



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Sinnott, RO;Beuschlein, F;Effendy, J;Eisenhofer, G;Gloeckner, S;Stell, A

Title:

Beyond a Disease Registry: An Integrated Virtual Environment for Adrenal Cancer Research

Date:

2016-12-01

Citation:

Sinnott, R. O., Beuschlein, F., Effendy, J., Eisenhofer, G., Gloeckner, S. & Stell, A. (2016). Beyond a Disease Registry: An Integrated Virtual Environment for Adrenal Cancer Research. *Journal of Grid Computing*, 14 (4), pp.515-532. <https://doi.org/10.1007/s10723-016-9375-x>.

Persistent Link:

<https://hdl.handle.net/11343/283240>

Beyond a Disease Registry: An Integrated Virtual Environment for Adrenal Cancer Research

¹Richard O. Sinnott, ²Felix Beuschlein, ¹Jemie Effendy, ³Graeme Eisenhofer,
¹Stephan Gloeckner, ¹Anthony Stell

¹*Department of Computing and Information Systems, University of Melbourne, Australia*
²*Ludwig-Maximilians-University Munich, Munich, Germany*

³*Department of Medicine and Institute of Clinical Chemistry and Laboratory Medicine,
Technische Universität Dresden, Dresden, Germany*

Contact Author: rsinnott@unimelb.edu.au

SUMMARY

Many biomedical research collaborations are focused on the establishment of web-based databases that capture phenotypic and in some cases genotypic information, targeted to specific diseases – so called *disease registries*. Such resources are often used for clinical *matchmaking* and allow information on patients and patient disorders to be shared by clinicians with wider biomedical research communities outside of a given hospital setting, and potentially with patients and/or patient advisory groups. Whilst addressing aspects of clinical collaborations through making targeted biomedical data accessible - such registries are really only a starting point for what can be achieved to support research collaborations. In particular, registries should ideally be augmented with a portfolio of additional service offerings that facilitate secure research collaborations: bio-banking and bio-sample data tracking capabilities; support for feasibility analysis on clinical trials and studies; offer seamless data transfer to/from clinical trials; provide search and analytical capabilities in a user-driven research environment. Such a feature rich, Internet-based virtual research environment (VRE) has been established as part of the European Union funded ENS@T-CANCER¹ project that has a particular focus on supporting research into four primary types of adrenal tumors. This paper provides an overview of the ENS@T-CANCER VRE, outlining its core capabilities and how it has galvanized previously and largely fragmented, country-specific database and registry efforts. To illustrate the utility of the ENS@T-CANCER VRE for clinical trials we focus on the full 4-phase, international, multi-centre Prospective Monoamine Tumor (PMT) clinical trial and its utilisation of capabilities provided by the VRE.

Keywords: Disease registry, adrenal cancer, virtual research environment.

1. INTRODUCTION

A ubiquitous problem that exists in undertaking biomedical research is access to clinical/biomedical data and especially access to and sharing of data across organisational and national boundaries. These challenges are multi-faceted and comprise information governance, ethics and privacy challenges [1-3], human/social and organisational factors [4], as well as a variety of technical implementation issues that must be overcome. On the latter challenge, a body of work on the realisation of solutions for secure, web-based biomedical data sharing now exists [5-7]. For many clinical/biomedical collaborations this is through support for targeted web-based *disease registries*. Such systems are used to aggregate clinical/phenotypic data that can be used to coalesce the common understanding and treatment/management of patients with those particular diseases at a clinical level, and provide a model for potentially sharing of physical biosamples of patients with a specific phenotype - subject to ethics and agreement of clinicians, patients and indeed the organisations involved. These biosamples can then be used for a range of bioinformatics analysis and –omics research. Many such registries have been established [5,8,9]. More often than not, such platforms focus on sharing of phenotypic data. Whilst important, such capabilities represent a starting point for the technical opportunities that can now be achieved.

The ENS@T-CANCER project has taken the basic idea of a web-based disease registry to a new technological level to support a complete virtual research environment (VRE) for integrated biomedical research into adrenal tumors. The primary focus of this paper is to highlight the core features of the

¹ www.ensat-cancer.eu

ENS@T-CANCER VRE and illustrate the way in which they collectively provide a step change in biomedical research capabilities for the international adrenal tumor research community.

The rest of the paper is structured as follows. Section 2 provides background to the ENS@T-CANCER project and identifies the previous international databases and community ecosystems from which it arose. This section also highlights core capabilities required by the biomedical collaborators from the ENS@T-CANCER VRE. Section 3 focuses on how these core capabilities have been implemented and provides information on the utility of the system. Section 4 focuses on the multicentre prospective study of biochemical profiles of monoamine-producing tumors (PMT Study) and how it leverages capabilities of the VRE. A range of other clinical trials and studies is also identified that are currently being supported through the ENS@T-CANCER VRE. Finally, section 5 draws some conclusions to the work as a whole and identifies plans for the future and further applications of the ENS@T-CANCER VRE model currently underway across Australia

2. ENS@T-CANCER BACKGROUND

Development of collaborative IT-based systems around particular disease areas is beset by numerous challenges that must be overcome. Not least of which is coordination and cooperation of the biomedical research community involved. As well as the human networking element of collaboration, a major challenge is the agreement of the data that should actually be collected. Even in disease-specific areas like adrenal tumors, hospitals and research groups alike collect a multitude of data sets and information. These data are typically heterogeneous, both from the actual information that is collected, the format of the data and coding systems used for classification, as well as from the IT systems within which this information resides. To tackle the coordination and harmonisation of data to be collected for adrenal tumor research, as well as the overarching cooperation and planning of European-wide research into adrenal tumor research, the ENS@T-CANCER project builds on the international cooperation and coordination efforts of the European Network for the Study of Adrenal Tumors (ENS@T²). ENS@T was founded in 2002 through the merging of several existing but largely independent adrenal tumor research networks: the French COMETE (COrtico- MEduillo- Tumeurs Endocrines) [10]; the National Italian Study Group on Adrenal Tumors (NISGAT); the German Adrenal Network Improving Medical Education (GANIMED) [11] and the United Kingdom Adrenal Cortical Tumor network (UK ACT). The central aim of the ENS@T consortium was to understand the molecular mechanisms of particular tumor types and in so doing, improve the management and treatment of patients with adrenal tumors.

Tumors of the adrenals arise from the cortical or the medullary components of the glands. Clinical manifestations of adrenal tumors often arise through symptoms caused by excess secretion of hormones by these tumors. This can include an excess of catecholamines or of steroid hormones such as cortisol and aldosterone. Malignant adrenal tumors can also manifest themselves through local mass effects or symptoms related to distant metastatic spread. Adrenal tumors can be benign or malignant. Often this separation is difficult to establish and long-term close follow up is necessary after surgical removal to detect recurrences early in patients who have adrenal cancer. While malignant tumors of the adrenal gland are rare, up to 3.5% of the population have so called *adrenal incidentalomas* - tumors of the adrenals often found incidentally during investigation for unrelated conditions. The majority of these do not secrete hormones [12]. Of particular focus for ENS@T-CANCER are four main tumor types:

- *Aldosterone Producing Adenoma (APA)*;
- *Pheochromocytomas and Paragangliomas (Pheo/PGL)*;
- *Non-aldosterone producing adrenal cortical adenomas (NAPACA)*, and
- *Adrenocortical Carcinomas (ACC)*

Among the above Pheo/PGL and ACC are rare [13,14], whereas APA and NAPACA are more commonly encountered adrenal tumors. Through establishing a critical mass of information on patients with these adrenal tumor types, ENS@T-CANCER aims to show how *coordinated* studies of the genetics and treatment of patients with adrenal tumors can reveal new molecular mechanisms of the growth of these tumor types and provide insight into associated clinical research areas, e.g. their role in hypertension [15].

