Harmonization of routine care data from hospitals in the Digital Oncology Network for Europe (DigiONE) into Observational Medical Outcomes Partnership (OMOP) databases reveals changes in the number of new primary cancers diagnosed and 12-month survival during COVID-19 lockdowns

S Theophanous*[1], H Fenton*[2], A Lobo Gomes[3], E Ross[4], J Thonnard[5], A Wolf[6], C Brandts[6], A-L Bynens[7], G Hall[1], S Bachir[8], E Bolton[1], O Bouissou[4], D Brucker[6], S Cheeseman[1], A Collard[5], A Dekker[3], P Galgane Banduge[3], L Halvorsen[9], I Kaczmarczyk[10], D Kadioglu[8], P Kalendralis[3], J Khan[11], P Mahon[2], T Schneider[6], L Schumann[6], A Traverso[12], A van Maanen[5], C van Marcke[5], A Vengadeswaran[8], J Wörmann[6], J Yeap[2], T Yousaf[4], R McDonald**[2], E Hallan Naderi**+[4]

*Joint first authors; **Joint last authors; +corresponding author: elinad@ous-hf.no. [1]Leeds Teaching Hospitals NHS Trust, United Kingdom. [2]IQVIA Ltd, United Kingdom. [3]Department of Radiation Oncology (Maastro), GROW Research Institute for Oncology and Reproduction, Maastricht University Medical Centre+, The Netherlands. [4]Oslo University Hospital, Norway. [5]Cliniques Universitaires Saint-Luc, Brussels, Belgium. [6]University Cancer Center, University Hospital Frankfurt, Germany. [7]Maastricht University Medical Center+, The Netherlands. [8]Goethe University Frankfurt, University Medicine, Institute of Medical Informatics (IMI), Frankfurt am Main, Germany. [9]edenceHealth NV, Belgium. [10]OMOP & PPG Solutions, IQVIA Ltd, London, United Kingdom. [11]IQVIA Ltd, India. [12]IRCCS San Raffaele Hospital, Milan, Italy.

INTRODUCTION

- The curation, standardization, and harmonization of routine care data across hospitals is complex due to its often unstructured, incomplete, and dissimilar nature.
- This study standardized and harmonized data in OMOP across 5 hospitals in Europe to investigate the impact of COVID-19 lockdowns on new cancer diagnoses and 12-month survival covering the period before the pandemic, during peak lockdowns, and the subsequent recovery phases.

METHODS

- Five academic, public, and general oncology treating DigiONE hospitals transformed a core dataset from their electronic medical records, into local OMOP databases to establish the first pan-cancer European hospital OMOP network for federated analysis[1] (Figure 1).
- Patients with a diagnosis of one of the 11 cancer groups (breast, prostate, lung, colorectal, upper gastrointestinal, hematological, gynecological, head and neck, non-prostate urological, melanoma, brain) between 01-Jul-2018 and 30-Sep-2023 were included.
- Interrupted time-series analysis using generalized least squares was applied to assess changes in the volume of cancer diagnoses at each hospital, as well as overall as a network for each cancer group. Data between Jul-2018 to Jun-2019 were used as the baseline.
- For patients with a 12-month follow up, 12-month overall survival was described by three-month (quarterly) cohorts based on index date.

Figure 1. Approach to harmonizing data across five hospitals

- Study design committee aligned on objectives and variable requirements
- Study documentation development – protocol & statistical analysis plan
- Ethics approval from participating hospitals
- Data source identification with each hospital data lake
- Data standardisation
  - Mapping of source diagnosis codes into standard OMOP concepts[2]
- Extract, transform, load (ETL) by hospitals following DigiONE network ETL handbook
- OMOP database quality check (QC) following DigiONE network QC points
- Medical code list development and validation by clinicians
- Patient identification from OMOP databases using SQL
  - Output a flat file with patient minimal dataset for analysis, which sits within respective hospital
- Analysis scripts sent to hospital databases for on-premise analysis against the minimal dataset
  - Output hospital-level aggregated results to be shared externally
- Meta-analysis of hospital-level results for network level results

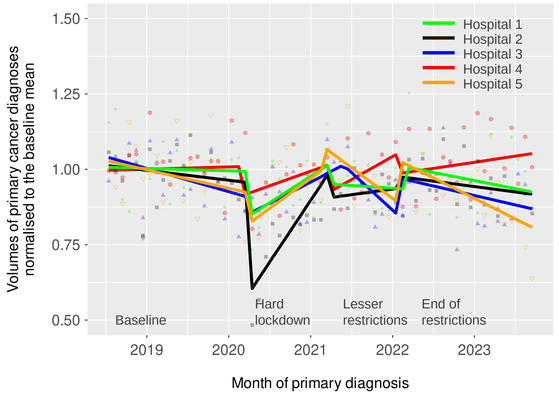
RESULTS

- The study included 124,682 patients across 5 hospitals. There was a four-fold difference in the average number of monthly primary cancer diagnoses between the smallest and the largest hospital (Table 1).
- The mean age at index was 63.9 years and 50.13% were male. There was no statistical significance change in age and sex groups over the study period.

