Aalborg Universitet

I-O Optimise

A novel multinational real-world research platform in thoracic malignancies

Ekman, Simon; Griesinger, Frank; Baas, Paul; Chao, David; Chouaid, Christos; O'donnell, John C.; Penrod, John R.; Daumont, Melinda; Lacoin, Laure; Mckenney, Alexia; Khovratovich, Masha; Munro, Robin E.J.; Durand-Zaleski, Isabelle; Johnsen, Søren Paaske

Published in:
Future Oncology

DOI (link to publication from Publisher):
10.2217/fon-2019-0025

Creative Commons License
CC BY-NC-ND 4.0

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Citation for published version (APA):
Ekman, S., Griesinger, F., Baas, P., Chao, D., Chouaid, C., O'donnell, J. C., Penrod, J. R., Daumont, M., Lacoin, L., Mckenney, A., Khovratovich, M., Munro, R. E. J., Durand-Zaleski, I., & Johnsen, S. P. (2019). I-O Optimise: A novel multinational real-world research platform in thoracic malignancies. *Future Oncology*, *15*(14), 1551-1563. https://doi.org/10.2217/fon-2019-0025

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

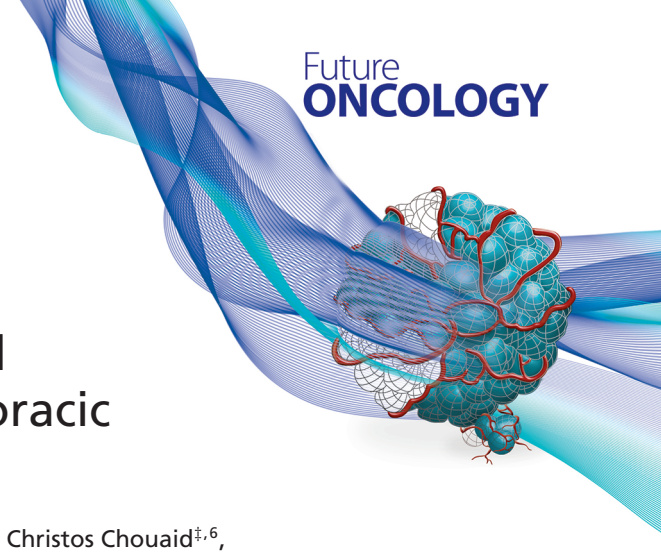
- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from vbn.aau.dk on: December 06, 2025

Methodology

For reprint orders, please contact: reprints@futuremedicine.com



I-O Optimise: a novel multinational real-world research platform in thoracic malignancies

Simon Ekman*[‡,1,2], Frank Griesinger[‡,3], Paul Baas[‡,4], David Chao[‡,5], Christos Chouaid[‡,6], John C O'Donnell[7], John R Penrod[7], Melinda Daumont[8], Laure Lacoin[8], Alexia McKenney[9], Masha Khovratovich[9], Robin EJ Munro[9], Isabelle Durand-Zaleski[‡,10] & Søren Paaske Johnsen[‡,11]

[1]Department of Oncology, Karolinska University Hospital, Stockholm, Sweden
[2]Department of Oncology-Pathology, Karolinska Institutet, Stockholm, Sweden
[3]Department of Haematology & Oncology, University Department Internal Medicine-Oncology, Pius-Hospital, Medical Campus University of Oldenburg, Oldenburg, Germany
[4]Department of Thoracic Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[5]Department of Oncology, Royal Free Hospital, London, UK
[6]Pneumology Unit, Centre Hospitalier Intercommunal de Créteil, Créteil, France
[7]Worldwide Health Economics & Outcomes Research, Bristol-Myers Squibb, Princeton, NJ, USA
[8]Worldwide Health Economics & Outcomes Research, Bristol-Myers Squibb, Braine-l'Alleud, Belgium
[9]Real-World & Analytics Solutions, IQVIA, London, UK
[10]URC Eco IdF, Unité de Recherche Clinique en Économie de la Santé d'Ile de France, AP-HP Paris, Paris, France
[11]Danish Center for Clinical Health Services Research, Aalborg University, Aalborg, Denmark
*Author for correspondence: Tel.: +46 725 721 111; simon.ekman@ki.se
‡Members of the I-O Optimise External Scientific Committee

Aim: To describe I-O Optimise, a multinational program providing real-world insights into lung cancer management. **Materials & methods:** Real-world data source selection for I-O Optimise followed a structured approach focused on population coverage, key variable capture, continuous/consistent data availability, record duration and data latency, and database expertise. **Results:** As of 31 October 2018, seven real-world data sources were included in I-O Optimise, providing data on characteristics, treatment patterns and clinical outcomes from more than 45,000 patients/year with non-small-cell lung cancer, small-cell lung cancer and mesothelioma across Denmark, Norway, Portugal, Spain, Sweden and the UK. **Conclusion:** The ongoing I-O Optimise initiative has the potential to provide a broad, robust and dynamic research platform to continually address numerous research objectives in the lung cancer arena.

Lay abstract: This article describes the design and creation of I-O Optimise, a project that will collect real-world evidence about patients with lung cancer across many different countries. A specific structured approach was followed to select real-world data sources for inclusion in I-O Optimise and, as of 31 October 2018, seven real-world data sources were included, providing information on more than 45,000 patients per year with lung cancer in Denmark, Norway, Portugal, Spain, Sweden and the UK. I-O Optimise is an ongoing project that has the potential to provide information that will be important in improving the care of patients with lung cancer.

First draft submitted: 15 January 2019; Accepted for publication: 19 February 2019; Published online: 11 March 2019

Keywords: clinical outcomes • database • mesothelioma • non-small-cell lung cancer • real-world evidence • small-cell lung cancer • thoracic malignancies • treatment patterns

Lung cancer is the leading cause of cancer mortality worldwide [1]. Up to the late 1990s, pharmacological treatments for advanced lung cancers were limited to platinum-based chemotherapy combinations, regardless of tumor histology and without available options for subsequent treatment [2]. Significant advances in the understanding of the complexity of lung tumor biology have resulted in the development of novel therapeutic agents for patients who

meet appropriate criteria (e.g., targeted tyrosine kinase inhibitors) [3,4]. Recently, the advent of immunotherapies has further transformed cancer management, including the treatment of lung cancer [5–7], with several immunotherapy agents targeting the PD-1/PD-L1 axis now approved for the treatment of non-small-cell lung cancer (NSCLC) and recommended in both European and US guidelines [8–10].

