

Visual Comparison of Guidelines: Method and Application to Potentially Inappropriate Medication Lists

Abdelmalek MOUAZER^{a,1} Karima SEDKI^a,
Rosy TSOPRA^{b,c} and Jean-Baptiste LAMY^a

^a *Université Sorbonne Paris Nord, LIMICS, INSERM, UMR 1142,
F-93000, Bobigny, France*

^b *INSERM, Université de Paris, Sorbonne Université, Centre de Recherche des
Cordeliers, Information Sciences to support Personalized Medicine,
F-75006 Paris, France*

^c *Department of Medical Informatics, Hôpital Européen Georges-Pompidou,
AP-HP, Paris, France*

Abstract. Therapeutic guidelines developed by experts are essential tools for improving therapy and drug prescription. Several guidelines often exist that target the same patient, from different organizations and countries. The case of lists for the detection of potentially inappropriate medications (PIMs) is an example which illustrates how these guidelines can be varied and multiple. In order to have an overview to the divergences and similarities between different lists of PIMs, we propose a visual method to compare PIMs lists, based on set visualization, and we apply it to 5 guidelines.

Keywords. Set visualization, Therapeutic guidelines, Potentially inappropriate medications, Drug classification, ATC classification.

1. Introduction

Therapeutic guidelines include recommendations for optimizing patient therapy, inspired by evidence-based medicine. However, several guidelines with the same objective can be available at the same time from different institutions and/or countries, and this can lead to potential conflicts between recommendations. The comparison of several guidelines is a complex and tedious task. In the literature, few approaches have been proposed for comparing guidelines [1,2], and they remain limited to two or three guidelines only. Regarding the detection of potentially inappropriate medications (PIMs), several guidelines have been published. The first and one of the best known is the BEERS List [3], published in its first version in 1991 in the United States, which focuses on American practice. Inspired by this one, several guidelines have been developed in different countries including criteria more suited to local practice. These

¹Corresponding Author, Abdelmalek MOUAZER, Université Sorbonne Paris Nord, LIMICS, INSERM, UMR 1142, F-93000, Bobigny, France; E-mail: malikmouazer@gmail.com

guidelines are presented in the form of a list of criteria. Each criterion is presented as: either a drug potentially inappropriate in all situations in the elderly, or a drug potentially inappropriate under certain conditions. These conditions may be an underlying disorder (contraindication), a high dose, or an associated drug (association not recommended). An example from Beers criteria is: “History of gastric or duodenal ulcers with aspirin>325 mg/d or COX-2 non selective NSAIDs.” This criterion indicates that aspirin at a dose greater than 325 mg per day as well as non-selective COX-2 NSAIDs are inappropriate in elderly patients with a history of gastric or duodenal ulcer. The objective of this work is to propose a method for the quantitative comparison of PIMs guidelines. The comparison will focus on the drugs mentioned in PIMs guidelines. It is aimed at facilitating the selection of the most consensual guidelines for future implementation in a decision support system. We will apply this method to the comparison of 5 PIMs guidelines: 4 European guidelines, LAROCHE [4], PRISCUS [5], GHEOPS [6], STOPP (Subset of stopp & start criteria version 2 [7]) and the BEERS List version 2019 [8], which remains an international reference and the first list to be published.

2. Method

2.1. Extraction and standardization of potentially inappropriate drugs

Our guidelines are lists of criteria in textual format containing the drug descriptions and additional information relating to the conditions for applying the criterion. Standardization of the drug descriptions is necessary in order to be able to compare them. We focused on the targeted drugs in a criterion, disregarding the eventual conditions under which the criteria were applied. We used automatic natural language processing (NLP) tools for extracting the ATC codes (Anatomical Therapeutic Chemical Classification System) associated with each criterion. We then refined this result by manually adding the missed codes. Drugs are expressed at various levels of granularity in PIMs guidelines, *e.g.*, the aforementioned criterion targets both a molecule, aspirin, and a drug class, COX-2 non selective NSAIDs. The ATC classification has a 5-level hierarchical structure. We expressed all drugs with one or more codes of the fifth level. Higher-level codes representing drug classes were replaced by the fifth-level codes of their descendants.

During this step, each molecule is associated with a single ATC code representing the molecule alone, without taking into account the association codes that exist in ATC. For molecules that have more than one ATC code even without associations, we have kept only one, ensuring that it is the same for all the guidelines. Some drugs (*e.g.*, ginkgo biloba) have no ATC code. So we assigned them an arbitrary code.

