Data Preparation Challenges in Analysis of Recurrence-Free Survival in Breast Cancer Using Saxony-Anhalt Registry Data

Ian Wittenberg[1], Ahmed Bedir[1,2], Alexander Kluttig[1], Maria Elena Lacruz[1], Saskia Thies[1]

[1]Clinical Cancer Registry Saxony-Anhalt, Halle (Saale), Germany, Contact: i.wittenberg@kkr-lsa.de
[2]University Clinic of Radiation Oncology, Medical Faculty of Otto-von-Guericke-University, Magdeburg, Germany

Objectives: Breast cancer can recur at various times after diagnosis, making recurrence-free survival an essential measure that provides valuable insights into prognosis and treatment effectiveness. (1) The reporting of recurrence and progression can be facilitated in the German Standardized Oncological Basic Data Set (oBDS) (2) through various reporting events. These include pathology follow-up reports with new UICC staging, clinical reports of new metastases or remission, or therapy reports targeting specific tumor parts. Determining the timepoint of recurrence is challenging due to the necessity of identifying a prior complete remission. However, the absence of such data may be mitigated by conditional supplementation of information from other reporting events.

Methods: For 48,598 breast cancer patients diagnosed between 1993 and 2022 194,565 follow-up reports have been documented in 461,677 person-years. Following an exploration of database dependencies, deterministic rules for cross-validation of different reporting events will be established. These rules are of particular importance in cases where multiple reports contain contradictory information (e.g., complete remission following a UICC IV pathological report or surgery shortly after complete remission report). Survival analyses will be conducted using the Kaplan-Meier method and the Cox proportional hazards model.

Results: The establishment of robust rules for conditional supplementation is expected, leading to more accurate outcome determination. This will increase the reliability of recurrence-free survival analysis and reduce the proportion of missing values in registry data.

Conclusions: The established conditional supplementation rules will improve the reliability of recurrence-free survival analyses using cancer registry data. This approach will improve the usability of cancer registry data and support better clinical decision-making and research in breast cancer management.

