D5.1

Retrospective Data Retrieval and upload

Related Work Package	**WP5 – Development of the Master models**
Related Task	Task 5.1. - Upload and annotation of retrospective exams
Lead Beneficiary	FCHAMPALIMAUD
Contributing Beneficiaries	FORTH, Radboudumc, HULAFE, UNIPI, IPC, HACETTEPE, IDIBGI, JCC, NCI, GAONA St Savvas, RMH, QUIRONSALUD, FPO, CNR
Document version	**v.1.0**
Deliverable Type	Report
Distribution level	Public
Contractual Date of Delivery	31/5/2022
Actual Date of Delivery	31/10/2022

Authors	**Dr. Nikolaos Papanikolaou**
Contributors	Dr. Stelios Sfakianakis, Valia Kalokyri, Ioannis Karatzanis, Dr. Katerina Nikiforaki, Dr. Nikolaos Tachos, Dr. Haridimos Kondylakis, Prof. Kostas Marias
Reviewers	M. Tsiknakis

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement **n° 952159**

Version history

Version	Description	Date completed
0.1	Initial configuration of the document and draft TOC (CF)	11.05.2022
0.2	Section2,3 Contribution (CF, FORTH)	15.5.2022
0.3	Section 4 contribution (FORTH)	20.5.2022
0.4	Section 5 contribution (CF, FORTH)	24.5.2022
0.5	Section 6 contribution (FORTH	26.5.2022
0.6	Review and comments updates (CF)	29.5.2022
0.7	Section 7 contribution (CF, FORTH)	15.7.2022
0.8	Updates on section 7 (FORTH)	17.10.2022
0.9	preFinal version updates after internal review (CF)	27.10.2022
1.0	Final version (CF)	31.10.2022

Statement of Originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgment of previously published material and of the work of others has been made through appropriate citation, quotation, or both.

Disclaimer

Executive summary

This deliverable describes the process performed to retrieve Retrospective Data and upload them to the ProCAncer-I platform, as it is defined in the relevant task 5.1 of the DoA. In more detail, the task 5.1 states that Retrospective examinations that fulfill inclusion criteria to the study will be mined from the local PACS systems of each clinical partner including clinical information that should accompany each imaging study (i.e. PSA, Gleason score, and others).

The latter will be uploaded after local dual anonymization using black listing and white listing methodology to the cloud-based ProCAncer-I platform to be used for developing the Master models. All clinical information will be inserted into the platform through secure access by each clinical partner.

The present Deliverable reports on the development of the ProCAncer-I eCRF – Data Upload Tool that was developed to ease the data uploading process and to support ProCAncer-I clinical partners in the process of compiling the required imaging and clinical information and following the defined protocols for uploading data to the project's cloud repository. The tool integrates with the CTP anonymization tool and with the ProCAncer-I repository services such as authentication, DICOM upload API, and eCRF upload API, enabling the anonymization and upload of data using methods that comply with privacy and security requirements.

The Deliverable also reports on all supporting tools and processes established aimed at assisting the Data Providers as much as possible, to quickly resolve any potential issues that arise during the uploading procedure.

Table of Contents

List of Abbreviations

Abbreviation	Explanation
DICOM	Digital Imaging and Communications in Medicine
CRF	Case Report Form
mpMRI	Multi-parametric Magnetic Resonance Imaging
PCa	Prostate Cancer
DWI	Diffusion Weighted Image
DCE	Dynamic Contrast Enhanced
PI-RADS	Prostate Imaging Reporting & Data System
PSA	Prostate Specific Antigen
PHI	Personal Health Information

List of Tables

List of Figures

1 Introduction

This Deliverable reports on all activities that have taken place to a) guide the Clinical sites regarding how to best organize their internal processes regarding the collection and preparation of imaging and related clinical data, b) the policies regarding the anonymization of the imaging and clinical data, and c) the support provided to the clinical sites regarding training with the tools and technologies developed to support the for executing the uploading of their data into the ProstateNET.

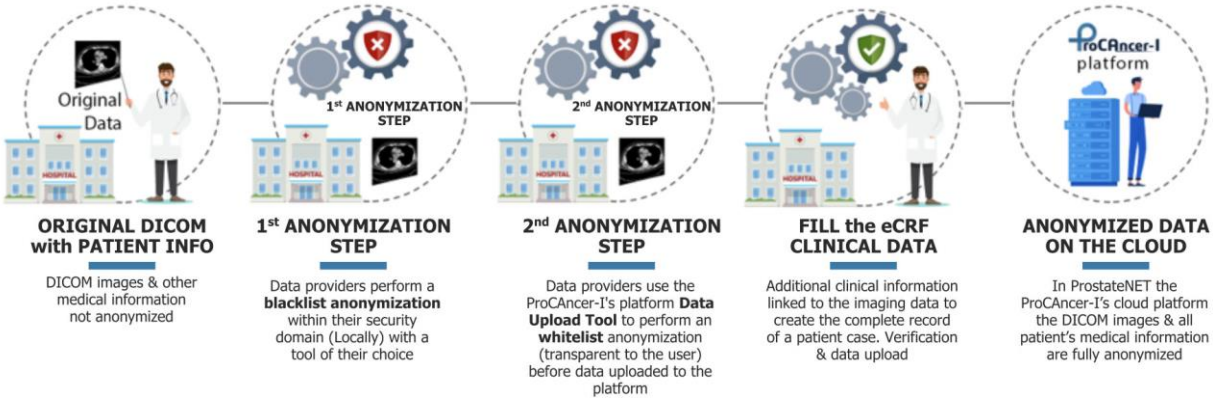


Figure 1: Overview of the Data Anonymization, Data Preparation, and Upload proceses ProCAncer-I

Considering the different levels of expertise among partners in handling and securely sharing data, as well as the number of expected retrospective cases to be uploaded, different needs regarding data processing became evident. Thus, the process was tailored to meet the needs of the clinical partners as expressed in the consortium meetings. Such changes regarded the easier integration of the upload and annotation process in the clinical workflow which in turn required modifications of the underlying structure and provided utilities.

1.1 Deliverable structure

The structure of the deliverable is as follows:

Section 1 is an introductory section that presents the purpose and the objective of this deliverable.

Section 2 presents the guidelines provided to the data providers (clinical partners of the ProCAncer-I consortium) for the data selection and their preparation before the uploading process starts.

Section 3 provides an overview of the anonymization process designed, developed, and utilized in ProCAncer-I in order to assure the privacy of the sensitive data included in the selected data.

Section 4 presents, in brief, the uploading tool (ProCAncer-I eCRF tool) which is developed to guide the uploading process, perform the anonymization, and executed as software at the

premises/IT infrastructure of the data providers. The presentation includes both versions of the tool aka the single mode and the batch mode.

Section 5 provides a description of the ProstateNet repository along with the necessary tools to be utilized along the activities of data uploading.

In Section 6 we included the supporting actions the ProCAncer-I consortium has established to support the needs of data providers.

Section 7 analyses the status of the retrospective data uploaded in the "staging area" of the ProstateNet platform along with the monitoring tools developed for the specific purposes. Finally, Section 8 concludes the Deliverable.

1.2 Reasons for the delayed submission of the Deliverable

The main reason for the delayed submission was the late initiation of the uploading process due to technical reasons related to the software used for uploading (eCRF tool) and the various technical problems that were different across the clinical partners, sometimes related to firewall rules, limited bandwidth, and others. For the latter reasons, there were several versions of the eCRF tool that were developed to overcome problems reported as feedback from the clinical partners.

The initial kick-off of the uploading process took place, as planned, in the 1st of Dec 2021, when the clinical coordinator shared instructions related to the local collection of data, and through several rounds of testing, eventually, the first case arrived in the platform in January 2021. Additional reasons for the delay in data collection and uploading were related to the wide variability across the different clinical sites related to examination protocols. In collaboration with the clinical coordinator, we defined by consensus technical requirements (MRI sequence parameters) that the data should be compliant with not to harm the model's performance. It was a fine balance to maintain diverse data simulating "real world data" and in the same time maintain high-quality standards of the ProstateNET repository.

However, the key reason for the delay in data uploading was the fact that the initial version of the anonymization and upload tool – based on initial functional requirement analysis – did support only one-case at a time uploading. Having experiences with initial data uploading and the time required to prepare and upload a case (images and accompanying clinical data) it became apparent that for partners having to upload large amounts of data, this was not a realistic scenario. The Consortium discussed the issue, and it was proposed that a new version of the upload tool was required, capable to support "batch uploading". The development of such functionality into the upload tool, and its alpha and beta testing, before delivering it for routine use, demanded significant additional effort and time.

In parallel, during our consortium meeting in May of 2022, a decision was made to relax several non-critical clinical variables, since the feedback from the clinical partners dictated such a measure due to a risk of a significant reduction in the number of eligible cases to be uploaded. Additional time was consumed to develop a version of the tool (eCRF) to support the reduced set

of clinical variables. As a result, the tested version of the upload tool, supporting both the reduced set of clinical variables as well as the batch uploading mode was made available in June 2022.

For all the previously stated reasons, the Consortium has decided to extend the uploading process till the end of October 2022, to give enough time to the clinical partners to upload all data identified. Currently – on the 26th of October 2022 - the platform hosts 7051 patients, 13.615 data points, and more than 3.7 million images that have been successfully uploaded, which will make feasible the finalization of the planned work in T5.3, i.e. the development of the master models based on the retrospective data. However, uploading of retrospective data will continue for as long as there are clinical centers that possess data that fulfill the inclusion criteria established. We estimate that small numbers of retrospective data will continue to be uploaded onto the platform for an additional period of two months.

2 Guidelines for Data Selection and preparation

Considering the variability among clinical sites regarding the imaging protocol used for MRI of the prostate, a critical step was to define the number of mandatory sequences that are required to characterize the uploaded data for a specific patient as complete concerning the quantity of the data. It is of note that this minimum set of MR requirements is common among all UCs. Along with the mandatory sequences, the clinical partners were encouraged to add some clinically relevant acquisitions that were available in the retrospective examination protocol.

After identifying the set of MRI examinations eligible to be uploaded to the ProstateNET repository, the clinical partners received guidance from the technical partners for the appropriate data preparation to follow a seamless workflow for the most efficient data upload. More specifically, the clinical site's personnel has to ensure access to a Workstation within the security domain of the clinical site with an internet connection where the data will temporarily reside to be prepared. A unique case folder should be created containing all the DICOM images for upload. If a certain case requires multiple studies (baseline and follow-up) for it to treated as been complete, the user is instructed to upload one single examination, preferably the one closest to the date of the performed biopsy in case of a negative MRI or the one acquired before treatment in case of a positive MRI.

DICOMData folder

The DICOMData folder should be created containing all the DICOM images for upload. The PatientID of those DICOM studies must be the same and should not be empty.

For each study consider the following criteria for series selection:

- Required series: T2W Axial, DWI Axial High b-value, ADC Map;
- Highly desirable series: DCE;
- Optional series: T2W (Coronal, Sagittal), T1W (Axial, Coronal, Sagittal).

- **Axial T2 (3D dataset):** The T2 weighted images are acquired by a sequence belonging to the fast spin echo sequence family, which is a robust method that is also able to alleviate image artifacts related to the presence of magnetic field inhomogeneity that are accountable for image intensity nonuniformity and distortions.

- **DWI / ADC (multiple 3D or 4D dataset / 3D parametric map):** The DWI images are acquired based on the EPI methodology and comprise of a low or zero diffusion contrast (low b value or zero b value contrast) and a high b value contrast that is the image where the diffusion contrast dominates the anatomy contrast. The b value parameter is a user defined parameter (confined by both the hardware and software capabilities of the scanner) that adjusts the degree of diffusion information present in the image. In detail, a b zero or low b value image is an image with no or minimal diffusion contrast

respectively and thus resembles very much an anatomical image, namely a T2. A high b value image is a b zero image multiplied by diffusion contrast which means high signal in areas of high cellularity and signal drop in areas with normal tissue architecture. Notwithstanding the low signal and spatial resolution, DWI images offer very important information as the degree of signal loss between the low/zero b value image and the high b value image is indicative of tissue cellularity and correlates thus with the degree of malignancy.

- **DCE (multiple 3D or a single 4D dataset)**: Dynamic contrast-enhanced series (DCE) comprise of a series of T1 weighted gradient echo series in order to capture the contrast agent extravasation over several minutes with low temporal resolution It has been considered to be a non-mandatory, although recommended series for the Prostate Net repository.

- **Other sequences:** T2 sagittal, T2 coronal, T1 axial.

Data curation: The EPI based DWI technique can scan the entire volume in a time-efficient while also including diffusion contrast but is a sequence very prone to artifacts, more frequently image distortion and chemical shift artifact. Most frequently, the co-registration of the high b value image to the anatomy is required, however, it suffers from a low signal to noise ratio, and very frequently it is challenging to use it directly to be co-registered to the anatomy. A workaround is to perform initially an inter-volume co-registration of the T2 weighted sequence to the b zero value series which is expected to succeed as they convey very similar contrast. As a next step to perform an intra-volume co-registration of the high b value sequence to the b zero value.

As T2 turbo (fast) spin echo sequences can be tailored to produce high spatial resolution representations of the anatomy without spatial distortions, they are chosen to serve as the image series of reference (static image) for any further intervolume co-registration steps. This implies that any volume from the 4D sequences (any b value volume from the DWI series or any dynamic volume from the DCE series) will be mapped onto the T2 weighted series. Moreover, the ADC parametric map will be also mapped onto the T2 weights series for distortion correction through the utilities offered at the curation stage.

The ability to perform intra-series motion correction, being the spatial alignment of 3D volumes within the same 4D series, will be only offered to 4D data stored in a single DICOM file. The reason for excluding the ability in the opposite case of a multitude of 3D volumes that comprise the 4D sequence is the mitigation of relative data misplacement.

Apart from the native images, one parametric map is required in the minimum number of sequences to be uploaded, the ADC map. The ADC map is expected to be calculated by the vendor's DWI specific software with the customization (i.e. noise level, appropriate DWI model) from the user.

Table 1: Step by step guidance for data uploading. Phrases in bold emphasize actions decided in the recent CM in Heraklion in May 2022.

Step	Action	Short description of the action
1	File preparation	DICOM series are inspected and saved in a folder at the same computer where the data upload tool is installed
2	Anonymization - Blacklisting anonymization	The DP reforms anonymization with a software tool of the site's preference
3	Clinical data	The DP ensures the availability of all necessary information depending on the UC.
4	Data upload- Whitelisting anonymization	The actor follows a 5-step process guided through the individual steps of the dedicated tool. White listing anonymization is performed within this process
5	Data Curation	The actor has the ability for intra-series and inter-series motion correction. **The co-registration among mandatory sequences is a required action (decided in CM 05/2022)**
6	Anatomy annotation	The user performs AI-aided anatomy registration. Available tools are used to correct the automatic result for the prostate gland, peripheral zone, and seminal vesicles.
7	Lesion annotation	The user provides lesion annotations for 5% of the uploaded data with the provided tool by the platform or 80% with a tool of their preference
8	Data forwarded to the ProstateNET	The user approves the set of necessary sequences (initial data and secondary series after curation and annotation) to be forwarded to the ProstateNET repository

3 Anonymization

As reported in D4.1, the ProCAncer-I's strategy regarding data handling

"The DICOM anonymization tool will be used to anonymize DICOM images at the clinical premises before they are pushed to the ProCAncer-I platform."

The DICOM anonymization tool implements a whitelist solution to cope with the personal health information that might be included in the DICOM tags. The whitelist has been defined by ProCAncer-I's partners involved in the anonymization task, and its purpose is filtering the DICOM tags keeping only the ones included in the whitelist to be kept (either with their original values, or modified, depending on each respective tag and based on a predefined set of rules) while removing or emptying the others. The DICOM anonymization tool is a standalone application that has no other automatic interaction with other tools/components, as the anonymization process is a manual procedure performed by the clinician. Besides tools for the anonymization of the DICOM data, there will also be available tools for the anonymization of the clinical data that will accompany the DICOM series to be uploaded.

Upon anonymization, the data upload tool is installed at the clinical sites to upload the data to the staging area of the platform (it will be explained in detail in the following sections). The sequence diagram of the aforementioned process is shown in Figure 2.

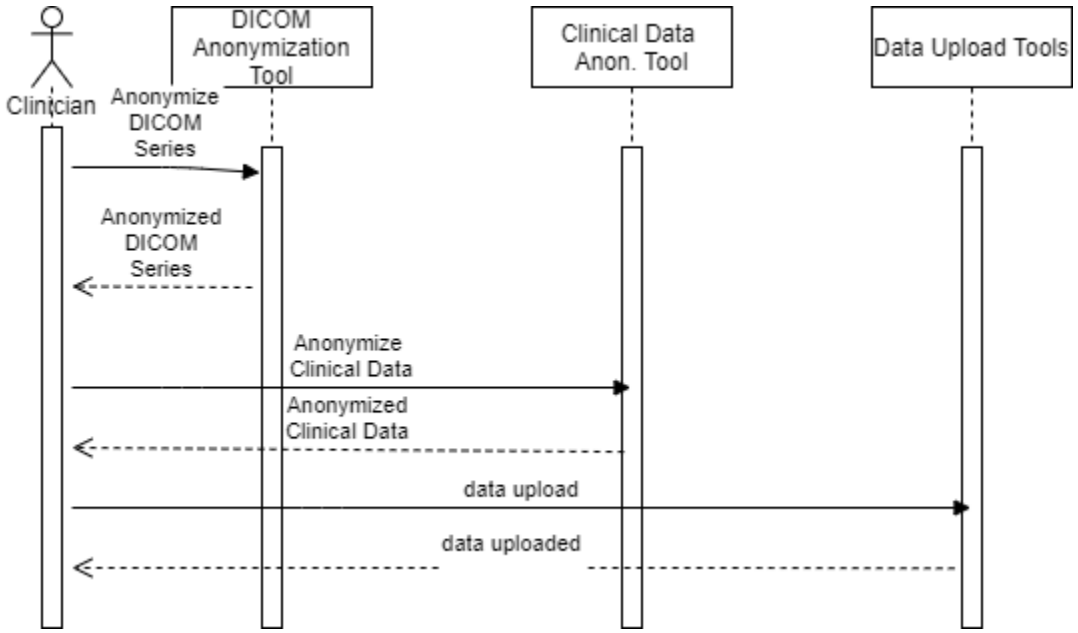


Figure 2: The sequence diagram for anonymization and uploading data to the platform.

3.2 ProCAncer – I Anonymization strategy

In order to maintain a high level of data security, the process of anonymization rather than pseudo-anonymization or deidentification was adopted. Anonymization is an irreversible processing operation that consists of using a set of techniques in such a way as to make it impossible, in practice, to identify the person by any reasonable means. The main difference between anonymization and the other two aforementioned techniques is that patient privacy has to be preserved even when other external sources are used, such as the clinical sites patient database or any other document held on the hospital premises. Attempts of patient identification that may entail illegal means, currently unavailable technology, or extremely high financial or time costs fall beyond the description of reasonable means.

The anonymization process is challenging when dealing with DICOM formatted data. The complexity lies in effectively performing an anonymization strategy while also preserving the value of the DICOM dataset as input for model development where the individual characteristics of each case have to be adequately described. The optimal compromise between data quality and safety was sought in extensive and multiple meetings with clinical and technical partners. Importantly, in the frame of a multi-centric data collection, the issues of DICOM header variability should also be considered concerning both security and future data deployment.

During the preparatory process, all participants in the dedicated online meetings were able to express their opinion based on their experience and daily practice as well as the active legislative framework of their country. During the discussions, a common excel sheet was shared among all partners, where FORTH anonymization team members listed the suggested list of PHI containing DICOM tags according to the DICOM Supplement on anonymization strategies. The most appropriate action for each DICOM tag was clearly stated to make sure that any decision taken would comply with the suggested guidelines and would contribute to the definition of an efficient and simple framework for this process that would be based on widely-accepted expertise.

The two main pillars of the expected outcome were simplicity on the one hand as not all ProCAncer-I clinical partners were familiar with the anonymization process and efficiency on the other hand as the security issue is rated very high in ProCAncer-I priorities. All discussions were held under the vigilant ear of the legal partner, who would address any aspect requiring clarification regarding the European legislation.

	cleaned tags	Property	Action Code	Action
2	(0008,0018)	SOP Instance UID	U	U - replace with a non-zero length UID that is internally consistent within a set of Instances
3	(0008,0020)	Study Date	Z	Z - replace with a zero length value, or a non-zero length value that may be a dummy value and consistent with the VR
4	(0008,0021)	Series Date	X/D	X/D - X unless D is required to maintain IOD conformance (Type 3 versus Type 1)
5	(0008,0022)	Acquisition Date	X/Z	X/Z - X unless Z is required to maintain IOD conformance (Type 3 versus Type 2)
6	(0008,0023)	Content Date	Z/D	Z/D - Z unless D is required to maintain IOD conformance (Type 2 versus Type 1)
7	(0008,0030)	Study Time	Z	Z - replace with a zero length value, or a non-zero length value that may be a dummy value and consistent with the VR
8	(0008,0031)	Series Time	X/D	X/D - X unless D is required to maintain IOD conformance (Type 3 versus Type 1)
9	(0008,0032)	Acquisition Time	X/Z	X/Z - X unless Z is required to maintain IOD conformance (Type 3 versus Type 2)
10	(0008,0033)	Content Time	Z/D	Z/D - Z unless D is required to maintain IOD conformance (Type 2 versus Type 1)
11	(0008,0050)	Accession Number	Z	Z - replace with a zero length value, or a non-zero length value that may be a dummy value and consistent with the VR
12	(0008,0090)	Referring Physician's Name	X	X - remove
13	(0008,1010)	Station Name	X/Z/U	X/Z/U* - X unless Z or replacement of contained instance UIDs (U) is required to maintain IOD conformance (Type 3 versus Type 2 versus Type 1 sequences containing UI
14	(0008,1030)	Study Description	X	X - remove
15	(0008,103E)	Series Description	X	X - remove
16	(0008,1050)	Performing Physician's Name	X	X - remove
17	(0008,1140)	Referred image sequence	X/Z/U	X/Z/U* - X unless Z or replacement of contained instance UIDs (U) is required to maintain IOD conformance (Type 3 versus Type 2 versus Type 1 sequences containing UI
18	(0008,1155)	Referenced SOP Instance UID	U	U - replace with a non-zero length UID that is internally consistent within a set of Instances
19	(0008,1150)	Referenced SOP Class UID	U	U - replace with a non-zero length UID that is internally consistent within a set of Instances
20	(0008,1155)	Referenced SOP Instance UID	U	U - replace with a non-zero length UID that is internally consistent within a set of Instances
21	(0008,1150)	Referenced SOP Class UID	U	U - replace with a non-zero length UID that is internally consistent within a set of Instances
22	(0008,1155)	Referenced SOP Instance UID	U	U - replace with a non-zero length UID that is internally consistent within a set of Instances

Figure 3: Indicative snapshot of the proposed actions for selected PHI containing tags according to the DICOM Committee

The Consortium partners, under the leadership and coordination of FORTH, examined several freely available software solutions that could potentially be adapted by clinical partners and simulated anonymization, either with their default configurations or with the optimal set of parameters in cases where this was an option.

An evaluation of the result concerning the suggested optimal practice was presented in the consortium meetings and the participants were urged to experiment with their datasets and report on the result and the ability to effortlessly use each software if they had not already had a preference for their own. If an established anonymization had already been established on their site, the partners were encouraged to check whether the result complied with the specific guidelines presented by DICOM Committee. The testing was performed using phantom data that could be shared among the consortium without running any danger of exposing patient PHI. During the discussions held in the virtual meetings, a demonstration of the DICOM Committee suggestions as presented in DICOM Committee Supplement 142, regarding the proposed actions for each DICOM tag, i.e. deleting, emptying, keeping, or modifying. Moreover, the DICOM tag type was clearly stated, too ensure that all partners are aware of the specific DICOM tags that are necessary to be preserved to maintain the DICOM structure of the processed data.

Thirty-seven tags required modification, and consequently, the evaluation of the selected tools was based on the number of PHI-containing tags accurately modified by the default of the most appropriately configured version of each tool, as shown in Fig. 3. Concerning other DICOM tags that were not included in DICOM Committee guidelines, an open discussion was held among

partners where the participants were encouraged to share their opinions concerning the most appropriate modification of each tag. The main challenge of these discussions was the large variability of scanners and software versions, each producing a different version of the anonymized DICOM header list. The part requiring attention was the private tag list, containing specific attributes in a non-well-defined fashion from each vendor, creating a point of high heterogeneity, even for datasets handled by the same tool and configuration. This heterogeneity was examined by gathering indicative results from each clinical partner that was willing to share them among the consortium.

3.3 The double-level anonymization approach

During the recursive meetings among technical and clinical partners, the two options of a more aggressive and a subtler anonymization approach were discussed. Blacklisting is the term used to define the former, while blacklisting is used to define the latter. The advantages and disadvantages of each approach were presented, as shown in Figure 4. In more detail, retaining more alleviates the danger of losing valuable information, i.e. attributes that will be necessary for any post-processing or modeling step. However, removing more creates an easier-to-handle and control DICOM tag list that is stripped from any tag even very remotely related to an attribute of patient identity. The outcome of the extensive collaboration of clinical, technical, and legal partners was to customize the process to the needs of the ProCAncer-I consortium and establish a double-level approach. The first stage is to perform on clinical site premises a basic anonymization approach, retaining more DICOM attributes, and then process again the data at a later stage, during the upload of clinical and imaging information with a command-line tool with common configurations for all partners to perform the second level of anonymization process. At this final stage, only necessary tags are kept, and the DICOM headers become homogeneous among partners while still the information is locally stored. So, fully anonymized data with a common configuration enter the platform.

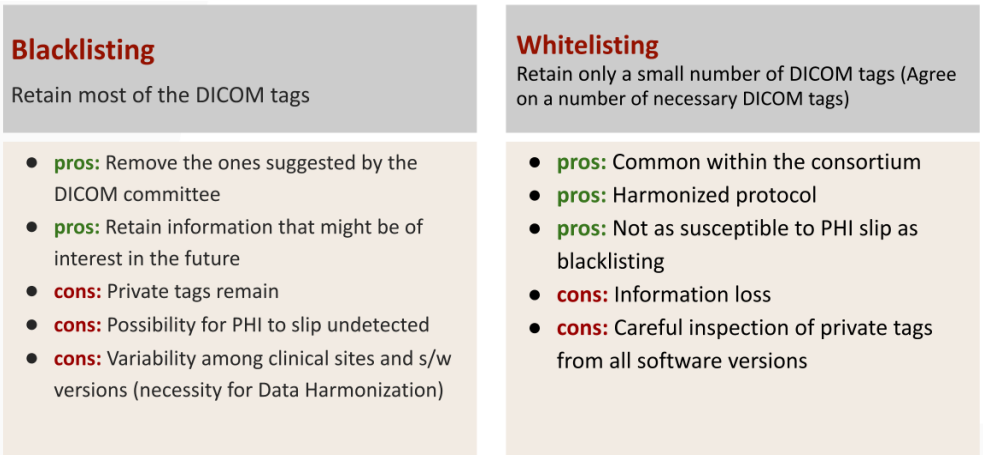


Figure 4: Snapshot from the presentations in the process of defining the most appropriate anonymization strategy.

