A qualitative research evaluation of a Portuguese computerized cancer registry

Cátia Santos-Pereira¹²³, Ricardo Cruz-Correia¹⁴⁵, António Carvalho Brito²³, Alexandre B. Augusto¹, Manuel E. Correia¹², Maria José Bento⁶, Luís Antunes⁶
HealthySystems - HLTSYS, Lda¹
Instituto de Engenharia de Sistemas e Computadores, Tecnologia e Ciência - INESC TEC²
Faculdade de Engenharia da Universidade do Porto – FEUP³
Center for Research in Health Technologies and Information Systems – CINTESIS⁴
Faculdade de Medicina da Universidade do Porto - FMUP⁵
Instituto Português de Oncologia Francisco Gentil do Porto⁶
Porto, Portugal
catiapereira@hltsys.pt; rcorreia@med.up.pt; acbrito@fe.up.pt; aaugusto@hltsys.pt; mcc@hltsys.pt; mjbento@ipoporto.min-saude.pt; luisantunes@ipoporto.min-saude.pt

Abstract — A cancer registry is a standardized tool to produce population-based data on cancer incidence and survival. Cancer registries can retrieve and store information on all cancer cases occurring in a defined population. The main sources of data on cancer cases usually include: treatment and diagnostic facilities (oncology centres or hospital departments, pathology laboratories, or imaging facilities etc.) and the official territorial death registry. The aim of this paper is to evaluate the north regional cancer registry (ORENO) of Portugal using a qualitative research. We want to characterize: the main functionalities and core processes, team involved, different healthcare institutions in the regional network and an identification of issues and potential improvements. RORENO links data of thirteen-two healthcare institutions and is responsible for the production of cancer incidence and survival report for this region. In our semi-structure interviews and observation of RORENO we identified a serious problem due to a lack of an automatic integration of data from the different sources. Most of the data are inserted manually in the system and this implies an extra effort from the RORENO team. At this moment RORENO team are still collecting data from 2011. In a near future it is crucial to automatize the integration of data linking the different healthcare institutions in the region. However, it is important to think which functionalities this system should give to the institutions in the network to maximize the engagement with the project. More than a database this should be a source of knowledge available to all the collaborative oncologic network.

Keywords – cancer registries; information system; qualitative research; BPMN, qualitative research, semi-structure interview, observation.

I. INTRODUCTION

Medical registries are described as a systematic collection of a clearly defined set of health and demographic data for patients with specific health characteristics, held in a central database for a predefined purpose [1]. A cancer registry is a standardized tool to produce population-based data on cancer incidence and survival [2]. Cancer registries can retrieve and store information on all cancer cases occurring in a defined population. Its data can be used in a wide variety cancer control ranging from etiological research, through primary and secondary prevention to health-care planning and patient care, so benefitting both the individual and society. Cancer registries are evolving to provide a high level of clinical details, and to improve their capability to provide an evaluation of health interventions in oncology. Diagnosis, stage, and treatment information is registered with increasing frequency and higher level of clinical detail. Thus, the cancer registry is evolving as a tool to support planning and evaluation of cancer control strategies. However, the traditional cancer registry is retrospective or historical in its nature since it is presently limited/bound to investigate variables routinely determined in health archives [1].

The storage of large quantity of cancer registries is possible. However, the real problem in expanding the cancer registry scope is the difficulty to access an increasing number of variables from different sources.

The Portuguese health Ministry are working on an unique cancer registry for Portugal – RON - Registo Oncológico Nacional [3]. This new system will integrate the data from the actual three regional cancer registries (ORENO, ROR-Centro and ROR-Sul), following the recommendations of the WHO (World Health Organization) [4]. So, in this moment of change we believe it is important to evaluate the actual systems understanding its needs. The outcome of this work could be an important source of knowledge for the definition of the next systems with the same scope.