The increasing pace of change in the lung cancer treatment landscape and the range of therapeutic options available necessitates complex clinical decision-making both for those responsible for delivering patient care (e.g., oncologists and other healthcare professionals) and those responsible for deciding how limited resources should be invested (e.g., payers and regulators) [11], and highlights a need for ongoing rapid insights that can inform these decisions. These insights are generally provided via data from randomized controlled trials (RCTs) and/or through the generation of real-world evidence (RWE). RWE is often able to provide information that would not be readily attainable through RCTs, generating data reflective of routine clinical practice in broader, real-life treatment populations. Within the lung cancer drug development arena, there are many examples of the usefulness of RWE in providing complementary data to reinforce and support clinical trial data. For example, recent real-world studies have evaluated the safety and/or effectiveness of lung cancer treatments, often in patient cohorts ineligible for clinical trials [12–15]; investigated the real-world burden of lung cancers and related treatment patterns and survivorship [13,14,16–18]; and assessed treatment-related costs/healthcare resource utilization (HCRU) [13,14,16,19]. As such, data from both observational real-world studies and traditional RCTs are important in informing best clinical practice, and regulatory bodies are now recognizing that these two methodologies are complementary in both the pre- and postauthorization stages of drug development [20,21].

Despite providing useful information, many existing RWE initiatives are limited in scope. Although the specific reasons vary, this may be due to a relatively short length of follow-up, restricted geographical coverage (i.e., only covering a single region or country), niche study populations with limited generalizability or long lead times from inception to data generation. Furthermore, many current real-world data sources (RWDS) were designed for auditing as opposed to research purposes and, therefore, are often unable to address more than a relatively small set of research objectives. With a rapidly evolving treatment landscape, these limitations may represent barriers to gathering timely insights on the treatment of lung cancers. One approach to mitigate these limitations, which has been adopted in other therapy areas [22,23], is to leverage multiple independent data sources (e.g., both electronic case report forms/electronic medical records [eCRF/EMR] and national registries) from different countries and regions with ongoing follow-up in order to develop a broader and more robust research platform that can address a wider array of research objectives and track changes over time.

This approach forms the basis of I-O Optimise, an ongoing collaborative initiative aimed at developing a multinational research platform that will leverage existing RWDS to provide continuous insights into the evolving thoracic malignancies landscape. In this article, we provide an overview of the development of a collaborative real-world research platform by describing the conception and creation of I-O Optimise, with a focus on the associated methodology and research prospects.

Materials & methods

I-O Optimise organization

Plans for research outputs from I-O Optimise are designed and analyses performed in a collaboration between the respective RWDS owners, the initiative facilitator (IQVIA), and the initiative sponsor (Bristol-Myers Squibb). Each RWDS has a specific contractual agreement with the initiative facilitator and sponsor; however, the respective RWDS retain full ownership of their data, with access permitted to the initiative facilitator and sponsor only in alignment with I-O Optimise research objectives and the related analyses.

As part of I-O Optimise, an external scientific committee comprised of a multinational, multidisciplinary team of experts (clinicians/oncologists, epidemiologists, health economists and RWDS owners) was convened. The primary role of this committee is to provide continuous insights and guidance on the latest medical and epidemiologic knowledge, to ensure deployment of the most rigorous methodological approaches during data analyses and to ensure that research outputs are independently verified. In addition, under the auspices of the I-O Optimise initiative, the external scientific committee meets regularly with other RWDS owners, leading clinicians and patient advocates to discuss the role of RWE in addressing challenges in lung cancer care, and to gain input on the initiative research priorities and objectives.

Table 1. Scope of I-O Optimise.

Topics	Research objectives
Epidemiology and clinical outcomes	• Understand thoracic malignancy landscape by describing **patient demographics** and clinical characteristics • Understand **real-world effectiveness** of treatments • Investigate impact of alternative **treatment patterns** on clinical outcomes • Explore **patient subpopulations** and associated outcomes
Treatment patterns	• Document current and emerging lung cancer **clinical care pathways** and treatment patterns • Explore context for **usage and duration of therapy (immunotherapies and other SACT)** • Investigate the **biomarker testing flow** and role in treatment decisions
Safety	• Understand the frequency of **AEs** for different treatments and for different patient subgroups • Evaluate the impact of AEs on **treatment duration**
Healthcare resource utilization	• Understand **costs** of treatment, care and patient management • **Assess economic value** of treatments (immunotherapies and other SACT)
Patient-reported outcomes	• Explore health-related **quality of life** of patients with lung cancer

AE: Adverse event; SACT: Systemic anticancer therapy.

Scope of I-O Optimise

The key research objectives of I-O Optimise are shown in Table 1 and were defined, following input from a broad range of stakeholders, to ensure that I-O Optimise insights are relevant to those involved in patient care.

Data sourcing process

To identify the most suitable data sources to address the defined research objectives, a three-step approach was followed: data source identification and selection; initial assessment of data sources; and full assessment of data sources (Figure 1).

Data source identification & selection

Different types of observational data from various countries (with an initial focus on Europe) were evaluated for inclusion in I-O Optimise. Data sources were identified via desk research (including a targeted literature review and background research) and interviews with disease area specialists. Data sources were shortlisted based on the size of the population catchment area, expected disease coverage (NSCLC, small-cell lung cancer [SCLC] and malignant pleural mesothelioma), number of new patients with lung cancer per year and prior research experience. Data sources were further prioritized (high, medium, low) in each country to ensure outreach efforts were focused on the RWDS most likely to have the ability to address the key research objectives.

Initial assessment of data sources

Initial assessment provided a preliminary view on the ability of the data sources to address I-O Optimise research objectives. Data source owners completed a set of semistructured questionnaires, based on variables needed to address the research objectives and on the nature of the RWDS. The short questionnaire, completed via email, was used to obtain a high-level overview of the data source (e.g., population coverage, data availability and ethics approval requirement). A longer, in-depth questionnaire, completed via telephone, was used to evaluate the finer details of the database (e.g., patient numbers per malignancy, number of patients added annually, presence or absence of desired variables, frequency of data collection and recording, restrictions to data access, existing infrastructure to support research and previous publications). Responses to the questionnaires were assessed against the five core research areas: clinical outcomes, treatment patterns, frequency of adverse events (AEs), HCRU and patient-reported outcomes. Data were not analyzed at this stage of the process. The ability of the data source to address research objectives was based on the self-reported presence/absence of the mandatory variables for each research area (i.e., the data source owner indicated which variables were covered by their database as part of the questionnaire responses). The decision on whether a data source should proceed to full assessment was based on the ability of a data source to address the research questions, the population coverage and representativeness, the need and possibility of linkage to other data sources, the level of engagement with the data source owners and the ability to address research questions that other data sources are less able to address (thereby increasing the strength of the portfolio).