3. RELATED WORK

Many applications domains have pursued the realization of web-based collaborative research environments to simplify or enhance their research efforts. This enhancement can be from many perspectives:

² www.ensat.org

- Providing user-oriented access to large-scale facilities such as HPC/Cloud resources to domain scientists and researchers without them necessarily having to be savvy in those specialist IT areas. One example of such a system was the UK e-Science Integrative Biology project. The project focused on development of an infrastructure to support computer simulation of whole bodily organs based on molecular and cellular models. It focused especially on supporting researchers studying cardiovascular disease and cancer in particular [26]. The VRE supported the research process from laboratory experimentation to mathematical model building and simulation, using distributed HPC resources. The results of the *in silico* experiments were subsequently fed back to laboratory experimentation as part of the research process.
- Providing seamless and secure access to diverse and heterogeneous data sets. One example of such a system is the Australian Urban Research Infrastructure Network (AURIN) system [27] which provides online (live) access to over 2000 data sets from 70 major and typically definitive data organisations across Australia to the urban and built environment research community. This solution offers a data rich environment that tackles many of the data access and heterogeneity issues facing urban researchers. It also provides an extensive range of analytical and visualization tools. This system is available through a single web-based portal that offers a complete research environment to support the urban and built environment research process.
- Supporting the research process through shareable workflows. One example of such a VRE is the myGrid project [28]. This allows researchers to define and share a wide variety of workflows using languages such as Taverna [29]. The primary application domain has been the biomedical/biological area. Once defined these workflows can be shared with other researchers using the myExperiment platform [30]. In so doing, this in principle means that it is possible to reproduce (re-enact) the scientific process. There are many open challenges in this endeavor however [31].
- In some domains, e.g. the arts and humanities, it is often the case that the content and form of the solutions is heavily tailored to remove the underlying complexity of the software infrastructure. The JISC funded ENROLLER platform established an e-infrastructure for Language and Literature researchers that simplified access to diverse language and literature resources [32].

A survey of web-based virtual research environments is given in [33]. They identify that VREs are beginning to change the ways in which researchers go about their work and how they communicate with each other and with other stakeholders such as publishers and service providers. They identify that these changes are driven by the changing landscape of data production, curation and re-use, by new scientific methods, by changes in technology supply and the increasingly interdisciplinary nature of research in many domains.

In the clinical and biomedical domain, web-based collaborative platforms have not yet truly fulfilled the possibilities that are now afforded by technology. A variety of solutions have been developed to establish targeted disease registries, e.g. [5,8]. Such solutions focus on sharing of phenotypic data. Clinical research in disease areas now demands far richer data to be collected covering the whole course of the disease: initial diagnosis, treatments, -omics data analysis, sending and receiving of biomaterials, engagement of the information with the patient communities, support for reuse of data in clinical trials and organisation of data related to clinical trials, e.g. ensuring that the same patient is not currently involved in two conflicting clinical trials. The ENS@T-CANCER VRE has grown to offer such a data and functionality rich environment that is now used globally.

4. ENS@T-CANCER KEY REQUIREMENTS AND THEIR REALISATION BY THE VRE

4.1 Data Standardisation and Consolidation

A key goal of the ENS@T-CANCER project has been to offer a primary source of information associated with adrenal tumor data. As noted, much of these data were originally in a diverse collection of national/centre-specific databases that shared some commonality of information, but these data sets were not sufficiently aligned to support direct comparisons or to aggregate data directly. These systems were also not web accessible but were typically of the form of Excel spreadsheets and/or Access databases that were distributed to clinicians/researchers through physical media (e.g. CDs).

Working closely with the ENS@T network, a common set of phenotypic data models were produced for the four adrenal tumor subtypes. These phenotypic data models were developed over an extensive period by the ENS@T community – predating the ENS@T-CANCER project. The first phase of the

ENS@T-CANCER project was to implement these data models as web-based databases. It should be noted that the ENS@T-CANCER data models are *community* defined and not directly aligned with any given hospital data models, e.g. those based on efforts such as health level 7 (HL7) [22], SNOMED-Clinical Terms [23] or International Classification of Diseases (ICD9/ICD10) [24]. A multitude of standards exist across partner organisations, both in describing the data sets they hold and the IT systems that they have implemented. There is little consensus on data inter-operability across health care organisations – in the sense of HL7 messaging for example. It is of course possible to implement solutions that overcome the diversity of hospital systems and integrate data *in situ*. However, a deliberate decision was made in ENS@T-CANCER to not attempt to integrate the adrenal registries with existing hospital systems for live data access/data feeds. The technical challenges and solutions for achieving access to data in hospital settings are largely well understood [6,16]; however the issues associated with processes for establishing secure access systems and opening of hospital firewalls remains a challenge.

A key requirement for the ENS@T-CANCER VRE was however to incorporate all available data from the ENS@T-CANCER partners. To this end, the ENS@T-CANCER VRE implemented a range of data ingestion and integration systems to support this endeavour. This includes scripts for extracting pre-existing datasets in a variety of databases and mapping them to the ENS@T-CANCER data model. Close coordination with the clinical community was required in this process. At present (Figure 1) the system provides access to 10,054 deeply phenotyped data sets from patients with a range of adrenal tumor types. Of these patients, there are 57,943 clinical annotations available (where each treatment, surgery, follow up etc is a single annotation) and at present 20,340 physical biosamples that are typically stored in distributed laboratories across the collaborating sites. It is noted that the ENS@T-CANCER-supported clinical trials include many more samples.

Summary

	ACC	Pheo	NAPACA	APA	Total
Records	2541	2758	2938	1809	10054
Biosamples	5134	1998	7449	5759	20340
Clinical Annotations	29018	14898	8387	5640	57943
Annotations Per Patient (Mean)	11.41	5.40	2.85	3.11	5.76
Biosamples Per Patient (Mean)	2.02	0.72	2.53	3.18	2.02
Active Centers	54	60	35	23	91

Figure 1: ENS@T-CANCER Summary Phenotypic Information (Screenshot August 2016)

Whilst data consolidation has been achieved and a critical mass of data now exists, a key aspect of consolidation is through the coordination of the community itself. As seen from Figure 1, the ENS@T-CANCER VRE is currently actively used by 91 centres - far in excess of the original 15 centres that were funded through the ENS@T-CANCER project. There are at present over 200 users actively using the ENS@T-CANCER VRE. As such, the VRE now provides a focal point for adrenal tumor research that has effectively galvanized the community. It should be noted that the VRE continues to evolve and grow with the adrenal tumor communities' needs and demands.

4.2 Ethics, Consent and Information Governance

It is the case that numerous institutions are involved in clinical service delivery and in biomedical research, and this is the case with adrenal tumors. Whilst some institutions have well-established local hospital IT systems that are used for all aspects of admission, management and treatment of patients, the integration of systems across organisational and national boundaries is far less mature. A key challenge in such data sharing is the process of obtaining and enforcing ethical access to data and its use for research purposes. As described in [20,21] the ENS@T-CANCER platform has been developed through close alignment and adherence to a range of international security and information governance efforts associated with patient

data and privacy. These include the EU Directive 95/46/EC of the European Parliament³, which focuses on protection of individual information including the processing and movement of personal data; the UK Data Protection Act 1998⁴ and the Data Protection Order 2000 and Directive 02/58/EC of the European Parliament which are concerned with the processing and protection of personal data in the communications sector⁵. To ensure alignment with these initiatives, the ENS@T-CANCER VRE project has adhered to the highest standards of data security in the development, and on-going use and management of the VRE. Specifically, all users of the ENS@T-CANCER VRE have specific privileges depending on their role within the project. These roles are subsequently used to restrict access to data and tools in the VRE. The default model is to deny access to all users except for individuals with authentic and valid credentials, i.e. digitally signed credentials recognised by the ENS@T-CANCER VRE service components.

Many of these security and information governance directives and acts make it clear that subjects that are included on any clinical research system have a right to know of their inclusion on those systems and a right to have their data removed from those systems. The practice of informing subjects of inclusion on clinical systems varies across the EU. In the UK, a system of opt-out consent is often used. In other countries opt-in models of consent are in place. Contributing clinical collaborators for ENS@T-CANCER thus need to ensure that they adhere to all policies (local, national, international) on patient consent (both initial consent from patients to add their data into the registry, and for on-going data collected during follow up visits). To tackle this, the ENS@T-CANCER VRE implements a range of data sharing policies defined and supported when a patient is added into the core registry. As shown in Figure 2 for ACC patients, patients can be added in to the ENS@T-CANCER VRE with access restricted to local (clinical site specific partners); to national partners, e.g. only collaborators from France; to European-wide ENS@T collaborators, or indeed to international researchers involved in adrenal tumor research. Clinicians and their local teams are completely autonomous in the level of data sharing that they decide to associate with a particular patient record.

Create ACC Record

Identification

The following data points are the minimum criteria required for creating a patient record.

