POSTER ID: EPH80

Epidemiology and Outcomes in a Linked Cohort of Cancer Data, Pathology Data, and Electronic Medical Records from the PHARMO Data Network

Authors: Eleonora Iob[1], Nikita Jeswani[1], Naomi Reimes[1], Rinus Voorham[2], Josine Kuiper[1]

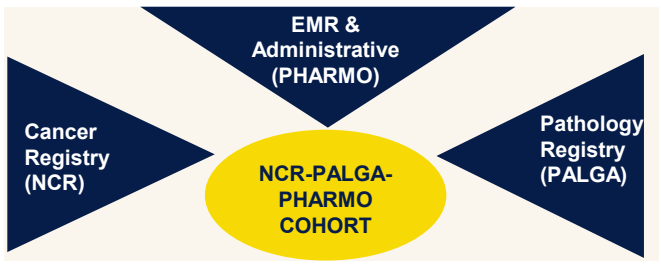
1. PHARMO Institute for Drug Outcomes Research, Lumanity, Utrecht, The Netherlands;
2. PALGA, Houten, The Netherlands.

INTRODUCTION

Leveraging real-world pathological data for studying targeted cancer therapies

- Understanding the intricate relationship between cancer histopathology and personalized treatment outcomes is crucial for advancing cancer care.
- The top five cancer types for which pathological data are essential for studying targeted therapies include:
  - Breast cancer
  - Colorectal cancer
  - Lung cancer
  - Skin cancer
  - Ovarian cancer
- These cancer types were chosen based on a combination of high prevalence and availability of established biomarkers with corresponding targeted therapies.
- Linked cancer data from the **Netherlands Cancer Registry (NCR)**, pathology data from the **Dutch Nationwide Pathology Databank (PALGA)**, and electronic medical records (EMR) from the **PHARMO Data Network** provide unique opportunities for research into new targeted cancer therapies and assessing their safety and effectiveness in real-world settings.

Figure 1. Linking Cancer, Pathology, and EMR data sources to create the NCR-PALGA-PHARMO cohort



OBJECTIVES

- Describe the **top five cancer types** requiring pathological data and identify their key biomarkers.
- Derive **cancer patient cohorts** specific to these types of cancer, using linked cancer data from **NCR**, pathological data from **PALGA**, and EMR data from the **PHARMO Data Network**.
- Estimate **the number of patients** captured in each cohort, their **survival rates**, and the prevalence of key **biomarkers**.
- Describe the most commonly used **biomarker-driven targeted therapies**.

METHODS

- This study was conducted using data obtained from **NCR**, **PALGA**, and the **PHARMO Data Network**.
- **NCR** is a national population-based registry with nationwide coverage maintained by the Comprehensive Cancer Centre of the Netherlands (IKNL) and comprises information on newly diagnosed cancer patients in the Netherlands. NCR contains information on the actual diagnosis of cancer and specific tumor and treatment characteristics.
- **PALGA** is a Dutch, nationwide network and registry of histo- and cytopathology in the Netherlands. PALGA contains (standardized) pathology reports of all Dutch inhabitants since 1991. The abstracts contain, amongst others, encrypted patient identification, demographic data, and a summary of the pathology report coded in accordance with SNOMED terminology (1).
- **PHARMO** is a population-based data source with combined anonymous electronic healthcare data from different primary and secondary healthcare settings in the Netherlands (2). Currently, the PHARMO Data Network covers over 7 million active persons out of 17 million inhabitants of the Netherlands.
- These three data sources were linked at the patient level, enabling each patient to be tracked across all sources.
- We estimated the number of patients captured in the linked cohort, along with their survival outcomes (across all stages) between 2015 and 2023.
- We also summarised key biomarkers and their prevalence within the cohort, as well as the most commonly used biomarker-driven targeted therapies in high-cost medicine data.

