ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ESMO REAL WORLD DATA AND DIGITAL ONCOLOGY

SHORT COMMUNICATION

AI-HOPE lung cancer: a multicenter real-world registry integrating artificial intelligence for metastatic non-small-cell lung cancer

F. R. Ogliari[1], M. Ferrara[1], J. Huijs[2,3], A. Traverso[4], S. Barbieri[4], D. Tiano[4], P. Scuri[4], D. Celada[4], L. Papotto[1,4], L. Mollica[1,4], S. Oresti[1], R. Ferrara[1,4], M. Steens[5], C. Steendam[6], A. Merlini[7], A. Luciani[8], D. Lang[9], S. Grisanti[10], C. Genova[11], M. Gemelli[12], M. Flospergher[13], G. Finocchiaro[14], M. Y. El-Soud[15], D. Cortinovis[16], G. Cerea[17], F. Caspani[18], A. Brouns[19], F. Bensch[20], L. Belluomini[21], C. Bareggi[22], J. Albuquerque[23], A. Esposito[4,24], C. Tacchetti[4,24], A. Bulotta[1], L. Hendriks[2†] & M. Reni[1,4†*]

[1]Department of Oncology, IRCCS Ospedale San Raffaele, Milan, Italy; [2]Department of Respiratory Medicine, Maastricht University Medical Center, GROW - Research Institute for Oncology and Reproduction, Maastricht; [3]Department of Radiation Oncology (Maastro Clinic), Maastricht University Medical Center, GROW - Research Institute for Oncology and Reproduction, Maastricht, The Netherlands; [4]University Vita-Salute San Raffaele, Milan, Italy; [5]Department of Pulmonary Diseases, VieCuri Hospital, Venlo; [6]Department of Pulmonary Diseases, Catharina Hospital, Eindhoven, The Netherlands; [7]Department of Oncology, University of Turin, AOU San Luigi Gonzaga, Orbassano; [8]Department of Oncology, Ospedale di Legnano, Italy; [9]Department of Pulmonology, Kepler University Hospital, Johannes Kepler University, Linz, Austria; [10]Department of Oncology, Spedali Civili di Brescia; [11]Department of Oncology, IRCCS Ospedale Policlinico San Martino, Genova; [12]Department of Oncology, IRCCS Multimedica Sesto San Giovanni; [13]Department of Oncology, Ospedale di Circolo e Fondazione Macchi, Varese; [14]Department of Oncology, IRCCS Humanitas Research Hospital, Rozzano, Milan, Italy; [15]Department of Pulmonology, Maxima MC, Veldhoven, The Netherlands; [16]Medical Oncology, Fondazione IRCCS San Gerardo dei Tintori, Monza; Department of Medicine Università Milano-Bicocca, Milan; [17]Department of Oncology, ASST Grande Ospedale Metropolitano Niguarda, Milan; [18]Department of Oncology, ASST Lariana, Como, Italy; [19]Department of Pulmonology, Zuyderland MC, Heerlen; [20]Department of Pulmonology, UMC Groningen, The Netherlands; [21]Department of Engineering for Innovation Medicine (DIMI), University of Verona School of Medicine and Verona University Hospital Trust, Verona; [22]Department of Oncology, Fondazione IRCCS Ca' Granda-Ospedale Maggiore Policlinico, Milan, Italy; [23]Department of Oncology, Hospital da Luz, Lisboa, Portugal; [24]Experimental Imaging Center, IRCCS San Raffaele Hospital, Milan, Italy

Available online xxx

Background: The AI-HOPE Lung Cancer study is a multicenter initiative designed to integrate artificial intelligence (AI) and real-world data to improve outcome prediction in patients with metastatic non-small-cell lung cancer treated with first-line immunotherapy-based regimens. AI-HOPE aims to leverage machine learning (ML) models to generate individualized predictions of progression-free survival (PFS), overall survival (OS), and treatment-related toxicity in a broad, unselected population.

