

Gnaeus: utilizing clinical guidelines for knowledge-assisted visualisation of EHR cohorts

P. Federico¹, J. Unger¹, A. Amor-Amorós¹, L. Sacchi², D. Klimov³, and S. Miksch¹

¹Vienna University of Technology, Austria ²University of Pavia, Italy ³Ben-Gurion University of the Negev, Israel

Abstract

The advanced visualization of electronic health records (EHRs), supporting a scalable analysis from single patients to cohorts, intertwining patients' conditions with executed treatments, and handling the complexity of time-oriented data, is an open challenge of visual analytics for health care. We propose an approach that, according to the knowledge-assisted visualization paradigm, leverages the domain knowledge acquired by clinical experts and formalized into computer-interpretable guidelines (CIGs), in order to improve the automated analysis, the visualization, and the interactive exploration of EHRs of patient cohorts. In this way, the analyst can get insights about the clinical history of multiple patients and assess the effectiveness of their health care treatments.

Categories and Subject Descriptors (according to ACM CCS): J.3 [Computer applications]: Life and medical science—Medical information systems

In recent years, the diffusion of Electronic Health Records (EHRs) has been growing, partially due to specific public health policies and legislative interventions, such as the National Programme for IT in the United Kingdom since 2002, the Health Information Technology for Economic and Clinical Health Act of 2009 in the United States, and the directive 24/2011/EU for cross-border healthcare in the European Union. Besides facilitating the online data transfer amongst hospitals and care providers, the increasing adoption of EHR systems has made available large amounts of data about patients' conditions and their care pathways. This data can be retrospectively analysed to assess the effectiveness of treatments and identify complex patterns, in order to improve the overall quality of health care. Several interactive information visualization techniques and systems have been proposed to visually explore EHR data, gain insights, form hypotheses, and validate them. Generally, the effective utilization of these systems requires analysts to rely on their domain knowledge, in order to interpret raw data, deduce the overall health status of a patient, and compare administered treatments with evidence-based best practices. In this context, we propose a solution for the visualization of EHR data of patients' cohorts. Our main contributions are:

- a knowledge-assisted visual analytics (VA) approach that leverages computer-interpretable clinical guidelines (CIGs) to drive analysis, visualization, and interaction;

- a prototypical system enabling the retrospective analysis of EHRs within two medical application scenarios: atrial fibrillation (AF) and gestational diabetes mellitus (GDM);
- an empirical evaluation with medical experts, demonstrating the advantages of our approach.

1. Related work

Rind et al. [RWA*13] conducted an extensive and systematic survey about interactive information visualization approaches to EHR exploration and querying. Evidence-based clinical practice guidelines (CPGs) are sets of statements and recommendations used to improve health care by providing a trustworthy comparison between treatment options in terms of risks and benefits according to patient's status [SGM*11]. They condense in a standardized narrative form the complex domain knowledge underneath the clinical practice. Their formalization as CIGs enables the implementation of guideline-execution engines and decision-support systems, assisting professional care providers during the daily practice [Pel13]. Several visualization techniques for CIGs have been proposed in the literature; they are generally aimed to visually support their acquisition and specification (e.g., AsbruView [MKSJ98], Gesher [SS05]). Techniques like CareVis [AM06] and CareCruiser [GAK*11] address the intertwined visualization of the executed CIG and the patient's health status, in order to assess the effects of the former onto

