automatic updates of context information during all navigation steps. The information presented includes access to online documentation as well as textual and visual summary information for the number of participants selected by the current class, as depicted in Figure 1 and using color to indicate the current class level. This allows the user to easily judge the specificity of the selection during the navigation step. Sorting the branches of the tree by the number of the participants reporting drug use, the most frequent used medication class appears near the top of the tree, whereas less frequently used medications are grouped closer to the bottom. Navigation from parent nodes to child nodes favors the top branches with the most frequently reported medication classes. The height of the selected location in the medication tree therefore has meaning in the context of the user task of finding an appropriate class for analysis.

4.1. Implementation

MedUse is implemented as a single-page web-application utilizing the DEAP user authentication and authorization framework. This ensures that users with valid access to the project data also have access to the application. The integration as a DEAP component allows the application to focus on a single workflow leading to a cleaner interface design. Other DEAP components are contacted by medUse to add derived data to the data dictionary and saving co-variate data accessible to the statistical analysis tool.

Whereas access to the study data is provided by the DEAP framework, the hierarchical ATC data is downloaded from the National Library of Medicines ATC web-API [Natb] as JSON encoded text and cached by the user's client browser to improve the
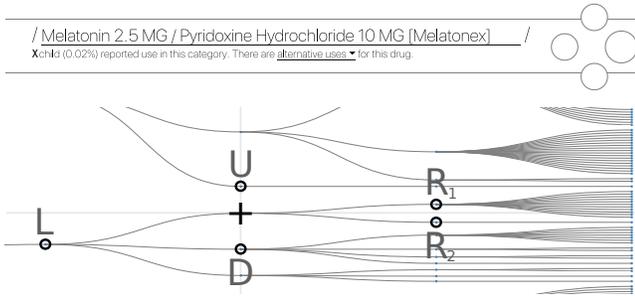
Figure 3: *(Top) Example focus information for Melatonin linking to search field and list of alternative uses and touch/mouse enabled four-point navigation fields (right). (Bottom) Navigation example omitting the display of actual node labels for clarity. Current location is marked with (+). The four node-linked neighboring positions in the tree are **L**eft (level L2), **R**ight (e.g. R_1, L4), and **D**own (L3). The **U**p position (L3) is local only relative to the selected ATC level 3, but not local given the ATC hierarchy.*

startup time of the application. All data manipulations are implemented client-side in JavaScript. This limits the requirements on the DEAP server infra-structure and allows the application to scale to a large number of simultaneous users. The tree visualization is implemented using the D3.js library [BOH11]. Focus-level information was added as additional HTML/CSS elements placed on top of the tree at a fixed screen location. During navigation the focus level information is updated and the tree view moves so that the new selection is in the center of the browser window. Animations are used to move the tree improving the user's location awareness.

5. User Study and Feedback

To test the usability of the application we asked 3 researchers to participate in a user test. They were familiar with the DEAP analysis system but unfamiliar with medUse. They had varying levels of expertise in analyzing medication use data. We instructed them to create a regression model analyzing the effects of brain structures on cognitive measures while correcting for reported medication use related to attention deficit hyperactivity disorder (ADHD). The task required them to create the summary measure for ADHD medication in medUse and to use the generated score in another DEAP application. All users did understand the hierarchical visualization using a node-linked tree and appreciated the level of detail it provides for individual medications in relation to the therapeutical subgroups.

Hypothesis generation in the DEAP statistical analysis involves the combination of multiple measures. We manually reviewed the variability of the user-generated hypotheses and observed a preference of experts for adding medUse generated medication groups at varying levels of specificity. Overall they tended to generate less than 10 alternative analysis models. MedUse supports the documentation of this user choice by generating fixed names for the exported medication use measures based on the name for the node in the tree. Sharing of the hypothesis specification on DEAP is there-

fore sufficient for other users to replicate ABCD findings related to medication use.

Our tool allowed the ABCD test users to visually explore hierarchical relationships on medication use for parents (during pregnancy) and youth by navigating from the most general class (drug/supplement or no-drug) down to individual medications. This is done in the context of a study participant cohort dataset that is interpreted and visualized with the help of the ATC ontology. This is in contrast to many previous approaches on graph visualization for ontological data [KHL*07] that focus on the ontology itself, without regard to a specific task in which that ontology is utilized.