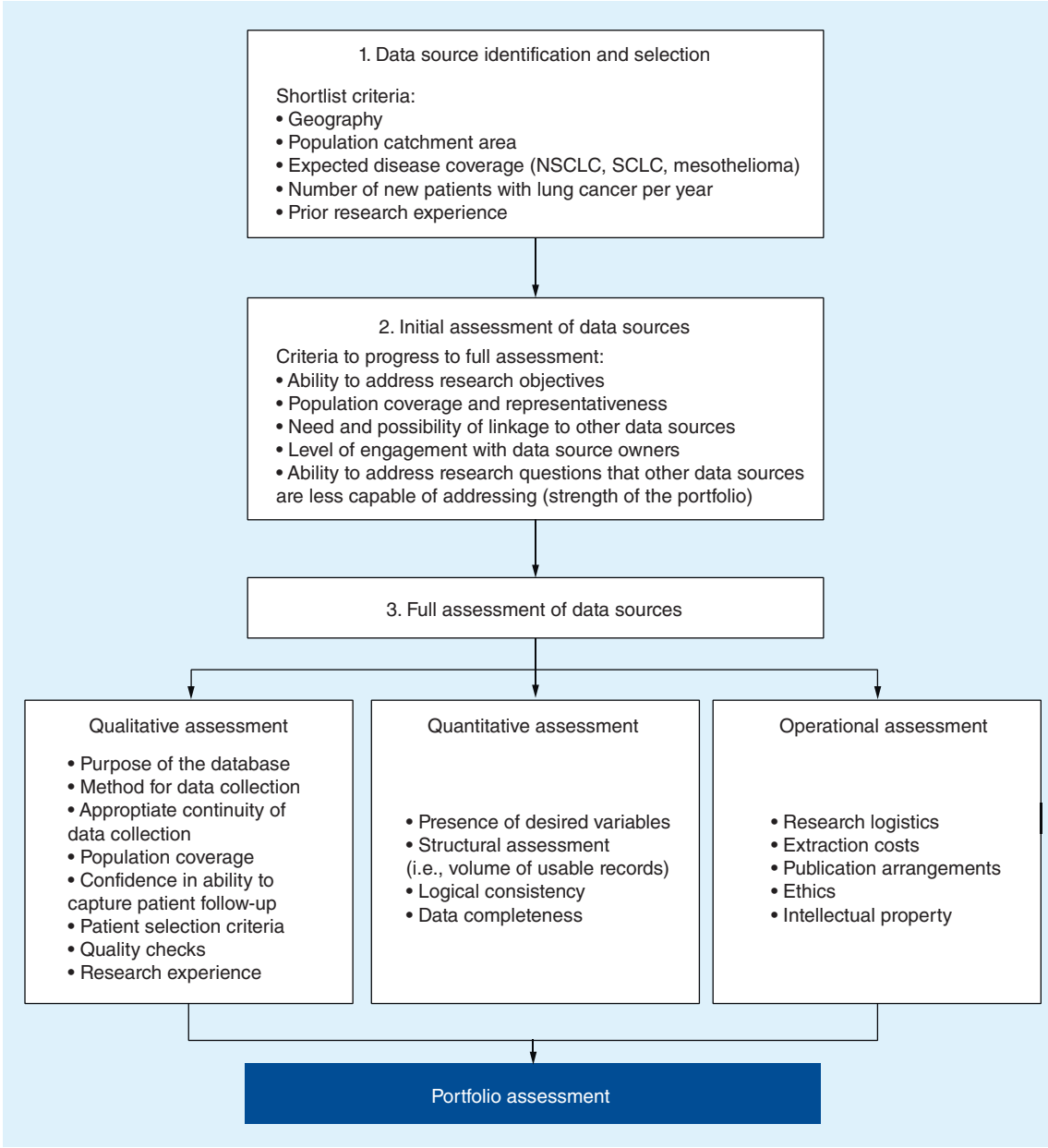


Figure 1. Data sourcing process.
NSCLC: Non-small-cell lung cancer; SCLC: Small-cell lung cancer.

Full assessment of data sources

The full assessment process involved qualitative, quantitative and operational assessments. Qualitative aspects assessed included the initial purpose of the database, methods for data collection, appropriate continuity of data collection (i.e., time between data collection and data availability for analysis), population coverage (representativeness), confidence in the ability of the database to appropriately capture patient follow-up, patient selection criteria, quality checks performed on the data source and the research experience of the individuals affiliated with the data source. The quantitative aspects assessed included: presence of desired variables, structural assessment, logical consistency and data completeness. It was conducted in four steps: assessment of variable coverage (120 variables related to the research objectives); preparation of a data set extract by the data source owners, covering all available variables; completion of the analyses; and review and discussion of analysis results with the data source owners. The following key analyses were conducted: linkage checks, to verify the total volume of usable records; logic checks,

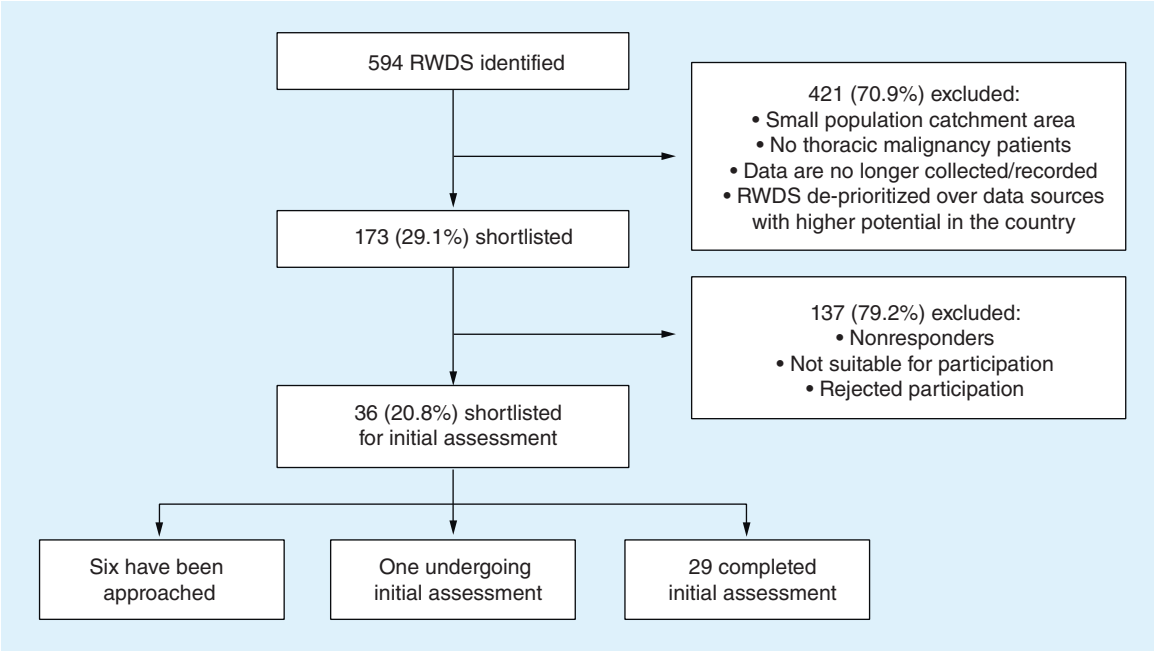


Figure 2. Data source identification and selection flow.
RWDS: Real-world data source.

to identify the number of illogical data entries; counts across years, to verify the number of new patients per year and the number of patients receiving active treatment per year; distribution checks, to provide insight into the frequency of distribution of key variables, allowing assessment of whether the content of these variables is logical and broadly as expected; data completeness checks; and patient follow-up checks, to assess the median length of follow-up for patients. The operational assessment evaluated various areas including research logistics, extraction costs, publication arrangements, ethics and intellectual property. The assessment was conducted via regular calls and face-to-face meetings with data source owners.