Since the response to the demand for feedback data regarding the result of the anonymization process was poor, the partners were encouraged to report in case of facing difficulties or compromised results. Bilateral talks were held with partners requiring help to complete this task. Considering the inability to capture the complete frame of variability among the DICOM tag lists from different workflows as they result from each clinical site, a new solution was sought as the DICOM tag homogeneity issue remained a desired attribute for the metadata repository. During discussions with the technical partners to address the problem of heterogeneity that was expected from the use of different anonymization software, the approach of adding a second anonymization layer was adopted. This second level of anonymization intends to further modify the already anonymized DICOM tags from the clinical site, and impose a next step of removing, erasing, or modifying the already anonymized data by a third party (technical partner) during the data upload step. A script file containing the configuration rules for the whitelisting is developed by technical partners and is part of the upload software, activated when the user combines clinical and imaging information, as described in the e-CRF section of the present deliverable.

It is important to note that both first-level and second-level anonymization take place locally, inside the site's clinical premises. A table containing indicative rules and their result in the anonymized files in their final version is shown in Table 2. This approach, which comprises a solution with unique characteristics of flexibility to promote a secure process in a user-friendly manner, is referred to as a double-level anonymization process in the project's documentation.

#	Tool	Total Nema Tags	Nema Tags found	Nema Tags removed	Tags with values	Empty Tags
1	ProSurgical3D	37	15	22	7	8
2	Strtvan	37	15	22	7	8
3	DicomPyler	37	38	-1	36	2
4	ModiCAS	37	26	11	24	2
5	DicomCleaner	37	32	5	28	4
6	DICAT	37	38	-1	37	1
7	RSNA	37	21	16	20	1
8	DICOM Browser					
9	grassroots					
10	MIViewnew	37	38	-1	38	0

Figure 5: Software tool evaluation method, based on the number of successfully modified PHI containing DICOM tags.

This approach was also considered more appropriate to address the issue of private DICOM tags, where useful information may reside, however vendors are not publishing the complete description of their content and consequently their presence in the final DICOM header may affect the success of the anonymization level. Partners were encouraged to preserve the private DICOM tags, and at the second anonymization level only selected private DICOM tags will be kept

under the condition of having a known and widely accepted value, with interest to the further use of the data.

The second stage of anonymization addresses the problem of DICOM tag list heterogeneity among partners using different scanner and software versions, as well as exploits the ability to access useful information from the private tag section. This action of homogenization is applied through the command line utility of RSNA Anonymized, after specific configuration performed by FORTH and B3D to search and retrieve useful private DICOM tags and then horizontally select the least number of useful and common among all partners DICOM tags that will be kept for the next stages of the project.

Table 2: Indicative actions and results produced at the second layer of anonymization

DICOM Tag	Content	Action	Indicative Result
0010, 0010	PatientName	param(PCa-)@hashptid(@SITEID,PatientID)	[PCa-175581055197336176830214458282127464338
0008, 1090	ManufacturerModelName	keep	Philips Medical Systems
0010, 1010	Patient's Age	">@round(this,5)	75
0008, 0080	InstitutionName	remove	-
0012, 0063	De-identification Method	ProCAncer-I Whitelist,ProCAncer-I Whitelist	ProCAncer-I Whitelist,ProCAncer-I Whitelist
0008, 0020	StudyDate	hashdate(this,PatientID)	20110219

Issues that were resolved with the help of legal partners and the STC concerned the option of defining age groups instead of specific ages and weight groups/intervals instead of integer inputs. This process requires defining the best compromise between efficient identification and adequate data descriptors for model developers. Population outliers that could be identified by a combination of tags not containing PHI information had to be considered as a possible danger.

To summarize, the partners agreed to adopt both a 'whitelisting' and a blacklisting anonymization strategy. As a first step, only PHI containing data were modified or removed inside each clinical partner's premises by a software tool of their own. At the second stage, during the upload process, an automated procedure running a second process of DICOM tag selection is performed, z process that retains the least number of necessary and common DICOM tags ensuring homogeneity among the DICOM tag list among all partners.

4 Data Upload Tool

The ProCAncer-I eCRF – Data Upload Tool aims to support ProCAncer-I clinical partners in the process of compiling the required imaging and clinical information and following the defined protocols for uploading data to the project's cloud repository. The tool integrates with the CTP anonymization tool and with the ProCAncer-I repository services such as authentication, DICOM upload API, and eCRF upload API, enabling the anonymization and upload of data using methods that comply with privacy and security requirements.

The tool is designed to follow a simple 5 steps workflow from the selection of the folder containing the DICOM files, to the anonymization, edition of the clinical information, and upload of DICOM and clinical information data.

The ProCAncer-I eCRF – Data Upload Tool is an application for the Windows® operating system that should be installed on a computer with Windows 10 or above and with an internet connection. In addition, ProCAncer-I eCRF – Data upload tool provides two modes of operation:

- Single Mode Operation
- Batch Mode Operation

The eCRF tool has been reported in the WP4-related deliverables. In the current document for complementarity purposes, a description of both modes is provided throughout the main document for "form 1" while all the functionalities of the tool are described in the ANNEX section. Since the first launch of the tool, the technical partners have refined and solved minor bugs of the tool based on the feedback from the data providers.

4.2 Single Mode Operation

This section explains how to use the eCRF Data Upload Tool (as a single-instance windows application) to select and anonymize DICOM studies, select the appropriate form and insert clinical data, and upload images and clinical data to the ProCAncer-I cloud repository.

4.2.1 eCRF – Data Upload Tool main window

The tool is designed to follow a simple workflow in 5 steps to compile and update the data to the ProCAncer-I cloud repositories:

1. Add the DICOM files
2. Anonymize the DICOM files
3. Edit the clinical information by selecting the appropriate Use Case
4. Upload the DICOM files
5. Upload the clinical information.

Note: The eCRF tool is specifically built for each clinical site to include authentication credentials assigned to the different ProCAncer-I clinical organizations.

4.2.2 Handling of clinical images

The workflow starts with "1: DICOM Add" for the selection of a case folder that contains the DICOM studies for the patient.

Step 1: DICOM Add

After clicking on the button "1: DICOM Add", a new window will open for selecting the appropriate folder. The selected folder or its subfolders must contain DICOM files otherwise it will not be accepted.

All relevant DICOM studies for the patient should be inside the selected folder.

Once a new case folder is added, it will be shown in the list of cases appearing below. Besides the DICOM folder, the list shows the Use Case and the status of Anonymization, DICOM Upload, eCRF completeness, and eCRF Upload. These fields will be updated as the user completes the different steps in the workflow.

To perform any of the following steps the user must select the case in the list.

Step 2: DICOM Anonymize

The second step "2: DICOM anonymize" runs the CTP anonymization tool with the pre-defined anonymization script that completely anonymizes the DICOM files as defined in the project. By pressing the DICOM anonymize button the anonymization process will start, showing the status of the process in the status area at the bottom of the page. At the end of the process, a message will display the result of the anonymization process.

The original files created at the preparatory step are not deleted or modified. The application creates a new folder for each case where the anonymized DICOM files, case information, and clinical information are saved.

If the DICOM files had already been anonymized, a new message will appear stating that the DICOM folder has already been anonymized.

Step 3: Form Edit

The third step "3: Form Edit" allows the user to select the appropriate form for the current case and insert the required clinical information. The edition of the clinical information can be made more than once until its completion. The eCRF Edit button will display a new window, the Use Case Form, with different tabs to select the appropriate Form. This manual contains a specific chapter describing the required information for each Use Case. After saving the information on

the Use Case Form, this window will close and the case will be updated on the list with the selected Form and the completeness of the information.

Once the clinical information is complete, the case can be uploaded to the staging area of the ProCAncer-I cloud repository.

Step 4: DICOM upload

The fourth step "4: DICOM upload" allows the user to upload the anonymized DICOM files of the selected case. It will only upload files if the case is already anonymized and has complete clinical information. This operation requires an internet connection to connect to the ProCAncer-I imaging APIs. By pressing the DICOM upload button the status bar will show the progress of the upload process. At the end of the process, the status bar will show the result of the upload process.

Step 5: Form Upload

The final step in the workflow is "5: Form Upload" which allows uploading the clinical information of the selected case to the ProCAncer-I cloud repository. It will only upload the information if the case was already anonymized and the clinical information is complete. This operation requires an internet connection to connect with the ProCAncer-I Clinical Data and Metadata APIs. By pressing the eCRF upload button the status bar will show the progress of the upload process. At the end of the process, the status bar will show the result of the upload process.

DICOM Remove

The user may remove a case from the list of cases by pressing the DICOM Remove button. This function will clear all the information created on the eCRF tool on the local machine related to that case. The original DICOM folder will remain. If the case was already uploaded the information on the cloud-related to the case will not be removed.

4.2.3 Clinical Data and Use Case Forms

The Use Case Form enables the user to select the most appropriate form regarding the Use Case or combination of Use Cases for the selected case.

For your reference, the ProCAncer-I Use Cases mapped in the prostate cancer management pipeline are shown in a pictorial form in Figure 6.

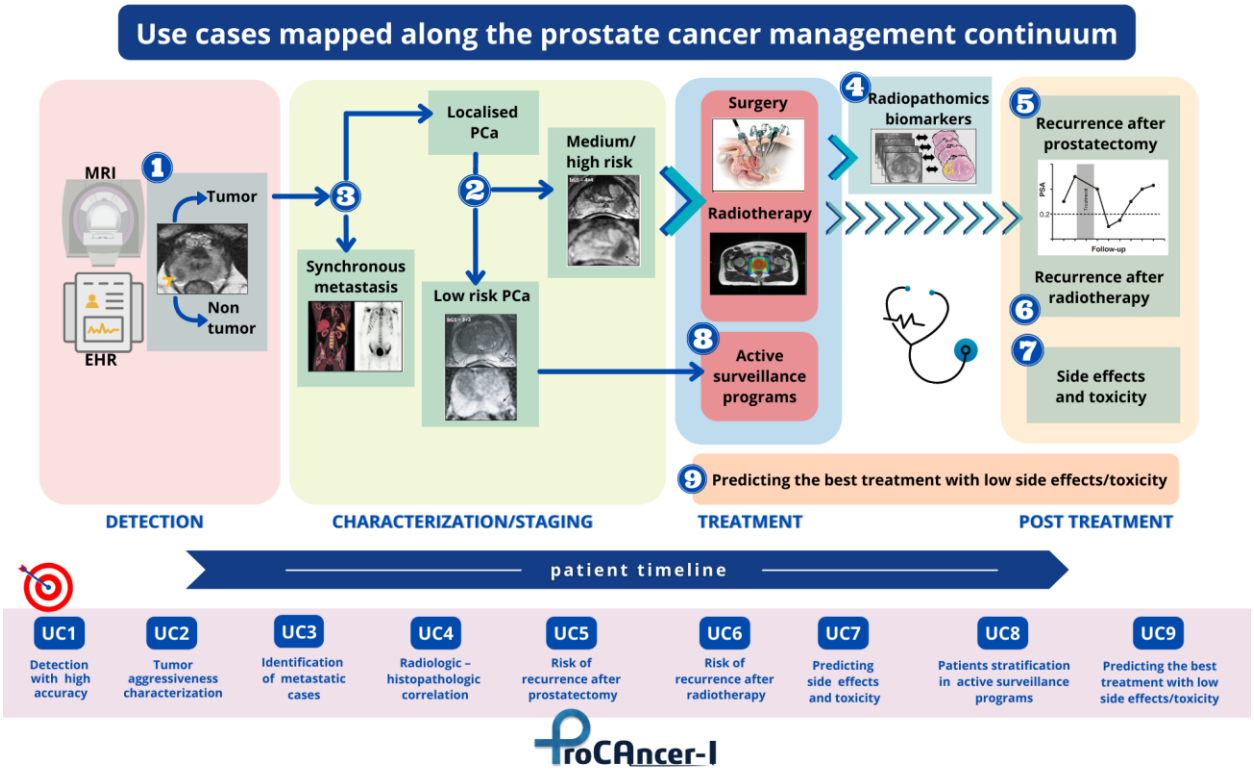


Figure 6: The ProCAncer-I Use Cases mapped into the prostate cancer management pipeline

To support the users (clinical experts) a decision flowchart has been prepared to assist them to understand which form(s) of the tool should be filled. The flowchart is depicted in the following figure.

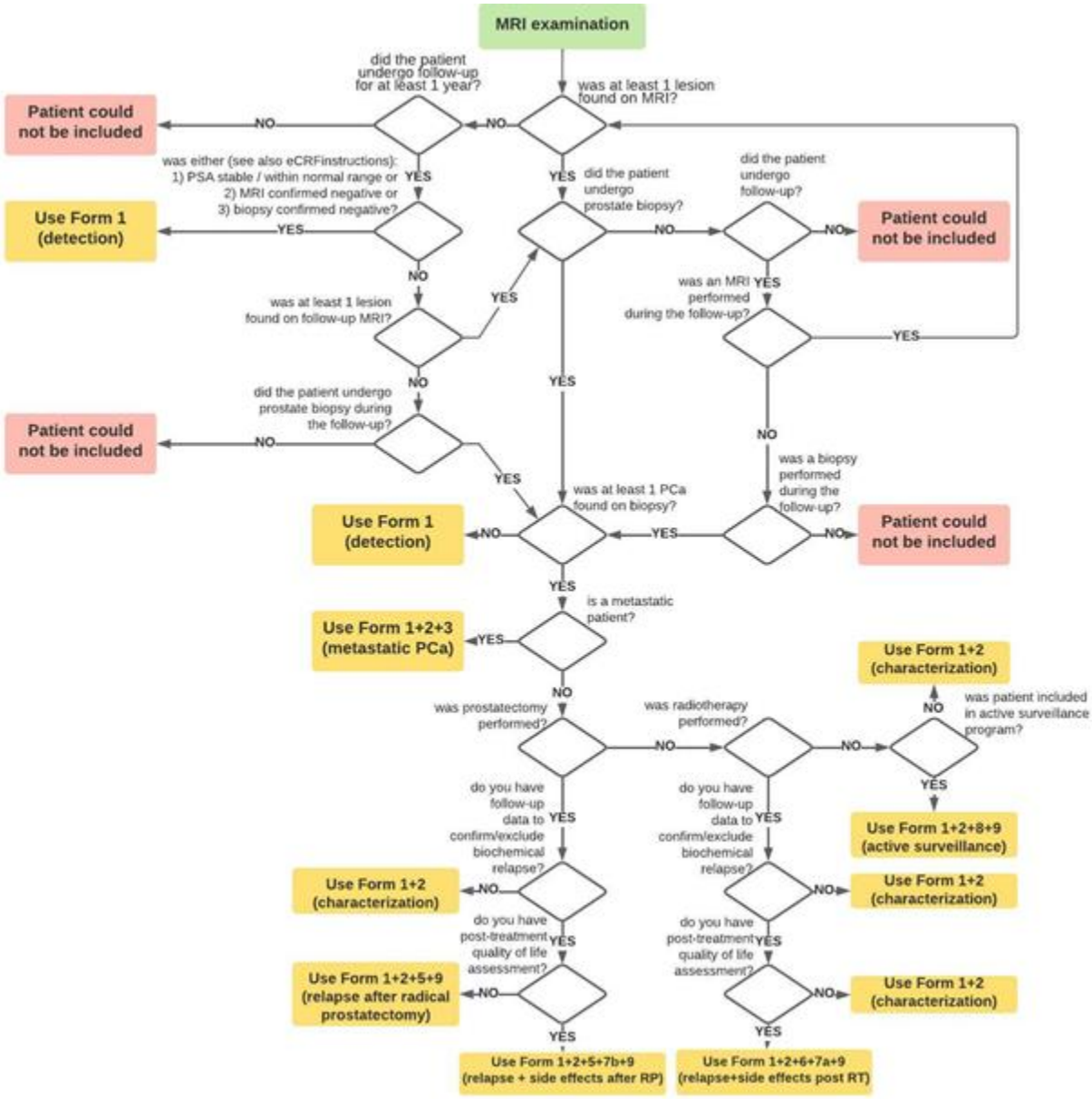


Figure 7: Upload process decision flowchart

Each tab in the form corresponds to one Use Case or combination of Use Cases as shown in the table below. 7 forms will be used to collect information for the 8 use cases. The user should verify in which use case(s) the current patient best fits and select the adequate form to include the clinical data.

We need to point out that UC4 is an exception to the rest since it will be performed only at partner FC with the data from 100 patients prospectively acquired starting on M36. UC4 is not related to the retrospective data at all.

Table 3: List of the 7 forms that will be used to collect information for the 8 use cases.

Input Forms / Use Cases	UC1	UC2	UC3	UC5	UC6	UC7	UC8	UC9
Form 1 (only negative PCa cases)	x							
Form 1 + 2 (positive PCa with Gleason score)	x	x						
Form 1 + 2 + 3 (positive + Gleason score + metastasis)	x	x	x					
Form 1 + 2 + 5 + 9 (positive + GS + prostatectomy)	x	x		x				x
Form 1 + 2 + 6 + 7a + 9 (positive + GS + radio + side effects)	x	x			x	x		x
Form 1 + 2 + 5 + 7b + 9 (positive + GS + prostatectomy + side effects)	x	x		x		x		x
Form 1 + 2 + 8 + 9 (positive + GS + follow-up active surveillance)	x	x					x	x

In each form, there is a button indicating the fields that are optional to facilitate the user with an efficient workflow.

By pressing the save button, the form will save the information on the currently selected form and return it to the main application window. The respective row on the table will be updated showing the status of the completeness of the clinical information. After completing the form, it is possible to upload the DICOM and clinical data to the ProCAncer-I repository staging area.

4.1.3.1 Form 1

Collection of patients without confirmed PCa by pathology (e.g. positive MRI but negative biopsy) or men with no PCa findings on MRI and confirmed negative at follow-up (at least 1 year).

This form only collects negative cases to Use Case 1.

This form has two information sections: Clinical and Follow-up.

In the Clinical section, the user should complete the following information:

- Age at baseline: insert an integer number;
- DRE: insert the result of the Digital Rectal Examination – Positive, Negative or Not Assessed – optional field;
- Biopsy before MRI: true (checked) or false (unchecked);

- Previous adenomectomy: true (checked) or false (unchecked);
- Insert one or more PSA tests and select the baseline in the list with the following fields:
    - PSA Total (ng/ml): decimal value > 0, 'dot' is used as a decimal separator;
    - PSA Free (ng/ml): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - PSA Ratio (%): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - Date: date of the PSA test;
    - Baseline: select which of the PSA tests in the list is the baseline – there must be one.

There is a delete button at the bottom of the table to remove the selected PSA. To select a PSA, click on any field of the row containing the appropriate PSA.

In the Follow-up section, the user should provide the following information considering at least 1 year of follow-up for negative MRI or negative biopsy:

- Follow-up PSA confirmed in the normal range: true (checked) or false (unchecked);
- MRI confirmed negative: true (checked) or false (unchecked);
- Biopsy confirmed negative: true (checked) or false (unchecked).

4.3 Batch Mode Operation

The eCRF batch mode is intended to support partners contributing a large volume of data to upload, who have already a database with the clinical information and also have the means to transform the clinical data into the defined JSON formats.

The batch mode provides a standard way for partners to implement their mechanism to export existing DICOM data and metadata to the eCRF format. Partners are required to extract the information in their current format (e.g., spreadsheet and databases) using their preferred scripting/programming language and generate a JSON along with a copy of all the DICOM data following a fixed directory structure.

A Windows-based binary program will read the corresponding files from the directories, verify their correctness and upload them without the need for the workflow of the single mode operation eCRF version. DICOM data are anonymized using the CTP anonymization tool from RSNA and following the whitelist script designed and provided by the ProCAncer-I project.

The batch mode provides a method to select a root folder containing all cases/patients to be processed. Each patient folder must contain a JSON file with the eCRF information and a folder named “DICOMData” where all the DICOM folders/files of that particular patient must be copied (see Figure 8 below).

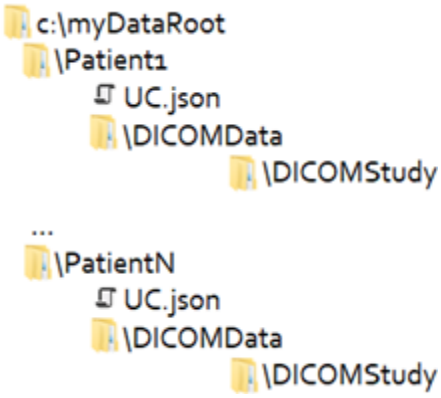


Figure 8: Root folder containing all cases/patients to be processed and the folder DICOMData

The content of each JSON file corresponds to one of the seven (7) available forms defined on the ProCAncer-I project (i.e., UC1, U1+2, UC1+2+3, UC1+2+5+7b+9, UC1+2+5+9, UC1+2+6+7a+9, UC1+2+8+9) referring to a single-use. The exact format of the cases is described in the following sections along with an accompanying set of JSON examples.

4.2.1 DICOMData folder

The DICOMData folder should be created containing all the DICOM images for upload. The PatientID of those DICOM studies must be the same and should not be empty.

For each study consider the following criteria for series selection:

- Required series: T2W Axial, DWI Axial High b-value, ADC Map;
- Highly desirable series: DCE;
- Optional series: T2W (Coronal, Sagittal), T1W (Axial, Coronal, Sagittal).

4.2.2 Formats of the JSON Files

The formats of the JSON files are a subset of the JSON files defined for the eCRF data upload tool, the Clinical Data, and Metadata APIs, containing only the clinical data. DICOM data required by the Metadata API will be automatically retrieved from the DICOM files by the eCRF tool.

4.2.2.1 Form 1

The following table describes the fields of the JSON file structure specifically for the cases belonging to “Form 1”.

Table 4 JSON fields description for the Form 1 cases

Field	Type	Optional	Description
age	string	no	Integer number converted to string

Field	Type	Optional	Description
dre	string	yes	Digital Rectal Examination. One possible value from: "Positive", "Negative", "Not Assessed"
biopsyBeforeMRI	bool	no	A biopsy was performed before MRI. true, false
previousAdenomectomy	bool	no	An Adenomectomy was performed before MRI. true, false
psas	List<PSA>	no	List of PSA objects. There must be one object with the 'baseline' field set to 'true' (see PSA tab)
fupPSAConfirmedNormal	bool	no	Follow-up PSA confirmed normal. true, false
fupMRIConfirmedNegative	bool	no	Follow-up MRI confirmed negative. true, false
fupBiopsyConfirmedNegative	bool	no	Follow-up biopsy confirmed negative. true, false
prospective	bool	no	Prospective or Retrospective data
useCaseType	string	no	Use case identifier. Mandatory value: "UC1"

Example:


```
{
 "age": "33",
 "dre": "Positive",
 "biopsyBeforeMRI": true,
 "previousAdenomectomy": false,
 "psas": [
 {
 "total": 7.2,
 "free": 0,
 "ratio": 0,
```

```
 "date": "2021-12-21T00:00:00Z",
 "baseline": true
 },
 {
 "total": 7.3,
 "free": 0,
 "ratio": 0,
 "date": "2022-01-04T00:00:00Z",
 "baseline": false
 }
 ],
 "fupPSAConfirmedNormal": false,
 "fupMRIConfirmedNegative": false,
 "fupBiopsyConfirmedNegative": false,
 "prospective": true,
 "useCaseType": "UC1"
}
```


PSA Object

Specifically for the PSA which is denoted as list of objects the following table describes the various fields.

Table 5 Description of the fields for the PSA object

Field	Type	Optional	Description
total	float	no	PSA Total (ng/ml), 'dot' is used as decimal separator
free	float	yes	PSA Free (ng/ml), 'dot' is used as decimal separator
ratio	float	yes	PSA Ratio (%), 'dot' is used as decimal separator
date	string/DateTime	no	PSA Date (e.g., "2021-12-21T00:00:00Z")
baseline	bool	no	Baseline (true, false). In a list of PSAs there should be 1 defined as the baseline.

5 The ProstateNET Image Repository

The ProCAncer-I Image Repository contains a staging area for each clinical partner to enable the verification, curation and annotation of cases before submitting them to the ProstateNET final repository area. Only the registered users from each clinical partner can have access to the staging area allocated to that partner.

5.2 Authentication and Authorization

5.2.1 User Accounts

All users of the platform need to create accounts to use any of the tools or parts of the platform. The ProCAncer-I platform has been linked with the ELIXIR Authentication Authorization Infrastructure (AAI) for single sign-on. This infrastructure provides a federation of identity providers of research institutions and companies around Europe and allows users to use the same credentials (e.g., their *home* university/organization username and password) to login to services linked with ELIXIR. The ProCAncer-I users can register or login at the ProCAncer-I AAI page at https://aai.procancer-i.eu/ as described in the following paragraphs (Accessed by the ProstateNet URL: https://www.prostatenet.eu).

Users need to visit https://aai.procancer-i.eu/login and they will be welcomed as shown in the following picture.

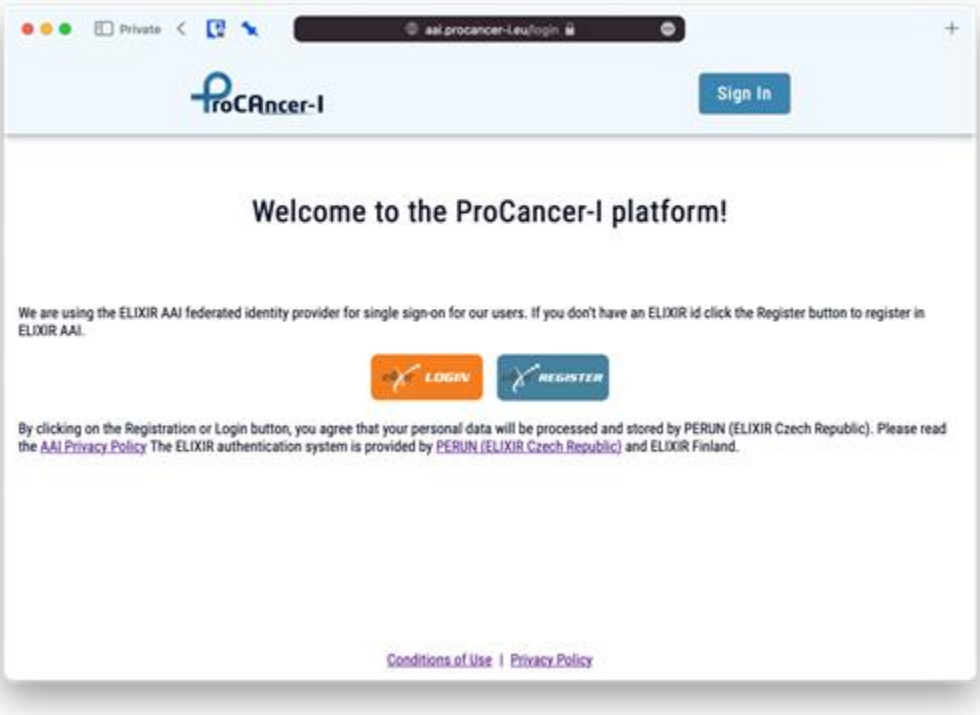


Figure 9. Screenshot of the Login welcome page of ProCAncer-I platform

On this screen, there are two buttons: **Login** and **Register**. If the user doesn't have an ELIXIR id, or if it's the first time the user is visiting ProCAncer-I, she/he should use the Register button to register in ELIXIR AAI and subsequently create a ProCAncer-I account. Choosing the Login button allows the user to enter the platform, but this is possible only after the user registration and the verification of the registration.

