The South Region Cancer Registry: an evaluation of its exhaustiveness in a cohort of lung cancer patients

Jessica Verdasca¹, Filipa A. Costa¹, Catarina Ramos², Rodrigo Murteira² & Ana Miranda²

¹ Centro de Investigação Interdisciplinar Egas Moniz (CiiEM), Instituto Universitário Egas Moniz (IUEM), Caparica, Portugal
² Instituto Portugues de Oncologia de Lisboa Francisco Gentil EPE, National Oncology Registry, Lisbon, Portugal

Keywords
Exhaustiveness; lung neoplasms; Portugal; registries; treatment effectiveness.

Correspondence
Filipa A. Costa, Instituto Portugues de Oncologia de Lisboa Francisco Gentil EPE, National Oncology Registry, Rua Prof. Lima Basto Lisbon 1099-023, Portugal.
Tel: +351 217 229 852
Fax: +351 217 229 806
Email: alvesdacosta.f@gmail.com

Received: 11 September 2018; Accepted: 21 October 2018.
doi: 10.1111/1759-7714.12915

Abstract
The regional cancer registry for Southern Portugal (ROR-Sul) is a population-based registry set up in 1988 to observe and monitor disease incidence, prevalence and survival. Recently, the need to monitor real-life use of early approved and high-priced medicines led to therapeutic effectiveness becoming an emerging area of interest. We aimed to evaluate the exhaustiveness of the ROR-Sul database, covering around 4.8 million inhabitants. We have used a retrospective cohort study comprising 3457 lung cancer cases diagnosed during 2014 and 2015 and extracted from ROR-Sul database. Descriptive analysis of missing data was undertaken using IBM SPSS software, v.24. Exhaustiveness of data registry was classified into high (missing values <1%), medium (missing values 1–15%) or low (missing values >15%). High exhaustiveness was found for patients demographic information, date of diagnosis, date of first medical appointment, topography, morphology, cancer differentiation, and surgery procedure. Medium exhaustiveness was found for biomarkers (ALK, KRAS, and EGFR) results, and immunotherapy regimens. Low exhaustiveness was found for performance status, chemotherapy regimen, and chemotherapy treatment response. The findings highlight the need to transform treatment variables into compulsory, so that the cancer registry may be used to support effectiveness studies. Education, training and behaviour changes must also be considered to foster the process.

Introduction
In the European context, Portugal is seen as an exemplar cancer registry due to its coverage and detailed information.¹ Until December 2017, there were four regional cancer registries, all population-based, hence enabling population-based nationwide coverage.²,³ The regional cancer registry for Southern Portugal (ROR-Sul) covered 4 800 000 inhabitants (46% of the population) through 24 hospitals.⁴ Registries were initially set up to observe and monitor disease epidemiology, enabling incidence, prevalence and survival estimates. Population monitoring is ensured by connecting information from various sources, namely primary care centers, hospitals (including care provided and diagnostic tests made), and ultimately death certificates. Whenever a cancer diagnosis is made, regardless of the setting, the case enters the registry and is then prospectively followed throughout the years, contacts made in every point of the health care system are captured, until death occurs. This implies data linkage is a prerequisite to the good functioning of the registry, which can occur by automatic data integration or by manual data entry, in both cases validated by a certified registrar. Automatic data integration is achieved by creating a full list of pending information generated every day to the registrar. This information, to be validated, originates from other sources and may include biopsies, death certificates, radiotherapy treatments, surgical procedures, etc. The registrar then either creates the case or accesses an existing case and updates it with the new information available.

As new challenges emerge, the registry must adapt to meet the needs of health care systems, researchers and policy makers. Population based registries have long been
used in various disease areas and countries to evaluate therapeutic effectiveness. Anticipating this need, therapeutic effectiveness monitoring became an area of investment for the ROR-Sul. The database, always in constant evolution, ensures accuracy, exhaustiveness and relevance of the information retrieved to support health policy. The number of variables is dynamic and some, initially optional, have been made mandatory, namely on information regarding treatment. Innovative medicines are associated with greater hope but also with concerns about costs for the individual hospitals and for the health care system as a whole. Moreover, the safety concerns associated with frequent drug-induced toxicity suggest benefits from intensive monitoring. For all these reasons, it is particularly important for the Portuguese health service to timely monitor the effectiveness of newly approved medicines.

The National Oncology Registry was created by law in 2017, to be implemented through 2018. This new legislation clearly states that it is mandatory for all health institutions to register new cases of cancer, regardless of the type of institution where the case is diagnosed and treated, ie, public or private, and moreover that information on treatments made available are necessary to serve one of the new primary aims of this register, which now includes providing support and evidence to guide effectiveness decisions. To fully achieve this new mission of therapeutic effectiveness monitoring, we started to investigate the exhaustiveness of data in a cohort of patients, so that improvement measures could be timely developed.