Adrenal Network ID Number:	[AutoGenerated]
Year of Birth:	1977
Sex:	M
Center ID:	AUSYZ
Referral doctor: Email:	Roderick Clifton-Bligh jclifton@med.usyd.edu.au
Date of AS Registration:	
Level of consent for clinical research:	[Select...]
Associated studies:	[Select...] Local National European ENSAT Partners International Collaborators
<input type="button" value="Confirm Details"/>	

Figure 2: Screenshot Consent Level Definition

It is important to emphasise that this level of consent is only for the clinical phenotypic information that is added to the registry and not for particular clinical trials that these patients might be involved in through ENS@T-CANCER, which have separate ethics requirements. Furthermore, the VRE is explicitly required to address information governance concerns of the collaborators and their associated organisations. A key element of this is to ensure the protection of patient identity and monitoring and tracking the use of patient information. As such, it is essential that no data that are included in the VRE can be used to potentially re-identify patients involved by others outside of the immediate (local) care providers at a given hospital. To address this, all data that are entered into the ENS@T-CANCER registry excludes any identifying data for patients, e.g. names, address, data of birth, local hospital identifiers. Instead a solution based on an automatically generated (and unique) identifier has been adopted. Furthermore, to avoid any potential inclusion of patient information, e.g. through manually entered text, the system has been designed

³ <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML>

⁴ <http://www.legislation.gov.uk/ukpga/1998/29/contents>

⁵ http://www.legislation.gov.uk/ukxi/2000/417/pdfs/ukxi_20000417_en.pdf

wherever possible with predefined menus containing values that have been agreed/standardised across the community. This harmonisation and agreement exercise is an essential component in the success of any disease registry.

The above data-sharing model also reflects a common challenge in clinical collaborations whereby clinical investigators are often unable or indeed unwilling to make *all* data available on *all* patients available to *all* collaborators. This can be due to the individual consent of patients, e.g. only wishing data to be available to a given centre, or to a given country; due to the differences in the ethics in place for a given country; due to the investigator's own research endeavours and/or the value proposition to the clinicians to share data for other researchers to publish results based on their patient data. With the ENS@T-CANCER data sharing model, clinical investigators are fully empowered to decide whether or not to share data. The VRE allows matches to be found but the access to further information depends completely upon the investigator involved to provide more information (including the sending of biosamples).

4.3 Patient Identification and Data Tracking Requirements

The ENS@T-CANCER VRE has been deliberately designed so that only the clinical investigators (or their local team members) who enter the patient data into the VRE are able to identify particular patients included in the VRE. Different hospitals and in some cases different countries, have established their own solutions for identification and tracking of patient information in/across the healthcare system. A unifying system was thus required. Simplistic solutions to this include the creation of a unique integer – incremented when a new patient is added to the system. Such a solution was incorporated into the I-DSD system [5]. However it was identified that an improved representation was required for ENS@T-CANCER that provided more structure reflecting the European-wide (and indeed the global!) nature of the project. To tackle this, a common global naming scheme was adopted utilising country-specific, centre-specific and patient-specific information. This scheme was designed to be scalable and allow for other countries and centres to be included.

The solution that has been rolled out capturing this naming structure is illustrated in Figure 3. Here the results of a sample search are depicted. As seen, the first column shows the patient identifier comprising the country code where GY-Germany and FR-France; the centre code where WU=Würzburg, NA=Nantes, MU=Munich, BO=Bordeaux, MA=Marseilles, and the final numeral (-7, -224 etc.) indicates that this is patient number 7 from Würzburg and patient 224 from Munich etc. It is important to note that only the local clinical investigators or their local staff are able to de-identify these patients. Furthermore, these individuals listed are responsible for the ethics and consent issues when adding these patients into the ENS@T-CANCER VRE. Clinical investigators maintain a local track of this record identifier on their own patient management systems and how it relates to an individual patient record in the ENS@T-CANCER VRE. Different sites adopt different approaches to recording this identifier locally; writing it on paper based hard copies of patient records or entering it into their IT systems. At no time will they ever be asked to reveal the actual identity of any individual patient to researchers outside of their immediate clinical care environment.

GYWU-7	Martin Fassnacht (fassnacht_m@medizin.uni-wuerzburg.de)	25 Oct 2011	10 Jun 2002	M	1962	National (Germany)	
FRNA-7	Delphine Drui (delphine.drui@chu-nantes.fr)	22 Apr 2011	13 Dec 2009	M	1955	National (France)	
GYMU-224	Felix Beuschlein (felix.beuschlein@med.uni-muenchen.de)	20 Jun 2011	21 Apr 2011	M	1962	European ENSAT Partners	Detail
FRBO-8	Antoine Tabarin (antoine.tabarin@chu-bordeaux.fr)	12 Jul 2011	10 Nov 2008	M	1941	National (France)	
FRMA-7	Patricia Niccoli (patricia.niccoli@ap-hm.fr)	12 Jul 2011	19 Sep 2008	M	1971	National (France)	

Figure 3: Screenshot Patient Identification and Data Sharing

Figure 3 also illustrates the way in which data sharing mechanisms are enforced. As seen here, several patients are listed with National data sharing only. In this case, only researchers from Germany will be able to see further details on patient GYWU-7, whilst patient GYMU-224 is accessible to all ENS@T-CANCER partners, i.e. partners named in the original grant proposal.

The ENS@T-CANCER supports a user-oriented search and export capability that allows for a range of simple to more advanced search and export capabilities. These can cover basic phenotypic/demographic information through to advanced disease or treatment-specific demands. The search interface directly reflects the core data model that was defined by the ENS@T network.

The use of a unique identifier for tracking patients in the ENS@T-CANCER VRE also allows for the association of other information that needs to be uniquely identified. Specifically, (physical) biosamples that are taken from patients, e.g. tissue samples (frozen/paraffin/DNA), plasma (EDTA/heparin), serum, urine (24 hour/spot), also need to be both uniquely identified and directly associated with particular patient records in the VRE. The ability to easily locate and track down all samples for a given patient is a key element of the VRE and is achieved through its biobanking support.

4.4 Biobanking Support

The ENS@T-CANCER VRE offers a range of biobanking capabilities, i.e. capabilities for managing and tracking physical (patient) bio-specimens such as blood, tissue, plasma, as an integral part of its functionality. Specifically the unique identification and tracking of samples across the collaboration is essential. As with the heterogeneity of data and healthcare IT systems across partner organisations, the utilisation of laboratory information management systems (LIMS) across partner sites is also highly varied. Whilst some ENS@T-CANCER sites have *local* integrated LIMS systems that provide *local* biosample management and tracking, other sites have no such systems, or have different and non-consistent solutions in place, e.g. with site-specific identification and naming/labeling protocols. To ensure that the needs of ENS@T-CANCER are met, a key requirement is to overcome this situation. The proposed solution allows for the unambiguous identification of physical samples; the association of particular samples with patients, and information as to the physical location of given samples.

Clinical partners themselves collect a range of physical bio-specimens. To support this targeted forms are supported that allow for information on what physical specimens have actually been collected. Each of these forms has an associated unique identifier that is associated with a particular patient in the VRE. Figure 4 (left) illustrates the range of associated information that can be associated with a particular patient. The left column indicates the patient identifier (patient 358 from Germany, Munich) with the form identifiers (775, 776, 773 etc) providing a range of additional information related to biomaterials amongst other things. Two of these forms are illustrated (on the right of Figure 4) with an indication of the biomaterials that are present including the volume of 24h urine that has been collected and the pathology information associated with that patient.

ENSAT ID	Form ID	Date	Record Information	Detail	Delete
GYMU-0222	20637	05 Dec 2013	Follow-Up	Detail	
GYMU-0222	20636	28 Nov 2012	Follow-Up	Detail	
GYMU-0222	639	09 Jul 2012	Biomaterial (Heparin plasma Serum 24h urine Spot urine)	Detail	
GYMU-0222	638	09 Jul 2012	Biomaterial (Leukocyte DNA Plasma Serum 24h urine)	Detail	
GYMU-0222	20634	29 Jun 2012	Follow-Up	Detail	Delete
GYMU-0222	19825	23 Apr 2012	Follow-Up	Detail	Delete
GYMU-0222	991	01 Mar 2012	Chemotherapy	Detail	Delete

ACC Update Follow Up Record 20637

Follow-up date: 05-12-2013

Lost to follow-up: No

Patient status: Alive with disease

Alive with disease (Ctrl-click to select):

- Local Recurrence
- Thoracic Lymph Nodes
- Remaining Adrenal Tumor
- Lung

ACC Update Chemotherapy Record 991

Chemotherapy date: 09-07-2012

Chemotherapy initiation: 01-03-2012

Chemotherapy end: 01-05-2012

Indication: Palliative

Best objective evaluation: Progressive

Evaluation date: 01-05-2012

Regimen (Ctrl-click to select):

- EOP
- Cisplatin (P)
- Carboplatin

Figure 4: Screenshots ENS@T-CANCER VRE Associated Record Information Forms

In scaling up the registry to thousands of patients and many trials and studies, clinical investigators and clinical centres more generally need capabilities for rapidly tracking down such information. As indicated in Figure 4, patients with adrenal tumors require a range of on-going information to be collected and shared. This can include the status of patients, e.g. whether they are alive and disease free; living with the disease or indeed have died and if so whether the death was related to the cancer; to the treatment of the cancer or for some other reason. It can also include information on the cancer itself, e.g. whether it has metastasised or gone into remission; the treatments the patients receive (drugs, chemotherapy, radiotherapy); the surgical procedures they have undergone; results of treatments/surgeries in a range of

forms, e.g. imaging information. Tracking the whole history of information for patients is a core demand of the ENS@T-CANCER VRE.