> *As mentioned in the screenshot above, by clicking any of the two buttons the user agrees with the processing and storage of personal data by PERUN (ELIXIR Czech Republic) which provides the ELIXIR authentication system.*

After pressing the Register button, the ELIXIR AAI web page is loaded and the next step is to choose the identity provider to use for the user authentication:

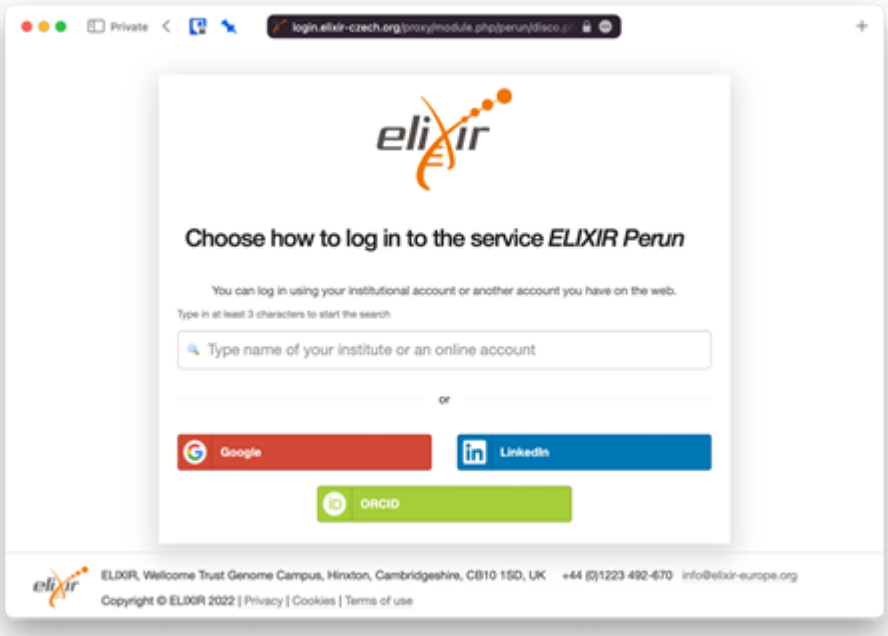


Figure 10. Screenshot of the ELIXIR AAI web page

If the user has an affiliation with the University or research institute in Europe, they can search using free text for their home institution as presented below or choose to authenticate through some social network like Google, LinkedIn, or ORCid.

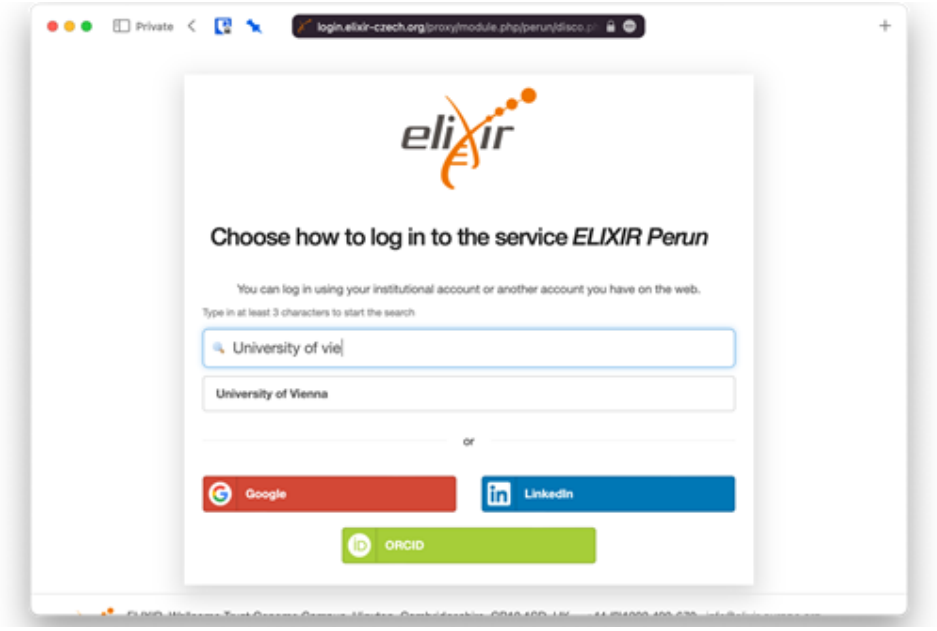


Figure 11. Screenshot of the ELIXIR AAI web page (Institution search)

The users will then be "transferred" to the Identity Provider chosen, where they can provide their credentials (typically their username and password). After successful authentication, the ELIXIR AAI website is loaded again and the user needs to consent to the sharing of the personal data shown with the ProCAncer-I platform:

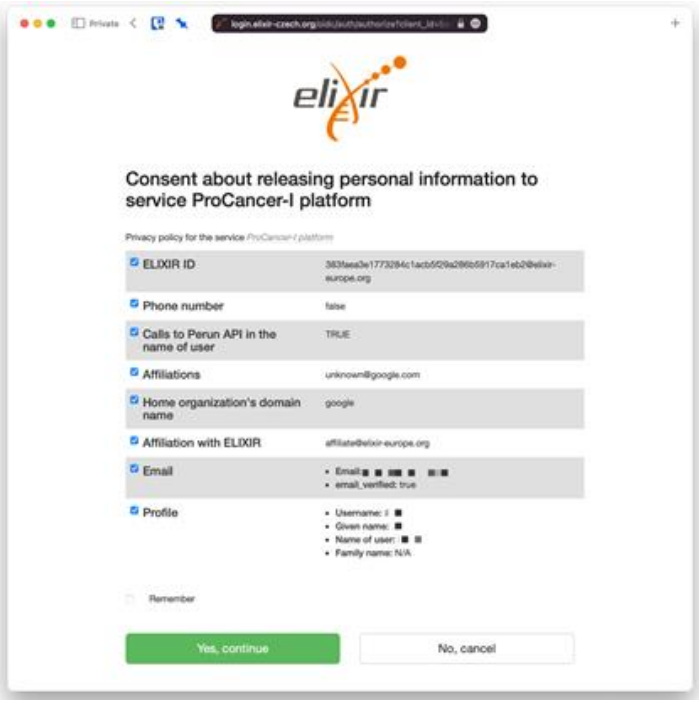


Figure 12. Screenshot of the ELIXIR AAI with the user personal stored information.

In the web page shown above there's also a "Remember" checkbox that, when checked, the users will not be asked again about their consent.

After this process, the users will automatically be forwarded to the ProCAncer-I AAI web page that they started with. At this stage, although the user's account has been created, *the user cannot yet use any tool of the platform: there's a manual verification process of the account created*, which involves the responsible security manager (currently a member of the FORTH team) contacting the user and validating his/her access request. In case the registration is not associated with an official e-mail account of the clinical partner, the user should notify FORTH by providing the e-mail and the associated organization of the created account, so it can be validated.

After the ProCAncer-I AAI team validates the new user, access to several services of the ProCAncer-I platform is enabled.

5.2.2 ProCAncer-I imaging repository login

Access to the ProCAncer-I imaging repository is available Through the main page of ProstateNet as depicted in the following figure.

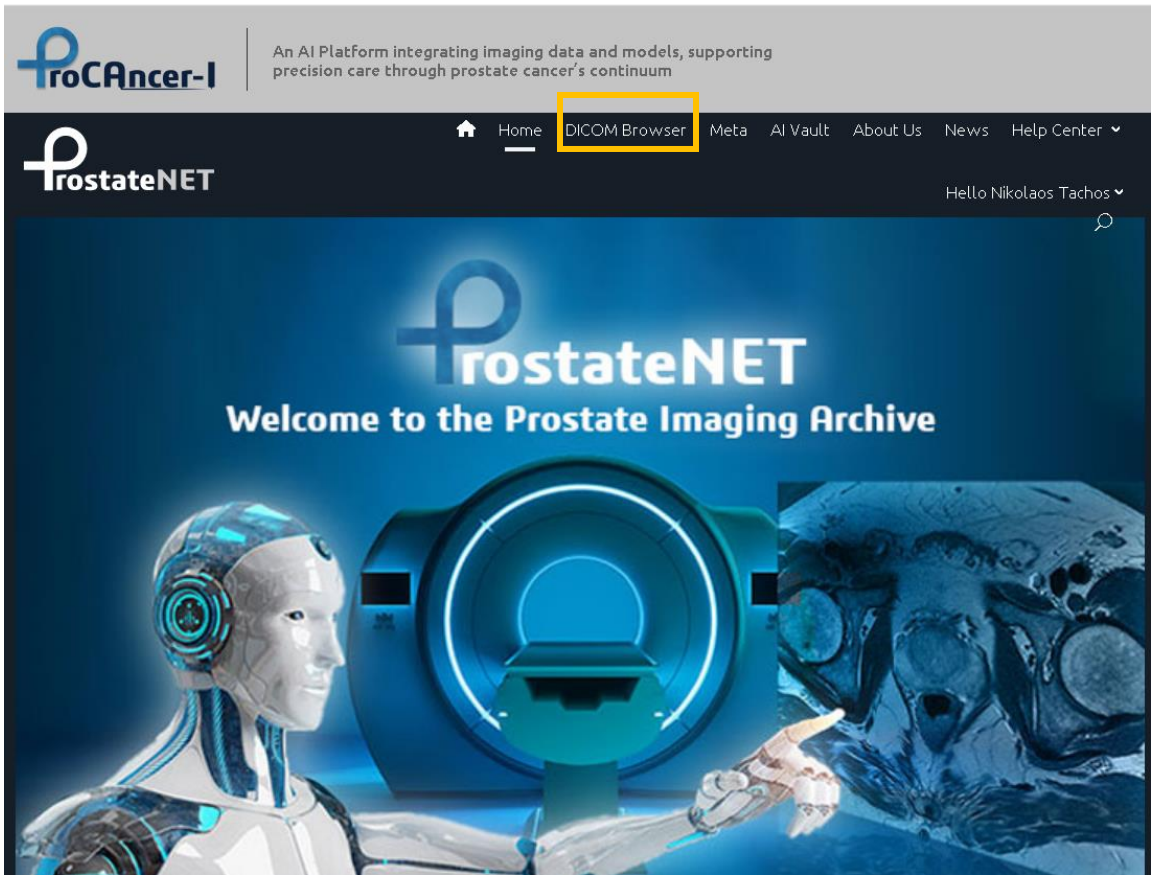


Figure 13. Screenshot of the ProstateNet landing page – Access to imaging repository

5.3 ProstateNET's Staging Area and Topology

The staging area allows the clinical partners to verify the completeness of their studies, perform co-registration and motion correction through the curation tools, and create segmentation. Once these tasks have been completed, the clinical partner can send the study to the final ProstateNET repository.

To send studies to the final ProstateNET repository, click the checkboxes of the respective studies and click the button “Move To” on the top of the list.

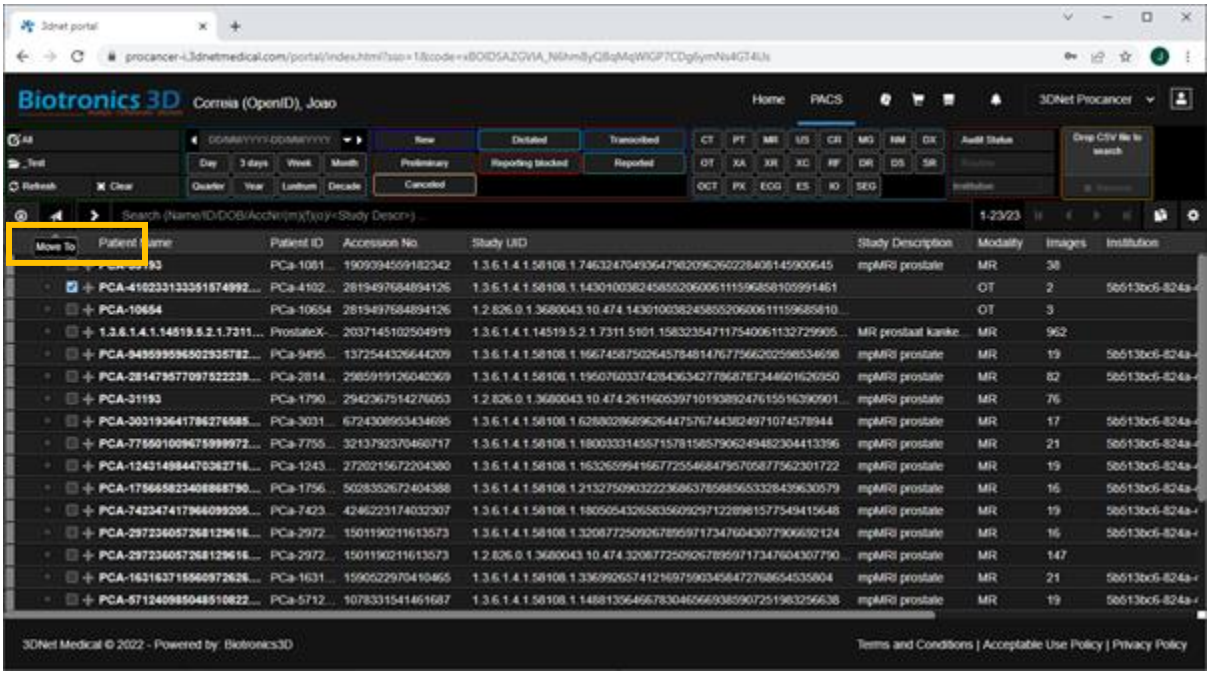


Figure 14. Screenshot of the Procancer-I imaging repository – Move operation from staging area to ProstateNet final repository area.

A new window will show the list of selected studies and a drop-down list to choose the destination folder. Select the folder “@ProstateNET” and click the “Move” button to complete the operation.

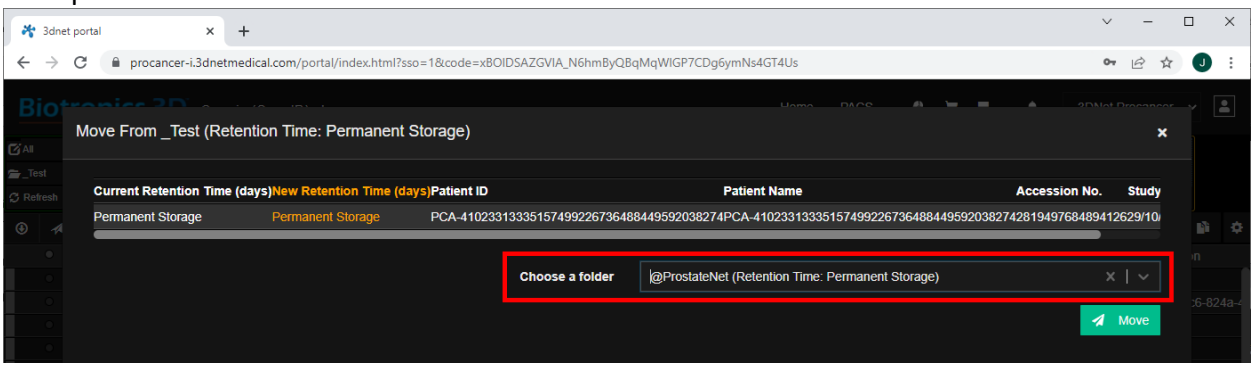


Figure 15. Screenshot of the Procancer-I imaging repository – Destination folder selection.

After moving the studies from the Staging Area to the ProstateNET repository these studies can be found in the folder @ProstateNET and become available to the entire ProCAncer-I community.

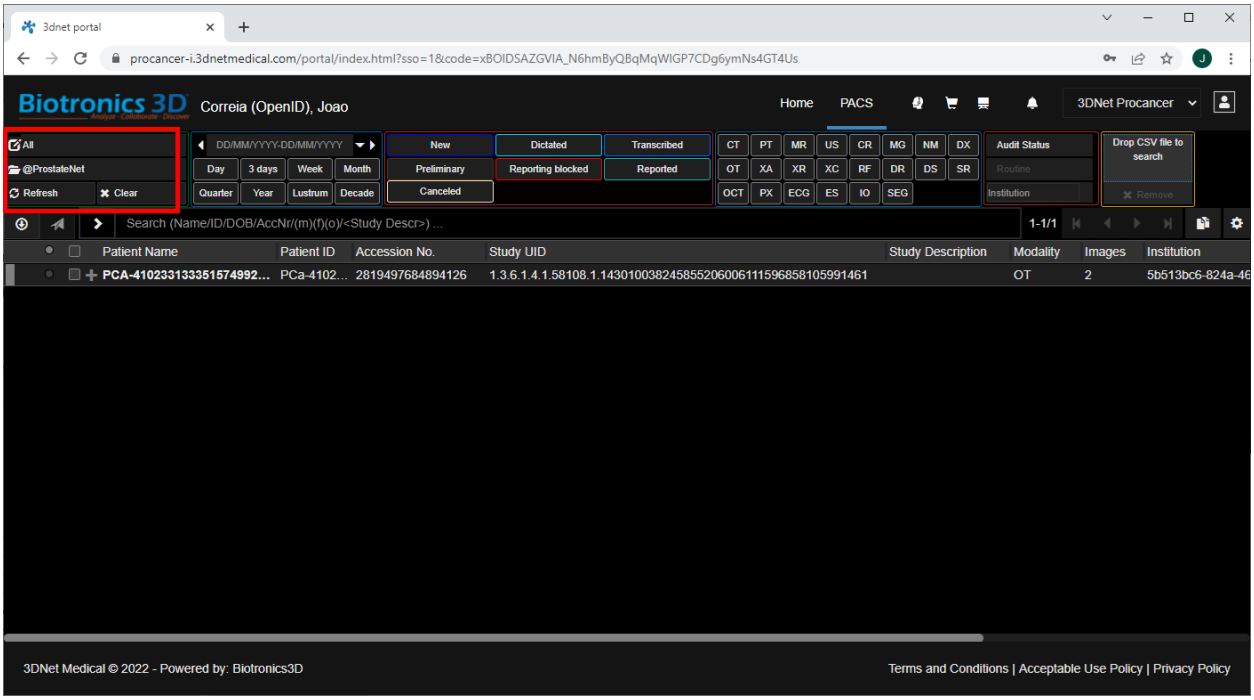


Figure 16. Screenshot of the ProstateNet final imaging repostory area.

5.4 Curation Tools

As described in the relevant WP4 deliverable the ProCAncer-I platform offers advance curation tools for motion correction and imaging co-registration.

To perform motion correction and co-registration click on the **"A" button** in the Study Viewer to launch Advantis Curation Tools. It will open a new browser tab displaying a list of series and available actions to perform.

As shown in Figure 6 below, the Curation Tool initially lists the series found in the selected study alongside various metadata, such as the image's plane of acquisition, shape, and zooms. The *STATIC* column denotes the static (3D) series to be used in the image co-registration phase, i.e. the image over which a second (moving) image will be displayed. A default static series is pre-selected if it is (a) 3D, (b) T2w, (c) axially acquired, and (d) not the result of a curation function. The "eye" icon to the left of each row can be used to view the corresponding image in a pop-up viewer. The colour-coded buttons to the right start the curation process: motion-correction, co-registration, and approval or rejection of the results. By approving the result of a curation function (motion-correction or co-registration), it is permanently stored in the Cloud Staging Area.

	NAME	ACQUISITION PLANE	SHAPE	ZOOMS	STATIC	
👁	3-Plane Loc	coronal	[512, 512, 9]	[0.8984000086784363, 0.8984000086784363, 20]	☐	
👁	SF SAG T2 3mm	sagittal	[512, 512, 24]	[0.4805000126361847, 0.4805000126361847, 3]	☐	
👁	Ax DWI RTr b1000	axial	[256, 256, 36, 2]	[1.5625, 1.5625, 6]		Motion-correct Coregister
👁	P Ax LAVA	axial	[512, 512, 40]	[0.8202999830245972, 0.8202999830245972, 5]	☐	
👁	COR T2	coronal	[512, 512, 16]	[0.4688000829696655, 0.4688000829696655, 3]	☐	
👁	SF AX T2 3mm	axial	[512, 512, 22]	[0.4688000829696655, 0.4688000829696655, 3]	☑	
👁	Ax DWI b1000	axial	[256, 256, 22, 2]	[0.9375, 0.9375, 3]		Motion-correct Coregister
👁	P Ax LAVA +C	axial	[512, 512, 40]	[0.8202999830245972, 0.8202999830245972, 5]	☐	

Figure 17: List of the series found in the selected study as displayed in the Curation Tool

The colour-coding indicates the state of the curation flow per series. A greyed-out button, e.g. co-register, in the figure above, cannot be clicked, unless the previous state is green. A green button indicates a complete step, which may be re-executed. A yellow button indicates a pending step. The colour-coding is also dependent on the STATIC column. For instance, if a series has been fully curated with Series A selected as static and then Series B is selected as static, the buttons' colors will change accordingly to indicate steps that need to be re-executed given the new static series.

Motion-correction application

The motion-correction application, which is instantiated by clicking the Motion-correct button, performs *inter-volume motion-correction of a DWI or DCE series*.

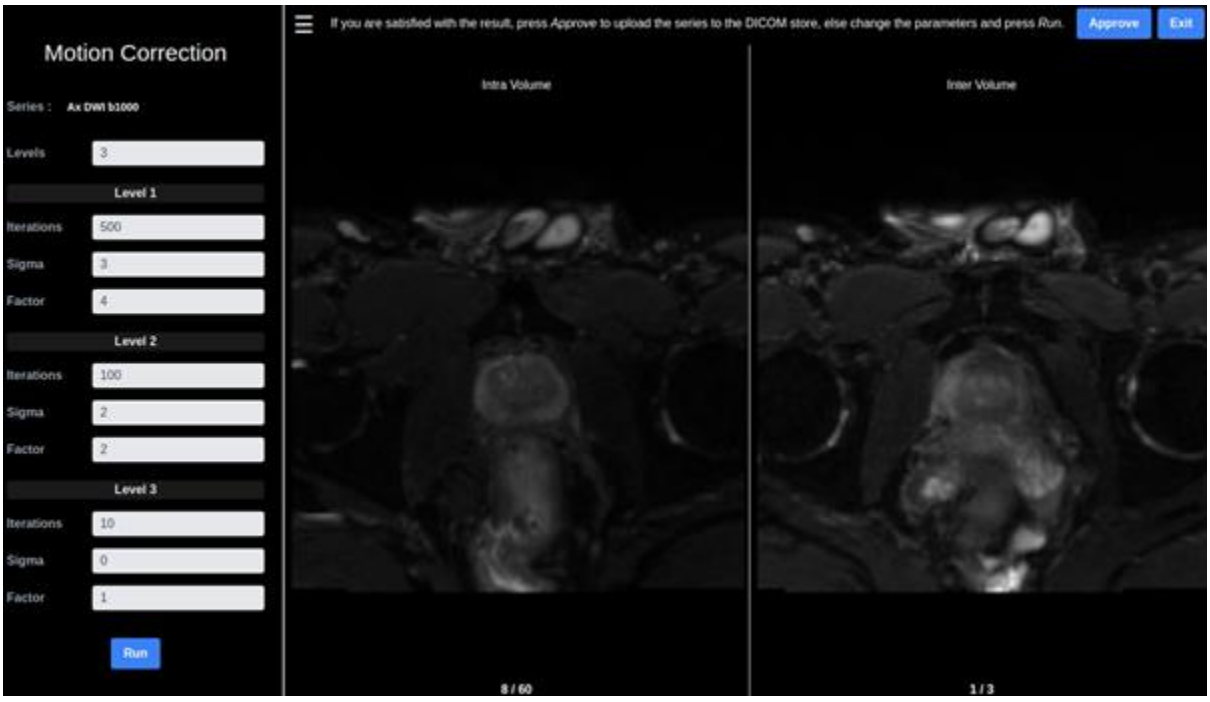


Figure 18: The motion-corrected series can be concurrently reviewed for intra- and inter-volume motion in two side-by-side viewers.

Details of the tool UI are presented in the previous figure. The UI provides the user the ability to change the parameters of the proposed image motion-correction workflow to output the desired result. Since the tool has been presented in the WP4-related deliverable for complementarity purposes in the ANNEX section there exists a more detailed description.

Co-registration application

The co-registration application co-registers the motion-corrected series to a T2w image. The moving (motion-corrected) and static (T2) series are also color-coded in green and red, respectively, as illustrated in the next figure.

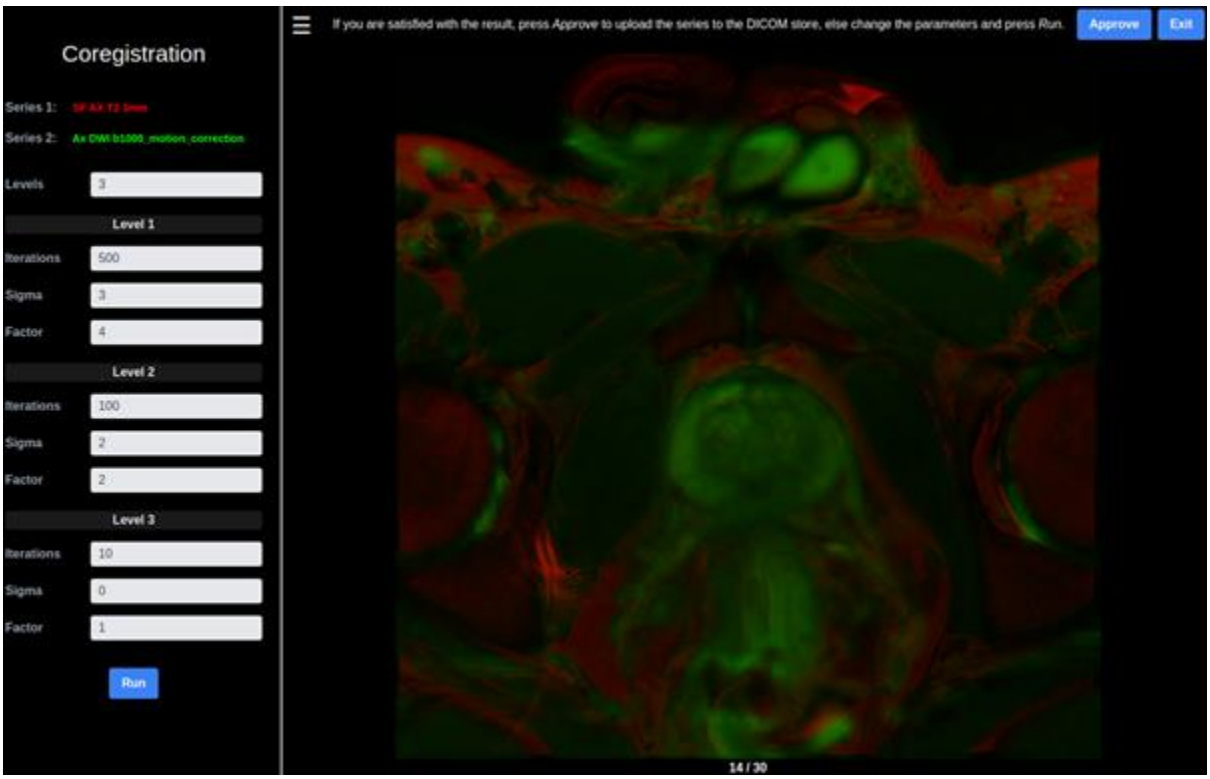


Figure 19: The moving (motion-corrected) and static (T2) series as depicted in the T2w image, color-coded in green and red, respectively.