**Methods**

A cohort of all lung cancer patients diagnosed during the years of 2014 and 2015 was extracted from ROR-Sul database, based on the topographical codes using the 10th International Classification of Disease for Oncology (C34.0-C34.3, C34.8-C34.9). The unit of analysis was the case or the treatment, depending on the type of exhaustiveness assessed.

**Statistical analysis**

Descriptive analysis of missing data was undertaken using IBM SPSS software, v.24. To assess the exhaustiveness of data entry, a 3-layer classification was developed: high (<1% missing values); medium (missing values [1–15%]); and low exhaustiveness (>15% missing values).

**Ethics**

Data was extracted anonymously by ROR-Sul. The study had auditing characteristics and did not interfere in service provision. Hence, no ethics submission was necessary.

**Results**

The selected cohort included 3457 cases. Patient identification variables (e.g. unique ID number, demographics, district of residence at diagnosis, date of last contact and vital state), as compulsory, yielded 100% exhaustiveness. Performance status was the sole variable where lower exhaustiveness was found and will therefore be made compulsory.

**Disease characterization**

Nine variables currently characterize disease and most showed high exhaustiveness. Biomarkers results, albeit mandatory since 2016, presented medium exhaustiveness in September 2017 (Table 1).

**Treatment received**

Treatment received is characterized by variables organized by treatment type. The initiation and discontinuation dates and Institution of administration are common to all types of treatment. Results show varying exhaustiveness depending on the type of treatment and on the characterizing variable. (Table 1).

Instituted treatment was stratified by disease staging to understand if missing data resulted from low exhaustiveness or from following of treatment guidelines (Table 2).

Table 2 shows that from the 1386 cases with local/locally advanced disease, 1010 (72.9%) had received at least one treatment option, with surgery and chemotherapy being the most frequent, with 555 and 546 cases, respectively. Among metastatic cases, 849 (49.9%) received chemotherapy, 459 (27.0%) radiotherapy and 38.2% (n = 649) had no treatment registered. Data show that 43 cases received treatment, even though the stage of disease was unknown. “Not applicable” staging was wrongly registered for 61 cases, as only 24 of the 85 represent non stageable morphologies, e.g. non-invasive neoplasies and sarcomas.

**Discussion**

As expected, variables never defined as optional show 100% exhaustiveness. The registry of biomarkers results which will enable the oncologist to anticipate the eligibility and expected efficacy of treatment have only been made compulsory by the end of 2016; this was achieved by programming the variables so that the case cannot be saved until this variable is filled in. This means that this variable, currently showing medium exhaustiveness, is expected to reach 100% exhaustiveness in cases diagnosed after this measure was implemented. At the time of the cohort, only molecular tests to identify EGFR mutations and ALK rearrangements were advised by the European Society for
Medical Oncology clinical guidelines for diagnosis and treatment, therefore the medium exhaustiness of KRAS is not worrisome.10

There were 2204 cases of unknown differentiation (63.8%), which may result from unavailability of that information at the time of registry or lack of exhaustiveness. We have detected 285 cases registered as "unknown stage," among which 44 received at least one type of treatment. Good clinical practice indicates that cancer should be staged before treatment institution, hence this information may have not been updated in the database. Cancer staging is recorded by the oncologist in the medical record and is then manually entered into ROR-Sul database with no predetermined frequency for case update. This means that for some cases "unknown" staging may result from no information being available on the medical record, whilst for others it may result from absent staging in practice. Moreover, among these with unknown stage, there were 72 cases dying within 60 days, suggesting these might not have been staged. Most of those (n = 68; 94%) did not receive treatment, reinforcing the quality of clinical practice and correct case registry.

Not all treatment related variables may be analysed similarly. The "end of treatment date" may be missing simply because the patient is still receiving treatment, and "treatment response" perhaps because not enough time has elapsed to enable evaluation. As an example, we have examined the cases where the date of ending chemotherapy was missing and found that treatment duration since the first dose to date of last contact was less than 30 days in 117 of them. Conversely, there were 362 cases (24.9%) referring to deceased patients (verified by death certificate). These cases correspond to an incomplete data update as the death information was integrated but the end of treatment was not registered. Currently the frequency of update varies widely. For example, recently diagnosed cases are

