

<https://doi.org/10.29289/259453942021V31S2044>

THE RECORD LINKAGE TECHNIQUE AS ALTERNATIVE FOR CAPTURING OF INFORMATION IN CANCER REGISTRY: APPLICATION IN A DATABASE OF BREAST CANCER CASES

Stela Verzinhasse Peres¹, Paola Engelmann Arantes², Alexandre Muxfeldt Ab'Saber³, Daniel Luiz Gimenes⁴, Maria Paula Curado², René Aloisio da Costa Vieira^{5,6}

¹Fundação Oncocentro de São Paulo – São Paulo (SP), Brazil.

²AC Camargo Cancer Center – São Paulo (SP), Brazil.

³Fundação Oncocentro de São Paulo – São Paulo (SP), Brazil.

⁴Grupo Oncoclínicas – São Paulo (SP), Brazil.

⁵Hospital de Câncer de Muriaé – Muriaé (MG), Brazil.

⁶Fundação Pio XII Hosp de Câncer de Barretos – Barretos (SP), Brazil.

Background: Studies that analyze biomarkers as prognostic factors are scarce due to the high cost of the examination in low-income populations. **Objective:** The aim of this study was to describe the sociodemographic variables, clinical, treatment characteristics, and molecular subtype from women with breast cancer through probabilistic and deterministic record linkage between Hospitalar Cancer Registry and Immunohistochemistry Laboratory database at the Oncocenter Foundation of São Paulo (FOSP). **Methods:** To obtain a complete follow-up of the patients, a link was made between the two bases of FO SP. The deterministic linkage, using the examination number, was applied between the database composed of hormonal receptors and the Ki-67 antigen versus the database formed by the HER-2 receptor, stored in the immunohistochemistry laboratory's (FO SP) information system. The probabilistic linkage was performed on the program OpenReclink III, version 3.1.615. The variables used in the matching process were the patient's name and date of birth. As for blocking, the variables were soundex of the first name, soundex of last name. Confirmatory variables for acceptance of a true match were, when available, mother's name, home address, Public Health System (SUS) card, and/or another personal identification document. **Results:** As result, 1,654 patients were matched. The average age was 56.8 (SD=13.2), with a median age of 56, varying between 22 and 96 years. Of these, 48% completed middle school. As for clinical characteristics, most tumors were of the ductal type (71.9%); and 7.2% of patients were presented with distant metastasis at diagnosis. Among the biomarkers, there was 25.9% for ER-, 33.9% for PR-, 85.5% negative HER-2, and 68.3% Ki-67 ($\geq 14\%$). Of the molecular subtypes, the Luminal B (Her2-) phenotype was more frequent among patients and 15.2% were triple negative. **Conclusion:** The linkage techniques contributed with the completeness of information, contributing in defining the vital status of the patient.

Keywords: Medical Record Linkage; Breast Neoplasms; Biomarkers.