In this case, the 1st volume of the previously motion-corrected series has been laid over the T2w series and the result for slice 14/30 is displayed. Again, if the result is satisfactory, the *Approve* button at the top-right of the screen accepts the result by sending it to the Cloud Staging Area and directs the user back to the beginning to select a new series to curate.

In either curation phase, if the result is unsatisfactory, the registration hyper-parameters on the left-hand side may be refined to re-run the curation function by clicking the *Run* button.

Details of the tool UI are presented in the previous figure. The UI provides the user the ability to change the parameters of the proposed image motion-correction workflow to output the desired result. Since the tool has been presented in the WP4-related deliverable for complementarity purposes in the ANNEX section there exists a more detailed description.

5.5 Annotation Tools

To perform segmentation, click on the **"Q" button** in the Study Viewer to launch Annotation Tools developed by QUIBIM. It will open a new browser tab displaying the images and the tools for annotation.

Figure 20: Annotation environment.

Table 6 shows the main buttons found in the annotation environment. The tool provides two modes of operation:

- Manual image segmentation
- Automatic image segmentation (based on an AI based DL model developed by QUIBIM)

Table 6: Icons present in the annotation environment and their main functionality.

Icon	Description
	Add notes to the study (not used for the moment)
	Browse through the slices with the mouse wheel.
	Activate/deactivate zoom and pan.

Icon	Description
	Center the image by resetting zoom and pan.
	Adjust the image contrast
	Set the initial windowing
	Select a predefined colormap
	Show/hide the information overlaid on the viewer area
	Show/hide the DICOM header tags
	Show MPR view
	Open a new view to load an additional series
	Save changes (overwrites previous DICOM Seg file)
	Save changes creating a new DICOM Seg file
	Download the DICOM Seg file locally
	Undo changes
	Redo changes
	Remove all ROIs
	Change label
	Activate brush tool
	Activate eraser tool
	Launch automatic prostate segmentation

Details about the functionalities of the tool are reported in a dedicated manual (IFU-instructions for use) which is provided to the users and available in the ProstateNet. For complementarity purposes part of the abovementioned is incorporated in the ANNEX section of the current deliverable.

Moreover, the tool offers the capability for a Multiplanar Reconstruction (MPR) view and opens in parallel secondary series as depicted in the following figure

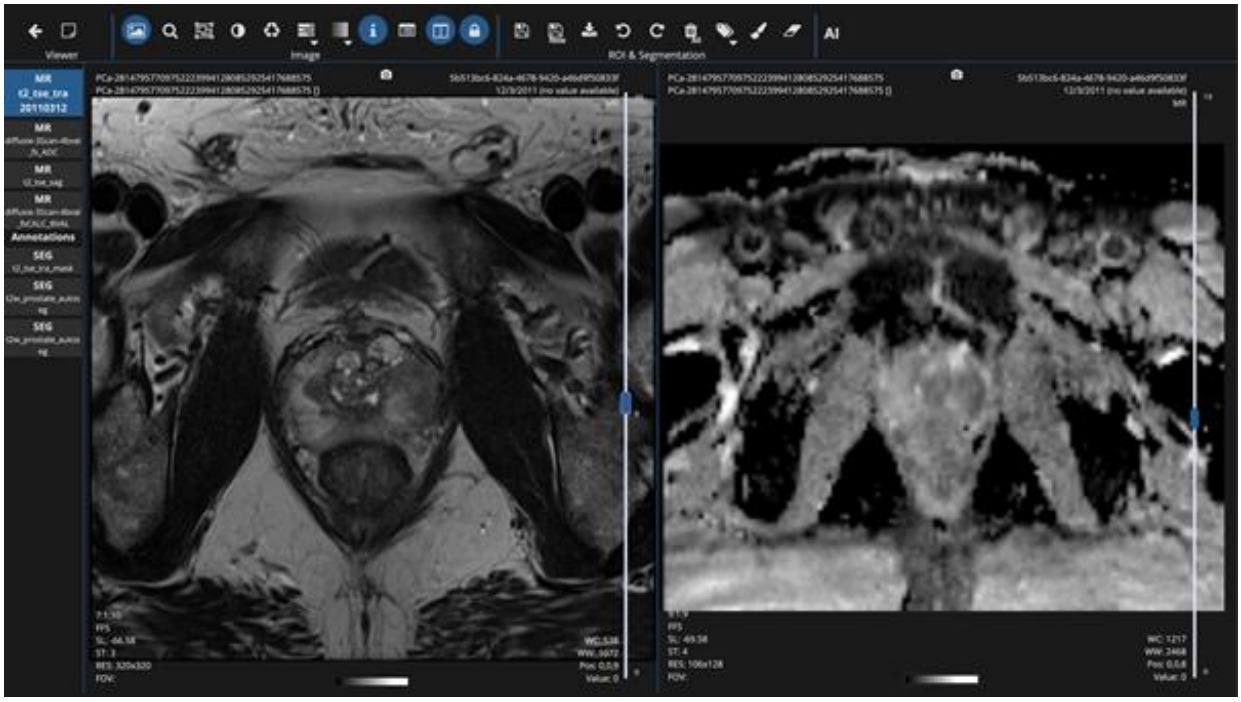


Figure 21. Screenshot from the annotation tool integrated in the ProstateNet.

From the tool, the user can select the label of the region to delineate. By default, these belong to the regions to delineate when doing the prostate gland segmentation:

- TZ+CZ: Transition Zone and Central Zone.
- PZ: Peripheral Zone
- SV: Seminal Vesicles

And for the lesions segmentation by default, three different labels are initialized. However, new labels can be created of new lesions that must be segmented by clicking the button “New Label”.

Execute automatic segmentation tool

When annotating the prostate gland, an automatic segmentation tool can be used as a pre-segmentation to ease the annotation workflow. It requires a T2w MRI sequence in the axial plane. To execute the analysis, click on the AI button. When the segmentation finishes with no errors, the following message will be shown (Figure 22).

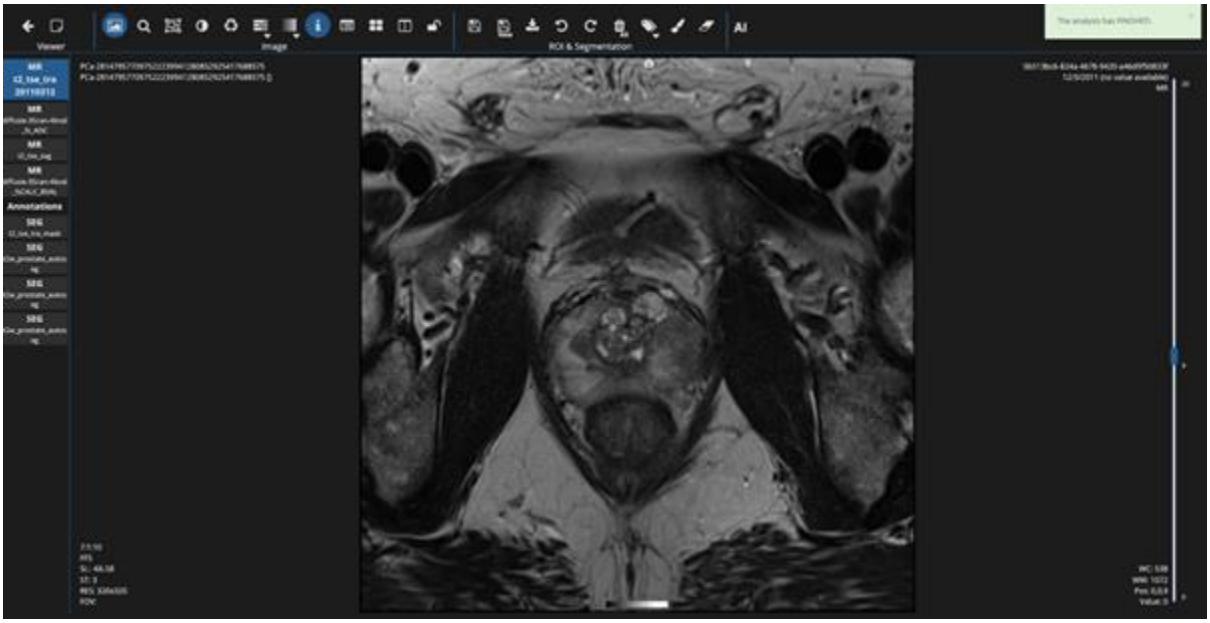


Figure 22: Prostate automatic segmentation finished successfully.

Once the segmentation finishes, a new segmentation series will appear in the bar on the left named “t2w_prostate_autoseg”. By clicking on it, it will be loaded over the T2w series (Figure 23). As described also in the ANNEX the user can edit the output segmentation. The tool provides a plethora of tools to edit or create new segmentation masks from MRI DICOM studies and save them on the ProstateNet DICOM Browser.

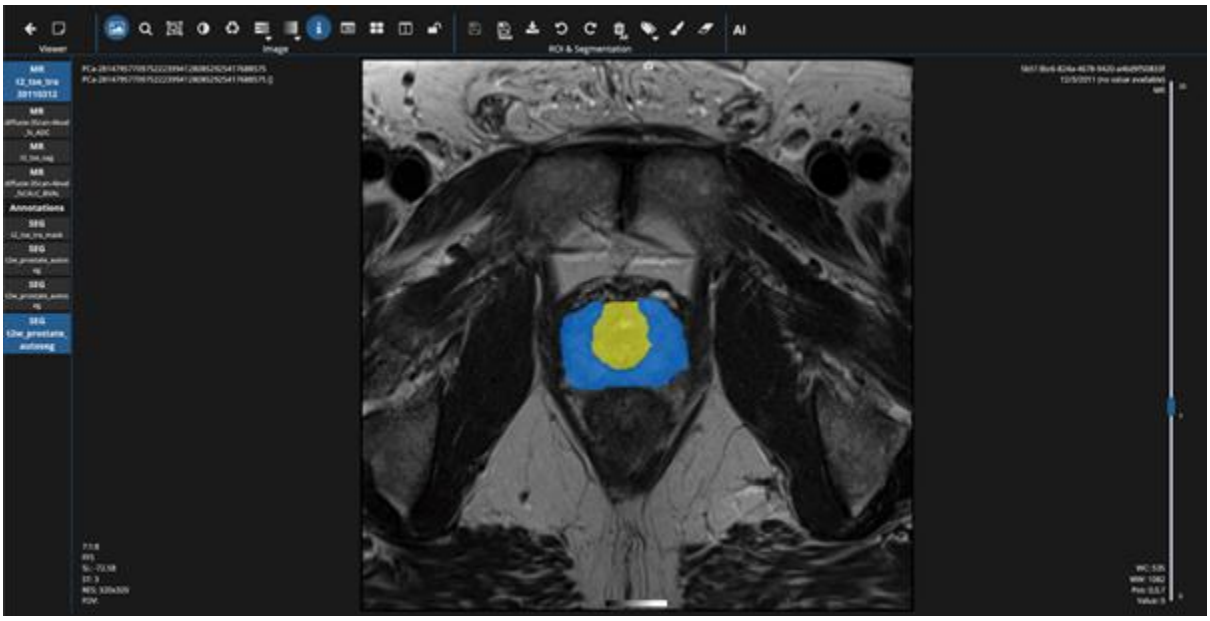


Figure 23: Automatic prostate segmentation overlaid over the T2w series.

5.4 Metadata Catalogue

The ProCAncer-I metadata catalogue is accessible from the main dashboard of the ProstateNet.. The homepage of the application is shown in Figure 24.

Figure 24: Homepage of the ProCAncer-I metadata catalogue.

5.4.1 Using the "Explore data" menu

When visiting the ProCAncer-I metadata catalogue, after having successfully signed into the ProstateNet platform, the user can navigate and explore the following data:

1. Clinical and imaging metadata per patient
2. Imaging and clinical metadata per series
3. Segmentation related metadata
4. Curation related metadata (coregistration/motion correction)

On the top-left part of the screen resides the "Explore Data" menu item (Figure 25), where the user can view, filter, and search the metadata that is currently available in the metadata repository. The above information is formatted in a tabular way, so any authorized user can "slice-and-dice" the data (by using the filtering/searching capabilities of the web interface) and download them into a single .csv file for future use (e.g., for AI model development). In addition, users can navigate the data in the OMOP-CDM data format by using the "Navigator" menu of the metadata catalogue.

Detailed description of how a user can exploit the capabilities of the ProCAncer-I metadata catalogue is provided in the ANNEX section (derived from the compiled manual for the end-users). In the main document only the scenario of "Use Cases" is presented.

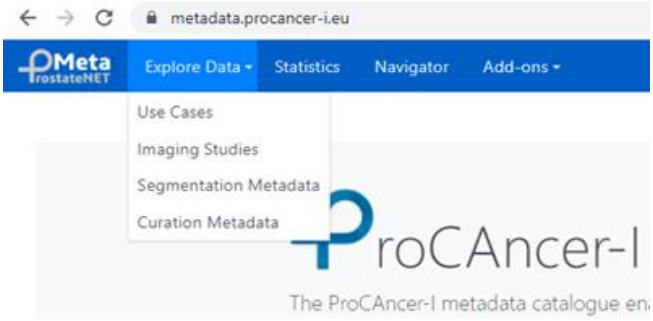


Figure 25: Explore data menu.

5.4.2 "Use Cases" exploration

By selecting the "Explore data" -> "Use Cases" menu item at the top left area of the metadata catalogue user interface, the user can explore the use case clinical data, as filled through the eCRF, along with the imaging metadata (Figure 26). In addition, apart from this information user can explore some metadata concerning the upload process itself, such as the institution uploading the current content, the time of the uploading, the method that the data was uploaded (single-mode or batch-mode) if it was validated by someone, and finally if the data refers to prospective or retrospective information. This information is under the attribute "Dataset Item" as depicted in the following figure.

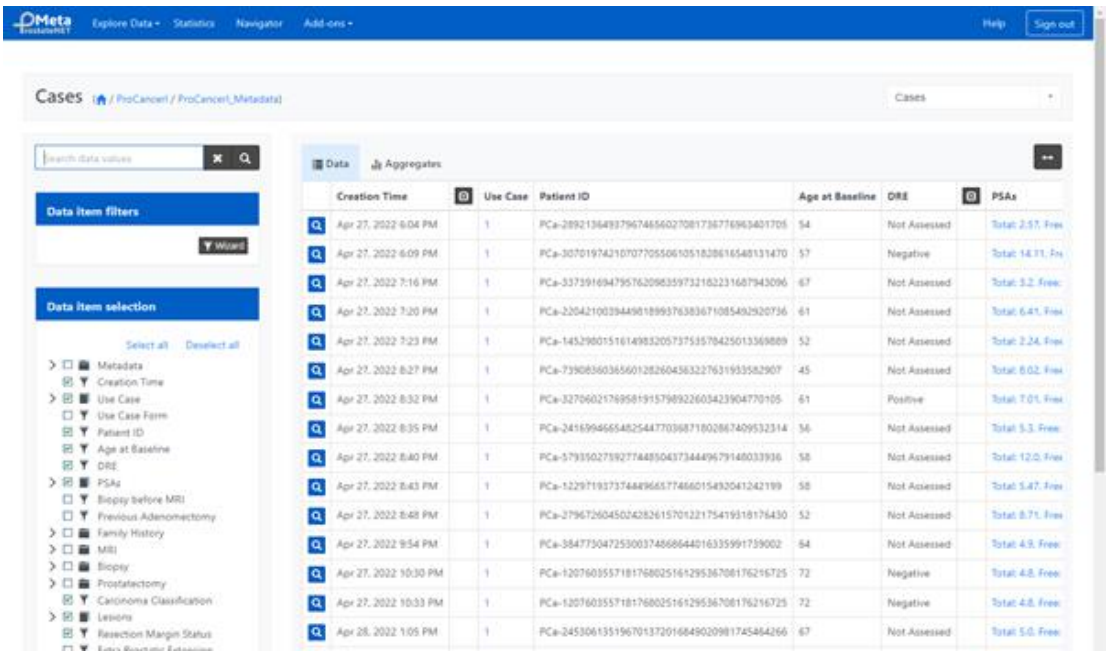


Figure 26: Explore Use Case Clinical and Imaging Metadata

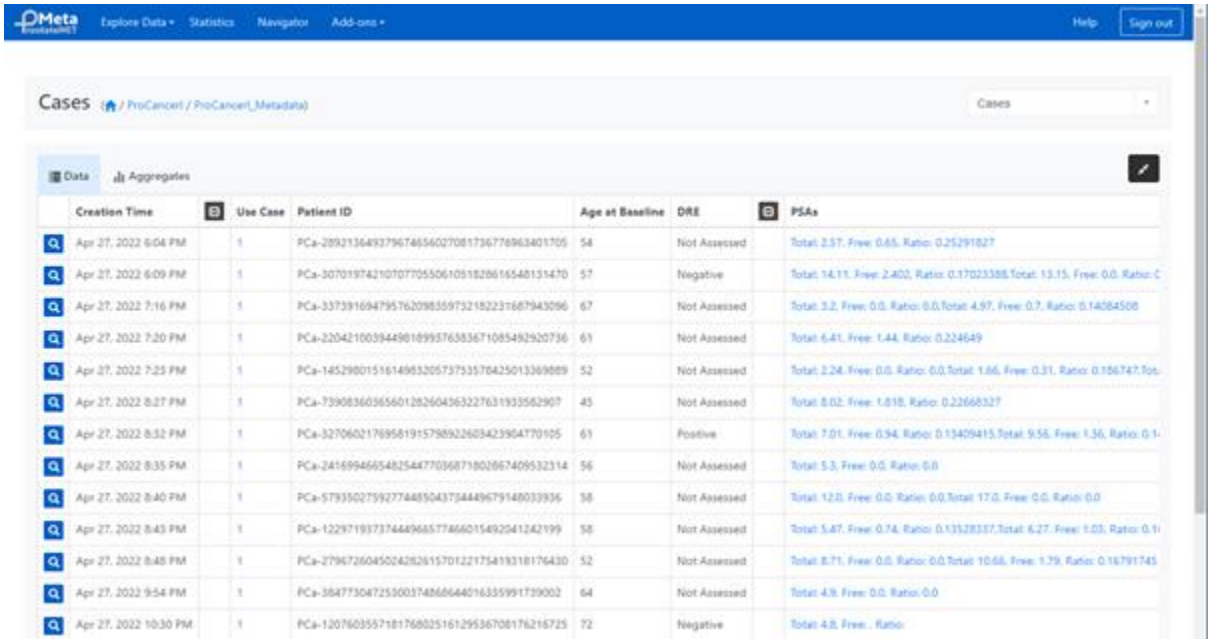


Figure 27: Expanded data view

The integrated Metadata catalogue service supports: Sort; Search; Filter; View tables; functionalities as described in details in the ANNEX section.

5.4.3 Uploaded data "Statistics" exploration

The technical partners have designed, developed, and integrated into the metadata catalogue an exploration of the statistics of the uploaded data to the ProstateNet from the data providers. Specifically, by selecting the **Statistics menu item**, the user can explore the statistics of the uploaded cases at the time point of triggering the abovementioned service. The numbers correspond to **unique patients** that have been uploaded so far, and not to the total number of uploads to the metadata repository (i.e. duplicates are not taken into consideration).

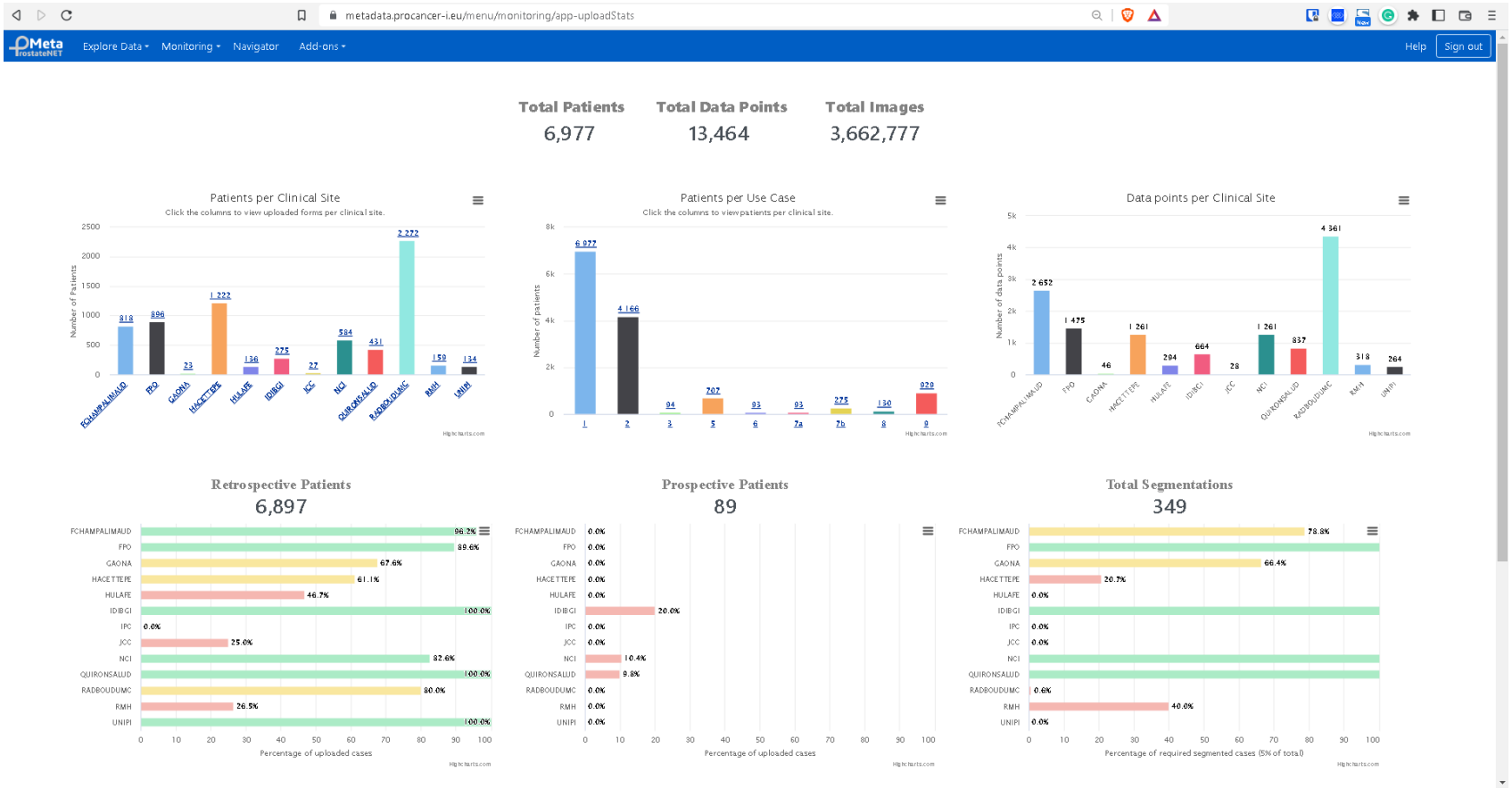


Figure 28: Upload statistics page

6 Supporting Actions

To assist the Data Providers as much as possible, multiple help sources have been created for them to quickly resolve any potential issues that arise during the uploading procedure. All the help sources are available from within the ProstateNET platform, gathered in the Help Center menu.

In the following sections, each of the help sources is presented in further detail.

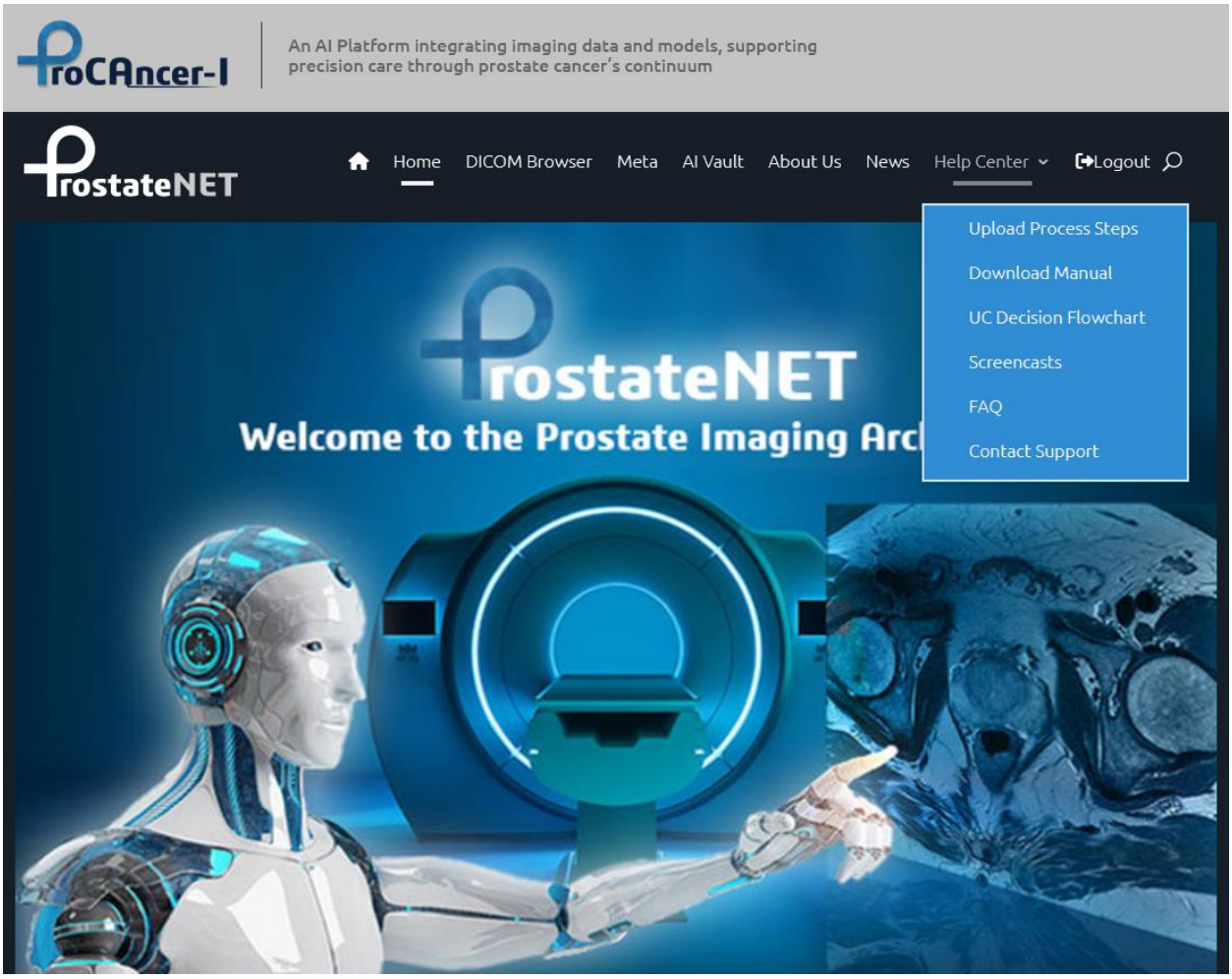


Figure 29: The Help Center menu

6.1 Upload Process Steps

The Upload Process Steps section provides a brief guide, like a before takeoff checklist which provides the Data Providers with a quick overview of the uploading steps that can be also used as a reference checklist at any point of the upload process.

Below follows the actual Upload Process Steps list, as it can be found in the corresponding section of the ProstateNET platform.

Prerequisites: Each data-providing partner has already reported a responsible Person as the contact point for the upload process.