The aim of this paper is to evaluate the computerized cancer registry implemented in “Instituto Português de Oncologia do Porto” using a qualitative research. The main goal is to characterize the environment of the system, such as the main functionalities and core processes, team involved and different
healthcare institutions in the regional network. Then identify the main problems and difficulties that arose in the last year.

II. BACKGROUND

A. Cancer registry content

The main source of data on cancer cases usually include: treatment, diagnostic facilities (oncology centres or hospital departments, pathology laboratories, or imaging facilities etc.) and the official territorial death registry. Additional data sources like ambulatory, private clinics, elderly care homes and general practitioners’ networks increase the completeness of data, but also the logistics and expenses [5].

Hospital-based cancer registries are more numerous and widespread than population-based cancer registries. The primary purpose of these registries is to contribute to patient care by providing accessible information on the patients with cancer, the treatment they received and its results. The data may also be used for clinical research and, to a certain extent, for epidemiological purposes. One of the main advantages of hospital registries are the availability and the completeness of medical records. The data collected by a hospital registry tend to be more extensive than those collected by a population registry [5].

The top benefits of a computerized cancer registry include: 1) more complete treatment information, 2) less time for case finding and data entry, 3) more available time for data retrieval and analysis, 4) improved completeness, accuracy, and timeliness, 5) better patient tracking for follow-up, and 6) improved workflow efficiency.

The main challenges of the utilization of a computerized system within a cancer registry identified by respondents are 1) lack of adequate funding, 2) lack of medical staff to support the system, 3) changing data standards, 4) lack of full-time commitments, and 5) lack of a standardized data exchange [5].

The minimum set of data recommended to be collected by all Cancer registries was formulated by the International Agency for Research on Cancer (IARC) [6] and by the European Network of Cancer Registries (ENCR) and are: cancer patient’s personal identification; tumour site; tumour histology, classified by the ICD-O; tumour stage; tumour diagnosis, related to the ICD9; tumour therapy; further treatment; follow up; individual history; family history; and death, including autopsy results, if any. These items correspond to sensitive personal and medical information [7].

B. European initiatives

Cancer registries have been expanding in Europe since the early 1900’s. Over the last three decades cancer registration has become an important element of the EU’s strategy against cancer, promoted within the framework of the European Action against Cancer Programme (1985–2008), the European Partnership for Action Against Cancer (EPAAC) (2009–2014) [8]. The last data shows that nearly 200 population-based cancer registries (PCRs) are active in Europe and they are members of European Network of Cancer Registries (ENCR) [9].

The quality registration coverage of population by national cancer registries are available in 22 European countries. High quality registration of 10–50% of the population are available in France, Italy, Switzerland, Spain, Germany, and Serbia [10]. High quality registration of <10% of the population are available in Poland and Portugal [10]. EU member states Romania, Greece, and Hungary had as per 2012 only regional or partial data although legislation is in place to allow national cancer registration [10].

C. European Legislation.

Cancer case reporting is mandatory by law in most of the European countries with high quality registers that included in the cancer incidence in five continents series [11].

Personal data protection legislation has a major impact on electronic cancer registry. In a pan-European survey within the EUROCOUSE project, 20–35% of responding cancer registries reported legal-related barriers to cancer registration across most of Europe, while in the South-West region these barriers amounted to up to 60% [12]. Particularly concerning are the nationally variable barriers in the linkage of cancer registries with other health-related databases like mass screening programmes, biobanks, vital status, and causes of death databases [12][13].

In 2014 the European Commission proposed to replace the Directive 95/46/CE [14] by the General Data Protection Regulation [15]. The overall intention of this reform is to protect personal data and to facilitate a free flow of data within the European Union (EU). This initiative will also help to overcome problems alluded by the research community concerning about data sharing across borders for research purposes.