Portfolio assessment

Following completion of the full assessment, a group of experts from across the program evaluated (and will continue to evaluate) the results of the assessments, and decided upon full engagement with the respective data sources.

Data protection

All data analyzed under the auspices of I-O Optimise are aligned with specified study protocols that are approved by the relevant authorities in compliance with the country-specific regulatory processes for each RWDS. Data are analyzed at a population level and are, therefore, anonymized. Publication and presentation of data from I-O Optimise will adhere to the relevant ethics agreements of the respective RWDS and will be compliant with the European Union General Data Protection Regulation.

Results

Data source identification & selection

As of the cut-off date for this article (31 October 2018), a total of 594 RWDS had been considered for inclusion in I-O Optimise, with 173 shortlisted and approached for participation in the initial assessment process (Figure 2). From these 173 RWDS, 36 were selected to participate. Reasons for excluding RWDS were: nonresponses (n = 60); data sources not being suitable for the project, for example, no continuous data collection, thoracic malignancies not covered, overlap with other selected data sources (n = 58); or the data source refusing to participate (n = 19). In total, 29 RWDS had completed the initial assessment process by 31 October 2018, with 12 selected for full assessment; five awaiting a decision; and the remaining 12 excluded due to limited sample size (n = 5), unsuitable

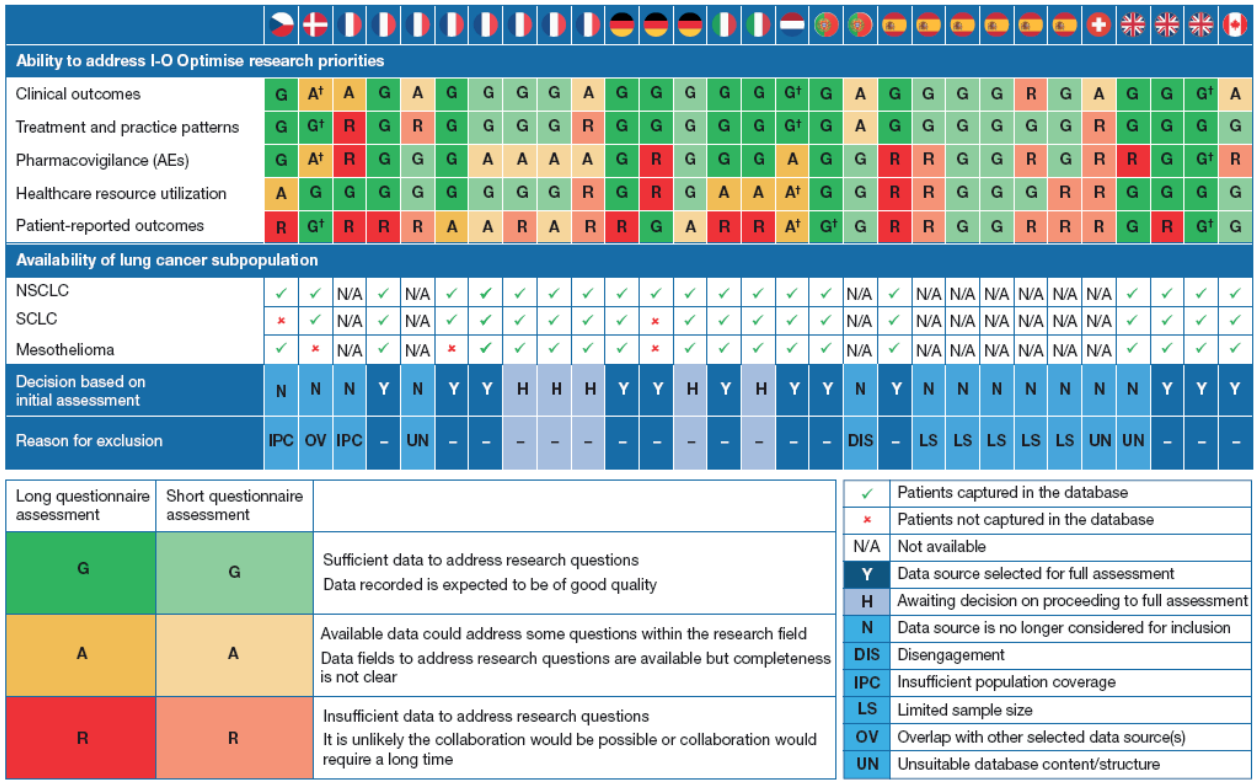


Figure 3. Initial assessment of results.
†View presented is based on enhanced data.
AE: Adverse event; NSCLC: Non-small-cell lung cancer; SCLC: Small-cell lung cancer.

database content/structure (n = 3), insufficient population coverage (n = 2), overlap with another selected data source (n = 1) or disengagement (n = 1) (Figure 3). In addition to the 12 RWDS selected via the initial assessment process, three linked data sources from Scandinavia were preidentified through an ongoing epidemiology project in NSCLC and selected for inclusion without initial assessment, giving a total of 15 data sources for full assessment. Following evaluation of the full assessment report for each data source on 31 October 2018, seven data sources were included in the initial build of I-O Optimise. The remaining eight data sources were still undergoing full assessment at this cut-off date.

Overview of RWDS selected for I-O Optimise

The seven data sources included in the initial build of I-O Optimise have a wide geographic coverage, including patients from Denmark, Norway, Portugal, Spain, Sweden and the UK (Table 2). The selected RWDS have a broad mix of data source type (registry data alone, linked EMR and registry data, hospital EMR data alone and data recorded in a hospital eCRF) and practice coverage (single hospital, all public hospitals in one region, multiple hospitals across a country and national registries). Thus, some sources will provide data on epidemiology and patterns of care for patients with lung cancer in the given country population, while others will provide a deep examination of patient and disease characteristics and treatments within a patient population referred to a cancer hospital or medical oncology practice. All data sources currently included have no formalized patient selection criteria, with patients being registered as soon as they have been seen at the sites covered by the data source. Where data are available, the average number of incident cases captured per year ranged from 435 to 30,900 for NSCLC, and from 49 to 4262 for SCLC. For mesothelioma, six of the seven data sources are suitable for inclusion in I-O Optimise based on sample size; the seventh has fewer than ten incident cases registered per year and is, therefore,

Table 2. I-O Optimise database overview: based on full assessments performed by cut-off date for this article (31 October 2018).