For each data providing partner ("Data Provider" or "DP" from now on) we need to perform the following:

1. The partner performs local data retrieval of the needed data according to the protocol agreed upon. In this step, the DP should use the available on-premise anonymization tool to anonymize the imaging data before using the Data Upload tool in the next step.
   Participants: **DP**
2. Responsible Radiologists or other people from DP register with the ProCAncer-I Id Provider at https://aai.procancer-i.eu/ and obtain an ELIXIR Id as documented in the User Manual. FORTH is responsible for verifying the users' registrations and their affiliation with a member organization of the consortium.
   Participants: **DP, FORTH**
3. A version of the Data Upload Tool needs to be sent to DP that is customized with the proper (partner-specific) credentials for uploading data. These credentials include a "client-id" and a "secret" password, known only to the responsible technical partners (B3D, FORTH).
   Participants: **DP, B3D, FORTH**
4. DP executes the eCRF and data upload tool, either in "single" or "batch" mode to submit the data to the "staging" area of the ProCAncer-I platform. During this step, imaging data are uploaded to the DICOM store managed by B3D and to the Metadata Repository managed by FORTH.
   Participants: **DP, B3D, FORTH**
5. Responsible Radiologists or other nominated individuals from DP login into https://prostatenet.eu/ using their ProCAncer-I linked accounts and visit the "DICOM Browser" area which is the access point for their uploaded ("staged") data. We need to point out that the Staging area is private and protected. As a result, in this staging area, they have access only to the data uploaded from their home organizations. Here, they can review and verify the correctness of the uploaded image sets and select the ones for which they will, subsequently, create segmentation masks.
   Participants: **DP, B3D, FORTH**
6. Responsible Radiologists or other nominated individuals from DP select the image series to be segmented in the DICOM Browser webpage and launch the QUIBIM Annotation Tool. Optionally, users can press the "AI" button on the Annotation Tool and launch execution of the Automatic Segmentation Model to get a good initial estimation that they can "tune" and improve later on. In its current implementation version, the Automatic

Segmentation Model requires approximately 30 seconds to complete. This amount of time increases if multiple users invoke the model at nearly the same time.
Participants: **DP, QUIBIM, FORTH**

7. Data remains in the private Staging area of each DP until DPs are notified by the project management to perform the final transfer of data to the ProstateNET repository. Authenticated Users from the DP overview the overall annotation process in the DICOM Browser. When 5% of representative imaging studies from the initial DP's dataset have been segmented and DP responsible people have validated the results, the press of the "MOVE" button will transfer the whole dataset to the "ProstateNET" production area where no further changes are possible.

   Participants: **DP, B3D, FORTH**

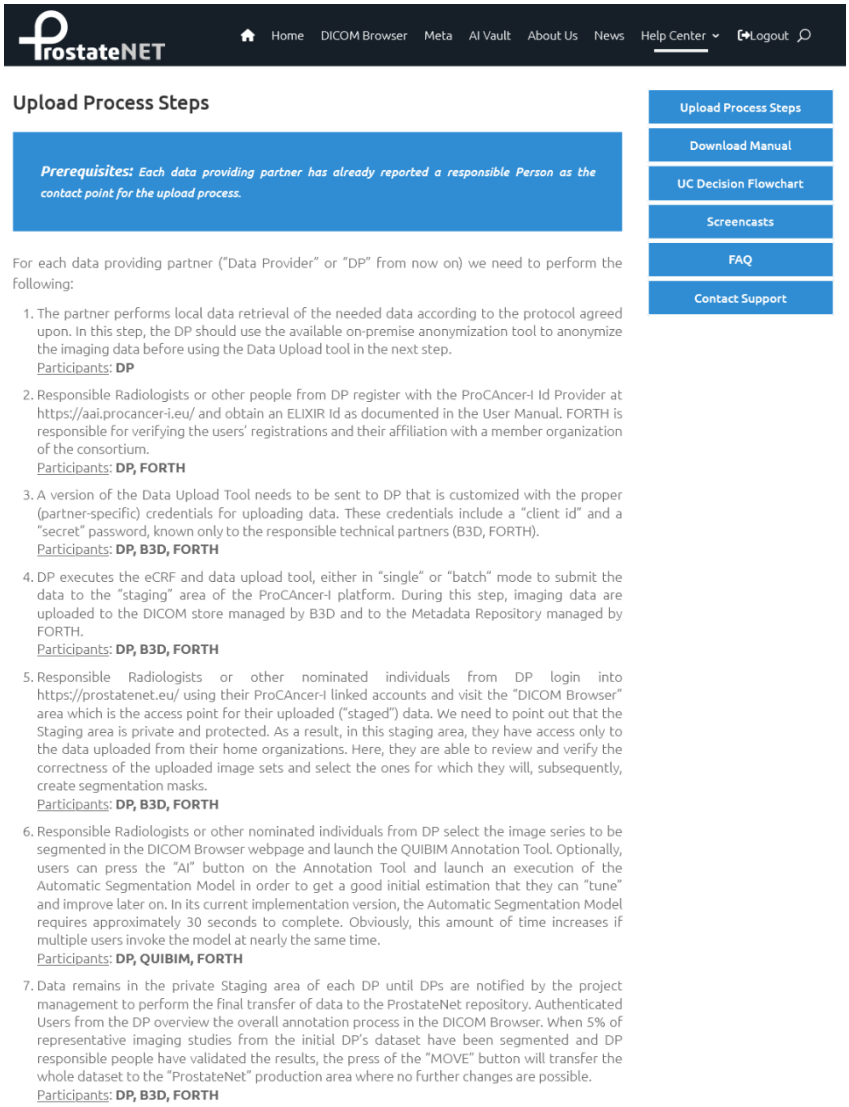
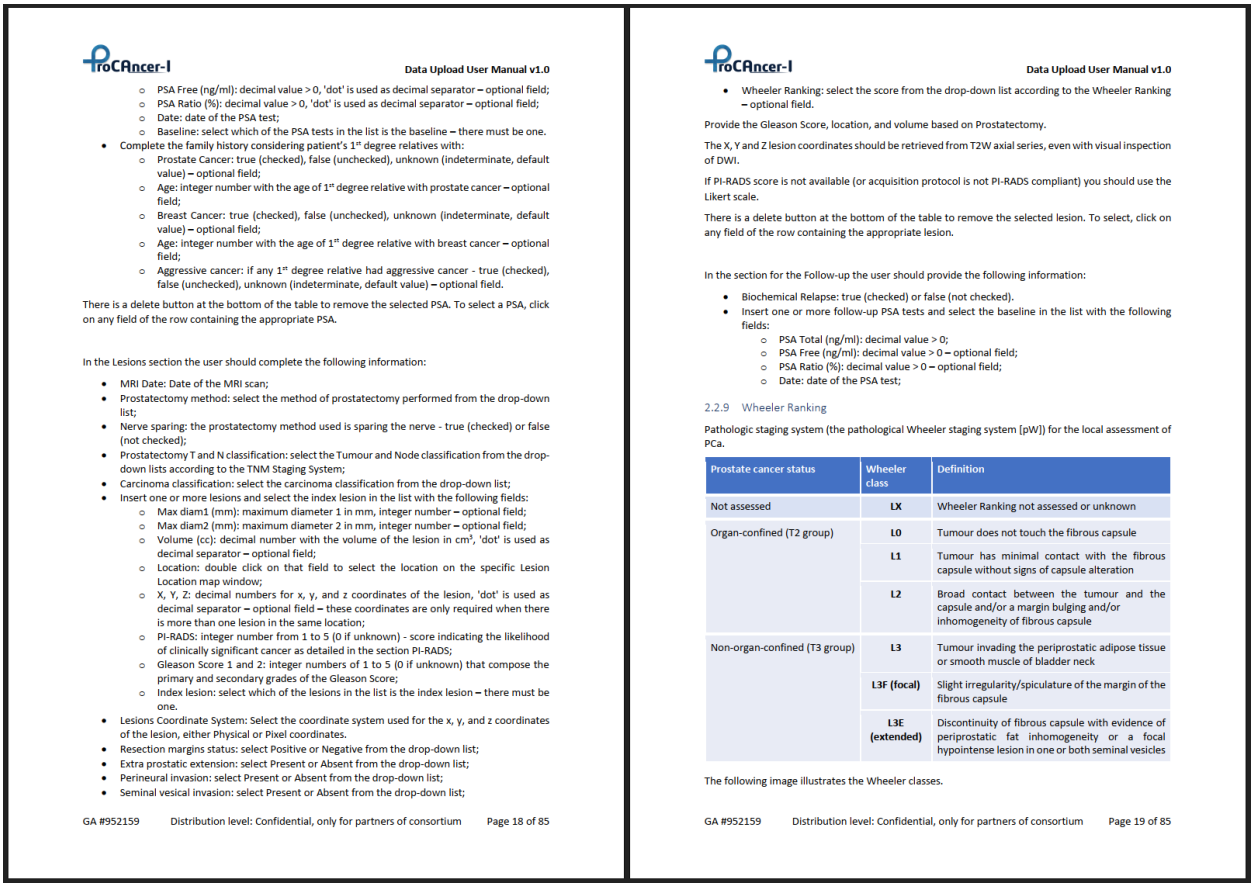


Figure 30: Screenshot of the Upload Process Steps section

6.2 Manual for the uploading process and the curation & annotation tools

A detailed manual has been compiled by the joint effort of all the technical partners of the project, which describes every function of the ProstateNET platform. The manual provides information from the beginning of the uploading process, by describing the data that have to be uploaded and by presenting in detail the Upload tool, the ProstateNET platform, the registration and authentication process, the staging area, the metadata repository, the PACS dashboard and finally the curation and the annotation tools. The manual has been circulated to the data-providing partners via the corresponding mailing list, and it is also available for download in the ProstateNET platform from the relative Help Center option.

The manual is constantly updated. Whenever an update of the tools (either a quick fix or an added feature) takes place or if further elaboration/clarification is needed in a certain section.



ProCAncer-I Data Upload User Manual v1.0

- PSA Free (ng/ml): decimal value > 0, 'dot' is used as decimal separator – optional field;
- PSA Ratio (%): decimal value > 0, 'dot' is used as decimal separator – optional field;
- Date: date of the PSA test;
- Baseline: select which of the PSA tests in the list is the baseline – there must be one.

- Complete the family history considering patient's 1st degree relatives with:
  - Prostate Cancer: true (checked), false (unchecked), unknown (indeterminate, default value) – optional field;
  - Age: integer number with the age of 1st degree relative with prostate cancer – optional field;
  - Breast Cancer: true (checked), false (unchecked), unknown (indeterminate, default value) – optional field;
  - Age: integer number with the age of 1st degree relative with breast cancer – optional field;
  - Aggressive cancer: if any 1st degree relative had aggressive cancer - true (checked), false (unchecked), unknown (indeterminate, default value) – optional field.

There is a delete button at the bottom of the table to remove the selected PSA. To select a PSA, click on any field of the row containing the appropriate PSA.

In the Lesions section the user should complete the following information:

- MRI Date: Date of the MRI scan;
- Prostatectomy method: select the method of prostatectomy performed from the drop-down list;
- Nerve sparing: the prostatectomy method used is sparing the nerve - true (checked) or false (not checked);
- Prostatectomy T and N classification: select the Tumour and Node classification from the drop-down lists according to the TNM Staging System;
- Carcinoma classification: select the carcinoma classification from the drop-down list;
- Insert one or more lesions and select the index lesion in the list with the following fields:
  - Max diam1 (mm): maximum diameter 1 in mm, integer number – optional field;
  - Max diam2 (mm): maximum diameter 2 in mm, integer number – optional field;
  - Volume (cc): decimal number with the volume of the lesion in cm^3, 'dot' is used as decimal separator – optional field;
  - Location: double click on that field to select the location on the specific Lesion Location map window;
  - X, Y, Z: decimal numbers for x, y, and z coordinates of the lesion, 'dot' is used as decimal separator – optional field – these coordinates are only required when there is more than one lesion in the same location;
  - PI-RADS: integer number from 1 to 5 (0 if unknown) - score indicating the likelihood of clinically significant cancer as detailed in the section PI-RADS;
  - Gleason Score 1 and 2: integer numbers of 1 to 5 (0 if unknown) that compose the primary and secondary grades of the Gleason Score;
  - Index lesion: select which of the lesions in the list is the index lesion – there must be one.
- Lesions Coordinate System: Select the coordinate system used for the x, y, and z coordinates of the lesion, either Physical or Pixel coordinates.
- Resection margins status: select Positive or Negative from the drop-down list;
- Extra prostatic extension: select Present or Absent from the drop-down list;
- Perineural invasion: select Present or Absent from the drop-down list;
- Seminal vesical invasion: select Present or Absent from the drop-down list;

GA #952159 Distribution level: Confidential, only for partners of consortium Page 18 of 85

ProCAncer-I Data Upload User Manual v1.0

- Wheeler Ranking: select the score from the drop-down list according to the Wheeler Ranking – optional field.

Provide the Gleason Score, location, and volume based on Prostatectomy.

The X, Y and Z lesion coordinates should be retrieved from T2W axial series, even with visual inspection of DWI.

If PI-RADS score is not available (or acquisition protocol is not PI-RADS compliant) you should use the Likert scale.

There is a delete button at the bottom of the table to remove the selected lesion. To select, click on any field of the row containing the appropriate lesion.

In the section for the Follow-up the user should provide the following information:

- Biochemical Relapse: true (checked) or false (not checked).
- Insert one or more follow-up PSA tests and select the baseline in the list with the following fields:
  - PSA Total (ng/ml): decimal value > 0;
  - PSA Free (ng/ml): decimal value > 0 – optional field;
  - PSA Ratio (%): decimal value > 0 – optional field;
  - Date: date of the PSA test;

2.2.9 Wheeler Ranking

Pathologic staging system (the pathological Wheeler staging system [pW]) for the local assessment of PCa.

Prostate cancer status	Wheeler class	Definition
Not assessed	LX	Wheeler Ranking not assessed or unknown
Organ-confined (T2 group)	L0	Tumour does not touch the fibrous capsule
	L1	Tumour has minimal contact with the fibrous capsule without signs of capsule alteration
	L2	Broad contact between the tumour and the capsule and/or a margin bulging and/or inhomogeneity of fibrous capsule
Non-organ-confined (T3 group)	L3	Tumour invading the periprostatic adipose tissue or smooth muscle of bladder neck
	L3F (focal)	Slight irregularity/spiculature of the margin of the fibrous capsule
	L3E (extended)	Discontinuity of fibrous capsule with evidence of periprostatic fat inhomogeneity or a focal hypointense lesion in one or both seminal vesicles

The following image illustrates the Wheeler classes.

GA #952159 Distribution level: Confidential, only for partners of consortium Page 19 of 85

Figure 31: Screenshot from the manual

6.3 Use Case Decision Flowchart

This is the section where two important figures reside.

- The visual representation of the ProCAncer-I defined UCs mapped along the PC management continuum (Section 4.1.3)
- Upload process decision flowchart (Section 4.1.3)

6.4 Screencasts

A screencast is a digital recording of a computer screen output, a screen recording, which presents a specific scenario/workflow to the user including either narration or subtitles. Screencasts are a helpful asynchronous mechanism to present in a short time the core functionality of each tool in ProstateNET's workflow. Screencasts are available from the **Help Center** menu → **Screencasts**

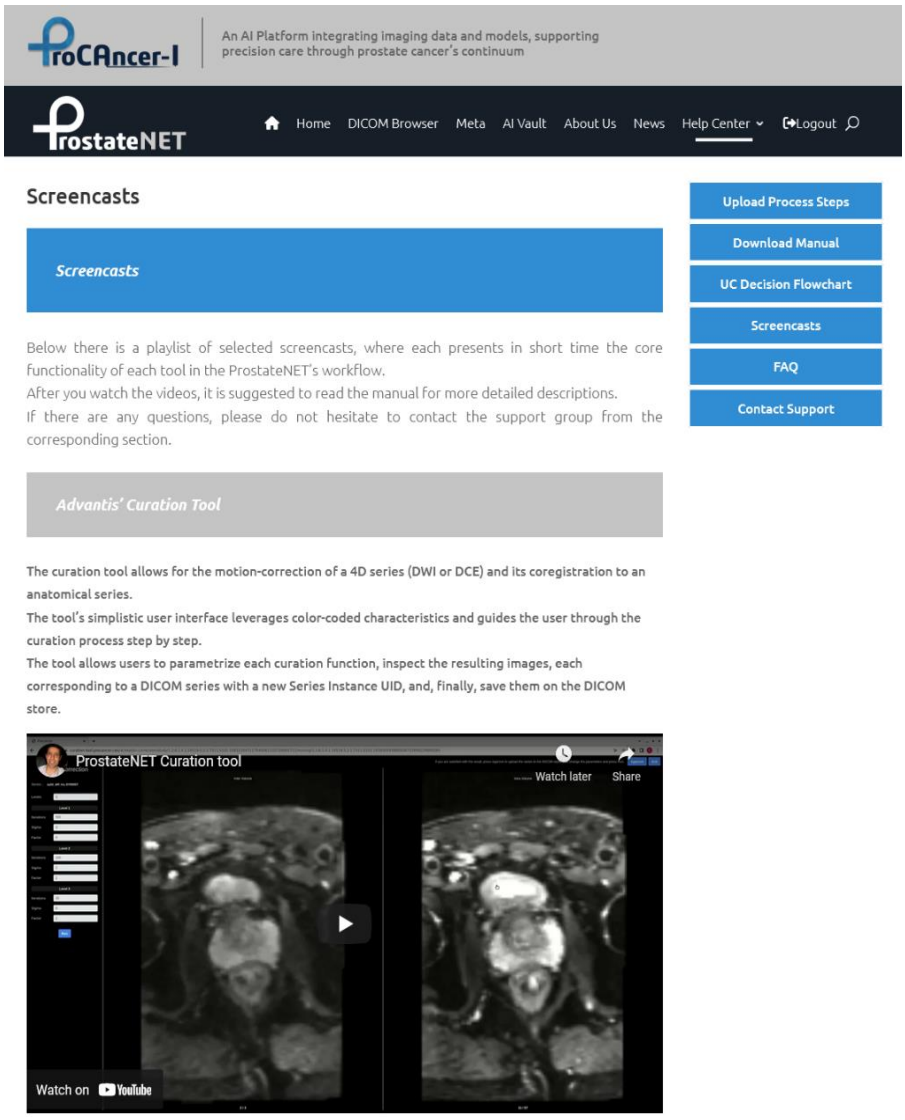


Figure 32: Screenshot of the Screencasts section

6.5 FAQ

The Help Center contains a Frequently Asked Questions (FAQ) section that gathers the most important questions risen during both the development and testing phase and also during the distribution and usage period. The questions are grouped and themed by the functionality/tool and each question is represented with an accordion control. The accordion control provides a compact presentation of all the questions in a stack and allows the answer to be visible as it expands when selected. The following figure depicts screenshots of the FAQ section in the Help Center.

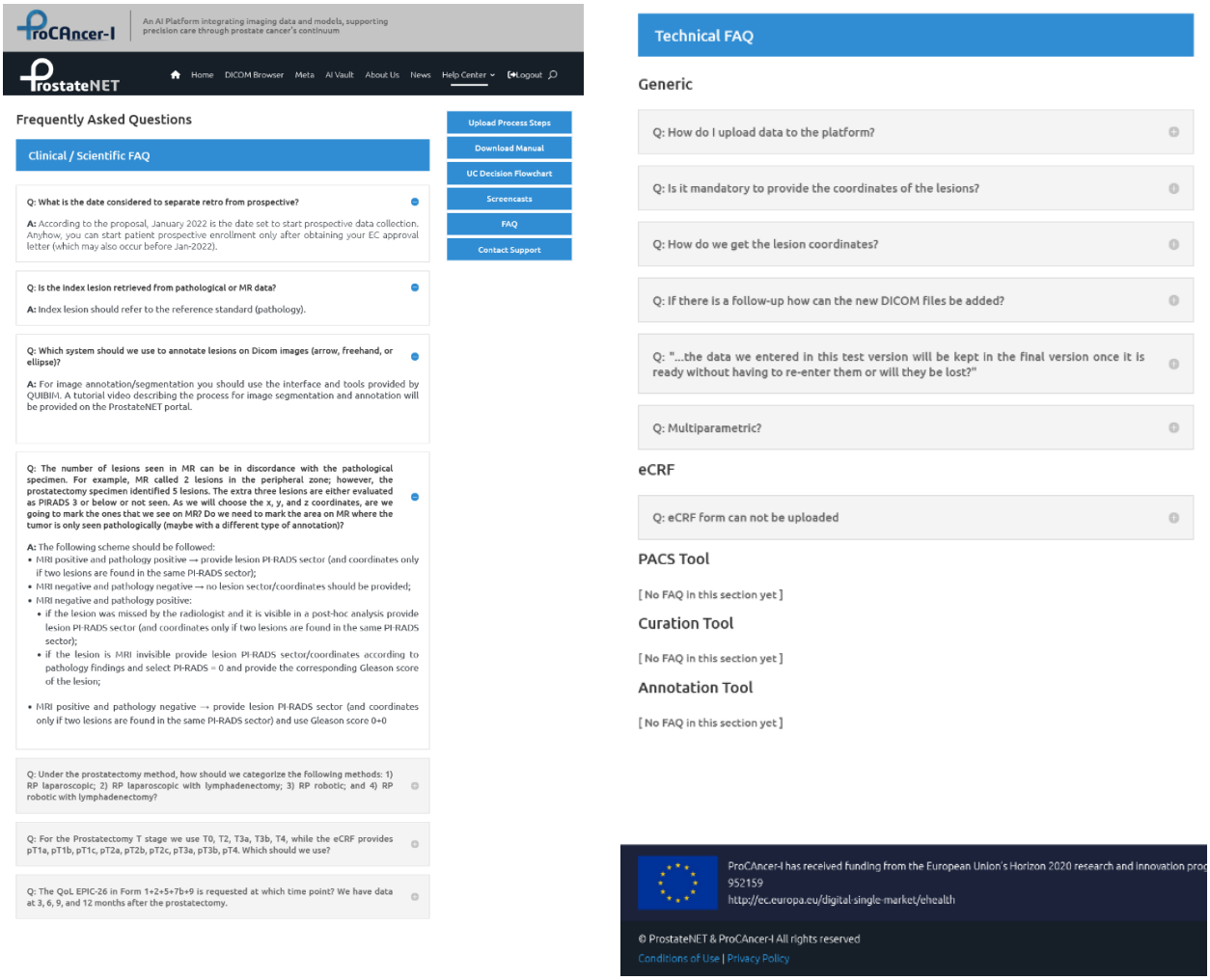


Figure 33: Screenshots of the Frequently Asked Questions section

6.6 Support channels

Numerous means to contact the support group members are provided to the Data Providers. The primary (and suggested) contact channel is the dedicated mailing list "procancer-i-support@procancer-i.eu". Each email sent to the list is read by all the responsible technical and

clinical partners which are assigned to provide support. The list can be used either directly or via ProstateNET's "Contact Support" form (https://prostatenet.eu/help-center/contact-support/). Aside from the support mailing list, there is also the dedicated support channel #procancer-i-support in the communication platform of the project, the CBMLChat.

Other support actions include Virtual Training Workshops and One-on-one support meetings with the data providers, where support is provided via screen sharing for the faster resolution of issues.

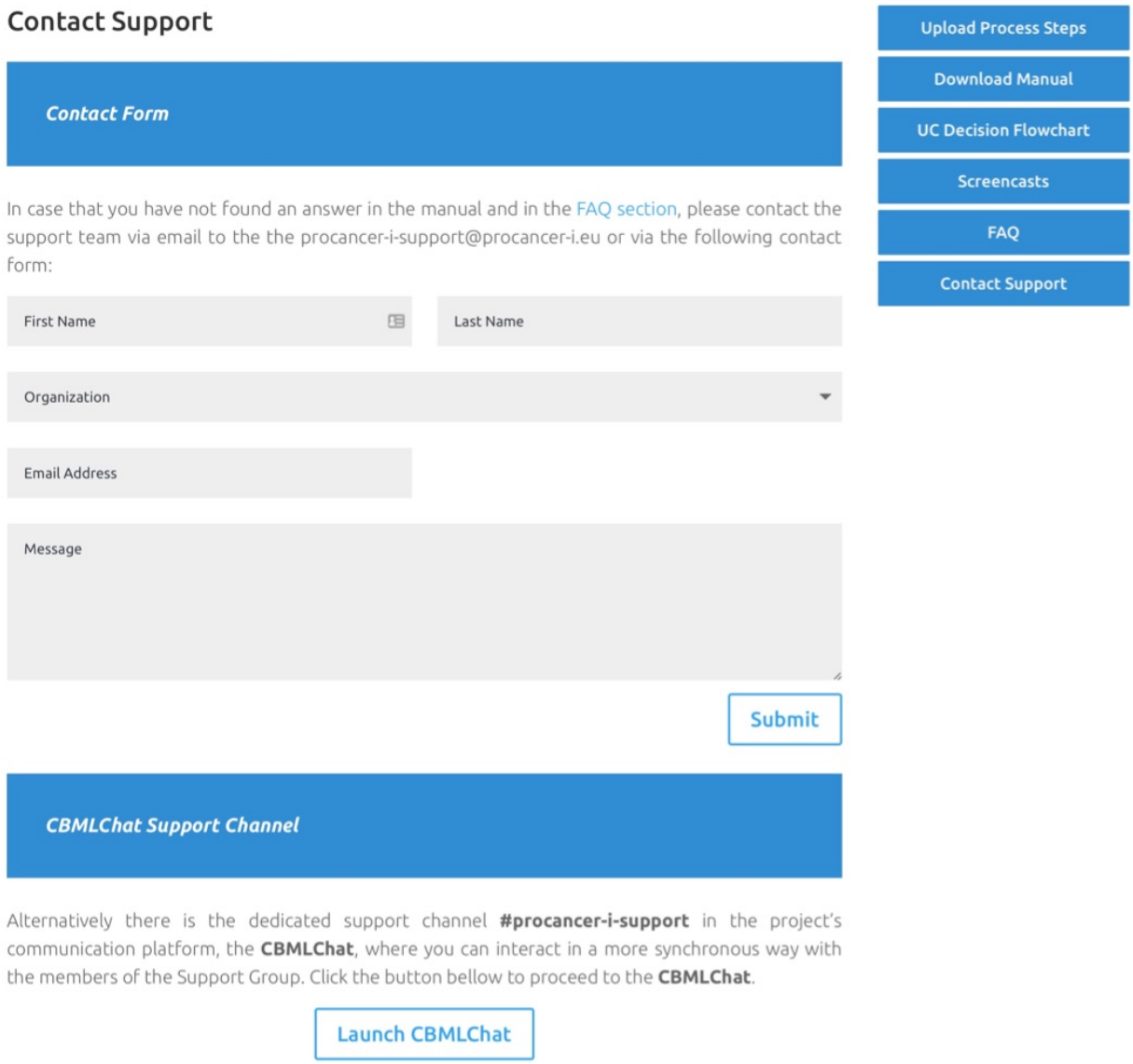


Figure 34: Screenshot of the supporting mailing list

7 Analysis of the retrospective uploaded data

One of the main aims of ProCAncer-I is to create one of the largest interoperable repository of PC Magnetic Resonance Images used for developing robust PC AI models. This repository will be populated from retrospective and prospective data derived by the clinical partners (data providers) of the ProCAncer-I consortium. The current deliverable reports on the collection of the retrospective data were carefully selected to fit the described project clinical UCs.