Materials and methods: Clinical and imaging data are harmonized and stored within a privacy-compliant infrastructure (San Raffaele Ai CEnter [S-RACE] platform), promoting FAIR (Findable, Accessible, Interoperable and Reusable) data principles and minimizing manual workload. The primary objective is the development of time-to-event models for PFS and OS. Complementary binary classification models will explore early progression, long-term survival, and clinically relevant toxicities.

Results: The study includes retrospective (from 2017) and prospective (until 2027) phases across 21 European centers. So far, 920 patients have been recruited for the study, of whom 621 have baseline imaging scans available for centralized analysis. In the AI-HOPE study, a flexible methodological approach integrates multiple ML models tailored to specific clinical questions, complemented by explainable AI tools. Multimodal models combining clinical variables with computed tomography and [^{18}F]2-fluoro-2-deoxy-D-glucose—positron emission tomography imaging features (when available) are supported through the S-RACE platform, which provides a partially automated imaging analysis workflow.

Conclusions: By combining structured clinical variables and multimodal imaging data, the AI-HOPE Lung Cancer study aims to support refined risk stratification and treatment personalization, ultimately facilitating the responsible integration of AI into routine thoracic oncology practice.

Key words: artificial intelligence, immunotherapy, metastatic non-small-cell lung cancer, real-world data

INTRODUCTION

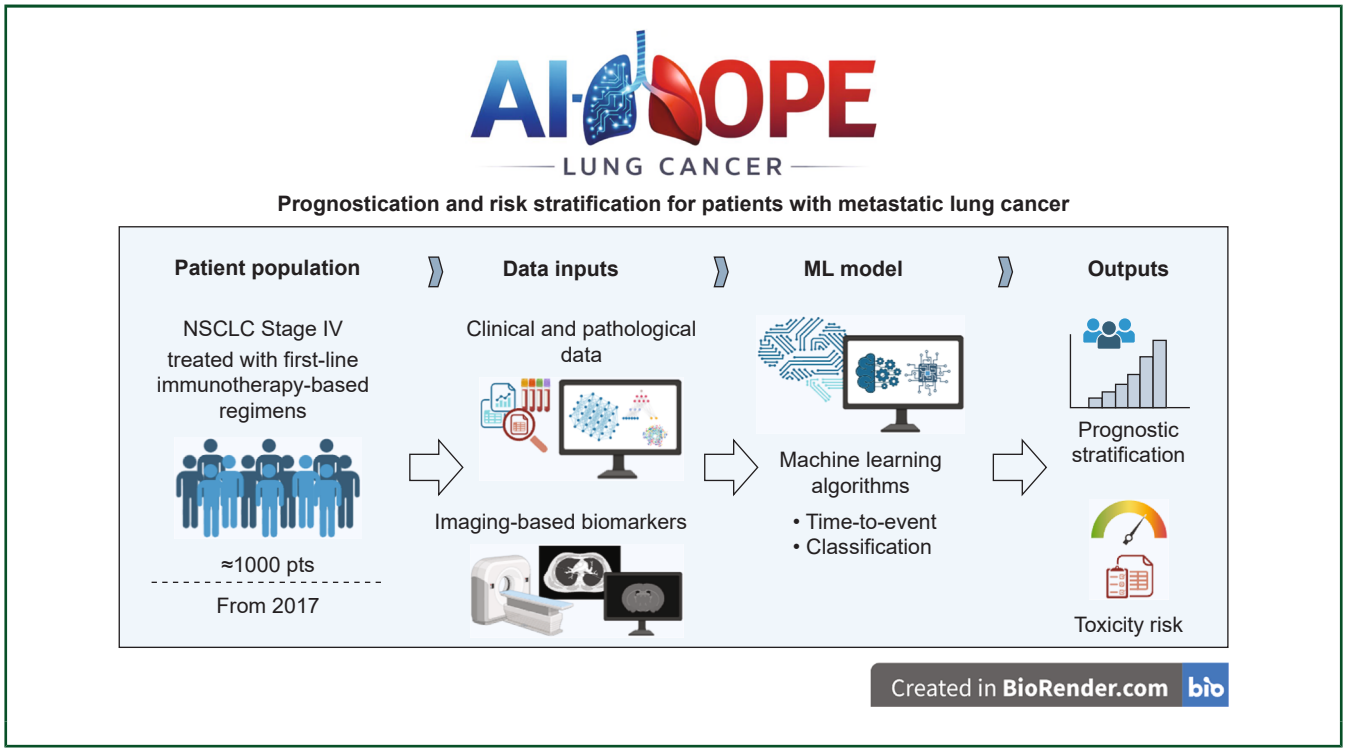
Real-world data (RWD) are a promising source of information in oncology because they provide insight into outcomes in a broader population compared with clinical trials.[1] Therefore, the amount of data that is generated every day during routine clinical practice is considered a 'gold mine' for artificial intelligence (AI) tools, thanks to