The outcome of the data protection reform is crucial to all epidemiological activities and clinical quality control in the EU. In contrast to a directive, a regulation is binding by itself and does not need implementing legislation by the Member States. It implies a harmonization of data protection measures across the EU, including use of data for public health purposes such as prevention and evaluation of screening programmes. However, on the one hand harmonisation, may facilitate valuable data sharing for research purposes, but on the other hand, excessive regulation can easily disable even simple monitoring of cancer, with disastrous consequences for public health information.

D. Portuguese computerized cancer registry in study

The RORENO [16] workgroup begins his activity in 1988 answering the governmental rule 35/88 [17]. This group is held in “Instituto Português de Oncologia do Porto Francisco Gentil, EPE – IPO-Porto”, a public reference oncolgic Hospital [18]. Since 1988 has been reporting cancer cases sent by different healthcare institutions of the north of Portugal. The geographical area comprehends the following districts: Porto, Braga, Viana do Castelo, Vila Real, Bragança, some cities in Aveiro district (Albergaria-a-Velha, Arouca, Castelo de Paiva, Espinho, Estarreja, Murtesa, Oliveira de Azeméis, Ovar, S. João da Madeira, Santa Maria da Feira and Vale de Cambra), in Viseu district (Cinfães, S. João da Pesqueira...
In 2005, this workgroup acquires a software named RORENO that facilitates the communication between this network. This system allows to accomplish the core competencies of the group:

- Collection of the most complete and up-to-date data on all new cases of tumours among residents of the Northern Region of Portugal.
- Production, analysis and interpretation of impact indicators of oncological disease and respective publication (national and international level).
- Provide information to all health professionals, researchers, policy makers, health care organizations in the effort to contribute to the prevention and control of cancer diseases [16].

The RORENO workgroup is expert in epidemiology and is constituted by a multi-disciplinary team of medical doctors, pathological anatomy technicians, statistical technicians and informatics professionals. The RORENO team is composed by thirteen professionals working in cancer registry of the North of Portugal.

III. METHODS

In this research, it was applied two kinds of qualitative techniques to acquire information about the health information system in study: semi-structured interview and observation. In the following sub-section, will be explained how these techniques were applied.

A. Semi-structured Interview

Semi-structured interviews are in-depth interviews where the respondents have to answer open-ended questions and thus are widely employed by different healthcare professionals in their research [19].

In this study, we collect data among RORENO team through semi-structured interviews. The main topics explored in the interviews were: roles and permissions characterization; main functionalities of the health information system; explanation of the core processes including how the data is collected and the which are the main sources (both institutions and information systems applications). It was realized five interviews with the duration of one hour each, between April and May of 2016.

B. Observation

Observation is a type of qualitative research method which not only included participant’s observation, but also covered ethnography and research work in the field. Observational data can be integrated as auxiliary or confirmatory research [19].

In this study, we decided to use observation to validate the information collected in the interviews, mainly the core processes workflow. For this together, with the end-user (mostly pathological anatomy technicians and statistical technicians), we observed how the system works and the different steps to perform a task. After this we designed core processes using BPMN in an iterative way until we had an agreement of the correct flow of steps.

IV. RESULTS

In this section, we presented the results of this research. This section is divided into the following sub-sections: description of the main architecture; main functionalities, data integration and roles characterization, data quality and problems identification.

A. RORENO main architecture

Figure 1 presents a general overview of the RORENO architecture. This system is stored in a reference oncologic hospital – IPO-Porto and managed by RORENO team. RORENO can be accessed by different healthcare institutions (e.g. hospitals, health centres…) in the North region and are within the national private healthcare network (RIS – Rede...
Informação de Saúde). Only the public institutions are within this private network. IPO-Porto is an important source of information to collect for RORENO since is the reference oncologic hospital in the North region and its also responsible to collect the cancer data between the north heath institutions to feed the RORENO system.