At the time being the following figure coming from the monitoring service (Section 5.4.3) visualizes the data that has been uploaded in the staging area of ProstateNet platform.

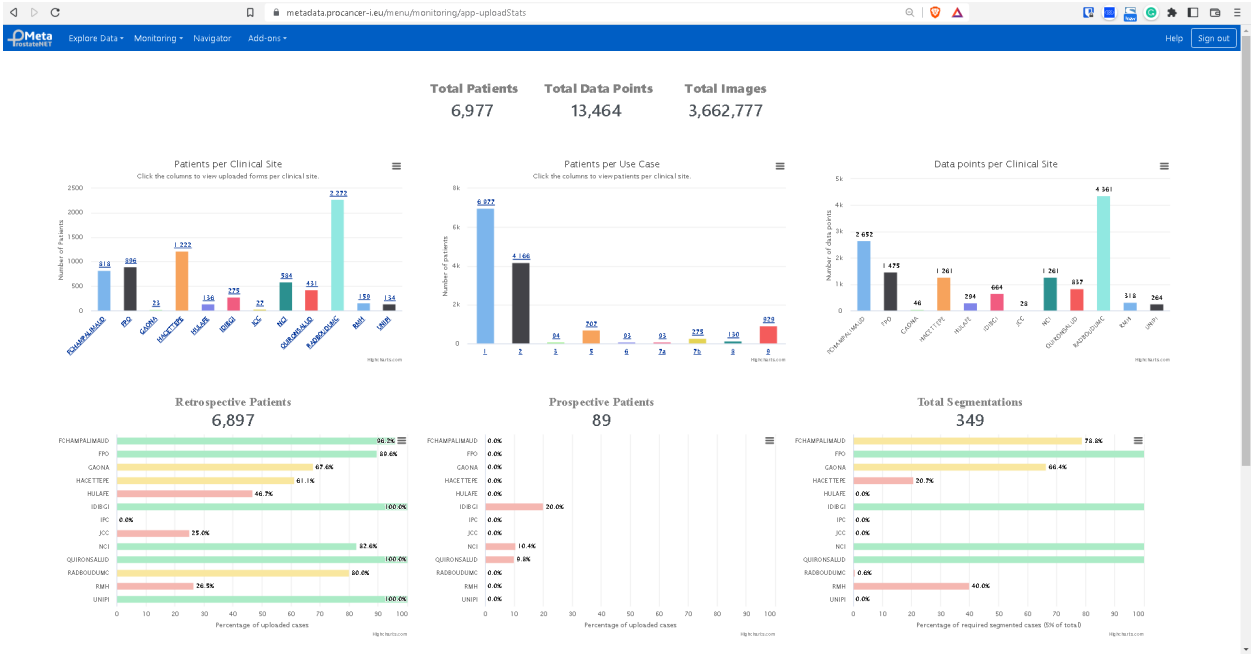


Figure 35: Screenshot of the statistics provided by the monitoring mechanisms of the upload process

Specifically, the following plot depicts the number of unique patients that have been uploaded to the repository of ProstateNet from each clinical site of ProCAcner-I project..

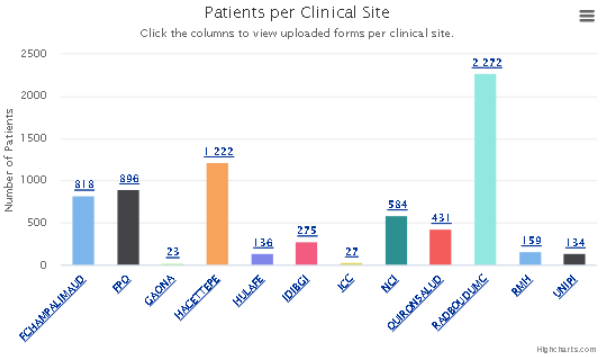


Figure 36: Plot of the number of patients uploaded on ProstateNet per clinical site

While next figure shows the unique number of patients belonging to each clinical use case as defined in ProCAncer-I.

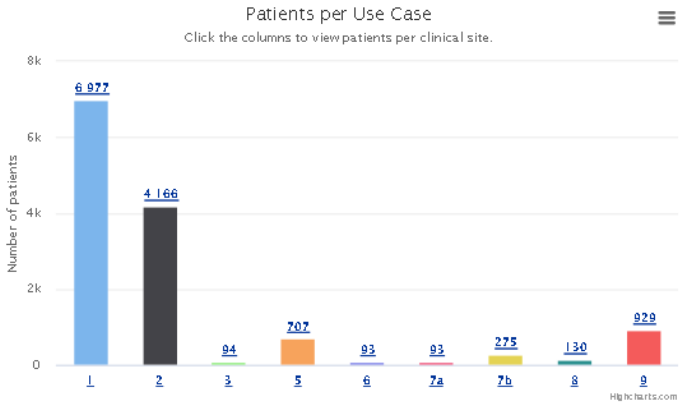


Figure 37: Plot of the number of patients uploaded on ProstateNet per clinical use case

The following figure presents in a bar plot the data points per clinical site uploaded in the ProstateNet and derived from the retrospective data.

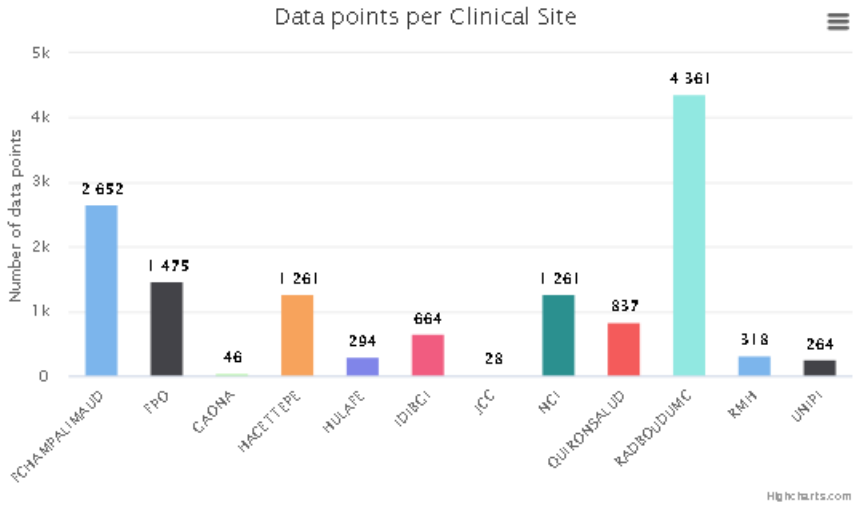


Figure 38: Plot of the number of data points uploaded on ProstateNet per clinical site

Note: Is known that in data analysis and statistics, a **data point** is a piece of information that describes one unit of observation, at one point in time, at the data collection level. In ProCAncer-I the term is utilized to identify a unique "observation" in a specific use case. The latter is related to a patient case, where the same patient can be used in multiple use cases.

So far, **6,980** patients have been uploaded into the platform along with **3,663,391** total images (90 patients belong to prospective cases), corresponding to **13,473** total data points.

Note: The uploading activities are “live” (during the compilation of the deliverable) and thus some discrepancies between several screenshots in the number of data may occur.

	Number of Cases
No PCa (negative cases)	2820 (41%)
PCa confirmed (positive cases)	4160 (59%)
Prostatectomy performed	1411 (20%)

Currently, there are **2,820 total negative cases** (no Prostate Cancer has been diagnosed) and **4,160 total positive cases** (Prostate Cancer has been confirmed). From those, 1,411 patients have been undergone prostatectomy.

	Min	Mean	Max
Age	32	64.8	89
PSA (ng/ml)	-	8.6	1135
Lesion Volume (cc)	-	12.4	193
Lesion Diameter (mm)	-	20.5	234

The average age of the uploaded patients is 64.8 years old, with an average PSA value equal to 8.6. The average lesion volume is 12.44cc and the average diameter is approximately 21mm. There are currently 8,128 total lesions registered in the platform, where most of them are in the peripheral zone (72%), followed by lesions in the transition zone (18%), Anterior Fibromuscular Stroma (4%), and Central Zone (2%) (4% of the lesions have no reported location).

	Percentage of Lesions
Peripheral Zone	72%
Transition Zone	18%
Anterior Fibromuscular Stroma	4%
Central Zone	2%

Finally, out of the 8,128 lesions, 28 % have a total Gleason score of 6, and 45% have a Gleason score of 7.

Gleason Score	Number of lesions
0	482
2	2
3	5
4	3
5	28
6	2,279
7	3,662

8	496
9	416
10	28

And,

Gleason1	Gleason2	Number of lesions
-1	-1	304
0	0	482
1	1	2
1	2	3
1	3	2
1	4	8
2	2	4
2	3	2
2	4	2
3	0	2
3	2	18
3	**3**	**2278**
3	**4**	**2506**
3	5	74
4	0	2
4	**3**	**1161**
4	4	408
4	5	327
5	3	14
5	4	90
5	5	28

7.1 Monitoring service for the uploaded data quality in the metadata catalogue

The technical consortium members of ProCAncer-I have developed a specific service integrated into the metadata catalogue component which monitors automatically the quality of the uploaded data at the staging area of ProstateNet. The latter provides insights into the quality, integrity, and potential issues in both the monitoring team and the data providers.

Specifically, the uploaded data are automatically based on specific rules tagged as:

- NOTICE
- LOW IMAGES
- EXCESS DATA
- INCOMPLETE
- ORPHAN
- WARNING

In order someone to get access to the feature, needs to login to the metadata catalogue and push the button “Quality Check” from the “Monitoring” tab in the menu bar ().

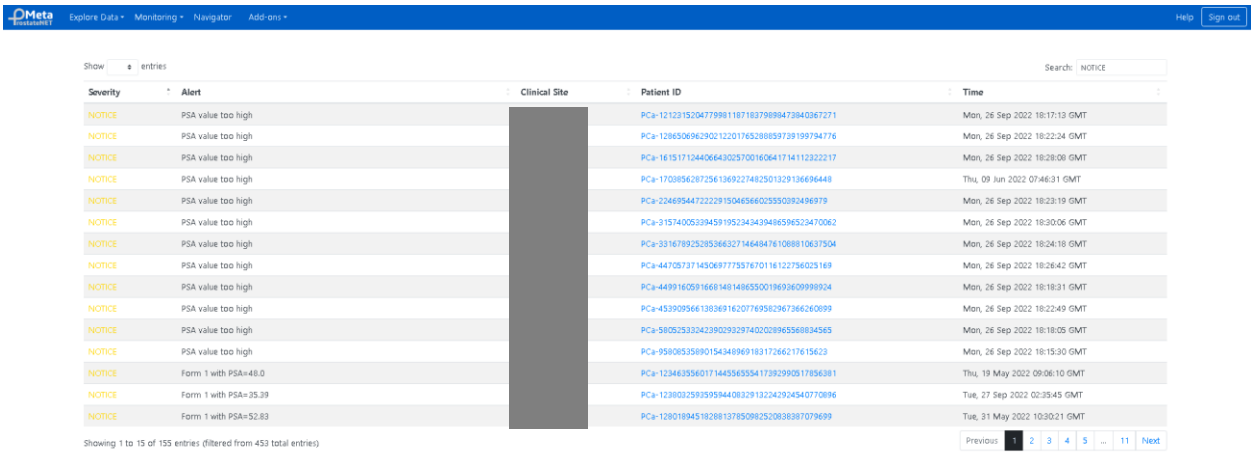


Figure 39 Screenshot from the “Quality Check” feature (blurred due to the public distribution level)

Specifically the following section describe the developed rules for the individual issues.

NOTICE

1. Patients with really high PSA values (PSA > 200 ng/ml)
2. Patients in Form 1 with PSA > 25 ng/ml
3. Patients in Form ‘1’ or ‘1+2’ or ‘1+2+3’ or ‘1+2+8+9’ with “MRI positive”=False but reported Lesions with “PIRADS” >= 3

LOW IMAGES

1. DICOM series with less than 10 images (per series).

EXCESS DATA

1. DICOM studies with more than 3500 images.

INCOMPLETE

1. Patients with less than 3 series (since T2, ADC, and high b-value are the minimum requirements). Less than 3 means that at least 1 series is missing, possibly failing to upload.

ORPHAN

1. DICOM study with no uploaded metadata. This means that there is a study in the DICOM browser, but there is no uploaded eCRF form accompanying this study.

WARNING

1. Series with suspicious secondary captures (e.g. screenshots).
2. Series has only 1 image (they might contain sensitive patient information).

The following section describes the instructions provided to the data providers for the issue resolving per case.

WARNING alert

Please delete manually from the "DICOM Browser" the series that contain sensitive information. Once this is done, you should resolve the alert manually through the "Resolve Alerts" menu of the metadata catalogue. However, only authorized users have the ability to resolve alerts.

NOTICE and ORPHAN alerts:

Since these correspond to errors in the ecrf form, you should just re-upload the ecrf (either from the ecrf tool manually by clicking the "3. Form Upload" button for that patient, or through the batch mode by using the "/j" flag)

INCOMPLETE alerts:

The data provider should re-upload the study from scratch with adding the missing series (both the DICOM study and the eCRF). The alert will be resolved automatically.

EXCESS DATA alert:

The data provider might consider deleting unnecessary information from the "DICOM Browser". Then you should resolve the alert manually, through the "RESOLVE ALERTS" menu (only if you have the right permissions).

LOW IMAGES:

if the data provider believes that the series is correct, then you can manually resolve the alert. If there was an upload error and the series was not uploaded properly, please re-upload the study along with the eCRF and the alert will be automatically resolved. If the series should not be there, you should delete the series manually from the "DICOM Browser" and resolve the alert manually as well.

8 Conclusions

The retrospective data collection and uploading is important to the success of ProCAncer-I, since it will be the backbone of ProstateNET providing a big dataset comprising of several millions of prostate representations visualized with different MRI contrasts (T2, DWI, ADC, DCE), coming from 13 geographically diverse clinical providers. In addition, the clinical institutions are either regional private diagnostic centers, public hospitals, specialized anti-cancer centers as well as university hospitals that are reference centers for prostate cancer. Apart from the size of the data, an equally important quality of ProstateNET is the data diversity based on 1.5T and 3T scanners, with or without the use of an endorectal coil, and many different sequence parameter combinations, which brings the ProstateNET very close to real-world data. The latter will have very positive effects towards developing not only high-performance AI models, but also generalizable which at the moment is the drawback of existing AI models that are not performing as promised in each and every setup (hospital, scanner, protocol, patient cohort).

Following a very careful standardization process related to data collected, the ProCAncer-I consortium defined the minimum non-imaging clinical variables required for each use Case as well as the type of MRI images that should be present to make a patient eligible for uploading to ProCancer-I platform.

The retrospective data uploading process has been concluded. Currently – on the 26th of October 2022 - the platform hosts 7051 patients, 13.615 data points, and more than 3.7 million images that have been successfully uploaded, which will make feasible the finalization of the planned work in T5.3, i.e. the development of the master models based on the retrospective data.

However, uploading of retrospective data will continue for as long as there are clinical centers that possess data that fulfill the inclusion criteria established. We estimate that small numbers of retrospective data will continue to be uploaded onto the platform for an additional period of two months.

At the same time, the tools and processes developed for managing retrospective data uploading will also be used for the collection, preparation, and uploading of prospective data, a process that has already begun.

ANNEX

The ANNEX section of the current deliverable (D5.1 Retrospective Data Retrieval and Upload) includes information derived from the detailed manual that has been compiled for the upload process of the data to the ProstateNet platform. Specifically, and for complementarity purposes the section includes details for the rest of the Forms of the eCRF ProCAncer-I upload tool (single & batch mode operation), detailed functionalities description for the curation tools and the annotation tool.

A.1 ProCAncer-I eCRF tool – single mode operation

Clinical Data and Use Case Forms

Form 1 + 2

Collection of patients with confirmed Prostate Cancer at biopsy and/or prostatectomy.

This form collects positive cases to Use Case 1 and Use Case 2.

This form has two information sections: Clinical and Lesions.

In the Clinical section, the user should complete the following information:

- Age at baseline: insert an integer number;
- DRE: insert the result of the Digital Rectal Examination – Positive, Negative or Not Assessed – optional field;
- Biopsy before MRI: true (checked) or false (unchecked);
- Previous adenomectomy: true (checked) or false (unchecked);
- Insert one or more PSA tests and select the baseline in the list with the following fields:
    - PSA Total (ng/ml): decimal value > 0, 'dot' is used as a decimal separator;
    - PSA Free (ng/ml): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - PSA Ratio (%): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - Date: date of the PSA test;
    - Baseline: select which of the PSA tests in the list is the baseline – there must be one.

There is a delete button at the bottom of the table to remove the selected PSA. To select a PSA, click on any field of the row containing the appropriate PSA.

In the Lesions section, the user should complete the following information:

- MRI Date: Date of the MRI scan;

- Biopsy performed: true (checked) or false (not checked);
- Type: select the type of biopsy performed from the drop-down list – required only if the biopsy was performed;
- Prostatectomy performed: true (checked) or false (not checked);
- Insert one or more lesions and select the index lesion in the list with the following fields:
    - Max diam1 (mm): maximum diameter 1 in mm, integer number – optional field;
    - Max diam2 (mm): maximum diameter 2 in mm, integer number – optional field;
    - Volume (cc): decimal number with the volume of the lesion in cm^3, 'dot' is used as decimal separator – optional field;
    - Location: double click on that field to select the location on the specific Lesion Location map window;
    - X, Y, Z: decimal numbers for x, y, and z coordinates of the lesion, 'dot' is used as decimal separator – optional field – these coordinates are only required when there is more than one lesion in the same location;
    - PI-RADS: integer number from 1 to 5 (0 if unknown)- score indicating the likelihood of clinically significant cancer as detailed in the section PI-RADS;
    - Gleason Score 1 and 2: integer numbers of 1 to 5 (0 if unknown) that compose the primary and secondary grades of the Gleason Score;
    - Index lesion: select which of the lesions in the list is the index lesion – there must be one.
- Lesions Coordinate System: Select the coordinate system used for the x, y, and z coordinates of the lesion, either Physical or Pixel coordinates.

Provide the Gleason Score, location, and volume based on Prostatectomy, otherwise location, and Gleason Score based on Biopsy.

The X, Y, and Z lesion coordinates should be retrieved from T2W axial series, even with visual inspection of DWI.

If PI-RADS score is not available (or the acquisition protocol is not PI-RADS compliant) you should use the Likert scale.

There is a delete button at the bottom of the table to remove the selected lesion. To select, click on any field of the row containing the appropriate lesion.

PI-RADS

The Prostate Imaging-Reporting and Data System (PI-RADS) is a structured reporting scheme for multiparametric prostate MRI in the evaluation of suspected prostate cancer in treatment-naive prostate glands. The ProCAncer-I uses version 2.1 (v2.1), published in 2019 and developed by an

internationally representative group involving the American College of Radiology (ACR), European Society of Urogenital Radiology (ESUR), and AdMeTech Foundation.

PI-RADS v2.1 Assessment Categories - Each lesion is assigned a score from 1 to 5 indicating the likelihood of clinically significant cancer:

- PI-RADS 1: very low (clinically significant cancer is highly unlikely to be present)
- PI-RADS 2: low (clinically significant cancer is unlikely to be present)
- PI-RADS 3: intermediate (the presence of clinically significant cancer is equivocal)
- PI-RADS 4: high (clinically significant cancer is likely to be present)
- PI-RADS 5: very high (clinically significant cancer is highly likely to be present)

Each lesion can be scored 1-5 on DWI and on T2W, as well as by the absence or presence of dynamic contrast enhancement. The contribution of these scores to the overall PI-RADS assessment differs depending on whether the lesion is located in the transition zone or peripheral zone of the prostate. For the transition zone, the PI-RADS assessment is primarily determined by the T2W score and sometimes modified by the DWI score. For the peripheral zone, the PI-RADS assessment is primarily determined by the DWI score and sometimes modified by the presence of dynamic contrast enhancement.

PI-RADS v2.1 full-text document can be found at the American College of Radiology site at https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/PI-RADS.

If PI-RADS score is not available (or the acquisition protocol is not PI-RADS compliant) you should use the Likert scale.

If PI-RADS score or Likert is not possible to determine set the value to zero (0).

Lesion Location map

The lesion location uses the 36 positions from PI-RADS:

- The right and left peripheral zones (PZ) at prostate base, midgland, and apex are each subdivided into three sections: anterior (a), medial posterior (mp), and lateral posterior (lp).
- The right and left transition zones (TZ) at prostate base, midgland, and apex are each subdivided into two sections: anterior (a) and posterior (p).
- The central zone (CZ) is included in the prostate base around the ejaculatory ducts.
- The anterior fibromuscular stroma (AS) is divided into right/left at the prostate base, midgland, and apex.

The user should check all the locations that apply for the selected lesion and click the save button to save and return to the previous window.

There is a reset button to uncheck all locations and start again.

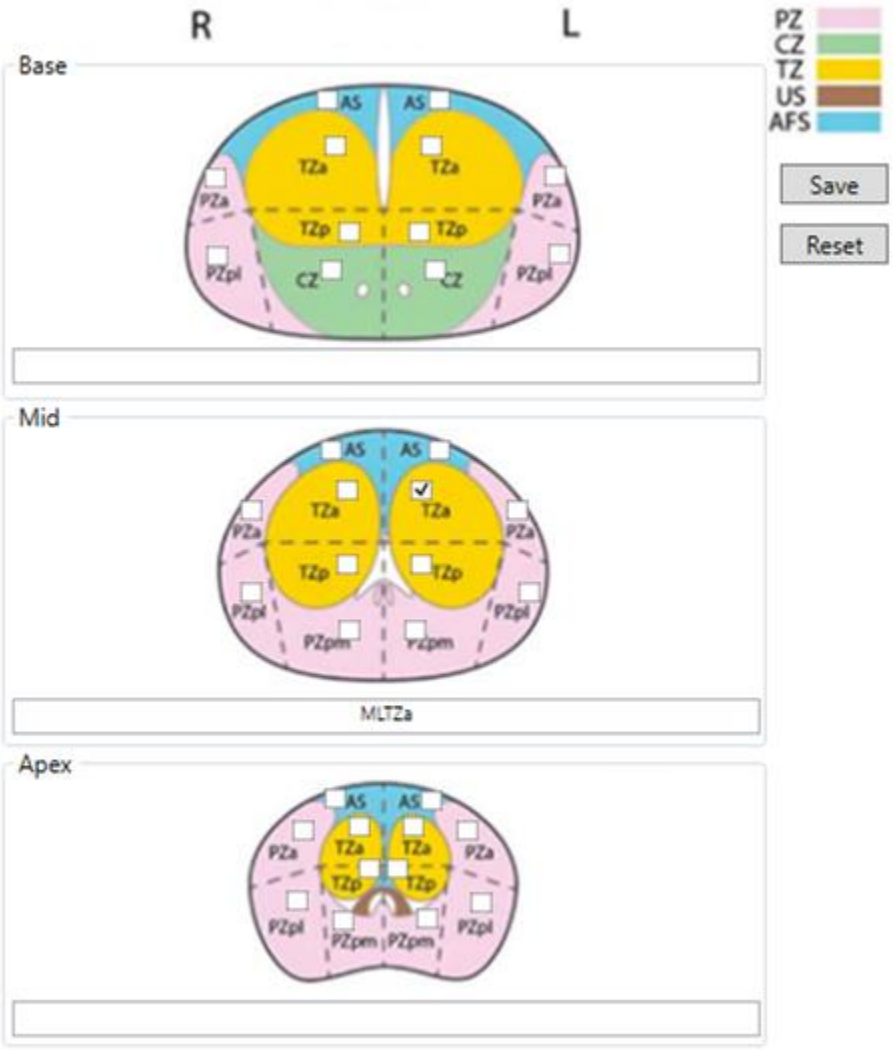


Figure 40: Lesion Location map

Gleason Score

The Gleason Score is the grading system used to determine the aggressiveness of prostate cancer.

The Gleason Score ranges from 1 to 5 and describes how much the cancer from a biopsy looks like healthy tissue (lower score) or abnormal tissue (higher score).

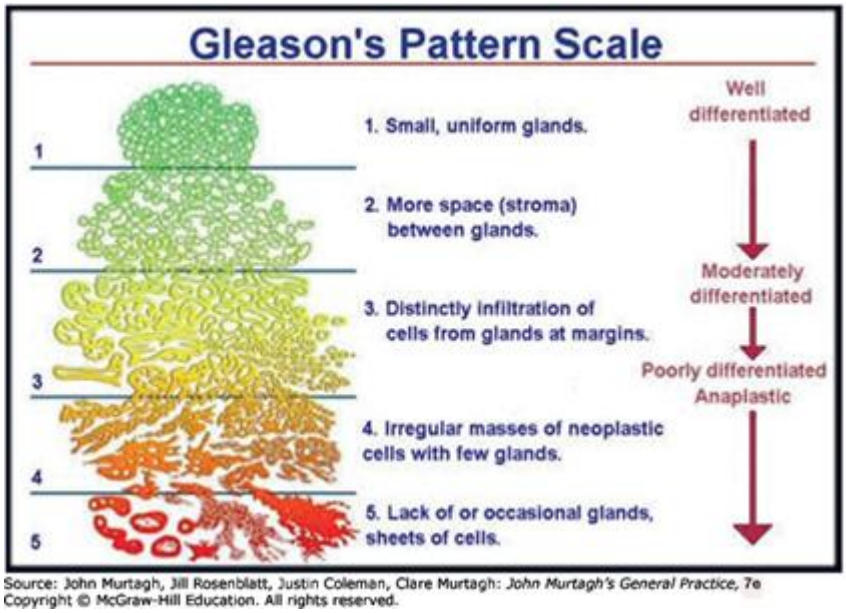


Figure 41: Gleason's Pattern Scale

Since prostate tumours are often made up of cancerous cells that have different grades, two grades are assigned for each patient:

- a primary grade is given to describe the cells that make up the largest area of the tumour,
- a secondary grade is given to describe the cells of the next largest area.

For instance, if the Gleason Score is written as 3+4=7, it means most of the tumour is grade 3 and the next largest section of the tumour is grade 4, together they make up the total Gleason Score. If the cancer is almost entirely made up of cells with the same score, the grade for that area is counted twice to calculate the total Gleason Score.

If the Gleason score is not possible to determine, then set the values to zero (0).

Form 1 + 2 + 3

Collection of patients with confirmed Prostate Cancer at biopsy and/or prostatectomy and with metastasis within 6 months from MRI.

This form collects positive cases to Use Cases 1, 2, and 3.

This form has three information sections: Clinical, Lesions, and Final diagnosis.

In the Clinical section, the user should complete the following information:

- Age at baseline: insert an integer number;
- DRE: insert the result of the Digital Rectal Examination – Positive, Negative or Not Assessed – optional field;
- Biopsy before MRI: true (checked) or false (unchecked);
- Previous adenomectomy: true (checked) or false (unchecked);
- Insert one or more PSA tests and select the baseline in the list with the following fields:
    - PSA Total (ng/ml): decimal value > 0, 'dot' is used as a decimal separator;
    - PSA Free (ng/ml): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - PSA Ratio (%): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - Date: date of the PSA test;
    - Baseline: select which of the PSA tests in the list is the baseline – there must be one.
- Complete the family history considering the patient's 1st degree relatives with:
    - Prostate Cancer: true (checked), false (unchecked), unknown (indeterminate, default value) – optional field;
    - Age: integer number with the age of 1st degree relative with prostate cancer – optional field;
    - Breast Cancer: true (checked), false (unchecked), unknown (indeterminate, default value) – optional field;
    - Age: integer number with the age of 1st degree relative with breast cancer – optional field;
    - Aggressive cancer: if any 1st degree relative had aggressive cancer - true (checked), false (unchecked), unknown (indeterminate, default value) – optional field.