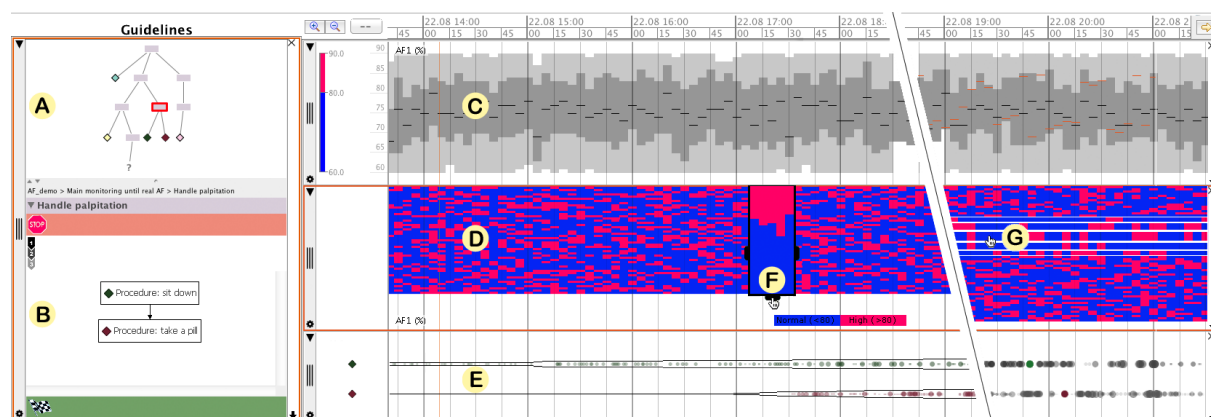


Figure 1: The GUI of *Gnaeus*, a guideline-based knowledge-assisted EHR visualization for cohorts. (A) The hierarchical structure of the clinical guideline comprising subplans and actions is shown as a tree with a layered top-down layout. (B) The procedural knowledge of the selected subplan is shown as a node-link hierarchical task network. (C) The raw data of the parameter that is relevant for the selected subplan is aggregated over the patients' cohort and shown as a streaming box-plot. (D) The raw data is also abstracted according to the declarative knowledge of the subplan and visualized as LifeLines. (E) The execution of clinical actions and their computed compliance with the guideline recommendations are shown in an aggregated visualization. (F) An interactive grouping lens reconfigures the abstractions to visualize their distribution as stacked bars. (G) A vertical fish-eye distortion enables a closer examination of the abstractions of a single patient, while other relevant data for the selected patient are also highlighted within the other views.

the latter. Bodesinsky et al. [BFM13] present an approach for visual analytics of compliance with clinical guidelines. Despite the wealth of visualization systems and techniques present in the literature, West et al. [WBH15] observe that the healthcare provider community has not yet significantly exploited these techniques to accelerate the use and understanding of EHR data. Many problems of VA for healthcare still need to be solved [AFG*12]. In this work we focus particularly onto three of these challenges: the simultaneous exploration of both single patients and cohorts data, the intertwined analysis of patients' conditions and treatment data, and an appropriate support for the time-oriented nature of EHR data.

2. A knowledge-assisted visualization approach

To address the aforementioned challenges, we present *Gnaeus* (Fig. 1), a guideline-based knowledge-assisted EHR visualization for cohorts. *Gnaeus* builds upon previous work [BFM13, GAK*11], extending it with a better scalability from single patients to cohorts and a tighter integration of the domain knowledge. According to the knowledge-assisted visualization approach [CEH*09], *Gnaeus* exploits domain knowledge to better support the VA process. It does not acquire clinical knowledge specifically for the visualization, since knowledge is usually acquired from medical experts into narrative-form CPGs to be used in daily practice, and also formalized into CIGs to be processed by decision-support systems. By placing the CIG at the core of *Gnaeus* (Fig. 2), we are able to leverage the domain knowledge to

inform the automated analysis, the visualization, and the interaction techniques for EHR data.

Clinical guidelines as a knowledge base. *Gnaeus* has been specifically designed to use guidelines written in Asbru, an intention-based and time-oriented language for CGI knowledge representation [MSJ97]. "Intention-based" means that an essential element of an Asbru guideline are the intentions, i.e. the goals expressed at various level of abstraction; intentions can be understood as patterns of actions or states to be achieved (or avoided), and can be temporally annotated by the means of complex time-oriented constructs. Additional elements of an Asbru guideline are effects, describing the functional dependencies between clinical actions and patients' parameters (e.g., administration of insulin lowers blood glucose). Intentions, temporal abstractions, and effects