Based on results from the user studies medUse is well suited to support visualization naive users in the selection of medication classification groups from the ATC ontology. Users tend to use a single navigation style and readily adopt the four-point navigation using the keyboard without additional training (see user study). Test users not familiar with the nomenclature used to identify medication use data or the classification system for organ specific or therapeutic uses of such medications have nevertheless been able to use medUse successfully. Compared to a manual re-coding of the medication information, the medUse approach is less error prone and can be repeated easily to include other medication classes. MedUse fully automates the mapping of data to the ATC ontology and can be used by users not trained in the classification of individual medications.

6. Discussion and Conflusions

MedUse makes assumptions that might limit its use for analyzing other medication related information. First of all, the restriction of only supporting a single classification system limits the choice for the user. One might argue that this reduces researcher degrees of freedom, but this is unhelpful if a required medication grouping is not supported. The application cannot be used to create arbitrary user defined classifications. A possible extension of medUse to support alternative ontologies are possible and requires further discussion. MedUse also limits itself to the display of ontology elements relevant for a specific data set. It cannot be used to explore branches in the ontology that are not supported by the current data in the study. There are also problems associated with the use of the ATC classification. According to the WHO [Wor], ATC codes are assigned frequently according to the mechanism of action rather than the therapy. Medication groups can therefore include medicines for different indications, and medicines with similar therapeutic use may be classified in different groups. MedUse tries to support these ambiguities with highlighting multi-use medications.

The need to zoom in order to localize visually the current class in relation to the overall ATC hierarchy can be problematic. At low zoom factors the hierarchy is visible, but font-sizes are scaled down and ATC text entries become unreadable. This favors desktop users with larger screen sizes and is not an issue for users that explore the hierarchy locally at higher zoom factors, or for users that use full-text search to navigate larger ATC distances in the tree. It would have been possible to add additional visualization components to provide two simultaneous level of detail or geometric distortions, but the benefits of such an addition may not outweigh the risk involved with making the user interface more complex, and therefore

less approachable for the target audience. It is worthwhile to explore dynamic labeling approaches in more detail as they can reduce visual complexity at different zoom levels.

As the ABCD study releases more data we plan to continue optimizing and extending this visualization tool to support futher research in developmental science and neuropsychology.

Acknowledgments

References

[ABC] ABCD Study data release 1.1. online. URL: http://abcdstudy.org, doi:10.15154/1460410. 2

[BOH11] BOSTOCK M., OGIEVETSKY V., HEER J.: D³ data-driven documents. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (Dec 2011), 2301–2309. doi:10.1109/TVCG.2011.185. 4

[CCC*18] CASEY B., CANNONIER T., CONLEY M. I., COHEN A. O., BARCH D. M., HEITZEG M. M., SOULES M. E., TESLOVICH T., DELLARCO D. V., GARAVAN H., ORR C. A., WAGER T. D., BANICH M. T., SPEER N. K., SUTHERLAND M. T., RIEDEL M. C., DICK A. S., BJORK J. M., THOMAS K. M., CHAARANI B., MEJIA M. H., HAGLER D. J., CORNEJO M. D., SICAT C. S., HARMS M. P., DOSENBACH N. U., ROSENBERG M., EARL E., BARTSCH H., WATTS R., POLIMENI J. R., KUPERMAN J. M., FAIR D. A., DALE A. M.: The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience 32* (2018), 43 – 54. The Adolescent Brain Cognitive Development (ABCD) Consortium: Rationale, Aims, and Assessment Strategy. URL: http://www.sciencedirect.com/science/article/pii/S1878929317301214, doi:10.1016/j.dcn.2018.03.001. 1

[Cui] CUI W.: A Survey on Graph Visualization. online. URL: http://www.cse.ust.hk/~weiwei/PQE/WeiweiPQE.pdf. 2

[DE01] DACHSELT R., EBERT J.: Collapsible cylindrical trees: A fast hierarchical navigation technique. In *InfoVis'01: Proceedings of the IEEE Symposium on Information Visualization* (2001), Andrews K., Roth S., Wong P. C., (Eds.), IEEE Computer Society, pp. 79–86. doi:10.1109/INFVIS.2001.963284. 2

[DEA] DEAP - Data Exploration and Analysis Portal, RRID:SCR_016158. online. URL: https://deap.nimhda.org/. 1

[GBP02] GROSJEAN J., BEDERSON B. B., PLAISANT C.: Spacetree: Supporting exploration in large node link tree, design evolution and empirical evaluation. In *Information Visualization, IEEE Symposium on* (Los Alamitos, CA, USA, oct 2002), IEEE Computer Society, p. 57. URL: https://doi.ieeecomputersociety.org/10.1109/INFVIS.2002.1173148, doi:10.1109/INFVIS.2002.1173148. 2

