Visualizing Patient Clusters and Symptom Development During Head and Neck Cancer Treatment

Carla Floricel*, Andrew Wentzel, Md Nafiul Alam Nipu, Naveen Kumar, G.E. Marai
University of Illinois at Chicago

Guadalupe Canahuate
University of Iowa

Lisanne Van Dijk, C. David Fuller
University of Texas

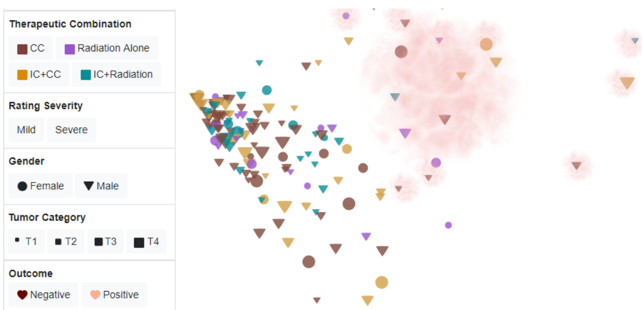

Figure 1: 2D projection of patients' features for a particular time point. The shaded areas shows patients associated with the severe rating cluster. Shape, size and color encode different aspects of the data as shown in the legend above.

Abstract

Approximately 7,500 cases of Head and Neck Cancer (HNC) are diagnosed in the US annually. Patients are increasingly likely to survive, but often experience acute and long term side effects [1]. Hence, great importance has been placed by clinicians on improving patient's quality of life (QoL) and reducing symptom burden during treatment. We introduce an interactive system which enables clinical and computational experts to visualize and assess medical data. Using novel combinations of visual encodings, our system provides context for new patients based on patients with similar features and symptom evolution, which could help oncologists to create better treatment plans.

1 Introduction

HNC patients often suffer strong symptoms and treatment-related side effects, which usually last long after treatment completion. Managing these symptoms constitutes a high priority for both patients and clinical oncologists. Precision medicine methods enable clinical researchers to leverage existing cohorts of similar patients in order to predict the QoL of a new patient. However, HNC patient cohort data is often large, multi-variate, and incomplete. Also, anatomical and dynamic temporal components influence the outcome of therapy, and the resulting patients' QoL. Properly utilizing the data requires close collaboration between clinical and computational researchers, which makes understanding and communicating the underlying anatomical and dynamic structure of the data essential. To find clusters of similar patients and discover related symptoms, thus paving the way towards improving the patients' QoL, we used collaborative design methods alongside domain experts. We implemented a visual analysis system to help researchers and clinicians analyze patient symptoms, and identify trends and outliers, while querying clinical data.

*e-mail: cflori3@uic.edu

2 Related Work

Our work is inspired by electronic medical records designed to keep track of patients' history. For example, Lifelines [5] navigates and analyzes patient records using a timeline visualization form. While it effectively visualizes events during treatment stage, it does not provide context for individual patient features. Nonetheless, a novel visualization encoding, the tendril plot [3], is a compact tool for sequential event data, showing outliers and trends in clinical trials but it does not give specific details about individual patients.

In the precision medicine domain, Gunn et al. [2] studied symptom burden for HNC patients. By clustering patients based on reported symptom ratings and their clinical covariates, they found similarities between symptoms associated with HNC. However, this study does not include time-series data nor symptom progression. In contrast, our system explores groups of similar patients while also capturing the temporal changes in their symptoms.

3 Data

We collected questionnaires from 157 patients treated for HNC at the MD Anderson Cancer Center completed at multiple time points regarding HNC symptoms. Each patient self-reported 28 symptoms on a 10 point scale ranging from "not present" to "as bad as you can imagine". The symptoms are classified into 13 'core' symptoms[1], 9 HNC specific symptoms[2], and 6 interference to daily life items[3]. Demographic and diagnostic data was also gathered about each patient's gender, clinical risk staging, and therapy type.

4 Design

We followed an activity centered design (ACD) paradigm [4], because of its proven success rate in the case of design projects that feature interdisciplinary collaboration. The paradigm is an extension of human-centered-design, with emphasis on user activities and workflow. Through a series of iterations, we met with the end users to define functional specifications, prototype the interface, evaluate prototypes, and decide on changes in the specifications.

Our proposed system incorporates multiple linked views that enable the user to get a thorough understanding of all aspects of the data, providing both overview and detail. We chose the encodings based on the data's multidimensional nature and incompleteness. The interface is separated into two side-by-side main views: **(1)** Symptom development, patients' characteristics and clustering, and **(2)** Symptom patterns and correlations. To our knowledge, there is

[1]fatigue, disturbed sleep, distress, pain, drowsiness, sadness, memory, numbness, dry mouth, lack of appetite, shortness of breath, nausea and vomiting

[2]difficulty swallowing, difficulty speaking, mucus in throat, difficulty tasting food, constipation, teeth/gum issues, mouth/throat sores, choking, and skin pain

[3]work, enjoyment, general activity, mood, walking, relationships

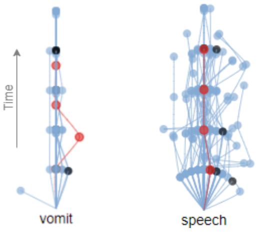

Figure 2: Tendril plots showing stable (vomit) vs volatile (speech) symptom development over time. Each plot shows all filtered patients as blue tendrils while the selected patient is highlighted in red. Dots encode time points and the black ones show the last time stamp of deceased patients.