		Data sources						
		1	2	3	4	5	6	7
Type of data source		Linked hospital EMR and regional cancer registry	Hospital eCRF	Hospital EMR	National cancer registry	National registries (including cancer registry)	Linked hospital EMR and national registries (including cancer registry)	National registries (including cancer registry)
Practice coverage		Single hospital	Hospitals from multiple sites across the country	All hospitals in one region	National coverage	National coverage	Hospitals from multiple sites; national coverage	National coverage
Inclusion/exclusion criteria		No	No	No	No	No	No	No
Average number of incident patients per year	NSCLC	435	N/A	1119	30,900	1881	2943 (registry) 478 (EMR)	3126
	SCLC	49	N/A	159	4262	419†	56†	626†
	Mesothelioma	3	N/A	71	355	47†	93†	57†
Data capture	Data availability	2012–present	2016–present	2006–present	2012–present	2005–present	2005–present	2005–present
	Data lag time	6 months	Variable‡	24 hours	>9 months	12 months	7–13 months	7–12 months

†For the registries, these represent indicative numbers per malignancy.
‡Records are updated via eCRF and input frequency will vary by hospital.
eCRF: Electronic case report form; EMR: Electronic medical record; N/A: Not available; NSCLC: Non-small-cell lung cancer; SCLC: Small-cell lung cancer.

deemed unsuitable for mesothelioma. Among these six data sources, the average number of incident mesothelioma cases per year ranged from 47 to 355. The data sources provide far-reaching data capture with data availability from 2005 up to the present day. Lag times for the data sources vary, and where available range from 24 h to 13 months (Table 2). All data sources included in the initial build of I-O Optimise have affiliated researchers with a high level of experience; most (six) have owners who have published an extensive list of relevant peer-reviewed journal articles in the area of oncology. In addition, a wide variety of researchers are involved in the RWDS, including clinical researchers, qualified epidemiologists and statisticians with oncology oversight.

Figure 4 shows the key variables covered by the seven RWDS. In terms of clinical characteristics, tumor histology and staging at diagnosis are captured in all data sources in a structured format (i.e., data were prospectively collected). The availability of data on other clinical characteristics varies between the RWDS, although the majority of variables are available in either a structured or unstructured (e.g., in EMR free text) format. In terms of treatment data, most (four) RWDS capture structured information related to systemic anticancer therapy (SACT), including drug names and dates of infusion (or prescription for oral treatment). Surgery information is captured through procedure codes in all except one data source; due to the specificity of the coding system, in most cases this will allow for the identification of the date, localization and type of surgery. For one data source, surgery information is captured in a structured format in the eCRF. Like surgery, radiotherapy information is mainly captured (in all but one data source) through procedure codes. However, unlike surgery codes, radiotherapy codes are less specific and only very limited information will be available on the localization, type, dose and fractionation of radiotherapy. Again, in one data source, radiotherapy information will be detailed in the eCRF. In terms of outcomes data, all seven data sources include information on overall survival. For all the data sources except one, date-of-death data are cross-checked with the national death registry. Structured data for other efficacy outcomes are currently only captured in a small number of RWDS. Regarding safety outcomes, data availability will be variable; for example, in one data source only those AEs leading to treatment change will be captured. However, AE information may not always be captured in patient records, and there may be a lack of consistency in the reporting terminology. None of the data sources currently included in I-O Optimise were initially collecting data on health-related quality of life, but one RWDS started capturing EuroQoL-5D quality of life questionnaire outcomes among all patients from 2018 onward. Data on HCRU are available in five RWDS, being limited to the care delivered in the single hospital covered for each data source. Actual cost data are likely to be limited, although generalized costs may be calculated from the available HCRU data.

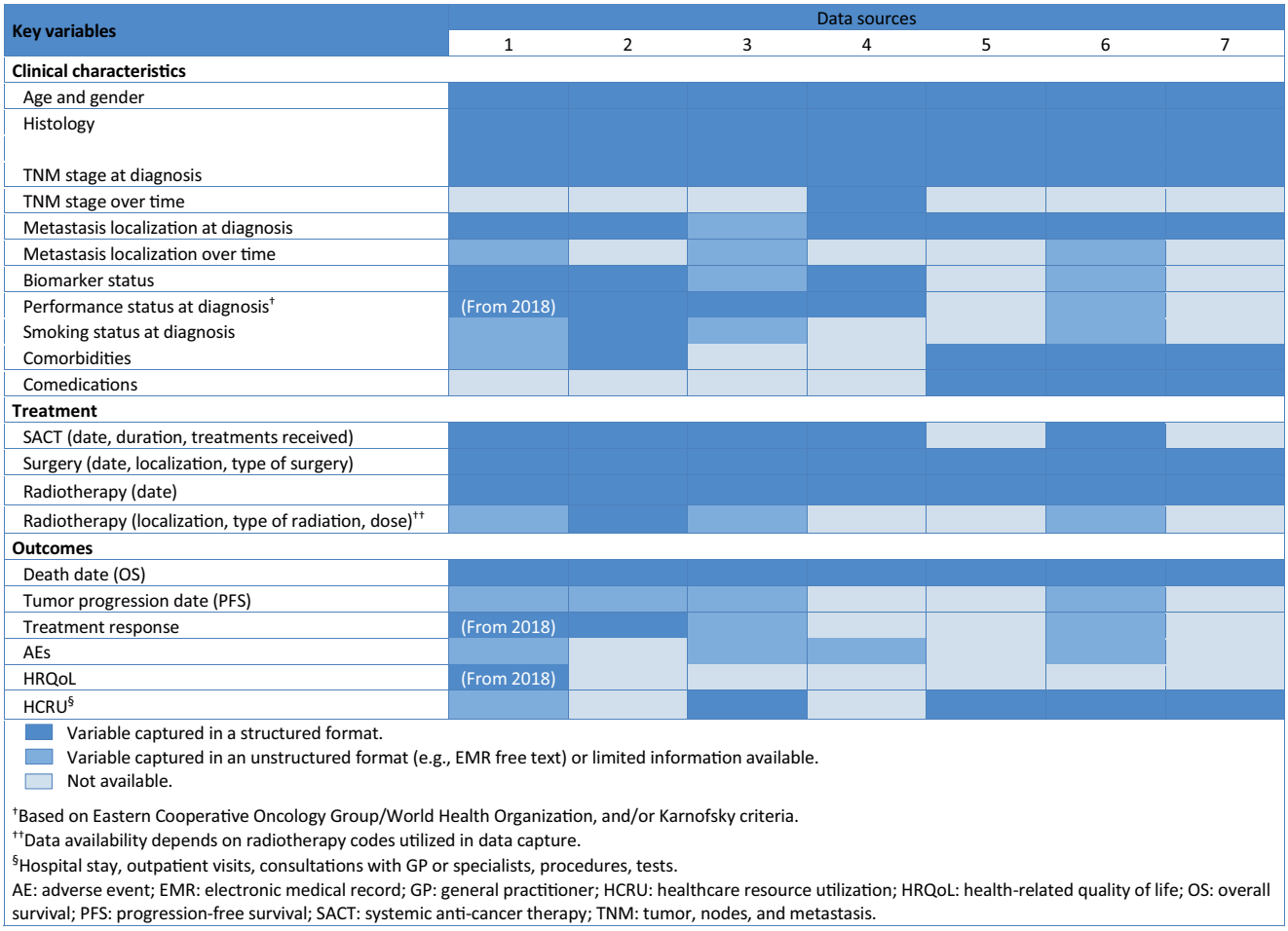


Figure 4. I-O Optimise key variables. Based on full assessments performed by cut-off date for this article (31 October 2018).