There is a delete button at the bottom of the table to remove the selected PSA. To select a PSA, click on any field of the row containing the appropriate PSA.

In the Lesions section, the user should complete the following information:

- MRI positive: true (checked) or false (not checked);
- MRI Date: Date of the MRI scan;
- MRI T classification: select the T classification from the drop-down list according to the TNM Staging System;
- MRI N classification: select the N classification from the drop-down list according to the TNM Staging System– optional field;

- Biopsy performed: true (checked) or false (not checked);
- Biopsy positive: true (checked) or false (not checked) – required only if the biopsy was performed;
- Biopsy T classification: select the T classification from the drop-down list according to the TNM Staging System– optional field;
- Type: select the type of biopsy performed from the drop-down list – required only if the biopsy was performed;
- Carcinoma classification: select the carcinoma classification from the drop-down list based on the biopsy or prostatectomy information;
- Prostatectomy performed: true (checked) or false (not checked);
- Prostatectomy T classification: select the T classification from the drop-down list according to the TNM Staging System;
- Prostatectomy N classification: select the N classification from the drop-down list according to the TNM Staging System– optional field;
- Insert one or more lesions and select the index lesion in the list with the following fields:
    - Max diam1 (mm): maximum diameter 1 in mm, integer number – optional field;
    - Max diam2 (mm): maximum diameter 2 in mm, integer number – optional field;
    - Volume (cc): decimal number with the volume of the lesion in cm^3, 'dot' is used as a decimal separator – optional field;
    - Location: double click on that field to select the location on the specific Lesion Location map window;
    - X, Y, Z: decimal numbers for x, y, and z coordinates of the lesion, 'dot' is used as a decimal separator – optional field – these coordinates are only required when there is more than one lesion in the same location;
    - PI-RADS: integer number from 1 to 5 (0 if unknown) - score indicating the likelihood of clinically significant cancer as detailed in the section PI-RADS;
    - Gleason Score 1 and 2: integer numbers of 1 to 5 (0 if unknown) that compose the primary and secondary grades of the Gleason Score;
    - Index lesion: select which of the lesions in the list is the index lesion – there must be one.
- Lesions Coordinate System: Select the coordinate system used for the x, y, and z coordinates of the lesion, either Physical or Pixel coordinates.

Provide the Gleason Score, location, and volume based on Prostatectomy, otherwise location, and Gleason Score based on Biopsy.

The X, Y, and Z lesion coordinates should be retrieved from T2W axial series, even with visual inspection of DWI.

If PI-RADS score is not available (or the acquisition protocol is not PI-RADS compliant) you should use the Likert scale.

There is a delete button at the bottom of the table to remove the selected lesion. To select, click on any field of the row containing the appropriate lesion.

In the section for the Final diagnosis established within 6 months, the user should provide the following information:

- Months after: integer number with a number of months to establish the final diagnosis;
- N and M classification: select the N and M from the drop-down lists according to the TNM Staging System;
- Metastasis: select all metastasis locations that apply;
- Image modalities: select all image modalities used for the final diagnosis.

TNM Staging System

The TNM system is the most widely used cancer staging system. In the TNM system:

- The T refers to the size and extent of the main tumour. The main tumour is usually called the primary tumour.
- The N refers to the number of nearby lymph nodes that have cancer.
- The M refers to whether cancer has metastasized. This means that cancer has spread from the primary tumour to other parts of the body.

According to the EAU-EANM-ESTRO-ESUR-ISUP-SIOG Guidelines on Prostate Cancer follows the 2017 Tumour Node Metastasis (TNM) classification used for staging:

T - Primary Tumour (stage based on digital rectal examination [DRE] only):

- TX Primary tumour cannot be assessed
- T0 No evidence of primary tumour
- T1 Clinically inapparent tumour that is not palpable
    - T1a Tumour incidental histological finding in 5% or less of tissue resected
    - T1b Tumour incidental histological finding in more than 5% of tissue resected
    - T1c Tumour identified by needle biopsy (e.g., because of elevated PSA)
- T2 Tumour that is palpable and confined within the prostate
    - T2a Tumour involves one half of one lobe or less
    - T2b Tumour involves more than half of one lobe, but not both lobes
    - T2c Tumour involves both lobes

- T3 Tumour extends through the prostatic capsule
    - T3a Extracapsular extension (unilateral or bilateral)
    - T3b Tumour invades seminal vesicle(s)
- T4 Tumour is fixed or invades adjacent structures other than seminal vesicles: external sphincter, rectum, levator muscles, and/or pelvic wall

N - Regional (pelvic) Lymph Nodes:

- NX Regional lymph nodes cannot be assessed
- N0 No regional lymph node metastasis
- N1 Regional lymph node metastasis

M - Distant Metastasis:

- M0 No distant metastasis
- M1 Distant metastasis
    - M1a Non-regional lymph node(s)
    - M1b Bone(s)
    - M1c Other site(s)

When more than one site of metastasis is present, the most advanced category is used. (p)M1c is the most advanced category.

Pathological staging (pTNM) is based on histopathological tissue assessment and largely parallels the clinical TNM, except for clinical stage T1c and the T2 substages. All histopathologically confirmed organ-confined PCas after radical prostatectomy (RP) are pathological stage T2 and the current UICC no longer recognizes pT2 substages.

Form 1 + 2 + 5 + 9

Collection of patients who have undergone prostatectomy, with or without biochemical relapse.

This form collects positive cases to Use Cases 1, 2, 5, and 9.

This form has three information sections: Clinical, Lesions, and Follow-up.

In the Clinical section, the user should complete the following information:

- Age at baseline: insert an integer number;
- DRE: insert the result of the Digital Rectal Examination – Positive, Negative, or Not Assessed – optional field;

- Biopsy before MRI: true (checked) or false (unchecked);
- Previous adenomectomy: true (checked) or false (unchecked);
- Insert one or more PSA tests and select the baseline in the list with the following fields:
    - PSA Total (ng/ml): decimal value > 0, 'dot' is used as a decimal separator;
    - PSA Free (ng/ml): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - PSA Ratio (%): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - Date: date of the PSA test;
    - Baseline: select which of the PSA tests in the list is the baseline – there must be one.
- Complete the family history considering patient's 1st degree relatives with:
    - Prostate Cancer: true (checked), false (unchecked), unknown (indeterminate, default value) – optional field;
    - Age: integer number with the age of 1st degree relative with prostate cancer – optional field;
    - Breast Cancer: true (checked), false (unchecked), unknown (indeterminate, default value) – optional field;
    - Age: integer number with the age of 1st degree relative with breast cancer – optional field;
    - Aggressive cancer: if any 1st degree relative had aggressive cancer - true (checked), false (unchecked), unknown (indeterminate, default value) – optional field.

There is a delete button at the bottom of the table to remove the selected PSA. To select a PSA, click on any field of the row containing the appropriate PSA.

In the Lesions section, the user should complete the following information:

- MRI Date: Date of the MRI scan;
- Prostatectomy method: select the method of prostatectomy performed from the drop-down list;
- Nerve sparing: the prostatectomy method used is sparing the nerve - true (checked) or false (not checked);
- Prostatectomy T and N classification: select the Tumour and Node classification from the drop-down lists according to the TNM Staging System;
- Carcinoma classification: select the carcinoma classification from the drop-down list;
- Insert one or more lesions and select the index lesion in the list with the following fields:

- Max diam1 (mm): maximum diameter 1 in mm, integer number – optional field;
- Max diam2 (mm): maximum diameter 2 in mm, integer number – optional field;
- Volume (cc): decimal number with the volume of the lesion in cm^3, 'dot' is used as decimal separator – optional field;
- Location: double click on that field to select the location on the specific Lesion Location map window;
- X, Y, Z: decimal numbers for x, y, and z coordinates of the lesion, 'dot' is used as decimal separator – optional field – these coordinates are only required when there is more than one lesion in the same location;
- PI-RADS: integer number from 1 to 5 (0 if unknown) - score indicating the likelihood of clinically significant cancer as detailed in the section PI-RADS;
- Gleason Score 1 and 2: integer numbers of 1 to 5 (0 if unknown) that compose the primary and secondary grades of the Gleason Score;
- Index lesion: select which of the lesions in the list is the index lesion – there must be one.

- Lesions Coordinate System: Select the coordinate system used for the x, y, and z coordinates of the lesion, either Physical or Pixel coordinates.
- Resection margins status: select Positive or Negative from the drop-down list;
- Extra prostatic extension: select Present or Absent from the drop-down list;
- Perineural invasion: select Present or Absent from the drop-down list;
- Seminal vesical invasion: select Present or Absent from the drop-down list;
- Wheeler Ranking: select the score from the drop-down list according to the Wheeler Ranking – optional field.

Provide the Gleason Score, location, and volume based on Prostatectomy.

The X, Y, and Z lesion coordinates should be retrieved from T2W axial series, even with visual inspection of DWI.

If PI-RADS score is not available (or the acquisition protocol is not PI-RADS compliant) you should use the Likert scale.

There is a delete button at the bottom of the table to remove the selected lesion. To select, click on any field of the row containing the appropriate lesion.

In the section for the Follow-up the user should provide the following information:

- Biochemical Relapse: true (checked) or false (not checked).
- Insert one or more follow-up PSA tests and select the baseline in the list with the following fields:

- PSA Total (ng/ml): decimal value > 0;
- PSA Free (ng/ml): decimal value > 0 – optional field;
- PSA Ratio (%): decimal value > 0 – optional field;
- Date: date of the PSA test;

Wheeler Ranking

Pathologic staging system (the pathological Wheeler staging system [pW]) for the local assessment of PCa.

Table 7 The wheeler ranking class definition

Prostate cancer status	Wheeler class	Definition
Not assessed	**LX**	Wheeler Ranking not assessed or unknown
Organ-confined (T2 group)	**L0**	Tumour does not touch the fibrous capsule
	L1	Tumour has minimal contact with the fibrous capsule without signs of capsule alteration
	L2	Broad contact between the tumour and the capsule and/or a margin bulging and/or inhomogeneity of fibrous capsule
Non-organ-confined (T3 group)	**L3**	Tumour invading the periprostatic adipose tissue or smooth muscle of bladder neck
	L3F (focal)	Slight irregularity/spiculature of the margin of the fibrous capsule
	L3E (extended)	Discontinuity of fibrous capsule with evidence of periprostatic fat inhomogeneity or a focal hypointense lesion in one or both seminal vesicles

The following image illustrates the Wheeler classes.

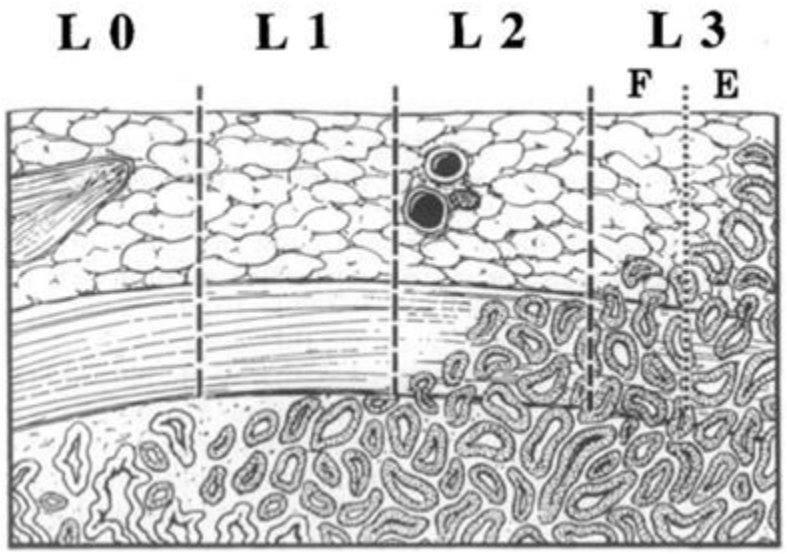


Figure 42 illustration of the wheeler classes

Form 1 + 2 + 6 + 7a + 9

Collection of patients who underwent radiation therapy, with/without biochemical relapse and with post-treatment (toxicity) data.

This form collects positive cases to Use Cases 1, 2, 6, 7a, and 9.

This form has four information sections: Clinical, Lesions, Treatment, and Follow-up.

In the Clinical section, the user should complete the following information:

- Age at baseline: insert an integer number;
- DRE: insert the result of the Digital Rectal Examination – Positive, Negative or Not Assessed – optional field;
- Biopsy before MRI: true (checked) or false (unchecked);
- Previous adenomectomy: true (checked) or false (unchecked);
- Insert one or more PSA tests and select the baseline in the list with the following fields:
    - PSA Total (ng/ml): decimal value > 0, 'dot' is used as a decimal separator;
    - PSA Free (ng/ml): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - PSA Ratio (%): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - Date: date of the PSA test;

  - Baseline: select which of the PSA tests in the list is the baseline – there must be one.
- Complete the family history considering the patient's 1st degree relatives with:
  - Prostate Cancer: true (checked), false (unchecked), unknown (indeterminate, default value) – optional field;
  - Age: integer number with the age of 1st degree relative with prostate cancer – optional field;
  - Breast Cancer: true (checked), false (unchecked), unknown (indeterminate, default value) – optional field;
  - Age: integer number with the age of 1st degree relative with breast cancer – optional field;
  - Aggressive cancer: if any 1st degree relative had aggressive cancer - true (checked), false (unchecked), unknown (indeterminate, default value) – optional field.

There is a delete button at the bottom of the table to remove the selected PSA. To select a PSA, click on any field of the row containing the appropriate PSA.

In the Lesions section, the user should complete the following information:

- MRI Date: Date of the MRI scan;
- MRI T and N classification: select the Tumour and Node classification from the drop-down lists according to the TNM Staging System;
- Biopsy T classification: select the Node classification from the drop-down list according to the TNM Staging System;
- Type: select the type of biopsy performed from the drop-down list;
- Carcinoma classification: select the carcinoma classification from the drop-down list;
- Insert one or more lesions and select the index lesion in the list with the following fields:
  - Max diam1 (mm): maximum diameter 1 in mm, integer number – optional field;
  - Max diam2 (mm): maximum diameter 2 in mm, integer number – optional field;
  - Volume (cc): decimal number with the volume of the lesion in cm^3, 'dot' is used as decimal separator – optional field;
  - Location: double click on that field to select the location on the specific Lesion Location map window;
  - X, Y, Z: decimal numbers for x, y, and z coordinates of the lesion, 'dot' is used as decimal separator – optional field – these coordinates are only required when there is more than one lesion in the same location;

  - PI-RADS: integer number from 1 to 5 (0 if unknown) - score indicating the likelihood of clinically significant cancer as detailed in the section PI-RADS;
  - Gleason Score 1 and 2: integer numbers of 1 to 5 (0 if unknown) that compose the primary and secondary grades of the Gleason Score;
  - Index lesion: select which of the lesions in the list is the index lesion – there must be one.
- Lesions Coordinate System: Select the coordinate system used for the x, y, and z coordinates of the lesion, either Physical or Pixel coordinates.

Provide the location and Gleason Score based on Biopsy.

The X, Y, and Z lesion coordinates should be retrieved from T2W axial series, even with visual inspection of DWI.

If PI-RADS score is not available (or the acquisition protocol is not PI-RADS compliant) the user should use the Likert scale.

There is a delete button at the bottom of the table to remove the selected lesion. To select, click on any field of the row containing the appropriate lesion.

In the section for the Treatment the user should provide the following information:

- Toxicity data:
  - Total dose (Gy): decimal number of total radiation dose in Gy, 'dot' is used as a decimal separator;
  - Fractions/week: integer number of fractions per week – optional field;
  - Dose/fraction (Gy): decimal number of dose per fraction, 'dot' is used as a decimal separator;
  - Radiotherapy Technique: select from the drop-down list;
  - Rectal Acute: select grade 1 to 5 from the drop-own list;
  - Rectal Chronic: select grade 1 to 5 from the drop-own list;
  - Genitourinary Acute: select grade 1 to 5 from the drop-own list;
  - Genitourinary Chronic: select grade 1 to 5 from the drop-own list;
- Hormonotherapy: true (checked) or false (not checked);
- Post-treatment QoL EPIC-26: integer number between 0 and 100 with a final score of the QoL EPIC-26 (Expanded Prostate Cancer Index Composite).

In the section for the Follow-up the user should provide the following information:

- Biochemical Relapse: true (checked) or false (not checked).
- Insert one or more follow-up PSA tests and select the baseline in the list with the following fields:
    - PSA Total (ng/ml): decimal value > 0, 'dot' is used as a decimal separator;
    - PSA Free (ng/ml): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - PSA Ratio (%): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - Date: date of the PSA test;

Form 1 + 2 + 5 + 7b + 9

Collection of patients with prostatectomy performed, with/without biochemical relapse and with post-treatment quality of life data.

This form collects positive cases to Use Cases 1, 2, 5, 7b, and 9.

This form has four information sections: Clinical, Lesions, Post-treatment, and Follow-up.

In the Clinical section, the user should complete the following information:

- Age at baseline: insert an integer number;
- DRE: insert the result of the Digital Rectal Examination – Positive, Negative, or Not Assessed – optional field;
- Biopsy before MRI: true (checked) or false (unchecked);
- Previous adenomectomy: true (checked) or false (unchecked);
- Insert one or more PSA tests and select the baseline in the list with the following fields:
    - PSA Total (ng/ml): decimal value > 0, 'dot' is used as a decimal separator;
    - PSA Free (ng/ml): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - PSA Ratio (%): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - Date: date of the PSA test;
    - Baseline: select which of the PSA tests in the list is the baseline – there must be one.
- Complete the family history considering the patient's 1st degree relatives with:

- Prostate Cancer: true (checked), false (unchecked), unknown (indeterminate, default value) – optional field;
- Age: integer number with the age of 1st degree relative with prostate cancer – optional field;
- Breast Cancer: true (checked), false (unchecked), unknown (indeterminate, default value) – optional field;
- Age: integer number with the age of 1st degree relative with breast cancer – optional field;
- Aggressive cancer: if any 1st degree relative had aggressive cancer - true (checked), false (unchecked), unknown (indeterminate, default value) – optional field.

There is a delete button at the bottom of the table to remove the selected PSA. To select a PSA, click on any field of the row containing the appropriate PSA.

In the Lesions section, the user should complete the following information:

- MRI Date: Date of the MRI scan;
- Prostatectomy performed: true (checked) or false (not checked);
- Method: select the method of prostatectomy performed from the drop-down list – required only if the prostatectomy was performed;
- Nerve sparing: the prostatectomy method used is sparing the nerve - true (checked) or false (not checked);
- T and N classification: select the Tumour and Node classification from the drop-down lists according to the TNM Staging System;
- Carcinoma classification: select the carcinoma classification from the drop-down list;
- Insert one or more lesions and select the index lesion in the list with the following fields:
  - Max diam1 (mm): maximum diameter 1 in mm, integer number – optional field;
  - Max diam2 (mm): maximum diameter 2 in mm, integer number – optional field;
  - Volume (cc): decimal number with the volume of the lesion in cm^3, 'dot' is used as a decimal separator – optional field;
  - Location: double click on that field to select the location on the specific Lesion Location map window;
  - X, Y, Z: decimal numbers for x, y, and z coordinates of the lesion, 'dot' is used as a decimal separator – optional field – these coordinates are only required when there is more than one lesion in the same location;
  - PI-RADS: integer number from 1 to 5 (0 if unknown) - score indicating the likelihood of clinically significant cancer as detailed in the section PI-RADS;

  - Gleason Score 1 and 2: integer numbers of 1 to 5 (0 if unknown) that compose the primary and secondary grades of the Gleason Score;
  - Index lesion: select which of the lesions in the list is the index lesion – there must be one.
- Lesions Coordinate System: Select the coordinate system used for the x, y, and z coordinates of the lesion, either Physical or Pixel coordinates.
- Resection margins status: select Positive or Negative from the drop-down list;
- Extra prostatic extension: select Present or Absent from the drop-down list;
- Perineural invasion: select Present or Absent from the drop-down list;
- Seminal vesical invasion: select Present or Absent from the drop-down list;
- Wheeler Ranking: select the score from the drop-down list according to the Wheeler Ranking – optional field.

Provide the Gleason Score, location, and volume based on Prostatectomy.

The X, Y, and Z lesion coordinates should be retrieved from T2W axial series, even with visual inspection of DWI.

If PI-RADS score is not available (or the acquisition protocol is not PI-RADS compliant) the user should use the Likert scale.

There is a delete button at the bottom of the table to remove the selected lesion. To select, click on any field of the row containing the appropriate lesion.

In the Post-treatment section, the user should provide the following information:

- QoL EPIC-26: integer number between 0 and 100 with a final score of the QoL EPIC-26 (Expanded Prostate Cancer Index Composite);
- Insert the answers to EORTC - QLQ - PR25 questions (score 1 – not at all, to 4 – very much):
  - Q52. To what extent was sex enjoyable for you? – optional field;
  - Q53. Did you have difficulty getting or maintaining an erection? – optional field;
  - Q54. Did you have ejaculation problems (e.g. dry ejaculation)? – optional field;
  - Q55. Have you felt uncomfortable about being sexually intimate? – optional field.

In the section for the Follow-up the user should provide the following information:

- Biochemical Relapse: true (checked) or false (not checked).
- Insert one or more follow-up PSA tests and select the baseline in the list with the following fields:

- PSA Total (ng/ml): decimal value > 0, 'dot' is used as a decimal separator;
- PSA Free (ng/ml): decimal value > 0, 'dot' is used as a decimal separator – optional field;
- PSA Ratio (%): decimal value > 0, 'dot' is used as a decimal separator – optional field;
- Date: date of the PSA test;

Form 1 + 2 + 8 + 9

Collection of patients with confirmed Prostate Cancer at biopsy enrolled in active surveillance programs.

This form collects positive cases to Use Cases 1, 2, 8, and 9.

This form has three information sections: Clinical, Lesions, and Follow-up.

In the Clinical section, the user should complete the following information:

- Age at baseline: insert an integer number;
- DRE: insert the result of the Digital Rectal Examination – Positive, Negative, or Not Assessed – optional field;
- Biopsy before MRI: true (checked) or false (unchecked);
- Previous adenomectomy: true (checked) or false (unchecked);
- Insert one or more PSA tests and select the baseline in the list with the following fields:
  - PSA Total (ng/ml): decimal value > 0, 'dot' is used as a decimal separator;
  - PSA Free (ng/ml): decimal value > 0, 'dot' is used as a decimal separator – optional field;
  - PSA Ratio (%): decimal value > 0, 'dot' is used as a decimal separator – optional field;
  - Date: date of the PSA test;
  - Baseline: select which of the PSA tests in the list is the baseline – there must be one.
- Complete the family history considering the patient's 1st degree relatives with:
  - Prostate Cancer: true (checked), false (unchecked), unknown (indeterminate, default value) – optional field;
  - Age: integer number with the age of 1st degree relative with prostate cancer – optional field;
  - Breast Cancer: true (checked), false (unchecked), unknown (indeterminate, default value) – optional field;

- Age: integer number with the age of 1st degree relative with breast cancer – optional field;
- Aggressive cancer: if any 1st degree relative had aggressive cancer - true (checked), false (unchecked), unknown (indeterminate, default value) – optional field.

There is a delete button at the bottom of the table to remove the selected PSA. To select a PSA, click on any field of the row containing the appropriate PSA.

In the Lesions section, the user should complete the following information:

- MRI positive: true (checked) or false (not checked);
- MRI Date: Date of the MRI scan;
- T and N classification: select the Tumour and Node classification from the drop-down lists according to the TNM Staging System;
- Biopsy positive: true (checked) or false (not checked);
- Type: select the type of biopsy performed from the drop-down list – required only if the biopsy was performed;
- Carcinoma classification: select the carcinoma classification from the drop-down list;
- Insert one or more lesions and select the index lesion in the list with the following fields:
  - Max diam1 (mm): maximum diameter 1 in mm, integer number – optional field;
  - Max diam2 (mm): maximum diameter 2 in mm, integer number – optional field;
  - Volume (cc): decimal number with the volume of the lesion in cm^3, 'dot' is used as decimal separator – optional field;
  - Location: double click on that field to select the location on the specific Lesion Location map window;
  - X, Y, Z: decimal numbers for x, y, and z coordinates of the lesion, 'dot' is used as decimal separator – optional field – these coordinates are only required when there is more than one lesion in the same location;
  - PI-RADS: integer number from 1 to 5 (0 if unknown) - score indicating the likelihood of clinically significant cancer as detailed in the section PI-RADS;
  - Gleason Score 1 and 2: integer numbers of 1 to 5 (0 if unknown) that compose the primary and secondary grades of the Gleason Score;
  - Index lesion: select which of the lesions in the list is the index lesion – there must be one.

Lesions Coordinate System: Select the coordinate system used for the x, y, and z coordinates of the lesion, either Physical or Pixel coordinates.

Provide the location and Gleason Score based on Biopsy.

The X, Y, and Z lesion coordinates should be retrieved from T2W axial series, even with visual inspection of DWI.

If PI-RADS score is not available (or the acquisition protocol is not PI-RADS compliant) the user should use the Likert scale.

There is a delete button at the bottom of the table to remove the selected lesion. To select, click on any field of the row containing the appropriate lesion.

In the section for the Follow-up, the user can add several follow-up records. After inserting the first the user can click the "Add" button to add new follow-up records. To navigate between different records, select the record from the drop-down list. To update the information on the current follow-up, click the "update" button. To delete the currently selected record, click the "remove" button. The reset "button" empties all fields of the follow-up to allow starting from a clean point, only affecting the existing records after using "add" for adding as a new record or "update" to update the previously selected record.

For each follow-up record the user should provide the following information:

- Insert one or more follow-up PSA tests and select the baseline in the list with the following fields:
    - PSA Total (ng/ml): decimal value > 0, 'dot' is used as a decimal separator;
    - PSA Free (ng/ml): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - PSA Ratio (%): decimal value > 0, 'dot' is used as a decimal separator – optional field;
    - Date: date of the PSA test;
    - Baseline: select which of the PSA tests in the list is the baseline – there must be one.
- Follow-up PSA confirmed within normal range: true (checked) or false (not checked);
- MRI performed: true (checked) or false (not checked);
- MRI confirmed negative: true (checked) or false (not checked);
- Date: Date of the MRI scan;
- Likelihood of progression: select from the Drop-down list the Likert value for the Assessment of the likelihood of radiologic progression;
- T and N classification: select the Tumour and Node classification from the drop-down lists according to the TNM Staging System;

- Biopsy Performed: true (checked) or false (not checked);
- Biopsy confirmed negative: true (checked) or false (not checked);
- Type: select the type of biopsy performed from the drop-down list – required only if the biopsy was performed;
- Date: Date of the Biopsy;
- T classification: select the Tumour classification from the drop-down list according to the TNM Staging System;
- Carcinoma classification: select the carcinoma classification from the drop-down list.
- Insert one or more lesions and select the index lesion in the list with the following fields:
    - Max diam1 (mm): maximum diameter 1 in mm, integer number – optional field;
    - Max diam2 (mm): maximum diameter 2 in mm, integer number – optional field;
    - Volume (cc): decimal number with the volume of the lesion in cm^3, 'dot' is used as a decimal separator – optional field;
    - Location: double click on that field to select the location on the specific Lesion Location map window;
    - X, Y, Z: decimal numbers for x, y, and z coordinates of the lesion, 'dot' is used as a decimal separator – optional field – these coordinates are only required when there is more than one lesion in the same location, using the same coordinate system selected for first MRI;
    - PI-RADS: integer number from 1 to 5 (0 if unknown) - score indicating the likelihood of clinically significant cancer as detailed in the section PI-RADS;
    - Gleason Score 1 and 2: integer numbers of 1 to 5 (0 if unknown) that compose the primary and secondary grades of the Gleason Score;
    - Index lesion: select which of the lesions in the list is the index lesion – there must be one.