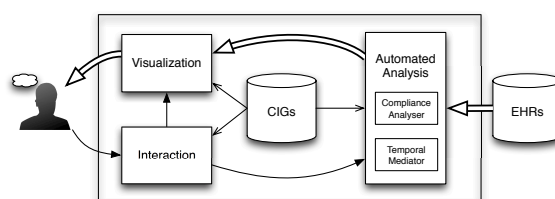


Figure 2: The guideline knowledge is the core of *Gnaeus* and informs the automated analysis, the visualization, and the interaction techniques for electronic health records.

can be seen as the declarative knowledge formalized in the CIGs, as they describe what can be observed and what must be achieved. An Asbru guideline also comprises the preferences, the conditions, and the plan body. They represent the procedural knowledge of the guideline and recommend how to proceed (activating or aborting subplans, performing specific actions) in order to accomplish the guideline intentions according to the patient's health status.

Automated analysis. In our design, both the declarative knowledge and the procedural knowledge are exploited to drive two analytical components: the temporal mediator and the compliance analyzer. The declarative knowledge, specified as guideline intentions, is exploited to process the input raw, time-stamped data, such as Blood Glucose (BG) values at particular times to produce a set of clinically meaningful summarizations and interpretations. The “BG monthly good pattern”, for example, is defined as if during the last month patient had up to one abnormal value of BG per week and no more than four abnormal values per month, while the BG abnormal values are defined in the context of pregnant diabetic patients according to taking insulin medication and fetus size. In our study, we apply the knowledge-based temporal-abstraction (KBTA) method [Sha97] to compute this kind of temporal abstractions. We implemented the principles of the KBTA method in an enhanced version of the IDAN temporal mediator [BS03], that mediates between decision-support applications and a time-oriented clinical database. AF and GDM are diseases that can be managed with a combination of right amount of physical activity, appropriate diet (for GDM only), and drugs (e.g. insulin for GDM, anticoagulants for AF). Thus, it is particularly important to assess not only the general efficacy of treatments, but also the compliance of patients and caregivers with the clinical guidelines for the management of these diseases. An executed treatment is compliant if the recommendations the patient was eligible for were fulfilled by performing the corresponding actions within the suggested response time windows. We developed an enhanced version of the RoMA (Reasoning on Medical Actions) rule-based engine [Qua08], that uses the procedural

knowledge to analyse the compliance of patients' treatments. While the original version produced a compliance summary at the patient's discharge, the enhanced one provides detailed information about the fulfilment of guidelines recommendations for large cohorts. Both the KBTA method and the compliance analysis underlying our work are domain independent, since the knowledge-based principles can be applied in any medical domain.

Visualization. *Gnaeus* adopts coordinated multiple views for visualizing EHR data as well as the procedural knowledge of the CIG. The hierarchical structure of the guideline is visualized as a tree diagram with a top-down layered layout, whose nodes represent subplans and leaves represent clinical actions. The logical structure of a subplan is shown as a node-link diagram of a hierarchical task network. Figure 1 A-B shows the hierarchical structure and the logical structure of a small illustrative guideline for the self-management of AF with a pill-in-the-pocket approach: in case of palpitation and measured AF probability above the 80% threshold, the patient is recommended to reduce physical activity, take a pill and, if the condition persists, call a doctor. The procedural knowledge of the guideline can be used to visually analyze the synchronization of executed subplans, aggregated over a cohort. Asbru subplans, indeed, can be characterized by various temporal constraints: *parallel* (to be started together), *sequential* (one after another, in a given order), *any order* (one after another, in any order), and *unordered* (no synchronization constraints). Figure 3 shows a modified tree layout encoding the synchronization constraints, while the edge color represent the average execution time; comparing the expected execution order with the actual execution time, the user can check the synchronization of the executed treatment, identifying temporal patterns and outliers. Since the layout integrating synchronization information is less space efficient than a simple top-down layered layout, it can be toggled on demand. The temporal abstractions, reducing a numeric (univariate or even multivariate) parameter into a nominal variable, enable a compact visualization of time-oriented EHR data. For small cohorts, *Gnaeus* supports Qualizon Graphs (Fig. 4), a space-efficient visualization combining quantitative data and qualitative abstractions for single patients [FHR*14]. Qualizon Graphs are based on the well-known horizon graphs [Rei08], but they extend them with non-uniform bands corresponding to the value ranges of state abstractions; they are as fast and accurate as horizon graphs for raw data, but also support the integrated visualization of state abstractions. For complex abstractions, or for large cohorts demanding a more compact visualization, *Gnaeus* provides a pixel-based LifeLines-like [PMR*96] visualization; Figure 1 D shows the qualitative abstractions of a parameter such as the probability of AF: each horizontal line corresponds to a patient and is coloured according to KBTA. The raw data for a numeric parameter can also be aggregated over all the patients of a cohort and visualized as a streaming box-plot (Fig. 1 C);