As we can see in Figure 1 RORENO is only integrated with IPO-Porto administrative database that includes administrative information about a patient such as demographic information and schedule (appointments, surgeries, vital state…). The others important sources of information such as departmental information systems, laboratory information system (LIS), radiologic information system (RIS) and drug information system doesn’t have an automatic integration with RORENO. So, in this case, RORENO team needs to access these systems one by one, patient by patient and collect manually the information needed.

In the case of other national relevant healthcare databases such as Plataforma de Dados de Saúde (PDS - Health Data Platform) and Registo Nacional de Utentes (RNU - national patients’ registration database) that have crucial information to complete the registries, RORENO also doesn’t have an automatic integration. The team needs to access these platforms and collect the needed data manually. After collected the data, RORENO team fulfil a paper form (example available in [20]) and then transcribe for RORENO electronic form.

The external institutions collection of data is explained better in sub-section data integration.

B. RORENO main functionalities

The main functionalities identified in the study are:

- Forms to insert information of an individual case;
- Suggestion for data aggregation (information from several institutions from the same patient and same diagnosis);
- Excel export option (important to perform data analysis and produce incidence and survival reports);
- Excel import option (allows to insert excel data files);
- Follow-up forms (possible to receive and send automatically this information from/to the hospital administrative system).
- Worklist of new patients accepted in the hospital (information sent by the administrative hospital system) that is needed to complete medical information.

Figure 2 describes through a process map the main functionalities important to complete activities and what is the trigger, what are the relations between them and the output. The four boxes in the left column represent four different triggers, and the four boxes in the right represent four outputs.

By analysing the Figure 2, the three first core processes could be improved with a full automatic integration with main databases in IPO-Porto, integration with PDS and RNU database and an integration with external institutions. The performance of these three core processes in terms of time is very time and resource consuming and incurs a high risk of errors (due to manually interaction).

C. RORENO data integration

The main goal of this system is to aggregate cancer registries from different healthcare institutions in the region. However, the system is almost a stand-alone system and professionals needs to insert manually the data in the system. Although the institutions had the possibility to use the forms in the systems to insert their own data, for some institutions is more convenient to send the information by email (through an excel file) or by post office (paper-based records) and ask some assistance to the professionals that work with the system to insert this data. Figure 3 shows a scheme that summarizes these three kind of processes selected by external Institutions. In a total of thirteen-three institutions that belongs to this region, six are a healthcare institution type 1, twenty-one are institutions type 2 and six are institutions type 3. These institutions include public hospitals and health centres and...
private pathological anatomy laboratories. Since private pathological anatomy laboratories cannot access the private network of the Portuguese Ministry of Health they cannot share their data through a direct integration. The only options they have is to send the information (through excel file or paper).

In this sub-section, we identified the main profiles of users and its permissions/tasks. These users belong to RORENO team. In terms of front-end management, we identified: administrator, responsible for entering data and external institution profile.

- **Administrator**: This role is responsible for the management of users, integration of files (Excel) provided by external institutions, data validation (manually and using IARC Check tool) and have the permission to create new variables in the system (allowing the adaptation of the system).

- **Responsible for entering data**: This role is performed by several users and has in charge the responsibility to insert data in the system. Due to interoperability problems, sometimes they need to insert variable by variable in the system manually. RORENO team need to insert this information from different institutions. They also have to search follow-up information about the patients (e.g. vital state). RORENO team search (manually) in national patient databases such as RNU, contact by phone others institutions, city responsible, families etc. They also use the IARC check tool [21] to validate the data quality. And have permissions to export and explore data to produce the incidence and survival reports.

- **External Institution**: Each external institution in the region has credentials to access this system and fulfil the cancer registry forms (manually or automatically). They can also insert information about the patients’ follow-up and can export and explore the data. The main difference between this role and the role before is that each external institution can only view information about their own patients.

In terms of back-end, the informatics department of the healthcare institution that holds the cancer registry ensure the availability, security of the system and the backup copies to use in case of failure.