2.2. Visualization

Comparing guidelines in terms of the inappropriate drugs they contain may seem simple. However, this is only true when the number of guidelines is very small. A given drug may belong to one or more guidelines. Thus, if we have n guidelines, we have $2^n - 1$ sets of drugs to consider, one per each non-empty subsets of the set of all

guidelines, called *intersections*. As the number of intersections increases exponentially with n , the comparison rapidly becomes very complex. Sophisticated visual techniques are thus required for presenting intersections visually. Since our data is in the form of sets, the appropriate method should be a set visualization approach. Here, we chose the RainBio approach, developed by Lamy & al. [9]. Compared to the well-known Venn diagram, Rainbio has the advantage to allow proportional visualization. In this approach, it is a matter of visually highlighting the membership relationships of a set to one or more elements. In our case, we consider PIMs guidelines to be sets and ATC codes as elements. For this purpose, we have created a data set with the syntax imposed by RainBio and which contains the data extracted from the PIMs guidelines.

3. Results

989 drugs were identified as potentially inappropriate in the elderly by at least one of our guidelines. Each guideline is associated with the drugs it detects as inappropriate. A drug can be detected as inappropriate by one or more guides which represents $2^5-1=31$ possible intersections between one or more guidelines. All 31 non-empty intersections were present, which suggests a great interconnection between the guidelines. We visualized these intersections using RainBio (Figure 1). Now when comparing the 5 guidelines, we can : (a) The BEERS and STOPP guides cover a larger number of drugs compared to the other guidelines. (b) There are 37 drugs for which there is an overall consensus as to their inappropriateness (red box at the bottom), this roughly represents only 4% of the total number of drugs and 5% to 20% of the total number of drugs in each guideline. This suggests some discrepancies between the guides when viewed together. (c) The dark blue boxes at the top show the specific drugs for each guideline. 175 and 213 of drugs present in STOPP and BEERS are specific, compared with 27, 29 and 15 drugs for GHEOPS, PRISCUS and LAROCHE. Furthermore, if we take a closer look at these last 3 guides, we notice that almost all of their content that is not specific to them is present in BEERS and/or STOPP (90% to 94% of their content). This suggests that BEERS and STOPP have a high of drugs that are absent from all other guidelines, and thus might be in excess. (d) 362 drugs are shared between BEERS and STOPP (sum of all the boxes that cross the two columns). This suggests that these two guidelines share some similarities, but which still have their specificity. (e) The STOPP guideline shares 452 drugs with the other guides (all of the boxes that cross the STOPP column), it is the guide that shares the most drugs with one or more other guides. It has been placed in the center of the figure by Rainbio's algorithm, which shows its central position. (f) While STOPP, BEERS and GHEOPS share two large yellow boxes, LAROCHE and PRISCUS are each excluded from one of these boxes, thus missing some drugs considered as inappropriate by all other guidelines. To conclude, STOPP and BEERS have a lot of specific drugs, and LAROCHE and PRISCUS miss a lot of drugs present in all other guidelines, so GHEOPS seems to be the more consensual one. Additionally, if we are interested in drugs, RainBio makes it easy to identify drugs in a given intersection by clicking on the corresponding box.

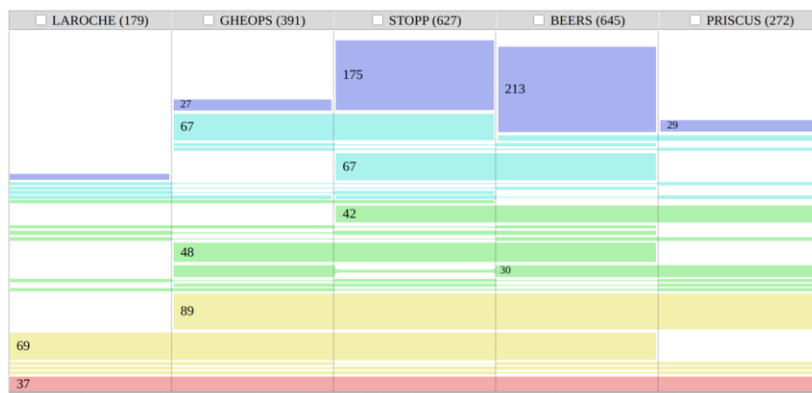


Figure 1. Visualizing Potentially Inappropriate Drugs From 5 Guidelines Using RainBio. RainBio represents guidelines by columns ordered by similarity (each is placed next to the ones with whom it shares the most drugs). The intersections between the guidelines (i.e. shared drugs) are represented by rectangular boxes. The color of the boxes indicates the degree of the intersection, it is a color gradient going from dark blue (intersection with a single guideline; the drugs placed in these boxes are thus specific to that guideline) to red (intersection between all guidelines, drugs in this box are present in all PIMs guidelines). The height of the box is proportional to the number of drugs in the intersection. The boxes are finally organized vertically by degree of intersection and therefore also by color, so we find the dark blue boxes at the top and the red boxes at the bottom.