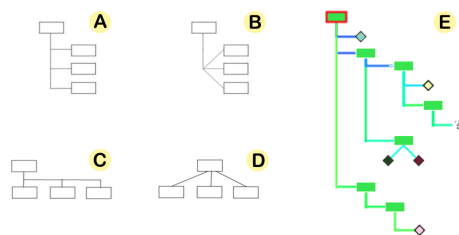


Figure 3: The layout of the guideline hierarchy reflects the prescribed synchronization of subplans: (A) sequential ordered, (B) sequential any order, (C) parallel, (D) unordered. (E) Subplans are colored according to their actual execution time (blue=earlier, green=later).

it shows the five-number statistical summary: the mean is mapped to the black line, the 25th and 75th percentile are mapped to the dark gray bars, the minimum and maximum are mapped to the light grey bars. Executed actions are aggregated over the cohort and visualized as transparent circles along the time-axis (Fig. 1 E): the number of occurrences of an action within an interval is mapped to the alpha channel, while the number of patients to which the action has been administered is mapped to the radius. The height of the bar represents the number of patients within the interval who were eligible for the recommendations: the white space between the action circle and the ends of the eligibility bar represents the number of non-compliant patients.

Interaction. Given the large scale, the multivariate nature, and the temporal complexity of EHR data, specific interaction methods are needed to support user's intentions and enable data exploration. A first set of interaction techniques provided by *Gnaeus* is aimed at facilitating the transition between analysis of single patients and cohorts. A magic lens [TGK*14] reconfigures the arrangement of the temporal view of qualitative data. Outside of the lens, each lifeline represents the history of a single patient; within the lens, the lines are grouped by abstraction, thus enabling a quick overview of the distribution of abstractions across the population in terms of bar charts (Fig. 1 F). Conversely, a fish-eye interaction allows the user to focus on a single patient. When hovering upon the lifeline of a patient in the qualitative view (Fig. 1 G), this line is magnified, the corresponding quantitative data is overlaid on top of the streaming box-plot and the corresponding treatment data is also highlighted in the context of the cohort. These interaction techniques enable a direct comparison between the selected patient and the rest of the cohort. The system provides also knowledge-assisted interactions, supporting specific tasks in the context of a guideline. Since an EHR can contain a large amount of multivariate time-oriented data for each patient, the guideline can be used as an index to browse the EHR data both across the different variables and along the time axis. When the users selects a subplan in the guideline views (Fig. 1 A-B), only relevant data is shown in the temporal views (Fig. 1 C-D-E), identified through the

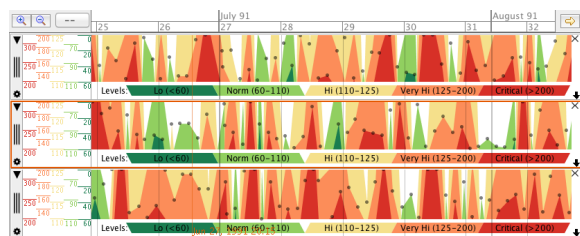


Figure 4: *The glycemia of three patients visualized as Qualizon Graphs, an integrated visualization of raw data and their knowledge-based temporal abstractions.*

plan-parameter dependency specified in the CIG declarative knowledge. Moreover, the user can switch from absolute to relative time, thus data of all patients are aligned according to the execution time of the selected subplan.