### E. Data quality

Since they receive a significant amount of data by excel files to integrate in the system, they use a tool (IARC check [21]) to scan the file and verify some inconsistencies between variables, for example [16]:

- Age, Incidence Date, Date of Birth
- Age, Site, Histology (ICD-O-3 classification)
- Site, Histology (ICD-O-3 classification)
- Sex, Site (ICD-O-3 classification)
- Sex, Histology (ICD-O-3 classification)
- Behaviour, Site (ICD-O-3 classification)
- Behaviour, Histology (ICD-O-3 classification)
- Grade, Histology (ICD-O-3 classification)
- Basis of diagnosis, Histology (ICD-O-3 classification)

### F. Main problems identified in RORENO

From the interviews with the end-users of the system, the IARC check [21] isn’t enough to verify the data quality, they need to make an extra effort to manually check if there are: duplicate data, incomplete data, and other types of errors. This data validation plus the insertion of almost all data manually is a process very slow that requires time from various professionals. One of the reasons that RORENO team think other institutions don’t use RORENO is because a lack of budget to have professionals dedicated to insert the data in the system since they do not have an automatically integration implemented.

By now RORENO team published the reports of regional cancer incidence and cancer survival of 2010 year [20] and are working on 2011 data. Since we are in 2017 this delay is a problem. However, this seems to be a world-wide problem. In this research it was searched for the newest cancer incidence report and survival reports in Europe and the most recent numbers are from 2013-2014 [22][23][10]. We don’t have clues about the problems in other countries but in our case-study an investment improving the actual RORENO version could be a path to attenuate the problem.

### V. DISCUSSION

In Europe, the progress that has been made deserves to continue, and the current momentum of attention and support for quality cancer registration should not be lost. The reduction of disparities in the quality and function of cancer registries is a cornerstone, as are the efforts to harmonise, standardise, and bring together in a comparable and
understandable way the wealth of cancer data across the continent [9].

In this study, we made a maturity diagnosis of a computerized cancer registry implemented in 2005. Since 2005 the informatics in the healthcare institutions has been evolving and the sources of data that were paper-based or even now are robust electronic systems that already supports standardized protocols for integration like HL7 (Health Level 7) and DICOM (Digital Imaging and Communications in Medicine). This scenario reflected not only the healthcare institutions but also the national projects promoted by the Portuguese Health Ministry like national patients’ database. Epidemiological research depends on the balance between preserving patients’ integrity and anonymity while also enabling important research to improve people’s health and the quality of care [13], until 2018 the healthcare institutions should be prepared for comply the changes in the data protection (GDPR), so this system should comply with these new privacy and security requirements.

VI. FUTURE WORK

After the diagnosis of RORENO as is model we pretend to study which improvements this system should implement. It is crucial to automatize the integration of data linking the different healthcare institutions in the region. However, it is important to think which functionalities this system should give to institutions in the network in order to maximize the engagement with the project. More than a database this should be a source of knowledge available to all the collaborative oncologic network.

Support tools like business intelligence (BI) and an I&D platform could be implemented within this new project. Some key performance indicators like effectiveness within screening programs, impact of new drugs, adherence of drugs, effectiveness of different approaches (surgical/ radiotherapy/ chemotherapy) for each type of neoplasia could be tracked almost in real-time. This monitoring could have a great impact in daily decision making.

ACKNOWLEDGEMENTS

This work was supported by project ODISSEIA - Oncology Disease Information System [POCI-02-0550-FEDER-012618] and project NanoSTIMA - Macro-to-Nano Human Sensing: Towards Integrated Multimodal Health Monitoring and Analytics [NORTE-01-0145-FEDER-000016]. It was also supported by FCT (Fundação de Ciência e Tecnologia) through the doctoral fellowship [SFRH/BDE/105533/2014].

REFERENCES

11. Bray F, Brewster DH, Malaba CA, Kohler B, Swaminathan R,
20. RORENO. Registo Oncológico Regional do Norte 2010. IPO-Porto. 2015;