| Table 1 Distribution of missing values for disease and treatment description variables and their exhaustiveness classification |
|--------------------------------------------------|-----------------|-----------------|
| **Disease Variables†**                           | Missing values (%) | Exhaustiveness Level |
| Date of diagnosis                                 | 0.00             | High            |
| Date of first medical appointment                 | 0.00             | High            |
| Cancer Topography                                 | 0.00             | High            |
| Cancer Morphology                                 | 0.00             | High            |
| Cancer differentiation                            | 0.00             | High            |
| Stage of disease at diagnosis                     | 0.00             | High            |
| ALK (Anaplastic lymphoma kinase) rearrangement    | 12.80            | Medium          |
| KRAS mutation                                     | 7.50             | Medium          |
| EGFR (epidermal growth factor receptor) mutation  | 2.30             | Medium          |
| **Treatment Variables‡**                          | Missing Values (%) | Exhaustiveness level |
| Radiotherapy (n = 893)                             | Institution 0.13 | High            |
| Type§                                             | 0.00             | High            |
| Chemotherapy (n = 1891)                           | Institution 0.05 | High            |
| Type§                                             | 0.00             | High            |
| Regimen                                           | 33.20            | Low             |
| Treatment response                                | 41.80            | Low             |
| Surgery (n = 718)                                 | Institution 0.40 | High            |
| Procedure                                         | 0.73             | High            |
| Immunotherapy†† (n = 145)                         | Institution 0.00 | High            |
| Regimen                                           | 6.90             | Medium          |

†The unit of analysis for disease variables is the case. ‡The unit of analysis for treatment variables is the treatment. §Including radiosurgery; chemo-radiotherapy; prophylactic radiotherapy; adjuvant, palliative and neoadjuvant radiotherapy. ¶Including chemo-radiotherapy; adjuvant, palliative and neoadjuvant chemotherapy. ††Including targeted therapies and immune checkpoint inhibitors (e.g. gefitinib, crizotinib, nivolumab).

| Table 2 Comparison with Stage of disease at diagnosis and Treatment received |
|--------------------------------------------------|-----------------|-----------------|
| Stage of disease at diagnosis                    | Radiotherapy (n = 764) | Chemotherapy (n = 1452) | Immunotherapy (n = 128) | Surgery (n = 681) | No treatment (n = 1325) |
| Local/locally advanced disease (n = 1386)        | 288             | 546             | 21              | 555              | 376               |
| Metastatic disease (n = 1701)                    | 459             | 849             | 106             | 111              | 649               |
| Unknown (n = 285)                                | 13              | 37              | 1               | 3                | 242               |
| Not applicable (n = 85)                          | 4               | 21              | 0               | 12               | 58                |

The unit of analysis is the case. This implies that for the treatment variables it was considered equivalent to have one or more treatments (dichotomous variable).
more often updated and conversely remission cases may be updated seldomly. Deceased cases in general are not updated, unless an inconsistency is detected in the course of cross-validation. Nonetheless, one may say that on average, cases are updated annually. This will need to be improved in the future by ensuring active cases are updated at least four times a year. So far, whenever a specific effectiveness study is in progress, the frequency is much higher, reaching every two months. There were 1701 cases of metastatic disease and 1386 cases with local/locally advanced disease, and in both groups 37–38% no treatment was registered. It should be expected that nearly all local/locally advanced cases received treatment, which does not necessarily happen in metastatic cases. As such, we may assume that in the local/locally advanced cases, absent treatment corresponds to a delay in case update. In all situations, the benefit-risk ratio of initiating treatment is taken into account, considering median life-expectancy, performance status, patient’s willingness, among others. Often in metastatic cases, only symptomatic treatment is applied. Nonetheless, we acknowledge that the proportion of “no treatment” is quite high. Some of this missing information may result from the fact that the treatment had not been initiated at the time of registry, as suggested by the finding that in 404 cases (30.5%) the difference between the date of diagnosis and the date of last contact was 60 days or less. Additionally, some patients may have died before being treated and others may have received best supportive care. Cases that received best supportive care could in theory be quantified by looking at the variable “treatment without anti-tumour activity”. However, as this was being the scope of the project, this variable was not made available for analysis.

We acknowledge some limitations in our analysis, namely the restriction to lung cancer cases, implying our findings are not necessarily valid for all tumours not even to other lung cancer cohorts diagnosed in different years. Nonetheless, the intention was to provide clues to improve our system and we believe many of these have external validity. The classification for exhaustiveness used resulted from an internal consensus method. We have developed this classification due to absent literature and because we felt it was worth evaluating if the variables considered fundamental (e.g. date of diagnosis) had different exhaustiveness from those considered desirable (e.g. ALK rearrangements) or those not constantly available (e.g. date of treatment cessation or treatment response).

This study highlighted areas for improvement of the cancer registry to achieve full potential to support future effectiveness studies foreseen by law. Some variables were made mandatory, such as biomarkers results, fundamental for treatment forecast, information about regimen and treatment response. A significant proportion of cases needs more frequent updates, a possibility intrinsically linked with human resources capacity or with alternative ways of functioning. The current law foresees that case registry and update is part of contract programmes established with the National Health System, where institutions and/or personnel will be awarded incentives (positive or negative) according to their performance. This implies a change in mentality where institutions must ensure that registry becomes embedded in the regular duties of the oncologist, or that they have enough manpower to subsequently update information.