[KH14] KIM K., HAN J.: Label embedded treemapping: A label overlap prevention technique for zoomable treemaps and a user interaction technique. In *Human Interface and the Management of Information. Information and Knowledge Design and Evaluation* (Cham, 2014), Yamamoto S., (Ed.), Springer International Publishing, pp. 44–53. 2

[KHL*07] KATIFORI A., HALATSIS C., LEPOURAS G., VASSILAKIS C., GIANNOPOULOU E.: Ontology visualization methods - a survey. *ACM Computing Surveys 39* (11 2007), 10. doi:10.1145/1287620.1287621. 2, 4

[KvK*19] KOUŘIL D., ČMOLIK L., KOZLIKOVÁ B., WU H., JOHNSON G., GOODSELL D. S., OLSON A., GRÖLLER M. E., VIOLA I.: Labels on levels: Labeling of multi-scale multi-instance and crowded 3d biological environments. *IEEE Transactions on Visualization and Computer Graphics 25*, 1 (Jan 2019), 977–986. doi:10.1109/TVCG.2018.2864491. 2

[LMS*18] LI D., MEI H., SHEN Y., SU S., ZHANG W., WANG J., ZU M., CHEN W.: Echarts: A declarative framework for rapid construction of web-based visualization. *Visual Informatics 2*, 2 (2018), 136 – 146. URL: http://www.sciencedirect.com/science/article/pii/S2468502X18300068, doi:10.1016/j.visinf.2018.04.011. 2

[LRP95] LAMPING J., RAO R., PIROLLI P.: A focus+context technique based on hyperbolic geometry for visualizing large hierarchies. In *CHI'95: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (1995), Katz I. R., Mack R., Marks L., Rosson M. B., Nielsen J., (Eds.), ACM Press/Addison-Wesley Publishing Co., pp. 401–408. doi:10.1145/223904.223956. 2

[Nata] re3data.org: National Institute for Mental Health Data Archive. *re3data.org - Registry of Research Data Repositories*. URL: http://data-archive.nimh.nih.gov, doi:10.17616/R3XV5P. 1

[Natb] National Library of Medicine. online. URL: https://www.nlm.nih.gov/research/umls/rxnorm/sourcereleasedocs/atc.html. 3

[PMNB13] PETERS L., MORTENSEN J., NGUYEN T., BODENREIDER O.: Enabling complex queries to drug information sources through functional composition. *Studies in health technology and informatics 192* (2013), 692–696. URL: https://www.ncbi.nlm.nih.gov/pubmed/23920645. 2

[SCF11] SONG H., CAI X., FU Y.: Rectangular tree browser: a navigation and visualization tool for large hierarchies. *Journal of Information and Computational Science 8*, 2 (2011), 354–361. URL: http://www.joics.com/publishedpapers/2011_8_2_354_361.pdf. 2

[Sch11] SCHULZ H.: Treevis.net: A tree visualization reference. *IEEE Computer Graphics and Applications 31*, 6 (Nov 2011), 11–15. doi:10.1109/MCG.2011.103. 2

[Shn92] SHNEIDERMAN B.: Tree visualization with tree-maps: 2-d space-filling approach. *ACM Trans. Graph. 11*, 1 (Jan. 1992), 92–99. doi:10.1145/102377.115768. 2

[SQX*06] SONG H., QI Y., XIAO L., ZHU T., CURRAN E. P.: Lenstree: Browsing and navigating large hierarchical information structures. In *16th International Conference on Artificial Reality and Telexistence–Workshops (ICAT'06)* (Nov 2006), pp. 682–687. doi:10.1109/ICAT.2006.81. 2

[U.S11] U.S. National Library of Medicine, National Institutes of Health, Bethesda, Maryland, USA, RxNorm API. online, 2011. URL: http://rxnav.nlm.nih.gov/RxNormAPI.html. 2

[Wor] World Health Organization Collaborating Center for Drug Statistics Methodology. online. URL: https://www.whocc.no/atc/structure_and_principles/. 3, 4

[ZBKN06] ZENG K., BODENREIDER O., KILBOURNE J., NELSON S.: RxNav: a web service for standard drug information. *AMIA ... Annual Symposium proceedings. AMIA Symposium 2006* (2006), 1156. URL: https://www.ncbi.nlm.nih.gov/pubmed/17238775. 2