4. Discussion & Conclusions

We proposed a visual method for the comparison of therapeutic guidelines. This method makes it easy to compare multiple PIMs guidelines. It allows for the exploration of their divergences and similarities. The comparison appears to be intuitive and very easy to understand, we have a clear and proportional overview. The results observed with this method are consistent with the development logic of the guidelines. Indeed, GHEOPS is based on 18 guidelines including those taken into account in this work, which explains its consensual nature. STOPP and BEERS also take into account drug interactions, contraindications and in particular those related to the renal function, which may explain the fact that they have a lot of specific drugs. That said, STOPP and BEERS are quite different concerning their specific drugs; this might be explained by the fact that BEERS is developed for American practice when STOPP is developed for European practice. The similarities observed are not to be taken as being absolute similarities, a drug may be deemed inappropriate without conditions in one guide while being associated with a condition in another guide. The comparison we propose should not lead to a judgment of the quality of one guideline compared to another. Our approach remains focused solely on a comparison based on the drugs to which the various guides are concerned. Comparing PIMs guidelines on the basis of the number of drugs they are interested in have some limits. Indeed, a guideline may contain many drugs rarely prescribed, or not available on a given market. There are also guidelines that offer alternatives while others do not. Taking these parameters into account seems relevant for an in-depth analysis and qualitative comparison. However, the quantitative comparison we proposed here is complementary to the usual qualitative comparison. Our perspectives include the improvement of the comparison by considering drug interactions and patient conditions present in some

guidelines. We are also thinking of improving our method by adding a filter on medicines available on a European market. We also intend to broaden our comparison by adding other guidelines found in the literature. Finally, our method can be extended to be applied to the comparison of clinical practice guidelines for the management of certain diseases. These guidelines provide recommendations for treating a disease based on the patient profile. We could therefore consider the patient profiles as elements and compare the guidelines on the basis of the patient profiles they support; or rather compare them on the basis of the recommended treatments by taking the recommendations as elements. In the future, our method will be implemented into a decision support system for helping health professionals to develop or revise guidelines and thus to improve care quality.

Acknowledgments

This work was partly funded by the French Research Agency (ANR) through the ABiMed project [grant number ANR-20-CE19-0017-02].

References

- [1] Galopin A, et al. Using an ontological modeling to evaluate the consistency of clinical practice guidelines: application to the comparison of three guidelines on the management of adult hypertension. *Stud. Health Technol. Inform* 2014;205:38–42.
- [2] Seroussi, B, et al. Using the Recommendations Inferred by a Decision Support System to Compare Breast Cancer Clinical Practice Guidelines. *Stud. Health Technol. Inform*. 2019;258:60–64.
- [3] Beers MH, et al. Explicit criteria for determining inappropriate medication use in nursing home residents. UCLA Division of Geriatric Medicine. *Arch. Intern. Med.* 1991;151: 1825–1832.
- [4] Laroche M-L, Charmes J-P, Merle L. Potentially inappropriate medications in the elderly: A French consensus panel list. *Eur. J. Clin. Pharmacol.* 2007;63:725–31.
- [5] Holt S, Schmiedl S, Thürmann PA. Potentially Inappropriate Medications in the Elderly: The PRISCUS List. *Dtsch. Ärztebl. Int.* 2010;107:543–551.
- [6] Tommelein E, et al. Older patients' prescriptions screening in the community pharmacy: development of the Ghent Older People's Prescriptions community Pharmacy Screening (GheOP3S) tool. *J. Public Health* 2016;38;e158-e170.
- [7] O'Mahony D, et al. STOPP/START criteria for potentially inappropriate prescribing in older people: version 2. *Age Ageing*. 2015;44:213–218.
- [8] By the 2019 American Geriatrics Society Beers Criteria® Update Expert Panel. American Geriatrics Society 2019 Updated AGS Beers Criteria® for Potentially Inappropriate Medication Use in Older Adults: 2019 AGS BEERS CRITERIA® UPDATE EXPERT PANEL. *J. Am. Geriatr. Soc.* 67, 674–694 (2019).
- [9] Lamy J-B, Tsopra R. RainBio: Proportional visualization of large sets in biology. *IEEE Trans. Vis. Comput. Graph.* (2019) doi:10.1109/TVCG.2019.2921544.