Table 1: Volume of monthly diagnoses of all eligible cancer groups over the study period

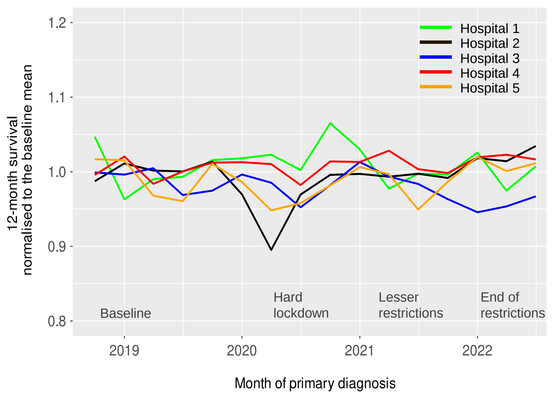
	Hosp 1	Hosp 2	Hosp 3	Hosp 4	Hosp 5
Average	197	356	240	763	172
Minimum	149	186	187	592	111
Maximum	255	428	290	907	220

Figure 2. Change in monthly primary cancers diagnoses in each hospital, across phase transitions using country-specific dates



- A statistically significant change in monthly volumes of primary cancers from baseline to hard lockdowns (all in March 2020) was observed in Hospitals 1 and 3 (p<0.05) and in Hospitals 2 and 4 (p<0.001).
- Hospitals 2 and 4 also showed a notable change in monthly volumes of primary cancers from hard lockdowns to lesser restrictions (end March to mid April 2021, p<0.005).

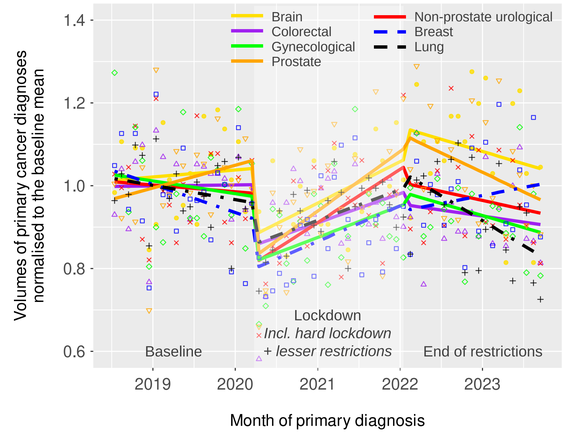
Figure 3. Change in proportion of patients surviving to 12 months from index in each hospital, across phase transitions using country-specific dates



- Proportion of newly diagnosed cancer patients surviving 12 months mirrored the drop in volume of primary cancer diagnoses from baseline to hard lockdown, particularly in Hospitals 2, 3 and 5.
- The trend of poorer survival during hard lockdowns which returns to pre-pandemic levels by the end of the study period was observed in four hospitals.
- Nevertheless, the lack of data collected on disease stage and cause of death, including COVID-19 related deaths, limits the interpretation of the current survival data and its association with the volume of primary cancer diagnoses.

Figure 4. Change in new diagnosis volumes in each cancer group as a network

Only shows cancer groups with a significant step change



- The overall number of primary cancer diagnoses significantly decreased from baseline to lockdown across the network (p<0.005).
- Statistically significant declines were observed in prostate (p<0.005), brain, colorectal, non-prostate urological, and gynaecological (p<0.05) cancers.
- Lung and breast cancer showed near-significant declines (p=0.057 and p=0.078, in dashed regression lines).

CONCLUSION

- This first pan-cancer study from the DigiONE network showcases the ability of 5 European hospitals to standardize and harmonize data in OMOP for over 120,000 patients collectively, collaborating on **data analysis in a privacy-preserving manner without transfer of patient level data**.
- The findings presented here showcase the use of the harmonized data for **comparative analyses of care practices and outcomes across different countries**, as well as **trends at the European network level**.
- The robust approach to OMOP data integration and harmonization at DigiONE hospitals not only **enhances the overall data reliability**, but also **expands the analyzable patient population**, facilitating faster collaborations for generating reliable real-world evidence in precision oncology research.

References

1. Mahon et al. Nat. Med., 30, 334–337 (2024)
2. Fenton et al. OHDSI Europe 2024 Poster

Acknowledgement

This work is part of the DigiCORE pan-European research network comprised of academic cancer centers and two industrial members (IQVIA and Illumina). The DigiONE project was funded by IQVIA and Illumina.