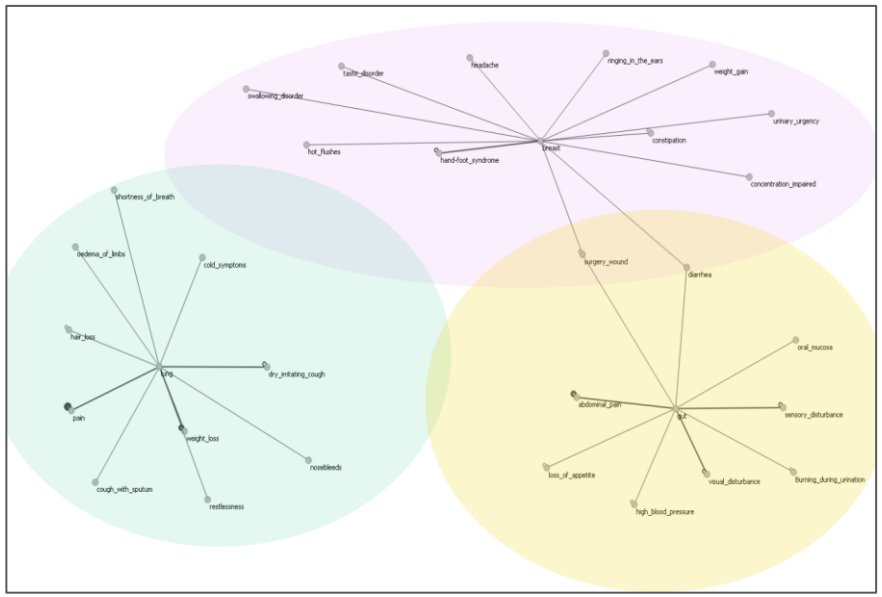
Dynamic electronic reporting of treatment related symptoms (ePROs) can reversely identify the type of underlying cancer

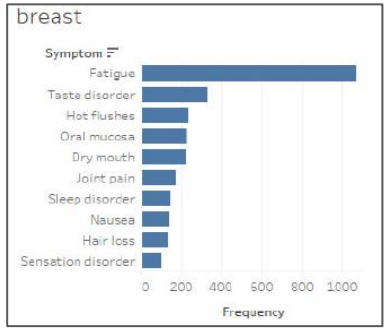
A. Trojan 1, HF Witschel 2, GA Kullak-Ublick 1, M Kiessling3 , N Asper 4

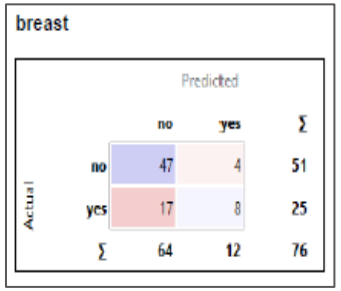
(1) Department of Clinical Pharmacology and Toxicology, University Hospital Zurich, University of Zurich, Zurich, Switzerland, (2) Fachhochschule Nordwestschweiz, Institut für Wirtschaftsinformatik, Olten, Switzerland, (3) Oncology, BreastCenter Zürichsee, 8810 Horgen, Switzerland, (4) University of Zurich, https://www.med.uzh.ch/en.html Zurich, Switzerland

Background: Digital symptom reporting of patients undergoing systemic treatment demonstrated early detection of symptoms, reduction of unplanned admissions, and may support Machine learning (ML) to predict when patients will require emergency treatments. We examined whether dynamic reporting of treatment related symptoms and high quality ePROs can reversely identify the type of underlying cancer and treatment.

Methods: 226 patients on treatment had self-reported on presence and severity (according to CTCAE) of more than 90 available symptoms via a medical device App. For a balanced analysis we used data from 25 patients treated for breast cancer, 19 for cancer of lung, 16 for colon, 12 for lymphoma and 7 for prostate cancer, respectively. Patients` symptoms over the entire study period were aggregated by counting the days on which a particular symptom was reported. Thus, each patient was represented by a vector of symptoms indicating how often the given symptom occurred. A human-interpretable ML logistic regression model was applied to predict the primary tumor of the patient from his/her respective symptom vector. All symptoms with positive coefficient above a certain threshold (0.1) were collected and then graphically displayed for association between symptoms and cancer type.







Results: With respect to ePRO Data, on average, participants reported 3.0 symptoms daily (with a range of 1.2 to 3.3 symptoms), resulting in a total symptom count of 43'430. The average duration of symptom tracking was 82 days (ranging from 14 to 225 days). The ML model was not able to recognize the prostate and blood-lymph patients in retrospect since their number was too small. Analysis for three remaining cancer types revealed a mean area under the curve (AUC) score of 0.72 (breast cancer AUC 0.74, CI: 0.62–0.85; gut cancer AUC 0.78, CI: 0.66–0.89; lung cancer AUC 0.63, CI: 0.50–0.77). Results indicate that ML performs "fair" and significantly better than random guessing (which would result in AUC = 0.5) for the reverse identification of the underlying cancer upon ePRO reporting from patients.

Conclusions: Cloud aggregation of patient reported symptoms and ML harbor the potential in identifying the type of cancer for which patients receive systemic treatment. Whether associations can be made from dynamic changes of reported symptoms, regarding the underlying cancer and adherence to oral medication shall be explored in prospective studies. Finally, ML and the anticipation of specific side effects might be a cost-effective tool in decentralized clinical trials and registries, enabling a more nuanced understanding of symptom associations with different cancer types.

Trial Registration: ClinicalTrials.gov NCT03578731;

Ref:- Submitted to: Front Dig Health 2025, and - Trojan et al. (Front Dig Health 2024) Towards an Early Warning System for Monitoring of Cancer Patients Using Hybrid Interactive Machine Learning

Presented at SOHC 2024 from 20 – 22 November 2024