To accurately track the physical specification of the biomaterial samples, key parameters to note are the aliquots that the samples are split into and their position in the local freezer. Sections in the form are programmed to open upon selection detailing the aliquot number, and the freezer, shelf, rack, box and position numbers therein. If the position is already occupied then a validation flag notes this and prevents re-entry of that specific position. When the sample is used and the position becomes free, this is noted and archived in an archive history of the local freezer. In this way it is possible to corroborate the transfer of material between centres to a high degree of accuracy. Figure 5 shows the aliquot and freezer information sections in the biomaterial forms for one patient in a centre (Munich, Germany).

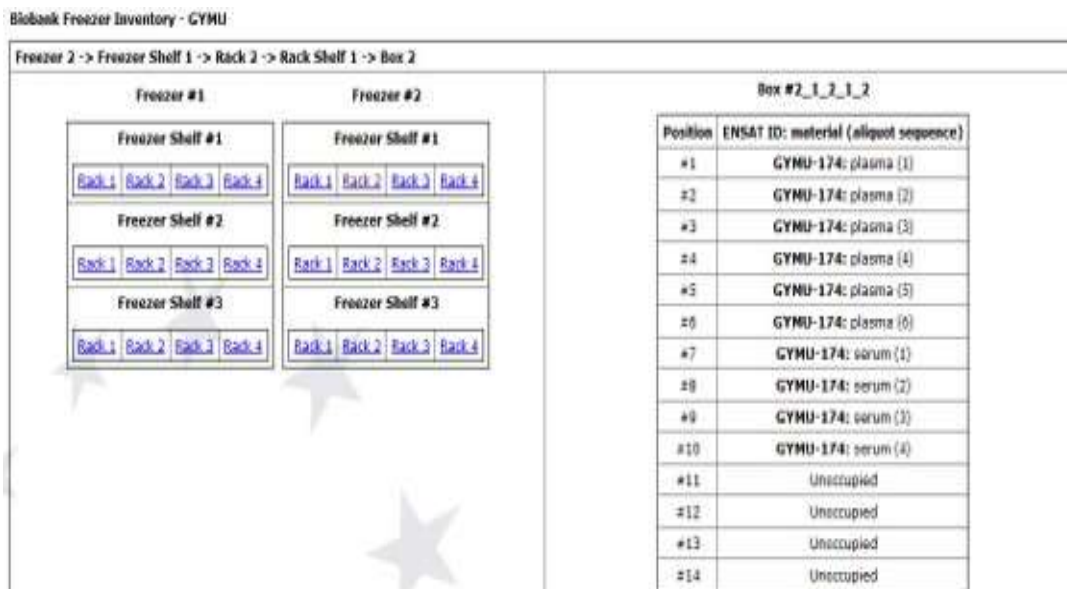


Figure 5: Screenshot Example of the Freezer Location of Samples

The production of labels is obviously of great significance when transferring human materials, along with the ability to provide information that is relevant to all of the participating centres. In this regard the sample specifications are stored permanently with the option for modification when it comes to the time of printing. The information captured includes the canonical identifier (including the centre code), the biomaterial form number, the aliquot number, and the sample name. This is also captured in barcodes for readers that wish to capture the information electronically. There are many barcode standards to choose from – QR codes were initially tried but rejected for biosamples stored in pathology laboratories due to the two-dimensional nature being unreadable on the curved surface of a typical sample tube. The project explored [16] solutions that catered for curved surfaces, however a standardised labelling convention was ultimately adopted for samples stored in the laboratories.

Supporting the tracking of physical biospecimens in transit between sites was identified as an important capability of the ENS@T-CANCER VRE. To support this mobile (both iOS and Android) applications (apps) have been developed that allow for scanning of QR codes encoding a range of detailed bio-sample information [17,18]. Use of these apps requires authentication codes generated by the ENS@T-CANCER VRE (Figure 6 left). The QR codes are included on the inventories associated with the boxes of biospecimens that are sent between collaborating partner sites – hence avoiding the problem of curved surfaces for labelling specimens on test tubes. The QR codes themselves are generated by the VRE. Upon receipt of these biospecimen boxes with their associated inventories (middle Figure 6), the QR codes are scanned using the app and the user prompted to verify that the samples contained therein have been received successfully. When this is the case, a notification is sent from the app to the ENS@T-CANCER database along with the actual location of the samples in the local fridge (Figure 6 right). Capabilities also exist to track these samples through parcel tracking services from couriers (if/when they are available). The app user can also identify that samples are missing or spoilt as appropriate through the app. This feature thus allows for constant monitoring of all *digital* and *physical* data used within the ENS@T-CANCER VRE.

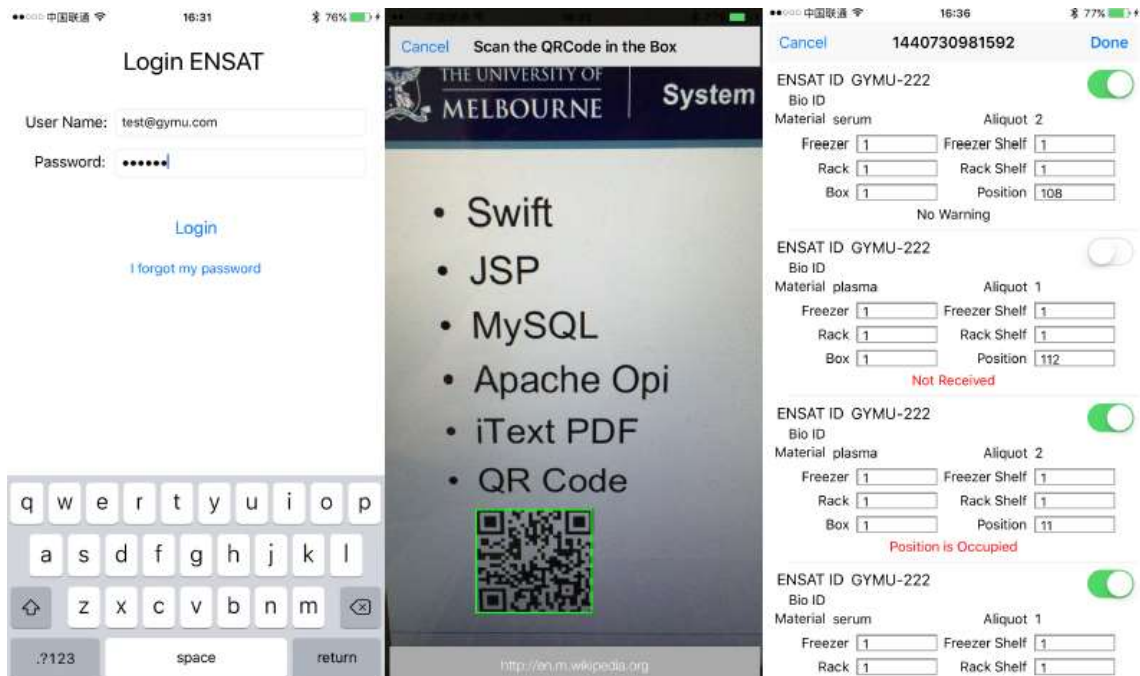


Figure 6: Biosample tracking and receipt notification mobile apps

4.5 Longitudinal Patient Visualisation

To aid in the understanding and visualisation of patient information, the ENS@T-CANCER VRE supports both tabular (e.g. Figure 4) and graph based visualisation of patient information. The graph-based visualisation provides a timeline through which all events related to the life course of a patient is displayed. Options exist to include further details for each time point associated with a patient event, e.g. the dosages of particular drugs that might be given over the treatment course. Figure 7 shows the typical timeline for a patient that has undergone a series of treatments.