*Correspondence to: Dr Francesca Rita Ogliari, Department of Oncology, IRCCS Ospedale San Raffaele, Via Olgettina, 60, 20132 Milan, Italy. Tel: +39 02 2643 4138

E-mail: francescaritaogliari@gmail.com (M. Reni).

†Equally contributing last authors.

GRAPHICAL ABSTRACT



their ability to delve into large datasets and discover hidden patterns.[2] Due to the intrinsic 'imperfectness' of RWD,[3] such as missing values, intercenter heterogeneity, and the manual workload required for data extraction, multicenter real-world studies are often biased by a lack of reproducibility and external validation, which hinder their development outside of academic institutions.

In thoracic oncology, for example, immunotherapy-based treatments have revolutionized the survival of patients with metastatic non-small-cell lung cancer (NSCLC), but robust predictive and prognostic factors beyond well-known clinical features are still lacking, and programmed death-ligand 1 expression, despite being the most widely used biomarker in clinical practice, remains insufficient to accurately predict treatment response and long-term outcomes.[4] Therefore, in recent years, there has been growing interest within the oncologic community in leveraging RWD to develop tailored predictions for broader patient populations. In this scenario, there is a substantial need for projects that bridge the gap between research in data science and hospital-based software used for patient care, and metastatic NSCLC represents a clear 'proof-of-concept' setting, given its high incidence and the wide availability of AI-based tools for imaging analysis.[5]

In this article, we describe the design, data structure, and digital infrastructure of the AI-HOPE Lung Cancer study, a real-world platform integrating clinical and AI tools for patients with metastatic NSCLC.

MATERIALS AND METHODS

Study design

The AI-HOPE study is a multicenter retrospective and prospective study (ClinicalTrials.gov identifier NCT06788366). The study was approved by the ethical committee of the sponsor site on 24 January 2024 and then locally as per local legislation. Included are all patients who are diagnosed with metastatic NSCLC and treated with first-line immunotherapy-based regimens as part of standard of care and in accordance with international guidelines.[6] Availability of baseline computed tomography (CT) or [^{18}F]2-fluoro-2-deoxy-D-glucose—positron emission tomography ([^{18}F]FDG—PET) scans is not mandatory for eligibility. Outcomes and follow-up are recorded as per clinical practice, and at each center, the reuse of local datasets is encouraged, in line with the Findable, Accessible, Interoperable and Reusable (FAIR) principles,[7] to reduce the manual workload for study personnel.

All data, once pseudo-anonymized, are transferred and stored on a universal data platform available at Università Vita-Salute San Raffaele (San Raffaele Ai CEnter [S-RACE] platform).[8] The S-RACE platform directly interacts with the electronic case report form (eCRF) on REDCap (https://projectredcap.org/) and has access to in-hospital software for internal patients' variables (such as laboratory software, imaging repository, etc.), as summarized in Figure 1. It also supports a cloud environment for external data ingestion, both in tabular format (via REDCap) and as Digital Imaging