Table 1. Patient counts and availability of biomarkers of predefined cancers from the NCR-PALGA-PHARMO cohort

Type of Cancer	N Patients in cohort	Pathological marker	N patients with marker
Breast	46,175	HER2	27,958
Colorectal	36,382	KRAS/EGFR	4,553/1,850
Lung	50,908	KRAS/EGFR	10,576/10,465
Skin	19,428	BRAF	1,853
Ovarian	4,154	BRCA1/2	1,213/1,147

Results

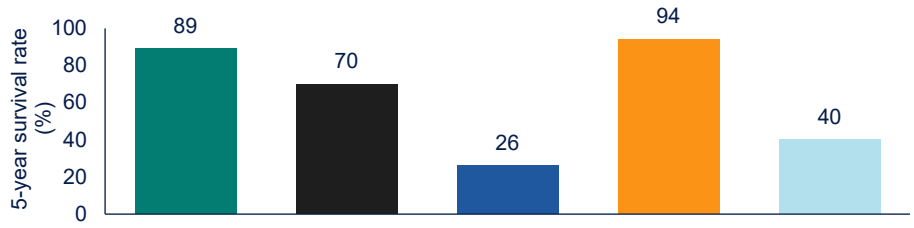
Breast Cancer:

- **Cohort size and survival:** 46,175 patients are captured, with a 5-year survival rate of ~89%.
- **Biomarkers:**
  - The human epidermal growth factor receptor 2 (HER2) mutation, found in ~4% of patients, is crucial as HER2-positive patients respond well to targeted therapies (e.g., trastuzumab).
  - In our linked cohort, 27,958 patients with HER2 status were identified based on structured data: 25,515 negative and 3,442 positive.

Colorectal Cancer

- **Cohort size and survival:** 36,382 patients are captured, with a 5-year survival rate of ~70%.
- **Biomarkers:**
  - Kirsten rat sarcoma viral oncogene homolog (KRAS) mutations are present in ~40% of patients, while epidermal growth factor receptor (EGFR) overexpression occurs in 60-80% of tumors. KRAS mutations predict the efficacy of EGFR inhibitors.
  - 1,294 patients have molecular data available, with 381 KRAS mutations and 15 EGFR mutations (structured data). Free-text searches of the patients' pathology reports provided 4,553 patients with KRAS mentions and 1,850 patients with EGFR mentions.

Figure 2. Five-year survival rates across cancer types (all stages) in the NCR-PALGA-PHARMO cohort



Lung Cancer:

- **Cohort size and survival:** 50,908 patients are captured. Only ~26% of lung cancer patients survive to year 5.
- **Biomarkers:**
  - EGFR mutations are associated with better responses to EGFR-tyrosine kinase inhibitors (TKIs), while KRAS mutations predict resistance.
  - 2,498 patients have molecular data available, with 191 EGFR mutations and 683 KRAS mutations (structured data). Free-text searches provided 10,576 patients with KRAS mentions and 10,465 patients with EGFR mentions.

Skin Cancer (Melanoma):

- **Cohort size and survival:** 19,428 patients are captured, with a 5-year survival rate of ~94%.
- **Biomarkers:**
  - The B-Raf proto-oncogene, serine/threonine kinase (BRAF) mutation is found in ~50% of melanomas. BRAF inhibitors are effective in treating BRAF mutation-positive melanoma.
  - 613 patients have molecular data available, with 218 BRAF mutations (structured data). Free-text searches provided 1,853 patients with BRAF mentions.

Ovarian Cancer:

- **Cohort size and survival:** 4,154 patients are captured, with a 5-year survival rate of ~40%.
- **Biomarkers:**
  - BRCA1 and BRCA2 gene mutations increase ovarian cancer risk and are associated with better responses to platinum-based chemotherapy and poly-ADP ribose polymerase (PARP) inhibitors.
  - 127 patients have molecular data available, with 2 BRCA1 mutations and 1 BRCA2 mutation (structured data). Free-text searches provided 1,213 patients with BRCA1 mentions and 1,147 patients with BRCA2 mentions.