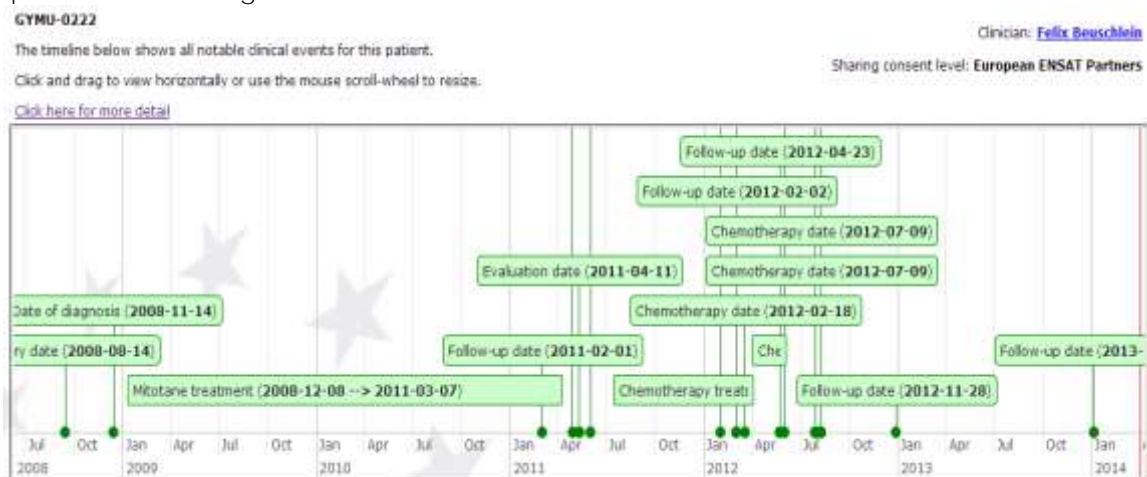


Figure 7: Visualisation of Longitudinal Patient Information Tracking

These longitudinal data provide a rich source of information that can be used to help determine the optimal treatment pathways for patients.

4.5 ENS@T-CANCER Technical Architecture

To realise the above functionality, the ENS@T-CANCER VRE has implemented an n-tier web application using a standard technology stack comprising MySQL for the database; JSP/Java for the business logic and connecting to the user interface that itself uses various supplementary JavaScript libraries such as jQuery.

4.5.1 User Interface

The main features that the user interacts with on the registry are the ability to create, edit and delete patient records (CRUD), to search those records and to export a snapshot of record subsets to a Microsoft Excel file at any given moment (both of these features are subject to security privilege). There are a variety of different parameters associated with each type of cancer, so a separate parameter database is used (similar to the idea of a descriptive “second schema”) to load the relevant parameters when a patient record or follow-up form is created, edited or searched on. This parameter list is loaded at the head of every relevant generated web page and is filtered using a PolicyHandler Java class, described in the Authorisation section below.

To enable a richer user experience, various JavaScript functions are implemented where appropriate. For example, jQuery libraries are used for date-pickers and the timeline function and are also used as client-side validation of parameter input, e.g. detecting when a number input has text in it. There are also a variety of server-side JSP controls for this that act as a fall back in case the client has JavaScript disabled. Statistical visualisations, such as completeness charts use the jQuery/JSON package HighCharts.

4.5.2 Business Logic

The JSP/Java implementation of business logic in the registry follows a standard Model-View-Controller paradigm (MVC), where one control class for each function mediates the view between the user and the back-end database. The most heavily used function throughout is the loading of parameters into the view in order that record details can be viewed, or manipulated in a view-confirm-process cycle. Other features have been added to the registry over time, which have their own individual functionality requirements. These include: biomaterial labelling, biomaterial freezer inventory management, record completeness statistics, geographical coverage, patient timeline visualisation, snapshots of specific patient status reports, transfer between other individual studies (e.g. PMT), and inclusion in studies that require use of mobile applications.

4.5.3 Data model

There are four databases that underpin the ENSAT-CANCER VRE. These include a parameter database (mentioned above in the User Interface section with the data model shown in in Figure 8); a main database to hold all the actual clinical data, a “center_callout” database which acts as an event clock to keep the ENSAT IDs synchronized across the registry and all associated studies (e.g. when a patient is created in PMT, it uses the next ID number from this database and increments this number as well), as well as a security database, which has a user account table, a membership table (which is updated according to whether the clinician is up-to-date with their ENSAT membership. The ENSAT network monitors this membership information. There are also several tables that enforce the role-based access control (RBAC) systems required to delineate domains within the ENSAT community (see Figure 9).

Field
param_id
param_name
param_type
param_text_size
param_label
param_order_id
menu
param_sub_param
param_table
param_optional

Figure 8: Descriptive “second schema” of parameters used. The features of each parameter allows them to be ordered, configured and validated in a standard way for all record creation, editing and searching.

The RBAC security policy allows parameters to be viewed depending on a user’s domain (e.g. Europe, America, Asia) as well as the role that they have been assigned (e.g. investigator), and the sharing consent level that has been assigned to the single patient record or grouping of listed patients on the tumor-type home-page (e.g. local, national, ENSAT, everyone). The RBAC system itself as shown in Figure comprises: Users (listed by ID in the user table), actions (five actions are shown) and targets (listed by ID in the parameter database). These are all tangible assets that are assigned to roles in various combinations and labelled accordingly (e.g. “ENSAT investigator”). These roles interact then further (control of edit or read-only) with the sharing consent level in the business logic – this separation was a deliberate design decision as the sharing consent level is a property of the patient record, not the user doing the viewing. In

this way, roles can be quickly created and modified depending on need and the development of policy through the ENSAT steering committee.

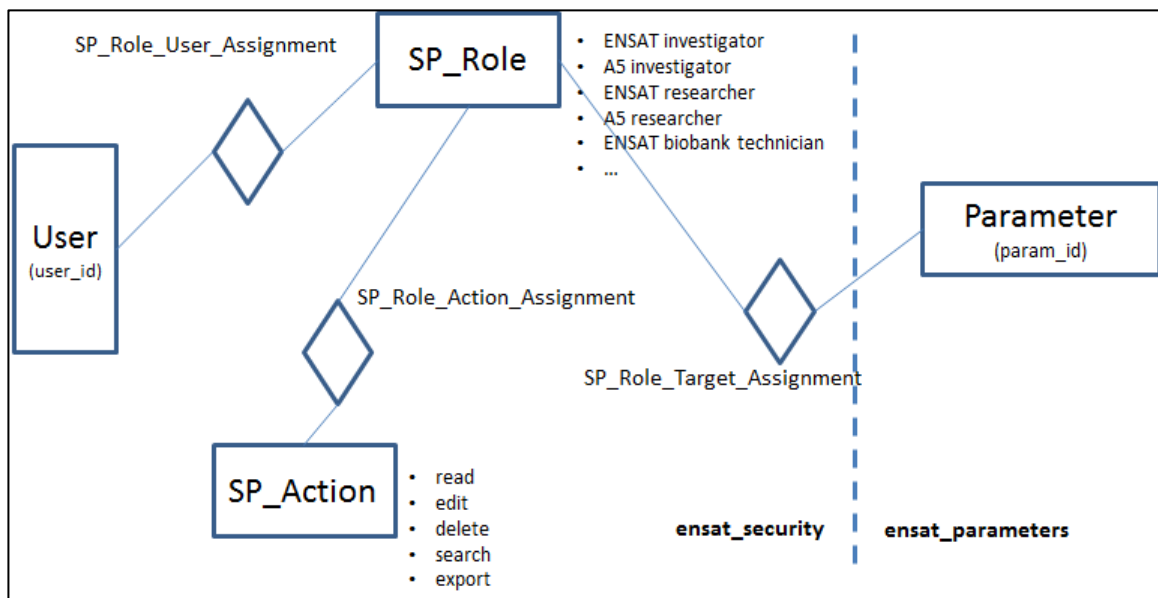


Figure 9: Underlying ENSAT-CANCER RBAC architecture as mediated by the SP_Role table.

4.5.4 Mobile Application API

Two specific applications in the ENSAT-CANCER VRE require the use of mobile phones to participate, and have apps specifically developed for the iOS and Android operating systems that integrate with the registry. These are the “NAPACApp” study, which is a patient-driven study that allows patients with a NAPACA tumor to enter a daily diary of symptoms, and a biomaterial application that allows clinicians to transfer biomaterials between specialist centres and efficiently create an inventory to catalogue the biomaterial information held in a manifest file. This is achieved through scanning a QR code that encodes this information and uploading it to their local freezer inventory in the registry. Both of these applications require an API connection to the registry to function. This API is a standard RESTful interface using *https* to a participating device (e.g. an iPhone). The authentication methods for both applications are slightly different due to the different nature of the target users of each application. The NAPACApp application uses a mechanism of a doctor enrolling a patient into the study, whereby they are given an access code (a 128-byte limited BigInteger generated through Java’s *SecureRandom()* method) that they then communicate to the patient out-of-band. The patient downloads the app from the Apple app or Google Play store, and upon installing and running the application, they input this code. That code authenticates the user by associating it with the appropriate ENSAT ID in the registry (transmitted as an *auth-header* in the http request), and the user can subsequently upload their daily diary information. There are two JSP files that enable these functional interactions between device and registry, with the authentication performed in the header of each. This authentication code – similar to all sensitive material held in the registry – is encrypted with a salted SHA-256 algorithm.