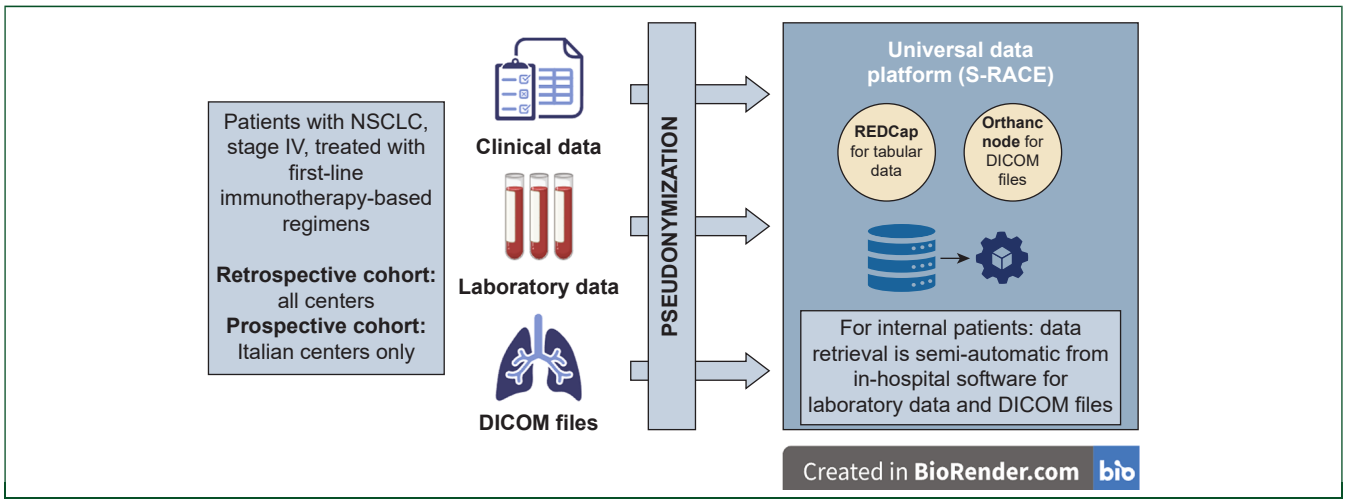


Figure 1. Data input in the universal data platform (in the S-RACE platform).
DICOM, Digital Imaging and Communications in Medicine; NSCLC, non-small-cell lung cancer.

and Communications in Medicine files, for CT or [^{18}F]FDG–PET scans, when available. The Orthanc node, which is the core of the imaging analysis, is stored within S-RACE to ensure privacy and security compliance and hosts both the imaging files and the AI tools for segmentation, analysis, feature extraction, and visualization. Moreover, as thoroughly described by Traverso et al.,[8] the S-RACE platform ensures the governance and data quality of the study, promoting quality checks, data validation, accessible audit trails, and full compliance with the General Data Protection Regulation.

Study objectives

The primary objective of the study is to build time-to-event models, which can achieve an individualized prediction of overall survival (OS) and progression-free survival (PFS) of this population, both overall and according to first-line treatment regimen (mono-immunotherapy compared with chemo-immunotherapy). Binary classification models for toxicity occurrence (with a special focus on moderate-severe toxicities according to Common Terminology Criteria for Adverse Events) and for survival risk stratification (early progressions and long-term survivors or responders) will also be built.

RESULTS

Registry overview and current dataset

The AI-HOPE Lung Cancer study is currently active in 21 European centers: 13 centers in Italy, 6 centers in the Netherlands, 1 center in Austria, and 1 center in Portugal. Each center has two active users, the local principal investigator and the local data manager personnel, who are allowed to log in to the eCRF and the Orthanc node. Access to the eCRF is by username and password only, whereas access to the Orthanc node is provided via Microsoft Azure secure log-in.

Before site activation, remote site initiation visits were conducted by the sponsor's staff to provide training on the study applications. In addition, each center is provided with a codebook containing the definitions of all variables collected in the eCRF, as well as access to remote support in case of issues related to data entry or other study procedures. All variables, including oncologic outcomes, are collected according to the investigator's judgment and routine clinical practice; no centralized monitoring or review procedures are planned. However, regular data cleaning and quality assessment cycles are carried out by specialized data analysts according to predefined rules (e.g. temporal consistency checks, outlier detection, etc.).