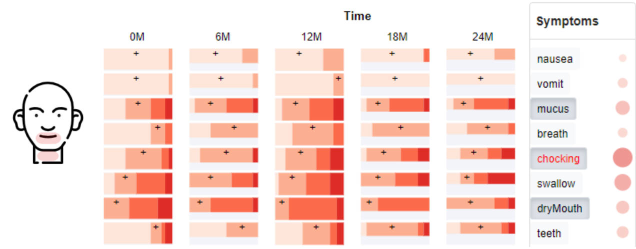

Figure 3: Stacked bar charts dividing patients into 4 rating groups shown using colors for all symptoms and each time point. Crosses point the selected patient's position among the groups. Alongside the symptom list is the correlation plot between symptom dry mouth and the other symptoms. The anatomical sketch (left) highlights the affected areas by the selected symptoms (mouth and neck).

no other visual system that incorporates all the presented functionalities of QoL data.

Given a time point, a custom scatterplot [Figure 1] in the left main view encodes the patients' features as a 2D projection determined using the Principal Component Analysis (PCA) of symptom ratings. The user can filter this view by: therapeutic combination before treatment, gender, tumor category, symptom severity, and outcome. Patients' outcome is conditioned on their survival and whether at least half of the symptoms improved from the first to the last time point. Using Ward's method, patients are clustered, based on their reported symptom ratings, into high and low severity groups. To emphasize these symptoms' impact on patient groups, we provide an option for dynamically calculating new PCA projections and patient clusters using predefined subsets of symptoms.

To compactly show trends in symptom evolution and identify outliers, 5 tendril plots [Figure 2] placed below the scatterplot encode the development of 5 selected symptoms over time. Rating evolution is segmented into time stamps, starting from the origin. The curvature degree for a tendril at each time step shows the relative change from the previous rating, where clockwise rotation indicates worsening symptoms (rating increase). This representation also facilitates discovering steady and variable progressions of symptoms.

The second main view visualizes correlations and similarities between symptoms [Figure 3]. A composite heatmap shows the distribution of individual symptoms at different time points. Bar graphs show the percentage of patients within a different rating group (0, 1-5, 6-9, or 10) for a given symptom at a given time point.

Related symptoms are listed next to the heatmap, allowing the user to select the 5 symptoms shown in the tendril plot. Correlations between a single target symptom and all other symptoms are shown to the right of the symptom list via circles, which encode Spearman's coefficient using size and color. Additionally, to support visual anchoring with patient anatomy, regions in the head and neck affected by the selected symptoms are highlighted in an anatomical sketch to the left of the heatmap [Figure 3].

Finally, a particular patient can be selected, which will highlight him in all plots, revealing individual characteristics among the overview. Moreover, we provide an option for highlighting the 3 patients who are most similar to the selected patient.

5 Evaluation

We conducted a qualitative evaluation with two end users, a data mining specialist and a clinical radiation oncologist. Due to limitations from the COVID-19 pandemic, these sessions were conducted online. The end users asked questions to direct the exploration, and provided feedback. The tendril plots, composite heatmaps, and anatomical sketch yielded remarkably enthusiastic feedback. In particular, the ability to show a current patient and the practicality of the anatomically-inspired layout of symptoms in the context of the heatmap was deeply appreciated. The explicit link between symptoms and the anatomical sketch was considered very useful, since patients often point to the location of their symptoms. For future work, the users expressed interest in the ability to add new patients to the system.

Clustering the patients based on symptom severity had revealed two main groups of patients for high and low overall rating severity [Figure 1], which was explored during the evaluation using various filtering operations. The collaborators focused on an unusual group of patients with low symptoms yet adverse outcomes, which was noted as an intriguing point of future investigation.

6 Discussion and Conclusion

Our interactive visual analysis system tackles a difficult problem in radiation oncology: relating dynamic patient QoL data to the anatomical location of the patient treatment and the therapeutic combination selected by the clinician. Our interface successfully links the QoL data to its underlying spatial and dynamic aspects. Our preliminary qualitative evaluation session shows that our interface is helpful in assisting domain experts in exploring the existing dataset, formulating new hypotheses, and potentially using the system in the clinic when a new patient comes for a visit.

Acknowledgments

This work is supported by the US National Institutes of Health, through awards NIH NCI-R01CA214825 and NIH NCI-R01CA2251. We thank all members of the Electronic Visualization Laboratory, members of the MD Anderson Head and Neck Cancer Quantitative Imaging Collaborative Group, and our collaborators at the University of Iowa and University of Minnesota.

References

[1] D. J. Adelstein et al. Head and neck squamous cell cancer and the human papillomavirus: summary of a national cancer institute state of the science meeting, november 9–10, 2008, washington, dc. *Head & Neck: Journal for the Sciences and Specialties of the Head and Neck*.

[2] G. B. Gunn, T. R. Mendoza, C. D. Fuller, I. Gning, S. J. Frank, B. M. Beadle, E. Y. Hanna, C. Lu, C. S. Cleeland, and D. I. Rosenthal. High symptom burden prior to radiation therapy for head and neck cancer: a patient-reported outcomes study. *Head & neck*, 2013.

[3] M. Karpefors and J. Weatherall. The tendril plot—a novel visual summary of the incidence, significance and temporal aspects of adverse events in clinical trials. *Journal of the American Medical Informatics Association*, 2018.

[4] G. E. Marai. Activity-centered domain characterization for problem-driven scientific visualization. *IEEE transactions on visualization and computer graphics*, 2017.

[5] C. Plaisant, B. Milash, A. Rose, S. Widoff, and B. Shneiderman. Lifelines: visualizing personal histories. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 221–227, 1996.