As shown in Figure 4, in many cases the data variables are captured prospectively in a structured format. However, when only unstructured data are available, additional efforts have been (and will continue to be) undertaken to successfully recover the information from EMR free text. Moreover, where other information is missing, relevant algorithms will be developed, validated and utilized for data capture. For example, treatments in oncology are generally described by line of therapy, driving treatment recommendations, but this parameter is often not captured in RWDS. In I-O Optimise, algorithms will be used to appropriately identify the different regimens received, and the sequencing of those different regimens, to define the line of therapy. A similar approach is being applied to describe the initial treatment given after diagnosis, taking into account the information on surgery, radiotherapy and SACT. As the initiative evolves, other algorithms may be developed.

Discussion

I-O Optimise, a novel, multinational research framework based on multiple RWDS, was established in response to a need for more robust, ongoing and timely RWE on the long-term outcomes with emerging treatments for thoracic malignancies. To maximize the utility of I-O Optimise, the RWDS selection process was rigorous. From almost 600 RWDS originally identified, only 15 were selected for full assessment. Moreover, the data source selection process followed for I-O Optimise adhered closely to the 'database selection' checklist proposed by Hall *et al.* [24], whereby a focus was placed on the population covered (i.e., unbiased patient selection), the capture of key variables of interest (i.e., good-quality data), the continuous and consistent capture of data (i.e., long-term data availability), the record duration and data latency (i.e., relatively consistent lag times) and the database expertise (i.e., evidence of peer-reviewed publication expertise).

As of the cut-off date for this article (31 October 2018), seven data sources were included in the initial build of I-O Optimise. The strengths of the current RWDS include a wide geographic reach, large thoracic malignancy populations, a variety of data source types and far-reaching data capture (from 2005 up to the present day). Among the seven data sources currently selected, all will provide a variety of data (see Figure 4) on clinical characteristics (including characteristics at baseline and, for some variables, how these characteristics evolve over time), treatment patterns and outcomes. Indeed, some of the data sources included in I-O Optimise have already produced interesting and insightful data [25–28]. An additional feature of I-O Optimise is that, through the additional processes of assessing eCRF/EMR free text for unstructured data and developing relevant algorithms for missing data, the quality of the research outputs can be further enhanced. Similar algorithms have been used successfully for the assessment of treatment patterns/sequences among patients with colon cancer, breast cancer and sarcoma [29–35]. Crucially, the algorithms developed as part of I-O Optimise will be used consistently across the different RWDS, further strengthening the generalizability of the data findings.

While the focus of this report is on the characteristics of the currently included RWDS, it is important to note that, although not all RWDS capture all variables of interest, the selection process was based on including a collection of different data sources to provide an overall balanced portfolio. The aim of this portfolio of data sources is to adequately cover the scope of the current research questions, with data from oncology-based referred populations addressing deep clinical questions and data from broader patient populations addressing questions related to epidemiology and public health. Furthermore, the I-O Optimise portfolio is also designed to be dynamic and constantly evolving. As such, the research objectives will be reviewed frequently and revised as needed to adapt to the changing therapeutic landscape for lung cancers. As a consequence, the RWDS included may change over time in order to constantly enhance the ability to address those research objectives. In addition, there is also the possibility of improving already enrolled data sources by working with the data source owners to continuously enhance and augment how the data are collected. Finally, although a comprehensive merger of data across the I-O Optimise RWDS will not be possible due to inherent methodological differences between the RWDS, there is the potential for conducting meta-analyses in specific patient populations across similarly structured data sources.

In establishing I-O Optimise, it was important to understand and consider the potential challenges related to the approach of bringing together disparate existing RWDS under a single initiative [36]. These challenges include those addressed by the aforementioned rigorous identification and selection process, such as the need to identify data sources with a sufficiently established infrastructure to provide data of required quality and completeness on a continuous basis and with appropriate information and operational governance (i.e., data privacy and site governance). Moreover, it was, and continues to be, important to acknowledge the challenges that could be faced when conducting analyses across these RWDS. Indeed, some of the main concerns related to the conduct of multi–data source initiatives appear to be related to incomplete/missing data and methodological differences [36,37]. For example, in the recently established Real world Outcomes across the Alzheimer's Disease spectrum for better care: Multi-modal data Access Platform (ROADMAP) initiative, designed to optimize real-world data generation in Alzheimer's disease, a noted challenge is the lack of standardized outcomes across the different data sources [22]. However, these challenges are not insurmountable and will be addressed in I-O Optimise via the development of appropriate algorithms, as discussed earlier, and through additional efforts to standardize data capture.

Conclusion

As the number of pharmacological treatments for patients with lung cancer increases, there is an increased need for RWE on long-term outcomes with the various therapeutic options available. Despite this need, many existing lung cancer RWDS are limited in scope and are unable to focus on more than a small number of research objectives. To address this challenge, the I-O Optimise initiative was established, enabling the use of multiple RWDS to provide a broad, robust and dynamic research platform that can continually address numerous research objectives in the lung cancer arena. Indeed, through ongoing engagement with RWDS owners, along with the ability to capture and analyze data continuously over a prolonged period, it is hoped that I-O Optimise will allow a thorough assessment of the evolution of the epidemiology of thoracic malignancies, practice patterns and clinical outcomes over time with the advance of new therapies, in particular with the increasing use of immunotherapies.

Supplementary data

An infographic accompanies this paper at the end of the references section. To download the infographic that accompanies this paper, please visit the journal website at: www.futuremedicine.com/doi/full/10.2217/fon-2019-0025

Acknowledgements

The authors are grateful to Hazel Jacobs and Grace Segall for their contributions to the development of this manuscript.

Author contributions

All the authors were involved in the conception and design of the I-O Optimise initiative described in this article. In addition, all the authors were involved in the development of the manuscript drafts, and approved the final manuscript for publication.