The retrospective phase collects data from 1 January 2017 to the activation of the center for the prospective data collection, and its target of accrual is 1000 patients. The prospective phase (conducted in Italian centers only due to approval limitations in non-Italian centers) collects data from the local approval of the protocol to 31 December 2027, with center-specific exemptions, and its target of accrual is 500 patients. External prospective validation sets are not covered by this study but may follow according to study results and overall feasibility considerations. The first planned analysis in 2026 will focus on data collected between 1 January 2017 and 31 December 2023, to ensure at least 2 years of follow-up.

At the data cut-off of 16 May 2026, 920 patients have already been included in the study in the retrospective cohort, and their data are fully available in the eCRF. Baseline CT scans have also been collected for 621 of 920 patients. For all the patients, variables derived from the clinical domain, pathology domain, laboratory values, and radiologic staging are collected, along with cancer-related outcomes, as summarized in Table 1.

Key methodological pillars: tailored modeling and integration of imaging-derived biomarkers

Time-to-event models are the best approach for oncology because they integrate both the occurrence and timing of clinically relevant events while appropriately handling

Table 1. List of domains and variables per domain, as collected in the eCRF of the AI-HOPE Lung Cancer study

Demographics
Age
Sex
Clinical variables
Smoking status
ECOG performance status
Body mass index
Symptoms at diagnosis
Comorbidities
Concomitant drugs
Pathology
Histology
Next-generation sequencing
PD-L1 percentage
Radiology
TNM[9]
Metastatic sites
Laboratory data
Blood cells count
White blood cells
Neutrophils
Lymphocytes
Platelets
Hemoglobin
Biochemistry
Aspartate aminotransferase
Alanine aminotransferase
Lactate dehydrogenase
Total bilirubin
Creatinine
Albumin
C-reactive protein
Electrolytes
Sodium
Potassium
Calcium
Therapy
Adjuvant therapy (if applicable)
First-line choice
Radiotherapy (if applicable)
Second-line (if applicable)
Toxicities
Outcomes
Toxicity outcome
Best RECIST1.1 response
Progression-free survival
Overall survival

eCRF, electronic case report form; PD-L1, programmed death-ligand 1; TNM, tumor-node-metastasis.

censoring, making them more informative for clinical decision making.[10] Binary classification models, on the contrary, are suboptimal for clinical practice because they ignore the continuous nature of survival outcomes,[11] but provide reliable information on secondary endpoints aimed at exploring the extremes of the distribution as well as on adverse event occurrence. Therefore, our methodological approach does not follow a 'one-size-fits-all' strategy; rather, it relies on a comprehensive framework that integrates multiple machine learning (ML) models, which can be selected according to specific clinical questions. Overall, we aim to combine powerful ML models, which may be limited by low interpretability, with explainable AI (XAI) tools to enhance user trust and facilitate their integration into clinical practice, as we thoroughly described in our first internal feasibility study.[12] In the models developed within the multicenter AI-HOPE Lung Cancer study, however, XAI approaches will be selected according to the type of model developed. SHapley Additive exPlanations-based methods will be applied to models based on tabular variables, including radiomic features, while Gradient-weighted Class Activation Mapping or attention maps will be used as an initial transparency assessment step for deep learning models developed on CT images.[13] In the multimodal models, dedicated XAI tools will summarize the contribution of each data modality, the specific approaches of which will be detailed in the corresponding original research papers.

Moreover, recent studies have demonstrated that multimodal models integrating [^{18}F]FDG—PET or CT imaging features with clinical and other data sources outperform single-modality approaches for key predictive tasks,[14] but reproducibility of results and manual work are unsolved issues when working with imaging. Radiomics, for instance, has been investigated for years in thoracic oncology, but it still lacks a standardized methodological framework, particularly with regard to automated segmentation and feature selection, as well as clear clinical implementation and impact.[15] Within the S-RACE platform, a partially automated imaging analysis workflow has been implemented, relying on open-source semiautomatic contouring tools and interchangeable coded pipelines for feature extraction, applicable across multiple oncologic use cases.