CANCER TREATMENTS

The most commonly used biomarker-driven targeted therapies in the NCR-PALGA-PHARMO cohort are shown in **Table 2.**

Table 2. Targeted cancer therapies by biomarker in the NCR-PALGA-PHARMO cohort

Cancer	Biomarker	Drug name	Stage/Treatment line
Breast	HER2	Trastuzumab	Early cancer, adjuvant/neoadjuvant
		Palbociclib	Advanced/metastatic cancer, first-line or second-line
		Pertuzumab	Early or locally advanced, neoadjuvant
		Paclitaxel	Early or locally advanced, neoadjuvant
		Capecitabine	Adjuvant combination therapy following neoadjuvant
Colorectal	KRAS	Cetuximab	Metastatic, second-line
		Irinotecan	Metastatic, first-line or later
	EGFR	Trifluridine comb.	Metastatic, later-line monotherapy
		Cetuximab	Metastatic, ≥ second-line
		Pembrolizumab	Metastatic, first-line
		Nivolumab	Metastatic, ≥ second-line
		Oxaliplatin	Metastatic, first-line
Lung	KRAS	Pembrolizumab	Metastatic, first-line
		Pemetrexed	Metastatic, first-line combination
		Osimertinib	Locally advanced/metastatic, first-line
		Erlotinib	Locally advanced/metastatic, first-line
		Afatinib	Locally advanced/metastatic, monotherapy
Skin	BRAF	Dabrafenib	Inoperable/metastatic, mono/combo (trametinib)
		Trametinib	Inoperable/metastatic, mono/combo (dabrafenib)
		Binimetinib	Inoperable/metastatic, combo (encorafenib)
		Encorafenib	Inoperable/metastatic, combo (binimetinib)
		Cobimetinib	Inoperable/metastatic, combo (vemurafenib)
Ovarian	BRCA1/2	Olaparib	Advanced/metastatic high-grade, mono after first-line platinum-based treatment
		Niraparib	Advanced/metastatic high-grade, mono after first-line platinum-based treatment

REFERENCES

1. Casparie M, Tiebosch ATMG, Burger G, Blauwgeers H, van de Pol A, van Krieken JHJM, et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. Anal Cell Pathol 2007;29:19–24.
2. Kuiper JG, Bakker M, Penning-van Beest FJA, Herings RMC. Existing Data Sources for Clinical Epidemiology: The PHARMO Database Network. Clinical Epidemiology. 2020;12:415-22.

CONCLUSIONS

- By linking cancer, pathology, and EMR data, we created a large **cohort of cancer patients for studying biomarkers and targeted cancer therapies**, providing extensive research opportunities into patient characteristics, treatment patterns, and long-term outcomes.
- The prevalence of cancers and biomarkers in the NCR-PALGA-PHARMO cohort underscores the **suitability** of these data for research on targeted therapies.
- **Lines of Therapy (LoT) algorithms** can be derived by capitalising on different sources of information, including data on cancer therapies at the time of diagnosis and during follow-up. This enables stratification of treatment response by biomarkers, facilitating the study of targeted therapies. Benefits of this approach include:

✓ **Comprehensive settings coverage:** Includes both inpatient and outpatient drug dispensings and high-cost medicines across the cancer care continuum.

✓ **Detailed dispensing information**: Duration, dose strength, and posology enable accurate identification of treatment start/stop dates, switches, and combinations.

✓ **Linkage with NCR**: Integration with the Netherlands Cancer Registry adds data on diagnosis, stage, and initial therapies for contextualising treatment pathways.

✓ **Accurate LoT derivation**: Supports precise reconstruction of treatment sequences and intervals.

✓ **Proxies for outcomes**: LoT transitions serve as real-world indicators of disease progression, recurrence, and progression-free survival (PFS).

- The NCR-PALGA-PHARMO cohort can shed light on understanding differences in survival and treatment response as they relate to the presence of different biomarkers.
- More evidence is needed on the effectiveness of treatments in populations exhibiting specific biomarkers, an area that can be investigated using linked cancer, pathology, and EMR data sources.

Poster presented at ISPOR Europe 2025; 9-12 November; Glasgow, Scotland, UK.