3. Evaluation

We designed *Gnaeus* by studying the two aforementioned real-world medical application scenarios (AF and GDM). In a user-centred design fashion, three expert users (i.e. medical doctors) were involved in different phases of design and development: requirements specification (as narrative text and use-case diagrams), visual mappings (as static mock-ups), and interaction design (as a proof-of-concept software implementation). Their feedback was analysed and incorporated in subsequent nested iterations [Mun09], for scoping the application, tailoring visual encodings and analytical algorithms to concrete medical needs, and improving the overall usability. We then performed a summative evaluation by a qualitative user study with two medical experts (distinct persons from the three doctors consulted during the previous phases, in order to get an unbiased final validation). Each subject was shown a short demo, then was allowed to directly interact with the system for a task-based session and a free data exploration session. They were asked to think aloud during the entire test, and at the end they answered a semi-structured interview. Both subjects were able to intuitively grasp most of the visual mappings; one subject needed further explanations about some interactions, but then was able to effectively use them to perform analytical tasks. They generally liked the prototype and stated it could support their daily routines. One subject pointed out that, even if she is as an expert practitioner and knows the CPGs of her medical speciality, the coordinated visualization of the EHR together with the CIG might be useful especially when dealing with updated CPGs or rarely applied branches. Moreover, also when considering a well-known CPG, for example when analysing previous patients retrospectively in order to find commonalities with the current case to be treated, the automatic computation of context-specific temporal abstractions and compliance values can alleviate her mental load and help her with making decisions.

4. Conclusion

We have presented a knowledge-assisted visualization for EHR cohorts that exploits the domain knowledge of CGIs to support the VA process and drive automated analysis, interaction, and visualization. A user-centred design and a summative evaluation involving expert users validate our approach for the retrospective analysis of EHR data within two real-world medical applications, such as AF and GDM.

Acknowledgements

Work funded by the EC grant #287811, MobiGuide project.