Financial & competing interests disclosure

S Ekman's Institution has received remuneration in relation to his role as a consultant on the I-O Optimise External Scientific Committee. F Griesinger, P Baas, D Chao, C Chouaid, I Durand-Zaleski and S Paaske Johnsen have received remuneration in relation to their role as consultants on the I-O Optimise External Scientific Committee. JC O'Donnell, JR Penrod, M Daumont and L Lacoin are employees of Bristol-Myers Squibb, the initiative sponsor. A McKenney, M Khovratovich and REJ Munro are employees of IQVIA, the initiative facilitator. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Professional writing and editorial assistance was provided by Richard Daniel, and Christina Campbell, of PAREXEL, and was funded by Bristol-Myers Squibb.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/

Summary points

- Although many existing real-world evidence initiatives can provide useful information on the management of lung cancers, most are limited in their ability to address the dynamic research needs related to a rapidly evolving treatment landscape.
- Leveraging multiple independent data sources from different countries and regions is one approach to mitigating these limitations, and can provide a broader and more robust research platform that can address a wider array of research objectives and track changes over time.
- This approach forms the basis of I-O Optimise, an ongoing collaborative initiative aimed at developing a multinational research platform that will leverage existing real-world data sources (RWDS) to provide continuous insights into the evolving thoracic malignancies landscape.
- As of the cut-off date for this article (31 October 2018), 36 data sources had been shortlisted for initial assessment, with 29 having completed that assessment. Of these 29, 15 were subsequently selected for full assessment.
- At the time of writing, seven RWDS had completed the full assessment and been included in I-O Optimise. These sources provide data from more than 45,000 patients annually with non-small-cell lung cancer, small-cell lung cancer and mesothelioma, from sites across Denmark, Norway, Portugal, Spain, Sweden and the UK.
- Additional strengths of the current RWDS include a variety of data source types (registry data alone, linked electronic medical records and registry data, hospital electronic medical records data alone and data recorded in an electronic case report form), variable practice coverage (single hospital, all public hospitals in one region, multiple hospitals across a country and national registries) and far-reaching data capture (from 2005 up to the present day).
- The I-O Optimise portfolio is designed to be dynamic and constantly evolving, and the research objectives will be reviewed frequently and revised as needed to adapt to the changing therapeutic landscape for lung cancers. As a consequence, the RWDS included may change over time in order to constantly enhance the ability to address those research objectives.
- The ongoing I-O Optimise initiative will provide a broad, robust and dynamic research platform that will continually address numerous research objectives in the lung cancer arena.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68(6), 394–424 (2018).

• **Latest statistics on global incidence and mortality of cancers, demonstrating the worldwide burden of lung cancers.**

2. Reck M, Heigener DF, Mok T, Soria JC, Rabe KF. Management of non-small-cell lung cancer: recent developments. *Lancet* 382(9893), 709–719 (2013).
3. Davidson MR, Gazdar AF, Clarke BE. The pivotal role of pathology in the management of lung cancer. *J. Thorac. Dis.* 5(Suppl. 5), S463–S478 (2013).
4. Boolell V, Alamgeer M, Watkins DN, Ganju V. The evolution of therapies in non-small cell lung cancer. *Cancers (Basel)* 7(3), 1815–1846 (2015).
5. Tsiatas M, Mountzios G, Curigliano G. Future perspectives in cancer immunotherapy. *Ann. Transl. Med.* 4(14), 273 (2016).
6. Blumenthal GM, Bunn PA Jr, Chaft JE *et al.* Current status and future perspectives on neoadjuvant therapy in lung cancer. *J. Thorac. Oncol.* 13(12), 1818–1831 (2018).
7. Passiglia F, Commendatore O, Vitali M, Conca R. Immunotherapy in non-small-cell lung cancer: a bridge between research and clinical practice. *Future Oncol.* 14(13 Suppl.), 41–60 (2018).
8. Planchard D, Popat S, Kerr K *et al.* Metastatic non-small cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* 29(Suppl. 4), iv192–iv237 (2018).
9. Ettinger DS, Wood DE, Aisner DL *et al.* Non-small cell lung cancer, version 5.2017, NCCN Clinical Practice Guidelines in Oncology. *J. Natl Compr. Canc. Netw.* 15(4), 504–535 (2017).
10. Ettinger DS, Aisner DL, Wood DE *et al.* NCCN guidelines insights: non-small cell lung cancer, version 5.2018. *J. Natl Compr. Canc. Netw.* 16(7), 807–821 (2018).
11. Berger ML, Sox H, Willke RJ *et al.* Good practices for real-world data studies of treatment and/or comparative effectiveness: recommendations from the Joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. *Value Health* 20(8), 1003–1008 (2017).

- **Guidance on good procedural practices for real-world data studies with a focus on methods to enhance decision makers' confidence in evidence derived from these studies.**

12. Kawachi H, Fujimoto D, Morimoto T *et al.* Clinical characteristics and prognosis of patients with advanced non-small-cell lung cancer who are ineligible for clinical trials. *Clin. Lung Cancer* 19(5), e721–e734 (2018).
13. Arunachalam A, Li H, Bittoni MA *et al.* Real-world treatment patterns, overall survival, and occurrence and costs of adverse events associated with second-line therapies for medicare patients with advanced non-small-cell lung cancer. *Clin. Lung Cancer* 19(5), e783–e799 (2018).
14. Bittoni MA, Arunachalam A, Li H *et al.* Real-world treatment patterns, overall survival, and occurrence and costs of adverse events associated with first-line therapies for medicare patients 65 years and older with advanced non-small-cell lung cancer: a retrospective study. *Clin. Lung Cancer* 19(5), e629–e645 (2018).

- **Recent US-based example of the breadth of real-world data that can be collected through appropriate databases, demonstrating the usefulness of the collected information.**

15. Fukui T, Okuma Y, Nakahara Y *et al.* Activity of nivolumab and utility of neutrophil-to-lymphocyte ratio as a predictive biomarker for advanced non-small-cell lung cancer: a prospective observational study. *Clin. Lung Cancer* doi:10.1016/j.cllc.2018.04.021 (2018) (Epub ahead of print).
16. Buck PO, Saverno KR, Miller PJ, Arondekar B, Walker MS. Treatment patterns and health resource utilization among patients diagnosed with early stage resected non-small cell lung cancer at US community oncology practices. *Clin. Lung Cancer* 16(6), 486–495 (2015).
17. Wallington M, Saxon EB, Bomb M *et al.* 30-day mortality after systemic anticancer treatment for breast and lung cancer in England: a population-based, observational study. *Lancet Oncol.* 17(9), 1203–1216 (2016).
18. Nadler E, Espirito JL, Pavilack M, Boyd M, Vergara-Silva A, Fernandes A. Treatment patterns and clinical outcomes among metastatic non-small-cell lung cancer patients treated in the community practice setting. *Clin. Lung Cancer* 19(4), 360–370 (2018).
19. Andreas S, Chouaid C, Danson S *et al.* Economic burden of resected (stage IB–IIIA) non-small cell lung cancer in France, Germany and the United Kingdom: a retrospective observational study (LuCaBIS). *Lung Cancer* 124, 298–309 (2018).
20. Cave A, Cerreta F. Use of real world data in development programmes. (2017). www.ema.europa.eu/docs/en_GB/document_library/Presentation/2017/05/WC500227703.pdf

- **European Medicines Agency presentation on opportunities for regulatory acceptability of real-world evidence in the drug development process.**

21. US FDA. Real world evidence. (2018). www.fda.gov/ScienceResearch/SpecialTopics/RealWorldEvidence/default.htm

- **US FDA opinion piece (with links to associated blogs, publications and guidance) on increasing role of real-world evidence in healthcare regulatory decision-making.**

22. Gallacher J, De Reydet De Vulpillieres F, Amzal B *et al.* Challenges for optimizing real-world evidence in Alzheimer's disease: the ROADMAP Project. *J. Alzheimers Dis.* 67(2), 495–501 (2019).
23. Vallejo-Vaz AJ, Akram A, Kondapally Seshasai SR *et al.* Pooling and expanding registries of familial hypercholesterolaemia to assess gaps in care and improve disease management and outcomes: rationale and design of the global EAS Familial Hypercholesterolaemia Studies Collaboration. *Atheroscler. Suppl.* 22, 1–32 (2016).