For this specific use case, chest CT scans (with contrast enhancement and a slice thickness of at least 3 mm) will be segmented directly within the Orthanc node using open-source algorithms. In each CT scan, the primary tumor and the lung parenchyma will be contoured by at least two automatic tools, with human review reserved for cases showing discrepancies. Differences between segmentations will be assessed using at least two metrics for each contour, such as the Dice coefficient for overlap and Volume Difference for volumetric assessment,[16] to avoid overlooking potential errors of the automated system before proceeding with the extraction of imaging features (radiomic or deep features). Standard preprocessing procedures and postsegmentation feature harmonization approaches will be implemented and described in detail in the original research articles, based on the specific requirements, which are beyond the scope of this manuscript.

DISCUSSION

The AI-HOPE Lung Cancer study is a multicenter initiative designed to integrate real-world clinical and imaging data to develop and validate ML models in metastatic NSCLC. By combining structured clinical variables with features derived from baseline imaging, the study aims to generate personalized predictions of survival and treatment-related toxicity. Once validated, this approach may support more refined risk stratification and treatment personalization in routine clinical practice and lay the foundation for future research in this field. The S-RACE platform and its

streamlined study design allow researchers to easily integrate additional data sources, such as extended genomic profiling, digitized pathology slides, and longitudinal clinical variables, as well as to develop 'mirror' case studies across different cancer types.

Therefore, the value of this registry lies both in its clinical setting (only real-world patients treated with first-line immunotherapy for metastatic NSCLC) and in the methodology supporting the study conduct, which is fully compliant with regulatory requirements while remaining flexible in terms of data sources and/or reproducible use cases that can be implemented using the same framework. The AI-HOPE Lung Cancer network will bridge the gap between purely academic and exclusively clinical institutions, contributing to the development of AI applications that genuinely aim to improve routine clinical care.

Visit ClinicalTrials.gov for more information (ClinicalTrials.gov ID NCT06788366).

FUNDING

This work was supported by the Italian Ministry of Research, Complementary Actions to the NRRP 'D34health - Digital Driven Diagnostics, prognostics and therapeutics for sustainable Health care' Grant [grant number PNC0000001]. This work was partially supported by the Forza Operativa Nazionale Interdisciplinare Contro il Cancro Polmonare [FONICAP] award (ed. 2023) for center activation. The funding source had no role in data collection, analysis, interpretation, writing, or the decision to submit the article for publication.

DISCLOSURE

FRO received honoraria for lectures, presentations, speakers' bureaus, manuscript writing or educational events from Roche Genentech, Amgen, support for attending meetings and/or travel from Takeda, Amgen, Johnson & Johnson, Daichii Sankyo, all outside the submitted manuscript. CS received consulting fees from BMS and Johnson & Johnson, honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational event from Medhealth, Treevents, Benecke, MSD, support for attending meetings and/or travels from Eli Lilly, AstraZeneca, BMS, MSD, and has a role in the Dutch Association of Pulmonologists, in the Dutch committee for the assessment of molecular oncological diagnostics, and is part of the program committee for Medhealth, Thoracic oncology, all outside the submitted manuscript. DL received consulting or advisory board fees from Amgen, AstraZeneca, BMS, BeOne, speakers' honoraria from Amgen, Roche, Daiichi Sankyo, Johnson & Johnson, AstraZeneca, BMS, PharmaMar, and travel/accommodation or education funding from AstraZeneca, Roche, MSD, PharmaMar, Daiichi Sankyo, Pfizer, all outside the submitted manuscript. DC received personal fees (as a consultant and/or speaker bureau) from Advisory Amgen, AstraZeneca, BMS, Boehringer Ingelheim, Catalym, Daiichi Sankyo, Lilly, Merck KGaA, MSD, Novartis, Pfizer, Roche, Jannsen, all outside the submitted manuscript. LB received speakers' fees from AstraZeneca, Pfizer, Takeda, Novartis, BMS, MSD and Roche; advisory board fees from AstraZeneca and Amgen; travel fees from Takeda, Servier, J&J, all outside the submitted manuscript. LH received research funding from Roche Genentech, AstraZeneca, Boehringer Ingelheim, Takeda, Merck, Pfizer, Novartis, Gilead, honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational event from AstraZeneca, Bayer, Eli Lilly, MSD, high5oncology, Takeda, Janssen, GSK, Sanofi, Pfizer, Medtalks, Benecke, VJOncology, Medimix (all payments were paid to the institution with the exception of Medtalks, Benecke, VJOncology, Medimix), advisory board fee from AbbVie, Amgen, Anhearth, AstraZeneca, Bayer, BMS, Boehringer Ingelheim, Daiichi, GSK, Janssen, Lilly, Merck, MSD, Novartis, Pfizer, Pierre Fabre, Roche, Sanofi, Summit Therapeutics, Takeda (all payments were paid to the institution), and is a member of the following committees: Dutch guidelines on NSCLC, brain metastases and leptomeningeal metastases (payment to self), European Society for Medical Oncology guidelines on metastatic NSCLC, non-metastatic NSCLC and small-cell lung cancer (non-financial), all outside the submitted manuscript. MR received personal fees (as a consultant and/or speaker bureau and/or travel support) from BMS, AliraHealth, Oncoscience, Incyte Biosciences, Oncoscience GmbH, LeoPharma, Viatris, SOTIO, Servier, MSD, AstraZeneca, Baxter, Daichii Sankyo, and Advanced Accelerator Applications, all outside the submitted manuscript. All remaining authors declared no conflicts of interest.