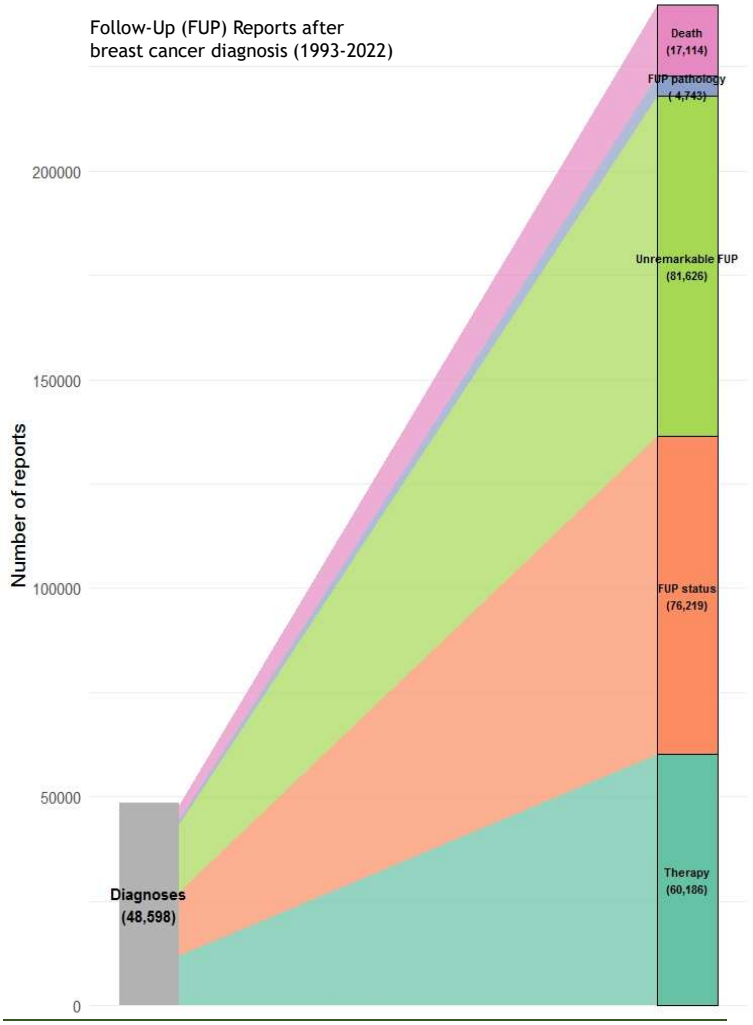

Figure 1. Diagram illustrating the distribution of therapy and follow-up (FUP) reports linked to 48,598 breast cancer patients diagnosed between 1993 and 2022

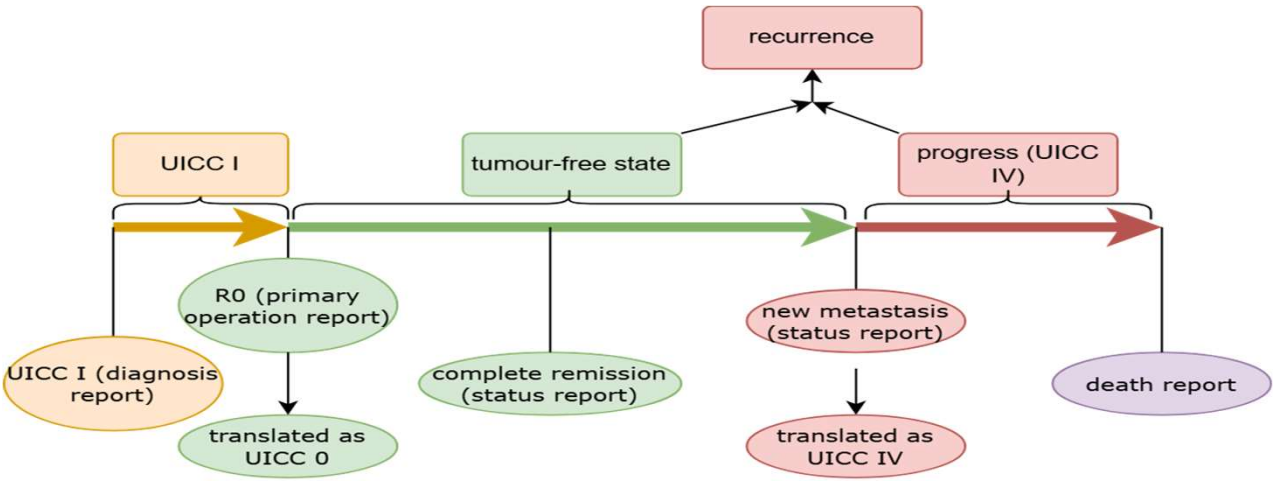

Figure 2. Temporal Progression of a Representative Breast Cancer Patient: This diagram illustrates the clinical course of an exemplar breast cancer patient over time. The lower section displays individual patient reports corresponding to various time points, with some reports conditionally supplemented based on the appropriate UICC stage—from tumor-free status (UICC 0) to advanced metastatic disease (UICC IV). The upper section depicts the phases associated with a specific UICC stage or tumor-free condition at each time point derived from individual reports. Notably, the emergence of new metastases, a progression to a higher UICC stage, or explicitly documented progression reports are classified as progression events. Finally, recurrences are defined as outcomes identified by an increase in UICC stage at the conclusion of a tumor-free interval.

Literature: (1) Jung, H., Lu, M., Quan, M.L., Cheung, W.Y., Kong, S., Lupichuk, S., Feng, Y. and Xu, Y. (2022). New method for determining breast cancer recurrence-free survival using routinely collected real-world health data. BMC Cancer, 22(1). doi:https://doi.org/10.1186/s12885-022-09333-6. (2) German Cancer Registries (GNCR). (2018). ADT/GEKID Basisdatensatz | DKR - Homepage. [online] Available at: https://www.dkr.de/adt-gekid-basisdatensatz [Accessed 23 May 2025].

www.kkr-lsa.de