

PL-10

Developing an interoperable advanced cancer surveillance system: the DARE model in Sicily

T. Lo Presti (1), E. Salamone (3), S. Fruscione (1), M. Zarcone (2), M. Guzzardo (1), S. Lo Verso (4), S. Pollina Addario (5), M. Ferrante (3,6), W. Mazzucco (1)

(1) University of Palermo, Palermo, Italy; (2) Azienda Ospedaliera Universitaria Policlinico di Palermo "Paolo Giaccone", Palermo, Italy; (3) Azienda Ospedaliera Universitaria Policlinico di Catania "G. Rodolico-San Marco", Catania, Italy; (4) ARPA Sicilia, Italy; (5) Dipartimento Attività Sanitarie con Osservatorio Epidemiologico, Assessorato della Salute Regione Siciliana, Italy; (6) Università degli Studi di Catania, Catania, Italy

Background. Fragmented health information systems limit the completeness, timeliness and interoperability of population-based cancer registries. Within the DARE (Digital Lifelong Prevention) project, two complementary Sicilian pilots aim to strengthen cancer surveillance through regional data integration and artificial intelligence (AI)-assisted case registration.

Materials and Methods. The Advanced Cancer Registry pilot designed an Institutional Repository to harmonize data from local registries with national specifications under the Italian Ministerial Decree of 1 August 2023, and an Advanced Research environment based on Giotto BigData to link cancer data with healthcare and environmental sources. Regional and national data fields were mapped with registry representatives and the Regional Health Department. The LINK pilot addresses integration of pathology reports and primary-care information into population-based registries, combining machine learning and Lean Six Sigma principles. An AI-assisted coding system processes free-text pathology reports according to AIRTUM and ICD-O-3 coding standards.

Results. For the Advanced Cancer Registry, the assessment identified field correspondences, discrepancies and harmonization requirements and defined a centralized regional dataset and standardized data-transfer workflow. The proposed architecture includes secure CSV ingestion, data validation and consolidation, national XML export and epidemiological dashboards. Within the LINK-related AI component, 1,050,952 free-text pathology reports from 2003–2018 were selected for algorithmic analysis. The best-performing algorithm achieved a reported precision of 98.9%, with processing speed of approximately 4 seconds per 1,000 reports. A 60–80% reduction in manual coding time is an expected benefit, not a demonstrated outcome.

Conclusions. The two DARE pilots offer complementary approaches to cancer surveillance: standardized regional interoperability and AI-supported identification and coding of cases. Further implementation and independent validation are needed to establish their effects on registry quality, timeliness and public-health decision-making.

Note. This research is co-funded by the Ministry of University and Research within the Complementary National Plan PNC-I.1 "Research initiatives for innovative technologies and pathways in the health and welfare sector" D.D. 931 of 06/06/2022, PNC0000002 DARE - Digital Lifelong Prevention

References. 1. Bray F, Parkin DM. Evaluation of data quality in the cancer registry: principles and methods. Part I: comparability, validity and timeliness. Eur J Cancer. 2009;45(5):747-55. 2. Giusti F, Martos C, Carvalho RN, Zadnik V, Visser O, Bettio M, et al. Facing further challenges in cancer data quality and harmonisation. Front Oncol. 2024;14:1438805. 3. Italian Ministry of Health. National Cancer Registry [Decree of 1 August 2023]. Gazz Uff Repub Ital. 2023 Aug 31;(203).