REFERENCES

1. Saesen R, Van Hemelrijck M, Bogaerts J, et al. Defining the role of real-world data in cancer clinical research: the position of the European organisation for research and treatment of cancer. *Eur J Cancer*. 2023;186:52-61.
2. Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol*. 2019;20(5):e262-e273.
3. Lever J, Cheng S, Quilodrán Casas C, et al. Facing & mitigating common challenges when working with real-world data: the Data Learning Paradigm. *J Comput Sci*. 2025;85:102523.
4. Memmott RM, Wolfe AR, Carbone DP, Williams TM. Predictors of response, progression-free survival, and overall survival in patients with lung cancer treated with immune checkpoint inhibitors. *J Thorac Oncol*. 2021;16(7):1086-1098.
5. Pellat A, Kather JN. Limited impact of real-world evidence in medical oncology: how can AI turn the tide. *Eur J Cancer*. 2025;228:115714.
6. Hendriks LE, Kerr KM, Menis J, et al. Non-oncogene-addicted metastatic non-small-cell lung cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol*. 2023;34(4):358-376.
7. Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016;3:160018.
8. Traverso A, Tiano D, Corvaglia A, et al. Powering responsible artificial intelligence with high-quality real-world data: the S-RACE platform for scalable, multi-specialty clinical research. *NPJ Digit Med*. 2026;9:6.
9. UICC. *TNM Classification of Malignant Tumours*. Hoboken, NJ: Wiley-Blackwell; 2016.
10. Le-Rademacher J, Wang X. Time-to-event data: an overview and analysis considerations. *J Thorac Oncol*. 2021;16(7):1067-1074.
11. Loh DR, Hill ED, Liu N, Dawson G, Engelhard MM. Limitations of binary classification for long-horizon diagnosis prediction and advantages of

a discrete-time time-to-event approach: empirical analysis. *JMIR AI*. 2025;4:e62985.
12. Ogliari FR, Traverso A, Barbieri S, et al. Exploring machine learning tools in a retrospective case-study of patients with metastatic non-small cell lung cancer treated with first-line immunotherapy: a feasibility single-centre experience. *Lung Cancer*. 2025;199:108075.
13. Mariappan R. Extensive review of literature on explainable AI (XAI) in healthcare applications. *Recent Adv Comput Sci Commun*. 2025;18(1): e200324228159.
14. Vanguri RS, Luo J, Aukerman AT, et al. Multimodal integration of radiology, pathology and genomics for prediction of response to PD-(L)1 blockade in patients with non-small cell lung cancer. *Nat Cancer*. 2022;3(10):1151-1164.
15. Limkin EJ, Sun R, Dercle L, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann Oncol*. 2017;28(6):1191-1206.
16. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging*. 2015;15: 29.