The low level of detail in treatment variables suggests that additional efforts are needed, by investing on education and training and by creating synergies within hospitals. Pharmacists and nurses could work closer with oncologists to improve the quality of information. Only then full potential may be reached in patient care and proactive effectiveness monitoring. Ideally an information system where various data entry points could be considered as compulsory and validated at the registry could be an option worth investing. Upscaling the regional registries into a National Oncology Registry per se is not enough to be able to evolve into an oncology database to be used to monitor therapeutic effectiveness. This is the beginning, but more frequent updates are needed, gradual awareness of the importance of evidence generation and additional automated data linkage systems will be detrimental to reach full potential. Furthermore, one key aspect in effectiveness studies will be the possibility to include the signalization of newly initiated treatments as one of the pending information to be validated and integrated. This will imply linkage of the pharmacy department and chemotherapy day care hospital units to the oncology registry. In some institutions, the latter is already implemented, albeit not nationwide; the linkage of the pharmacy department is to our knowledge yet to be accomplished. Our aim is to reach these ambitious goals within a timeline of three years and count on Government support to ensure legislation is respected.

Acknowledgment

The authors acknowledge Ms. Alexandra Meyer for data extraction and all individuals contributing to the oncology registry.

Disclosure

The authors declare that no funding was received for the present study. No authors report any conflict of interest.
References

1 Steliarova-Foucher E, O’Callaghan M, Ferlay J, Masuyer E, Forman D, Comber H, Bray F. European Cancer Observatory: Cancer Incidence, Mortality, Prevalence and Survival in Europe. Version 1.0 (September 2012) European Network of Cancer Registries, International Agency for Research on Cancer. [Cited 7 Jul 2017.] Available from URL: http://eco.iarc.fr/eureg.

2 Lunet N, Pimentel P. Registo Oncológico de base Populacional em Portugal: Reflexão sobre a Situação Atual e Perspetivas Futuras. [Population-based oncology registries in Portugal: Reflection on the current situation and future perspectives]. Arquivos de Medicina 2012; 26 (3): 124–8.

3 Portaria n°35/88 publicado em Diário da República I série, n° 13, 16 Janeiro 1988. Cria os Registos Oncológicos Regionais de Lisboa, do Porto e de Coimbra nos respectivos Centros Regionais do Instituto Português de Oncologia de Francisco Gentil [Decree 35/88 creating the regional Oncology registries of Lisboa, Porto and Coimbra in the respective Regional Centres of the Portuguese Oncology Institutes Francisco Gentil]. [Cited 12 Nov 2017.] Available from URL: https://dre.pt/web/guest/legislacao-consolidada/-/lc/107671385/201712171942/exportPdf/normal/1/cachleLevelPage?_LegislacaoConsolidada_WAR_drefrontofficeportlet_rp=indice.

4 Registo Oncológico Regional Sul. Incidência, Sobrevivência e Mortalidade por cancro na região sul de Portugal [Regional Oncology Register of the South. Incidence, Survival and Mortality attributable to cancer in the south region of Portugal].– ISM 2010|2011. Registo Oncológico Regional Sul, Lisboa 2017.

5 Eichler HG, Oyer K, Baird LG et al. Adaptive licensing: Taking the next step in the evolution of drug approval. Clin Pharmacol Therapeut 2012; 91 (3): 426–37.

6 EMA. Adaptive Pathways: Key Learning and Next Steps. Press Release EMA/484516/2016. [Cited 12 Nov 2017.] August 2016 Available from URL http://www.ema.europa.eu/docs/en_GB/document_library/Press_release/2016/08/WC500211522.pdf.

7 Deliberação 139/CD/2014 [Authorises the Use of Medicines under a Special Access Programme in Portugal, as a Supplement to the Decree 176/2006]. [Cited 12 Nov 2017.] INFARMED available from URL http://www.infarmed.pt/documents/15786/17838/Delib_139_2014.pdf/52417539-7520-496f-b81e-c93b95439309?version=1.0.

8 Bach P. Limits on Medicare’s ability to control rising spending on cancer drugs. N Engl J Med 2009; 360: 626–33.

9 Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin DM. International Classification of Diseases for Oncology, 3rd edn. World Health Organization, Geneva 2013.

10 Novello S, Barlesi F, Califano R et al. ESMO guidelines committee. Metastatic non-small-cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment, and follow-up. Ann Oncol 2016; 27 (Supplement 5): v1–v27.