For the biomaterial application, the API is designed for the clinicians that already have accounts in the registry, so this is how the authentication is performed directly. When a sample is transferred in this application, the biomaterials that are shipped are marked in the registry and a manifest is created in Excel listing the ENSAT ID, the sample material and the aliquot sequence numbers. This spreadsheet is given an ID, stored in the registry database and the spreadsheet ID is encoded as a QR code. Upon receipt of the shipment, the clinician logs into their mobile app and scans the QR code. The number is decoded and the mobile app interrogates the registry for the details of this spreadsheet manifest. Once returned the user can confirm receipt (or otherwise) of the materials – with the app displaying suggested freezer positions based on its knowledge of the occupancy of the target freezer. Upon confirmation, the freezer inventories are quickly updated to reflect the new sample and location information.

4.5.5 Infrastructure and Security

Given the sensitive nature of the data held, a strong emphasis is on security in the registry, with features programmed throughout the application, as well as a general awareness of security throughout the infrastructure stack that the registry is built upon. As already mentioned, all sensitive information such as passwords or other authentication codes are encrypted using SHA-256 algorithms (built into MySQL version 5.5 onwards) and are salted using the Java *SecureRandom()* pseudo-random number generator. The application is “gated” by having a number of default-deny JSP capture pages in the headers of each page accessed, which include checking credentials (user is authorised to view the page), and checking input (the requesting http does not include any parameters that aren’t approved and the values are not suspicious in nature, such as SQL or possible byte-code).

The registry is hosted on two virtual machines (VMs) - one for web-server and one for database. These are hosted in the University of Melbourne data centre. Nightly backup snapshots are taken of the data and encrypted with GPG, are backed up to a third volume that is stored on a physically separate site to aid with disaster recovery. All code is versioned using Git and stored in a private GitHub repository. Log files are managed using log4j – a rolling daily text log of every page-load and record update is held and emailed to the registry administrator every night. This is also copied to another “audit” VM and loaded into a CouchDB database to allow querying of application usage.

Infrastructure security is enforced by ensuring that the VM software is regularly patched and updated according to best-practice standards, and is regularly checked and monitored by separate sys-admin staff. Network security is enforced by configuring firewall access to only specific machines within the University network. Only port 443 is exposed for public general access. All other ports are locked down and are further protected by port-forwarding techniques (the data centre has services that specifically allows easy configuration of such setups).

Ultimately, the ENS@T-CANCER VRE collates data on patients with adrenal tumors or suspected adrenal tumors. However, much of the research associated with ENS@T-CANCER is conducted within the context of clinical trials and studies. At present over 25 major international clinical trials are supported through the ENS@T-CANCER VRE. To understand the relationship between the registry and these studies we consider one particular on-going clinical trial.

5. CLINICAL TRIAL SUPPORT

In the original ENS@T-CANCER proposal, several major international trials were identified. These included a study on the efficacy of adjuvant mitotane treatment in prolonging recurrence-free survival in patients with Adrenocortical Carcinoma at low-intermediate risk of recurrence submitted to radical resection (ADIUVO⁶); a randomized study in malignant progressive pheochromocytoma/paraganglioma (PPGL) (FIRSTMAPPP⁷), and an international prospective study focused on the diagnostic and prognostic value of urine steroid/biomarkers in patients with adrenocortical tumors (EURINE-ACT⁸). However it was recognised that numerous other trials should also be supported by the registry.

From the outset it was identified that several capabilities were required from the VRE to support this activity: patient recruitment to clinical trials and studies; the seamless flow of information between the core registry databases and the clinical trial-specific databases for patients matching/screening and recruitment needs of those trials, and indeed the reverse flow of information of data collected in clinical trials back to the central registry databases. Knowing how many patients meet the criteria for a particular trial can be important regarding the potential viability and associated cost of that trial. Knowing which patients are, or were, involved in which trials and studies is also important to avoid potentially invalidating trials and studies by having patients in multiple studies undergoing different courses of treatments.

The Prospective Monoamine Tumor (PMT) study⁹ is one of the larger clinical trials associated with ENS@T-CANCER.

5.1 Prospective Monoamine Study (PMT Study)

The PMT study is a full four-phase clinical trial that tracks patients who exhibit clinical indications of *suspected* pheochromocytoma [20]. This can be through their signs and symptoms; their therapy-resistant

⁶ <https://www.epidini.it/adiuvo>

⁷ www.firstmapp.fr

⁸ <http://www.ensat-cancer.eu/clinical-studies/eurine-act>

⁹ <https://pmt-study.pressor.org>

hypertension; through incidental findings, e.g. through imaging used for other conditions; through routine screening due to known mutations or hereditary syndromes, or indeed through routine screening due to a previous history of pheochromocytoma. Patients are tracked through all four phases of the trial. This includes the initial screening; clonidine tests; their molecular characterisation, and subsequent follow-up – and beyond up to five years after study. The current recruitment status of the PMT study is shown in Figure 10. As seen there are at present 2106 patients that have been recruited to phase 1 of the study. Associated with these patients there are 24,480 physical bio-specimens including plasma, serum, urine and frozen tumor samples. Importantly the PMT systems utilise the biobanking systems that have been developed through ENS@T-CANCER including the freezer storage location and sample tracking capabilities.

	Full Study	Phase 1 (Screening)	Phase 2 (Clonidine)	Phase 3 (Pheo Characterization)	Phase 4 A (Excluded Follow-Up)	Phase 4 B (Pheo Follow-Up)
Patients	2106	1181	49	8	659	209

Figure 10: Screenshot PMT Study Recruitment Status (August 2016)

The clinical path of patient progress through the PMT study is outlined in Figure 11. It is expected that 2400 patients will be recruited – specifically those with suspected pheochromocytoma or paraganglioma. These patients will undergo initial diagnostic screening, biochemical confirmation, disease characterisation followed by disease verification and subsequent follow-up for a much smaller set of patients (200). As seen from Figure 1, within the ENS@T-CANCER VRE there are at present 2386 patients that have been diagnosed with pheochromocytoma or paraganglioma at present. However many of these have been included from retrospective series and are not eligible for the PMT Study or indeed are willing/able to consent to be involved in the PMT study. Thus, the registry also includes many patients who have since deceased.

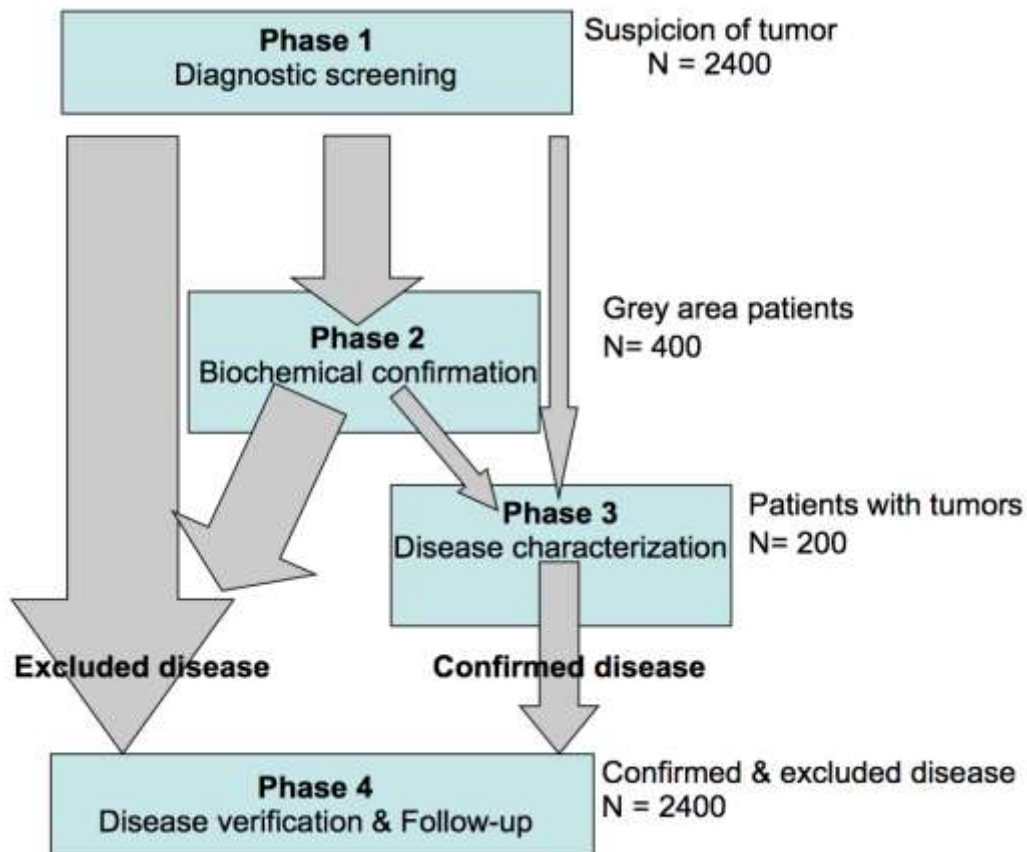


Figure 11: PMT Clinical Pathway [20]

The PMT study itself requires collection of an extensive collection of information on patients – far beyond the core information found in the ENS@T-CANCER registry. To support the utilisation of existing data as identified above, the electronic case report forms (eCRFs) for the PMT study and in turn the associated data models have been defined explicitly to overlap with the pheochromocytoma and paraganglioma data models provided in the registry. Specifically, the PMT Study tables that map directly to the ENS@T-CANCER include Identification, Demographics, Tumor Details and Genetics. These all form part of phase 1 of the PMT Study. They were designed with similar parameter fields so that transfers between both the VRE and the PMT Study database could be as smooth as possible.