References

- [AFG*12] AIGNER W., FEDERICO P., GSCHWANDTNER T., MIKSCH S., RIND A.: Challenges of time-oriented data in visual analytics for healthcare. In *Proceedings of the IEEE VisWeek Workshop on Visual Analytics in Healthcare* (2012), Caban J. J., Gotz D., (Eds.), IEEE, IEEE, p. 4. 2
- [AM06] AIGNER W., MIKSCH S.: Carevis: Integrated visualization of computerized protocols and temporal patient data. *Artificial Intelligence in Medicine* 37, 3 (2006), 203 – 218. doi:http://dx.doi.org/10.1016/j.artmed.2006.04.002. 1
- [BFM13] BODESINSKY P., FEDERICO P., MIKSCH S.: Visual analysis of compliance with clinical guidelines. In *Proceedings of the 13th International Conference on Knowledge Management and Knowledge Technologies* (New York, NY, USA, 2013), i-Know '13, ACM, pp. 12:1–12:8. doi:10.1145/2494188.2494202. 2
- [BS03] BOAZ D., SHAHAR Y.: Idan: A distributed temporal-abstraction mediator for medical databases. In *Artificial Intelligence in Medicine*, Dojat M., Keravnou E., Barahona P., (Eds.), vol. 2780 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2003, pp. 21–30. doi:10.1007/978-3-540-39907-0_3. 3
- [CEH*09] CHEN M., EBERT D., HAGEN H., LARAMEE R., VAN LIERE R., MA K.-L., RIBARSKY W., SCHEUERMANN G., SILVER D.: Data, information, and knowledge in visualization. *Computer Graphics and Applications, IEEE* 29, 1 (Jan 2009), 12–19. doi:10.1109/MCG.2009.6. 2
- [FHR*14] FEDERICO P., HOFFMANN S., RIND A., AIGNER W., MIKSCH S.: Qualizon graphs: Space-efficient time-series visualization with qualitative abstractions. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces* (New York, NY, USA, 2014), AVI '14, ACM, pp. 273–280. doi:10.1145/2598153.2598172. 3
- [GAK*11] GSCHWANDTNER T., AIGNER W., KAISER K., MIKSCH S., SEYFANG A.: CareCruiser: Exploring and visualizing plans, events, and effects interactively. In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis)*, 2011 (March 2011), pp. 43–50. doi:10.1109/PACIFICVIS.2011.5742371. 1, 2
- [MKSJ98] MIKSCH S., KOSARA R., SHAHAR Y., JOHNSON P. D.: AsbruView: Visualization of time-oriented, skeletal plans. In *Proceedings of the Fourth International Conference on Artificial Intelligence Planning Systems* (1998), Association for the Advancement of Artificial Intelligence, pp. 11–18. 1
- [MSJ97] MIKSCH S., SHAHAR Y., JOHNSON P. D.: Asbru: A Task-Specific, Intention-Based, and Time-Oriented Language for Representing Skeletal Plans. In *7th Workshop on Knowledge Engineering: Methods & Languages (KEML-97)* (1997), Motta E., van Harmelen F., von Pierret-Golbreich C., Filby I., Wijngaards N., (Eds.). 2
- [Mun09] MUNZNER T.: A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (Nov. 2009), 921–928. doi:10.1109/TVCG.2009.111. 4
- [Pel13] PELEG M.: Computer-interpretable clinical guidelines: A methodological review. *Journal of Biomedical Informatics* 46, 4 (2013), 744 – 763. doi:http://dx.doi.org/10.1016/j.jbi.2013.06.009. 1
- [PMR*96] PLAISANT C., MILASH B., ROSE A., WIDOFF S., SHNEIDERMAN B.: LifeLines: Visualizing personal histories. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 1996), CHI '96, ACM, pp. 221–227. doi:10.1145/238386.238493. 3
- [Qua08] QUAGLINI S.: Compliance with clinical practice guidelines. In *Computer-based Medical Guidelines and Protocols: A Primer and Current Trends*, ten Teije A., Miksch S., Lucas P., (Eds.), vol. 139 of *Studies in Health Technology and Informatics*. IOS Press, 2008, pp. 160–179. doi:10.1007/978-3-540-47527-9_8. 3
- [Rei08] REIJNER H.: The development of the horizon graph. In *Proc. of the Vis08 Workshop from Theory to Practice* (2008). 3
- [RWA*13] RIND A., WANG T. D., AIGNER W., MIKSCH S., WONGSUPHASAWAT K., PLAISANT C., SHNEIDERMAN B.: Interactive information visualization to explore and query electronic health records. *Foundations and Trends in Human-Computer Interaction* 5 (02 2013), 207–298. doi:10.1561/1100000039. 1
- [SGM*11] STEINBERG E., GREENFIELD S., MANCHER M., WOLMAN D. M., GRAHAM R., ET AL.: *Clinical practice guidelines we can trust*. National Academies Press, 2011. 1
- [Sha97] SHAHAR Y.: A framework for knowledge-based temporal abstraction. *Artificial Intelligence* 90, 12 (1997), 79 – 133. doi:http://dx.doi.org/10.1016/S0004-3702(96)00025-2. 3
- [SS05] SHALOM E., SHAHAR Y.: A graphical framework for specification of clinical guidelines at multiple representation levels. In *AMIA Annual Symposium Proceedings* (2005), vol. 2005, American Medical Informatics Association, p. 679. 1
- [TGK*14] TOMINSKI C., GLADISCH S., KISTER U., DACHSELT R., SCHUMANN H.: A survey on interactive lenses in visualization. In *EuroVis-STARs* (2014), Borgo R., Maciejewski R., Viola I., (Eds.), The Eurographics Association, pp. 43–62. doi:10.2312/eurovisstar.20141172. 4
- [WBH15] WEST V. L., BORLAND D., HAMMOND W. E.: Innovative information visualization of electronic health record data: a systematic review. *Journal of the American Medical Informatics Association* 22, 2 (2015), 330–339. doi:10.1136/amiajnl-2014-002955. 2