24. Hall GC, Sauer B, Bourke A, Brown JS, Reynolds MW, Locasale R. Guidelines for good database selection and use in pharmacoepidemiology research. *Pharmacoepidemiol. Drug Saf.* 21(1), 1–10 (2012).

●● **Guidance on the optimal process for selecting and using real-world databases in pharmacoepidemiological studies.**

25. Snee M, Cheeseman S, Thompson M *et al.* Clinical characteristics and overall survival of patients diagnosed with non-small cell lung cancer (NSCLC) between 2006 and 2017 in routine clinical practice: a retrospective analysis of a UK hospital database. *Lung Cancer* 115(Suppl. 1), S38 (2018).

26. Snee M, Cheeseman S, Thompson M *et al.* Treatment patterns in patients with stage IIIB–IV NSCLC in clinical practice: retrospective analysis of a UK Trust database. *J. Thorac. Oncol.* 13(10), S700 (2018).

27. Ekman S, Planck M, Sorensen JB *et al.* Treatment patterns in patients (pts) with stage IIIB–IV non-small cell lung cancer (NSCLC) in Sweden: the SCAN-LEAF study. *Ann. Oncol.* 29(Suppl. 8), viii493–viii537 (2018).

28. Sorensen JB, Ekman S, Horvat P *et al.* Initial treatment in patients (pts) diagnosed with non-small cell lung cancer (NSCLC) in Denmark from 2005–2015: the SCAN-LEAF study. *Ann. Oncol.* 29(Suppl. 8), viii536–viii537 (2018).

29. Reese ES, Onukwugha E, Hanna N, Seal BS, Mullins CD. Clinical and demographic characteristics associated with the receipt of chemotherapy treatment among 7951 elderly metastatic colon cancer patients. *Cancer Med.* 2(6), 907–915 (2013).

30. Zheng Z, Onukwugha E, Hanna N, Bikov K, Seal B, Mullins CD. Cost–effectiveness of second-line chemotherapy/biologics among elderly metastatic colon cancer patients. *Adv. Ther.* 31(7), 724–734 (2014).

31. Bikov KA, Mullins CD, Seal B, Onukwugha E, Hanna N. Algorithm for identifying chemotherapy/biological regimens for metastatic colon cancer in SEER-Medicare. *Med. Care* 53(8), e58–e64 (2015).

32. Liang C, Li L, Fraser CD *et al.* The treatment patterns, efficacy, and safety of nab®-paclitaxel for the treatment of metastatic breast cancer in the United States: results from health insurance claims analysis. *BMC Cancer* 15, 1019 (2015).

33. Bikov KA, Mullins CD, Hung A, Seal B, Onukwugha E, Hanna N. Patterns of biologics use across treatment lines in elderly (age >65) medicare patients with metastatic colon cancer. *Oncologist* 21(6), 676–683 (2016).

34. Li N, Hao Y, Kageleiry A *et al.* Time on treatment of everolimus and chemotherapy among postmenopausal women with hormone-receptor-positive/human-epidermal-growth-factor-receptor-2-negative metastatic breast cancer: a retrospective claims study in the US. *Curr. Med. Res. Opin.* 32(2), 385–394 (2016).

35. Villalobos VM, Byfield SD, Ghate SR, Adejoro O. A retrospective cohort study of treatment patterns among patients with metastatic soft tissue sarcoma in the US. *Clin. Sarcoma Res.* 7, 18 (2017).

36. Maissenhaelter BE, Woolmore AL, Schlag PM. Real-world evidence research based on big data: motivation-challenges-success factors. *Onkologe (Berl).* 24(Suppl. 2), 91–98 (2018).

37. Hampson G, Towse A, Dreitlein WB, Henshall C, Pearson SD. Real-world evidence for coverage decisions: opportunities and challenges. *J. Comp. Eff. Res.* 7(12), 1133–1143 (2018).

Article details

Title of article

I-O Optimise: A novel multinational real-world research platform in thoracic malignancies

Authors

Simon Ekman, Frank Griesinger, Paul Baas, David Chao, Christos Chouaid, John C O'Donnell, John R Penrod, Melinda Daumont, Laure Lacoin, Alexia McKenney, Masha Khovratovich, Robin EJ Munro, Isabelle Durand-Zaleski & Søren Paaske Johnsen

Article URL

www.futuremedicine.com/doi/10.2217/fon-2019-0025

Associate website

www.io-optimise.com

Rationale & objective

Rationale

The increasing pace of change in the lung cancer treatment landscape highlights the need for ongoing rapid insights from routine clinical practice that can inform clinical and reimbursement decisions

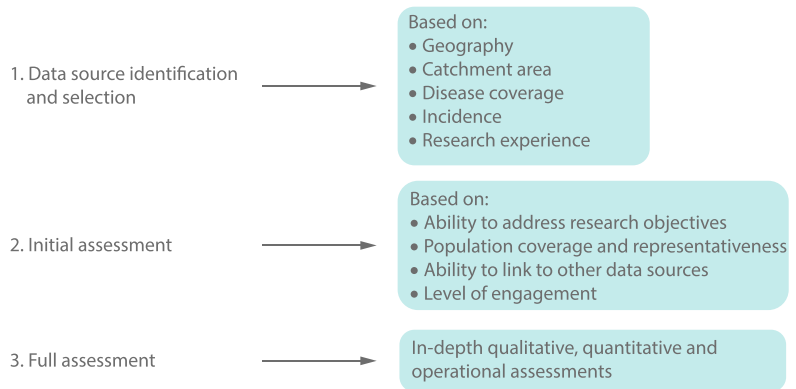
Objective

I-O Optimise is an ongoing collaborative initiative aimed at developing a multinational research platform that will leverage existing real-world data sources to provide continuous insights into the evolving lung cancer treatment landscape

Overarching research topics

- Epidemiology and clinical outcomes
- Treatment patterns
- Safety
- Healthcare resource utilization
- Patient-reported outcomes

Methodology flow



Current status (as of 31 October 2018)

- 7 Data sources completed full assessment and onboarded
- Data from >45,000 patients with non-small cell lung cancer, small cell lung cancer and mesothelioma per year
- Sites across Denmark, Norway, Portugal, Spain, Sweden and the UK
- Broad spread of variables captured (clinical characteristics at baseline and over time, patterns of treatment used, clinical outcomes)
- Variety of data source types (registry data alone, linked electronic medical record and registry data, hospital electronic medical record data alone, and data recorded in an electronic case report form) and practice coverage (single hospital, all public hospitals in one region, multiple hospitals across a country, and national registries)
- Far-reaching data capture (from 2005 up to the present day)

© 2019 Future Medicine Ltd