It is also noted that it is often the case that patients are recruited into the PMT Study outside of the registry. As such, the systems need to allow seamless flows of information from the registry to the study and from the study to the registry.

It is important to note that many of the patients in the Phase 1 (screening) of the PMT study were originally entered into and subsequently recruited from the ENS@T-CANCER registry. The recruitment itself has specific needs and demands based on the requirements of PMT. Thus the protocol identifies a range of data that are essential for the study and these data should ideally be available for the study to utilise. To support this, the ENS@T-CANCER VRE supports a multitude of clinical trial specific availability and completeness measures that give an indication of the availability of data sets of relevance to any particular study. In the case of the PMT study, Figure 12 shows the overall completeness of the data that is available in the registry that is required for the PMT study.

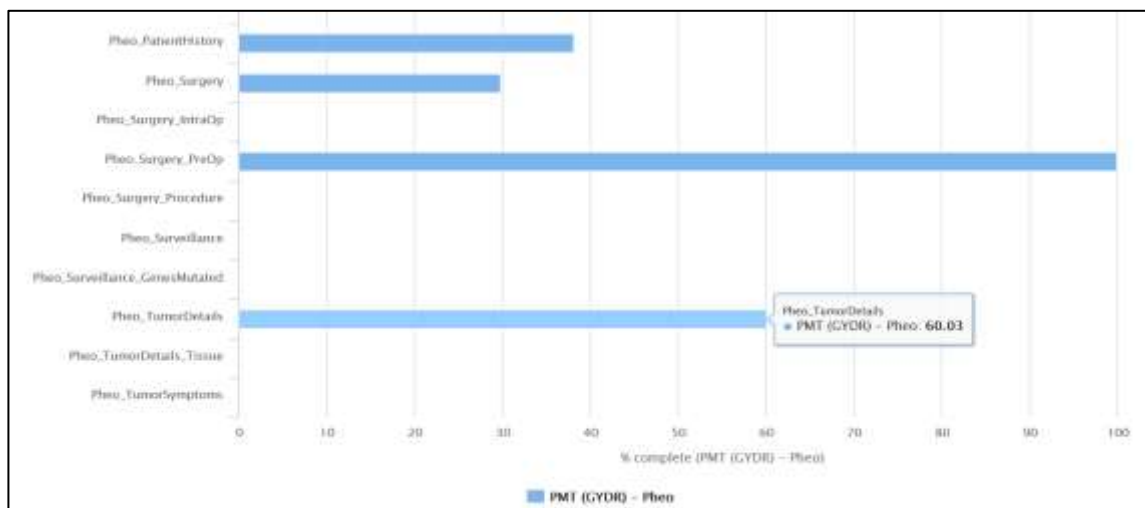


Figure 12: ENS@T-CANCER Registry PMT Study Data Completeness (December 2015)

It is important to note that the ENS@T-CANCER registry does not mandate that all data is compulsory and must therefore be entered fully, since many centres do not collect all data or face challenges with the availability of personnel to complete such information. Such issues are explored in [19]. As seen from Figure 13, the PMT study has an extensive set of data requirements on the data to be collected on patients with a suspected pheochromocytoma/paraganglioma. Many of these data are completely provided (<100%) in the registry whilst others are not available at all. This does not mean that the data cannot be collected - indeed it is often the case that the data are collected as part of the clinical trial. Interactions between the PMT Study and the contributors of data to the ENS@T-CANCER registry can be used to improve the quality of the overall data that are collected in the registry. Such interactions form part of the on-going collaboration between the protagonists involved.

The PMT Study database and many of the associated eCRF forms have been designed to collect far more data than is collected in the ENS@T-CANCER registry – a typical PMT Study patient can have over 10,000 data points. This has both data complexity challenges and implications on the data formats that are used. As an example of this complexity, Phase 1 of the PMT Study required a biochemical test matrix (shown in Figure 13). The units were input and converted between SI and imperial measures (e.g. pg/mL versus nmol/L). Originally these values were stored in the database in one format – and subsequently converted and rendered in both units when displayed. However, the accuracy loss during conversion was deemed unacceptable and the information was subsequently stored as separate data points, interchangeable if precision loss was tolerable, but losing none of the accuracy of the original input data units. It should be

noted that the range of numbers was large – with values of three decimal places for some measurements (e.g. 0.003 nmol/L), ranging to five significant figures for other measurements (e.g. 50,000 nmol/L).

Phase 1 - Screening Biochemical Tests [Print](#)

Update the screening biochemical tests for this patient by entering the data manually in the fields below

PLEASE NOTE: only local Dresden users should be entering information on this form, for the correct processing of LC-MS/MS data. All other users will not have access - for further information please contact [Prof. Grottel Eisenhofer](#).

Summary					
	Plasma Metanephrines		Urine Free Metanephrines		Urine Deconj. Metanephrines
Date of test	2010-11-25		2010-11-25		
Normetanephrine	73.7 (pg/mL)	0.402 (nmol/L)	21.5 (µg/day)	117.3 (nmol/day)	314.0 (µg/day) 1713.9 (nmol/day)
Metanephrine	29.8 (pg/mL)	0.151 (nmol/L)	14.7 (µg/day)	74.5 (nmol/day)	110.2 (µg/day) 558.8 (nmol/day)
Methoxytyramine	18.3 (pg/mL)	0.109 (nmol/L)	25.9 (µg/day)	154.8 (nmol/day)	188.0 (µg/day) 1124.3 (nmol/day)
Creatinine (urine only)			1549 (mg/day)		13.697 (mmol/day)

Figure 13: PMT Phase 1 screening biochemical test matrix

Various features are implemented in both the PMT Study application and the ENS@T registry, to support the information flow and connectivity across the virtual environments. Patients that have been recruited into the PMT study (or indeed any of the 25 clinical trials and studies that are currently on-going) are identified as such in the registry. See Figure 4 top right for a patient that has been recruited to the EURINE-ACT study for example.

The analysis of the data that is captured in the registry and the PMT study is a key factor to the success of the work. The solution that has been adopted thus far is to allow secure export of data into comma-separated-value (CSV) files that can subsequently be incorporated into a range of statistical software packages used by the research community. This can be non-trivial since both the registry and the associated studies contain tables that repeat or have one-many relationships, e.g. follow-up visits or screening tests. To support the export into CSV, the data is flattened into a 2D Excel spreadsheet-like structure. These structures can include over 10,000 columns per patient with data that can be sparsely populated throughout. The adrenal tumor research community have adapted to the complexity of the data that the registry makes available and are used to working with such data formats.

5. CONCLUSIONS

The ENS@T-CANCER VRE is a powerful web-based environment that holds together an active research community that is growing with the addition of multiple studies in the adrenal tumor field. The VRE is now globally adopted with groups in Brazil, Russia, USA and many other centres around the world who are actively using this platform as part of their daily research into adrenal tumors. The success of the system has occurred as a result of a range of factors. Firstly, it has leveraged agreements on the data models from the international community. The role of ENS@T was pivotal in this regard. Secondly, the data sharing models that are supported empower clinicians and researchers as opposed to researchers benefiting from clinical samples. Clinical investigators are not obliged to share data. They can decide if they wish to share samples with researchers based on a variety of factors: the patients interests; their own interests as well as any ethics and governance concerns that might exist. This *modus operandi* has allowed collaborations to flourish. Thirdly, the system has evolved with community needs. It leverages best practice software engineering and is moving the community forward through careful introduction to a range of data sharing mechanisms and associated technologies. The introduction of mobile applications is one such recent example.

Ultimately the success of ENS@T-CANCER can be considered from several perspectives. Is it being used? Figure 1 clearly demonstrates that this is the case. Is it used to produce important research results? At the time of writing over 90 high impact journal publications have been published including the Nature and New England Journal of Medicine¹⁰. Will it be sustained? Through the work on the VRE, a 5-year Horizon2020 project¹¹ investigating hypertension and whether it is possible to develop more personalised

¹⁰ <http://www.ensat-cancereu/publications>

¹¹ <http://www.ensat.hteu/>.

approaches to treatment and disease management. Multiple grants and follow on projects have been proposed that will hopefully ensure that this resource continues to serve the needs and demands of the international adrenal research community.

Acknowledgements

ENS@T-Cancer - The work leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 259735. This work has also received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 633983. The University of Melbourne Master students, Purathani Gunanayagam and Chao Lin undertook the work on the ENS@T-CANCER mobile applications and acknowledged for their efforts here. We would also like to thank the ENS@T community for their continued support and involvement in ENS@T-CANCER, and their adoption and usage of the VRE.

REFERENCES

- [1] Durugbo, C., Tiwari, A., Alcock, J.R., Modelling information flow for organisations: A review of approaches and future challenges, *International Journal of Information Management*, Feb 2013, <http://dx.doi.org/10.1016/j.ijinfomgt.2013.01.009>
- [2] Kaye, J., The Tension Between Data Sharing and the Protection of Privacy in Genomics Research, *Annual Review of Genomics and Human Genetics*, Vol. 13: 415-431, Sep 2012, <http://DOI:10.1146/annurev-genom-082410-101454>
- [3] Hickey, G.L., Grant, S.W., Cosgriff, R., Dimarakis, I., Pagano, D., Kappetein, A.P., Bridgewater, B. (2013) Clinical registries: governance, management, analysis and applications, *Eur J Cardiothorac Surg*, <http://doi:10.1093/ejcts/ezt018>
- [4] Kim, Y., Stanton, J.M., Institutional and Individual Influences on Scientists' Data Sharing Practices, *Journal of Computer Science Education*, Vol 3, Issue 1, June 2012.
- [5] Ahmed, S.F., Rodie, M., Jiang, J., Sinnott, R.O., The European DSD Registry – A Virtual Research Environment, *International Journal on Sexual Development, Special issue on Disorders of Sex Development*, “New concepts for human disorders of sex development”, *Sex Dev.* 2010; 4:192-198 (<http://DOI:10.1159/000313434>).
- [6] Sinnott, R.O., Stell, A.J., Jiang, J., Classifying Architectural Data Sharing Models for e-Health Collaborations, *Proceedings of the International HealthGrid Conference*, Bristol, UK, June 2011.
- [7] Stell, A.J., Sinnott, R.O., Jiang, J., Enabling Secure, Distributed Collaborations for Adrenal Tumor Research, *Proceedings of the International HealthGrid conference*, Paris, France, June 2010.
- [8] Bellgard, M. I., Macgregor, A., Janon, F., Harvey, A., O'Leary, P., Hunter, A., & Dawkins, H. A modular approach to disease registry design: Successful adoption of an internet-based rare disease registry. *Human Mutation*.
- [9] Kleophas, W., Bieber, B., Robinson, B. M., Duttlinger, J., Fliser, D., Lonnemann, G., and Reichel, H. (2012). Implementation and first results of a German Chronic Kidney Disease Registry. *Clinical nephrology*.
- [10] Plouin P.F., Gimenez-Roqueplo A.P., Bertagna X., COMETE, a network for the study and management of hypersecreting adrenal tumors, *Bull Acad Natl Med.* 2008 Jan;192(1):73-82.
- [11] Koschker A.C., Fassnacht M., Hahner S., Weismann D., Allolio B., Adrenocortical carcinoma - improving patient care by establishing new structures. *Exp Clin Endocrinol Diabetes.* 2006 Feb; 114(2):45-51.
- [12] Nieman, L.K., Approach to the patient with an adrenal incidentaloma *J Clinical Endocrinology and Metabolism* 2010; 95:4106-4113
- [13] Golden, S.H., Robinson, K.A., Saldanha, I., Anton, B., Ladenson, P.W., Prevalence and incidence of endocrine and metabolic disorders in the United States: a comprehensive review. *J Clin Endocrinol Metab* 2009; 94:1853-1878
- [14] Kebebew E., Reiff, E., Duh, Q.Y., Clark, O.H., McMillan, A., Extent of disease at presentation and outcome for adrenocortical carcinoma: have we made progress? *World J Surg* 2006; 30:872-878
- [15] Beuschlein, F., EJE PRIZE 2013: Regulation of Aldosterone secretion - from physiology to disease. *Eur J Endocrinol*, April 8, 2013 EJE-13-0263, <http://doi:10.1530/EJE-13-0263>.
- [16] Stell, A.J., Sinnott, R.O., The ENSAT Registry: A Digital Repository Supporting Adrenal Cancer Research, *Health Informatics Conference (HIC)*, Sydney, Australia, July 2012.
- [17] Gunanayagam P., *Biomaterial Information Tracker*, MSc Dissertation, University of Melbourne, November 2015.
- [18] Lin, C., *Development of an iPhone Biomaterial Tracking Application for the ENSAT Registry*, MSc Dissertation, University of Melbourne, November 2015.
- [19] Glöckner, S., Arit, W., Bancos, I., Stell, A.J., Sinnott, R.O., Improving Data Quality in Disease Registries and Clinical Trials: A Case Study from the ENSAT-CANCER Project 8th Australasian Workshop on Health Informatics and Knowledge Management, Sydney, Australia, January 2015.
- [20] Stell, A.J., Sinnott, R.O., e-Enabling International Cancer Research: Lessons Being Learnt in the ENS@T-CANCER Project, *IEEE e-Science Conference*, Beijing, China, October 2013.
- [21] Sinnott, R.O., Bruns, L., Duran, C., Hu, W., Jayaputera, G., Stell, A.J., Development of an Endocrine Genomics Virtual Research Environment for Australia: Building on Success, 13th International Conference on Computational Science and its Applications (ICCSA 2013), Ho Chi Minh City, Vietnam, June 2013.

- [22] Dolin, R.H., Alschuler, L., Boyer, S., Beebe, C., Behlen, F.M., Biron, P.V. and Shabo, A., 2006. HL7 clinical document architecture, release 2. *Journal of the American Medical Informatics Association*, 13(1), pp.30-39.
- [23] Donnelly, K., 2006. SNOMED-CT: The advanced terminology and coding system for eHealth. *Studies in health technology and informatics*, 121, p.279.
- [24] Organisation, T.W.H., 2010. International Classification of Diseases (ICD)-10.
- [25] Voss, A. and Procter, R., 2009. Virtual research environments in scholarly work and communications. *Library Hi Tech*, 27(2), pp.174-190.
- [26] Mascord, M., Jirotko, M. and Carusi, A., 2007. Integrative biology VRE. *Digital Pen Evaluation Report*.
- [27] Sinnott R.O., Bayliss, C., Bromage, A., Galang, G. Grazioli, G., Greenwood, P., Macauley, A., Morandini, L., Nino-Ruiz, M., Pettit, C., Tomko, M. Sarwar, M., Stimson, R., Voorsluys, W., Widjaja, I., *The Australian Urban Research Gateway*, Journal of Concurrency and Computation: Practice and Experience, April 2014, doi: 10.1002/cpe.3282.
- [28] Stevens, R.D., Robinson, A.J. and Goble, C.A., 2003. myGrid: personalised bioinformatics on the information grid. *Bioinformatics*, 19(suppl 1), pp.i302-i304.
- [29] Oinn, T., Li, P., Kell, D.B., Goble, C., Goderis, A., Greenwood, M., Hull, D., Stevens, R., Turi, D. and Zhao, J., 2007. Taverna/myGrid: aligning a workflow system with the life sciences community. In *Workflows for e-Science* (pp. 300-319). Springer-London.
- [30] Goble, C.A., Bhagat, J., Aleksejevs, S., Cruickshank, D., Michaelides, D., Newman, D., Borkum, M., Bedhofer, S., Roos, M., Li, P. and De Roure, D., 2010. myExperiment: a repository and social network for the sharing of bioinformatics workflows. *Nucleic acids research*, 38(suppl 2), pp.W677-W682.
- [31] Karwal, S., Lonie, A., Sinnott, R.O., Anderson, C., Challenges of Large-scale Biomedical Workflows on the Cloud – A Case Study on the Need for Reproducibility of Results, 28th IEEE International Conference on Computer Based Medical Systems, Sao Paulo, Brazil, June 2015.
- [32] Sarwar, M.S., Sinnott, R.O., Anderson, J.G., Alexander, M., Green, J., An e-Research Infrastructure for Language and Literature Research, UK e-Science All Hands Meeting, September 2010.
- [33] Voss, A. and Procter, R., 2009. Virtual research environments in scholarly work and communications. *Library Hi Tech*, 27(2), pp.174-190.