Provide the location and Gleason Score based on Biopsy.

The X, Y, and Z lesion coordinates should be retrieved from T2W axial series, even with visual inspection of DWI.

If PI-RADS score is not available (or the acquisition protocol is not PI-RADS compliant) you should use the Likert scale.

Assessment of the likelihood of radiologic progression

The Table below provides guidance and examples regarding the assessment of the likelihood of radiological progression.

Table 8 Table of likelihood of radiological progression

Likert	Assessment of the likelihood of radiologic progression	Example
1	Resolution of previous features suspicious on MRI	Previously enhancing area no longer enhances
2	Reduction in volume and/or conspicuity of previous features suspicious on MRI	Reduction in size of previously seen lesion that remains suspicious for clinically significant disease
3	Stable MRI appearance: no new focal/diffuse lesions	Either no suspicious features or all lesions stable in size and appearance
4	Significant increase in size and/or conspicuity of features suspicious for prostate cancer	The lesion becomes visible on diffusion-weighted imaging; a significant increase in the size of the previously seen lesion
5	Definitive radiologic stage progression	The appearance of extracapsular extension, seminal vesicle involvement, lymph node involvement, or bone metastasis

A.2 ProCAncer-I eCRF tool – batch mode operation

Form 1+2

Table 9 Form 1+2 JSON objects description

Field	Type	Optional	Description
age	string	no	Integer number converted to string
dre	string	yes	Digital Rectal Examination. One possible value from: "Positive", "Negative", "Not Assessed"
biopsyBeforeMRI	bool	no	A biopsy was performed before MRI. true, false

Field	Type	Optional	Description
previousAdenomectomy	bool	no	An Adenomectomy was performed before MRI. true, false
psas	List<PSA>	no	List of PSA objects. There must be one object with the 'baseline' field set to 'true' (see PSA tab)
mriPositive	bool	no	MRI confirmed cancer. true, false
mriDate	string/DateTIme	no	Date of the MRI (e.g., "2021-12-21T00:00:00Z")
biopsyPerformed	bool	no	Biopsy was performed. true, false
biopsyPositive	bool	no	Biopsy confirmed cancer. true, false
biopsyType	string	no	One possible value from: "Systematic", "Fusion", "Systematic+Fusion", "In-Bore"
prostatectomyPerformed	bool	no	Prostatectomy was performed. true, false
lesions	List<Lesion>	no	List of Lesion objects. There must be one object with the 'index_lesion' field set to 'true' (see Lesion tab)
lesionsCoordinateSystem	string	yes	One possible value from: "Physical", "Pixel"

Field	Type	Optional	Description
lesionLocationBasedOn	string	no	One possible value from: "Prostatectomy", "Biopsy"
prospective	bool	no	Prospective or Retrospective data
useCaseType	string	no	Use case identifier. Mandatory value: "UC1+2"

Example:


```
{
 "age": "63",
 "dre": "Positive",
 "biopsyBeforeMRI": true,
 "previousAdenomectomy": false,
 "psas": [
 {
 "total": 9,
 "free": 0,
 "ratio": 0,
 "date": "2022-01-04T00:00:00Z",
 "baseline": true
 },
 {
 "total": 9.2,
 "free": 0,
 "ratio": 0,
 "date": "2022-01-04T00:00:00Z",
 "baseline": false
 }
 ],
 "mriPositive": true,
 "mriDate": "2022-01-04T00:00:00",
 "biopsyPerformed": true,
 "biopsyPositive": true,
 "biopsyType": "Systematic",
```

```
 "prostatectomyPerformed": false,
 "lesions": [
 {
 "volume": 0,
 "gleason1": 4,
 "gleason2": 4,
 "diam1": 0,
 "diam2": 0,
 "location": [
      "BLTZp"
 ],
 "x": 0,
 "y": 0,
 "z": 0,
 "pi_rads": 4,
 "index_lesion": true
 }
 ],
 "lesionsCoordinateSystem": "Dicom",
 "lesionLocationBasedOn": "Prostatectomy",
 "prospective": true,
 "useCaseType": "UC1+2"
}
```


Lesion Object

The Lesion object is used in several forms, to record the lesion data.

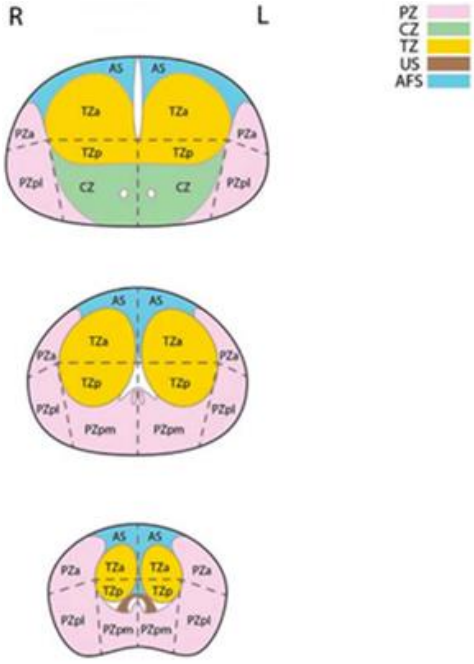
Table 10 Lesion JSON objects description

Field	Type	Optional	Description
volume	float	yes	Volume in mm3, 'dot' is used as decimal separator
gleason1	int	no	Gleason Score part 1
gleason2	int	no	Gleason Score part 2
diam1	int	yes	Maximum diameter 1 in mm
diam2	int	yes	Maximum diameter 2 in mm

Field	Type	Optional	Description
location	List<string>	no	Possible values: each string corresponds to a possible Base, Mid and/or Apex location (see "Locations" below)
x	float	yes	x coordinate, 'dot' is used as decimal separator
y	float	yes	y coordinate, 'dot' is used as decimal separator
z	float	yes	z coordinate, 'dot' is used as decimal separator
pi_rads	int	no	PI_RADS v2.1
index_lesion	bool	no	Index lesion. true, false

Lesion coordinates XYZ are based on MRI using either DICOM or Image coordinates.

Locations



Locations are used in the lesion object.

Locations are based on the PI-RADS location map.

The first letter is B, M, or A to specify the Base, Mid, or Apex.

The second letter is R or L to specify Right or Left.

The following letters are the different zones identified in the PI-RADS map: AS, TZa, TZp, PZa, PZpl, PZpm, and CZ.

Base BLAS, BLCZ, BLPZa, BLPZpl, BLTZa, BLTZp, BRAS, BRCZ, BRPZa, BRPZpl, BRTZa, BRTZp

Mid MLAS, MLPZa, MLPZpl, MLPZpm, MLTZa, MLTZp, MRAS, MRPZa, MRPZpl, MRPZpm, MRTZa, MRTZp

Apex ALAS, ALPZa, ALPZpl, ALPZpm, ALTZa, ALTZp, ARAS, ARPZa, ARPZpl, ARPZpm, ARTZa, ARTZp

Form 1+2+3

Table 11 Form 1+2+3 JSON objects description

Field	Type	Optional	Description
age	string	no	Integer number converted to string
dre	string	yes	Digital Rectal Examination. Possible Values: "Positive", "Negative", "Not Assessed"
biopsyBeforeMRI	bool	no	A biopsy was performed before MRI. true, false
previousAdenomectomy	bool	no	An Adenomectomy was performed before MRI. true, false
psas	List<PSA>	no	List of PSA objects. There must be one object with the 'baseline' field set to 'true' (see PSA sheet)
relativePCa	string/bool	yes	First degree relative with Prostate Cancer. One possible value from: "True", "False", "Unknown"
relativePCaAge	string	yes	Age of first degree relative when diagnosed with Prostate Cancer.
relativeBCa	string/bool	yes	First degree relative with Breast Cancer. One possible value from: "True", "False", "Unknown"
relativeBCaAge	string	yes	Age of first degree relative when diagnosed with Breast Cancer.
relativeAggressiveCancer	string/bool	yes	First degree relative with Aggressive Cancer. One possible value from: "True", "False", "Unknown"

Field	Type	Optional	Description
mriDate	string/DateTIme	no	Date of the MRI (e.g., "2021-12-21T00:00:00Z")
mriPositive	bool	no	MRI confirmed cancer. true, false
biopsyPerformed	bool	no	A biopsy was performed. true, false
biopsyPositive	bool	no	Biopsy confirmed cancer. true, false
biopsyType	string	no	One possible value from: "Systematic", "Fusion", "Systematic+Fusion", "In-Bore"
carcinomaClass	string	no	One possible value from: "Acinar adenocarcinoma", "Intraductal carcinoma", "Ductal adenocarcinoma", "Urothelial carcinoma", "Adenosquamous carcinoma", "Squamous cell carcinoma", "Basal cell carcinoma", "Adenocarcinoma with neuroendocrine differentiation", "Small cell neuroendocrine carcinoma", "Large cell neuroendocrine carcinoma", "Not Assessed"
prostatectomyPerformed	bool	no	A prostatectomy was performed. true, false
lesions	List<Lesion>	no	List of Lesion objects. There must be one object with the 'index_lesion' field set to 'true' (see Lesion sheet)
lesionsCoordinateSystem	string	yes	One possible value from: "Physical", "Pixel"

Field	Type	Optional	Description
monthsFinalDiagnosis	int	no	Number of months to confirm the final diagnosis
T_MRI	string	no	One possible value from: "cTX", "cT2a", "cT2b", "cT2c", "cT3a", "cT3b", "cT4"
N_MRI	string	yes	One possible value from:"cN0", "cN1"
T_Biopsy	string	yes	One possible value from:"pTX", "pT1", "pT2"
T_Prostatectomy	string	no	One possible value from: “pT1”, "pT1a", "pT1b", "pT1c", “pT2”, "pT2a", "pT2b", "pT2c", “pT3”, "pT3a", "pT3b", "pT4"
N_Prostatectomy	string	yes	One possible value from:"pNX", "pN0", "pN1"
N_Diagnosis	string	no	One possible value from:"cNX", "cN0", "cN1"
M_Diagnosis	string	no	One possible value from:"MX", "cM0", "cM1a", "cM1b", "cM1c"
metastasis	List<string>	no	Possible values: "Bone", "Lymph nodes", "Liver", "Lung", "Brain", "Digestive system", "Retroperitoneum", "Kidney and adrenal gland", "Other"
modalities	List<string>	no	Possible values: "MRI", "CT", "chloline-PET", "fluciclovine-PET", "PSMA-PET","SPECT-CT", "NaF-PET", "Bone scan"
prospective	bool	no	Prospective (true) or Retrospective (false) data

Field	Type	Optional	Description
useCaseType	string	no	Use case identifier. Mandatory value: "UC1+2+3"

Example:


```
{
 "age": "63",
 "dre": "Positive",
 "biopsyBeforeMRI": true,
 "previousAdenomectomy": false,
 "psas": [
 {
 "total": 10,
 "free": 0,
 "ratio": 0,
 "date": "2022-01-05T00:00:00Z",
 "baseline": true
 },
 {
 "total": 10.2,
 "free": 0,
 "ratio": 0,
 "date": "2022-01-05T00:00:00Z",
 "baseline": false
 }
 ],
 "relativePCa": "Unknown",
 "relativePCaAge": "0",
 "relativeBCa": "Unknown",
 "relativeBCaAge": "0",
 "relativeAggressiveCancer": "Unknown",
 "mriDate": "2022-01-01T00:00:00",
 "mriPositive": true,
 "biopsyPositive": true,
 "biopsyPerformed": tue,
 "biopsyType": "Systematic",
 "carcinomaClass": "Intraductal carcinoma",
 "prostatectomyPerformed": false,
```

```
 "lesions": [
 {
 "volume": 0,
 "gleason1": 4,
 "gleason2": 5,
 "diam1": 0,
 "diam2": 0,
 "location": [
      "MLTZp",
      "MRTZp"
 ],
 "x": 0,
 "y": 0,
 "z": 0,
 "pi_rads": 4,
 "index_lesion": true
 }
 ],
 "lesionsCoordinateSystem": "Physical",
 "monthsFinalDiagnosis": 3,
 "T_MRI": "cT2b",
 "N_MRI": "",
 "T_Biopsy": "",
 "T_Prostatectomy": "pT1a",
 "N_Prostatectomy": "",
 "N_Diagnosis": "cN0",
 "M_Diagnosis": "cM1a",
 "metastasis": [
 "Bone",
 "Lymph nodes"
 ],
 "modalities": [
 "MRI",
 "CT",
 "Bone scan"
 ],
 "prospective": true,
 "useCaseType": "UC1+2+3"
}
```

Table 12 Form 1+2+5+9 JSON objects description

Field	Type	Optional	Description
age	string	no	Integer number converted to string
dre	string	yes	Digital Rectal Examination. Possible Values: "Positive", "Negative", "Not Assessed"
biopsyBeforeMRI	bool	no	A biopsy was performed before MRI. true, false
previousAdenomectomy	bool	no	An Adenomectomy was performed before MRI. true, false
psas	List<PSA>	no	List of PSA objects. There must be one object with the 'baseline' field set to 'true' (see PSA sheet)
relativePCa	string/bool	yes	First degree relative with Prostate Cancer. One possible value from: "True", "False", "Unknown"
relativePCaAge	string	yes	Age of first degree relative when diagnosed with Prostate Cancer.
relativeBCa	string/bool	yes	First degree relative with Breast Cancer. One possible value from: "True", "False", "Unknown"
relativeBCaAge	string	yes	Age of first degree relative when diagnosed with Breast Cancer.
relativeAggressiveCancer	string/bool	no	First degree relative with Aggressive Cancer. One possible value from: "True", "False", "Unknown"

Field	Type	Optional	Description
mriDate	string/ DateTime	no	Date of the MRI (e.g., "2021-12-21T00:00:00Z")
prostatectomyMethod	string	no	Method of prostatectomy. One possible value from: "RP retropubic", "Laparoscopic", "Robotic-assisted laparoscopic", "RP perineal", "Other"
prostatectomyNerveSparing	bool	no	Nerve Sparing Prostatectomy. true, false
carcinomaClass	string	no	One possible value from: "Acinar adenocarcinoma", "Intraductal carcinoma", "Ductal adenocarcinoma", "Urothelial carcinoma", "Adenosquamous carcinoma", "Squamous cell carcinoma", "Basal cell carcinoma", "Adenocarcinoma with neuroendocrine differentiation", "Small cell neuroendocrine carcinoma", "Large cell neuroendocrine carcinoma", "Not Assessed"
lesions	List<Lesion>	no	List of Lesion objects. There must be one object with the 'index_lesion' field set to 'true' (see Lesion sheet)
lesionsCoordinateSystem	string	yes	One possible value from: "Physical", "Pixel"
T_Prostatectomy	string	no	One possible value from: “pT1”, "pT1a", "pT1b", "pT1c", “pT2”, "pT2a", "pT2b", "pT2c", “pT3”, "pT3a", "pT3b", "pT4"

Field	Type	Optional	Description
N_Prostatectomy	string	no	One possible value from:"pNX", "pN0", "pN1"
resectionMarginsStatus	string	no	One possible value from:"Positive", "Negative", "Not Assessed"
extraProstaticExtension	string	no	One possible value from:"Present", "Absent", "Not Assessed"
perineuralInvasion	string	no	One possible value from:"Present", "Absent", "Not Assessed"
seminalVesicalInvasion	string	no	One possible value from:"Present", "Absent", "Not Assessed"
wheelerRanking	string	yes	One possible value from:"LX", "L0", "L1", "L2", "L3F", "L3E"
biochemicalRelapseFUp	bool	no	Biochemical Relapse on Follow-Up. true, false
psasFUp	List<PSA>	no	List of PSA objects (see PSA sheet)
prospective	bool	no	Prospective or Retrospective data
useCaseType	string	no	Use case identifier. Mandatory value: "UC1+2+5+9"

Example:


```
{
 "biopsyType": null,
 "age": "59",
 "dre": "Positive",
```

```
"biopsyBeforeMRI": true,
"previousAdenomectomy": false,
"psas": [
{
"total": 13,
"free": 0,
"ratio": 0,
"date": "2019-02-01T00:00:00",
"baseline": true
}
],
"relativePCa": "True",
"relativePCaAge": "62",
"relativeBCa": "True",
"relativeBCaAge": "50",
"relativeAggressiveCancer": "True",
"mriDate": "2019-03-12T00:00:00",
"prostatectomyPerformed": false,
"prostatectomyMethod": "Robotic-assisted laparoscopic",
"prostatectomyNerveSparing": true,
"carcinomaClass": "Acinar adenocarcinoma",
"lesions": [
{
"volume": 630,
"gleason1": 4,
"gleason2": 5,
"diam1": 20,
"diam2": 15,
"location": [
    "BLTZa",
    "BLTZp",
    "BRTZa",
    "MLTZa"
],
"x": 0,
"y": 0,
"z": 0,
"pi_rads": 4,
```

```
 "index_lesion": true
 }
 ],
 "lesionsCoordinateSystem": "Physical",
 "T_Prostatectomy": "pT1",
 "N_Prostatectomy": "pN0",
 "resectionMarginsStatus": "Negative",
 "extraProstaticExtension": "Present",
 "perineuralInvasion": "Absent",
 "seminalVesicalInvasion": "Absent",
 "wheelerRanking": "LX",
 "biochemicalRelapseFUp": true,
 "psasFUp": [
 {
 "total": 15,
 "free": 0,
 "ratio": 0,
 "date": "2019-03-01T00:00:00",
 "baseline": false
 }
 ],
 "prospective": false,
 "useCaseType": "U1+2+5+9"
}
```


Form 1+2+6+7a+9

Table 13 Form 1+2+6+7a+9 JSON objects description

	Type	Optional	Description
age	string	no	Integer number converted to string
dre	string	yes	Digital Rectal Examination. Possible Values: "Positive", "Negative", "Not Assessed"
biopsyBeforeMRI	bool	no	A biopsy was performed before MRI. true, false

	Type	Optional	Description
previousAdenomectomy	bool	no	An Adenomectomy was performed before MRI. true, false
psas	List<PSA>	no	List of PSA objects. There must be one object with the 'baseline' field set to 'true' (see PSA sheet)
relativePCa	string/bool	yes	First degree relative with Prostate Cancer. One possible value from: "True", "False", "Unknown"
relativePCaAge	string	yes	Age of first degree relative when diagnosed with Prostate Cancer.
relativeBCa	string/bool	yes	First degree relative with Breast Cancer. One possible value from: "True", "False", "Unknown"
relativeBCaAge	string	yes	Age of first degree relative when diagnosed with Breast Cancer.
relativeAggressiveCancer	string/bool	yes	First degree relative with Aggressive Cancer. One possible value from: "True", "False", "Unknown"
mriDate	string/DateTIme	no	Date of the MRI (e.g., "2021-12-21T00:00:00Z")
biopsyType	string	no	One possible value from: "Systematic", "Fusion", "Systematic+Fusion", "In-Bore"
carcinomaClass	string	no	One possible value from: "Acinar adenocarcinoma", "Intraductal carcinoma", "Ductal adenocarcinoma", "Urothelial carcinoma", "Adenosquamous carcinoma", "Squamous cell carcinoma", "Basal cell

	Type	Optional	Description
			carcinoma", "Adenocarcinoma with neuroendocrine differentiation", "Small cell neuroendocrine carcinoma", "Large cell neuroendocrine carcinoma", "Not Assessed"
lesions	List<Lesion>	no	List of Lesion objects. There must be one object with the 'index_lesion' field set to 'true' (see Lesion sheet)
lesionsCoordinateSystem	string	yes	One possible value from: "Physical", "Pixel"
T_MRI	string	no	One possible value from: "cTX", "cT2a", "cT2b", "cT2c", "cT3a", "cT3b", "cT4"
N_MRI	string	yes	One possible value from:"cNX", "cN0", "cN1"
T_Biopsy	string	no	One possible value from:"pTX", "pT1", "pT2"
totalDose	float	no	Treatment total dose in Gy, 'dot' is used as decimal separator
fractionsWeek	int	yes	Number of fractions per week
doseFraction	float	no	Treatment dose per fraction in Gy, 'dot' is used as decimal separator
radiotherapyTechnique	string	no	One possible value from: "IMRT", "IGRT", "3DCRT", "HSR", "Other"
hormonotherapy	bool	no	Hormonotherapy treatment. true, false
rectalToxAcute	string	no	One possible value from: "Grade 1", "Grade 2", "Grade 3", "Grade 4", "Grade 5"

	Type	Optional	Description
rectalToxChronic	string	no	One possible value from: "Grade 1", "Grade 2", "Grade 3", "Grade 4", "Grade 5"
genitourinaryToxAcute	string	no	One possible value from: "Grade 1", "Grade 2", "Grade 3", "Grade 4", "Grade 5"
genitourinaryToxChronic	string	no	One possible value from: "Grade 1", "Grade 2", "Grade 3", "Grade 4", "Grade 5"
epic26	int	no	Post-treatment QoL EPIC-26
biochemicalRelapseFUp	bool	no	Biochemical Relapse on Follow-Up. true, false
psasFUp	List<PSA>	no	List of PSA objects (see PSA sheet)
prospective	bool	no	Prospective or Retrospective data
useCaseType	string	no	Use case identifier. Mandatory value: "UC1+2+6+7a+9"

Example:


```
{
 "age": "55",
 "dre": "Not Assessed",
 "biopsyBeforeMRI": true,
 "previousAdenomectomy": false,
 "prostateVolume": 0,
 "psas": [
 {
 "total": 32,
 "free": 0,
 "ratio": 0,
 "date": "2020-02-01T00:00:00",
 "baseline": true
 },
```

```
{
"total": 22,
"free": 0,
"ratio": 0,
"date": "2020-02-14T00:00:00",
"baseline": false
}
],
"relativePCa": "Unknown",
"relativePCaAge": "60",
"relativeBCa": "Unknown",
"relativeBCaAge": "60",
"relativeAggressiveCancer": "Unknown",
"mriDate": "2020-02-06T00:00:00",
"biopsyType": "In-Bore",
"carcinomaClass": "Adenosquamous carcinoma",
"lesions": [
{
"volume": 858,
"gleason1": 5,
"gleason2": 5,
"diam1": 32,
"diam2": 20,
"location": [
     "BRPZa",
     "BRTZa",
     "BRTZp",
     "MRPZa",
     "MRTZa"
],
"x": 0,
"y": 0,
"z": 0,
"pi_rads": 5,
"index_lesion": true
}
],
"lesionsCoordinateSystem": "Pixel",
"T_MRI": "cT2a",
"N_MRI": "cN0",
"T_Biopsy": "pT2",
```

```
 "totalDose": 80,
 "fractionsWeek": 2,
 "doseFraction": 4,
 "radiotherapyTechnique": "IMRT",
 "hormonotherapy": true,
 "rectalToxAcute": "Grade 3",
 "rectalToxChronic": "Grade 4",
 "genitourinaryToxAcute": "Grade 4",
 "genitourinaryToxChronic": "Grade 4",
 "epic26": 60,
 "biochemicalRelapseFUp": true,
 "psasFUp": [
 {
 "total": 45,
 "free": 0,
 "ratio": 0,
 "date": "2020-03-06T00:00:00",
 "baseline": true
 }
 ],
 "prospective": false,
 "useCaseType": "UC1+2+6+7a+9"
}
```


Form 1+2+5+7b+9

Table 14 Form 1+2+5+7b+9 JSON objects description

Field	Type	Optional	Description
age	string	no	Integer number converted to string
dre	string	yes	Digital Rectal Examination. Possible Values: "Positive", "Negative", "Not Assessed"
biopsyBeforeMRI	bool	no	A biopsy was performed before MRI. true, false
previousAdenomectomy	bool	no	An Adenomectomy was performed before MRI. true, false

Field	Type	Optional	Description
psas	List<PSA>	no	List of PSA objects. There must be one object with the 'baseline' field set to 'true' (see PSA sheet)
relativePCa	string/bool	yes	First degree relative with Prostate Cancer. One possible value from: "True", "False", "Unknown"
relativePCaAge	string	yes	Age of first degree relative when diagnosed with Prostate Cancer.
relativeBCa	string/bool	yes	First degree relative with Breast Cancer. One possible value from: "True", "False", "Unknown"
relativeBCaAge	string	yes	Age of first degree relative when diagnosed with Breast Cancer.
relativeAggressiveCancer	string/bool	yes	First degree relative with Aggressive Cancer. One possible value from: "True", "False", "Unknown"
mriDate	string/DateTIme	no	Date of the MRI
prostatectomyMethod	string	no	Method of prostatectomy. One possible value from: "RP retropubic", "Laparoscopic", "Robotic-assisted laparoscopic", "RP perineal", "Other"
prostatectomyNerveSparing	bool	no	Nerve Sparing Prostatectomy. true, false

Field	Type	Optional	Description
carcinomaClass	string	no	One possible value from: "Acinar adenocarcinoma", "Intraductal carcinoma", "Ductal adenocarcinoma", "Urothelial carcinoma", "Adenosquamous carcinoma", "Squamous cell carcinoma", "Basal cell carcinoma", "Adenocarcinoma with neuroendocrine differentiation", "Small cell neuroendocrine carcinoma", "Large cell neuroendocrine carcinoma", "Not Assessed"
lesions	List<Lesion>	no	List of Lesion objects. There must be one object with the 'index_lesion' field set to 'true' (see Lesion sheet)
lesionsCoordinateSystem	string	yes	One possible value from: "Physical", "Pixel"
T_Prostatectomy	string	no	One possible value from: “pT1”, "pT1a", "pT1b", "pT1c", “pT2”, "pT2a", "pT2b", "pT2c", “pT3”, "pT3a", "pT3b", "pT4"
N_Prostatectomy	string	no	One possible value from:"pNX", "pN0", "pN1"
resectionMarginsStatus	string	no	One possible value from: "Positive", "Negative", "Not Assessed"
extraProstaticExtension	string	no	One possible value from: "Present", "Absent", "Not Assessed"

Field	Type	Optional	Description
perineuralInvasion	string	no	One possible value from: "Present", "Absent", "Not Assessed"
seminalVesicalInvasion	string	no	One possible value from: "Present", "Absent", "Not Assessed"
wheelerRanking	string	yes	One possible value from: "LX", "L0", "L1", "L2", "L3F", "L3E"
epic26	int	no	Post-treatment QoL EPIC-26
EORTC_Q52	int	yes	EORTC - QLQ - PR25 (score 1 – not at all, to 4 – very much): Q52. To what extent was sex enjoyable for you?
EORTC_Q53	int	yes	EORTC - QLQ - PR25 (score 1 – not at all, to 4 – very much): Q53. Did you have difficulty getting or maintaining an erection?
EORTC_Q54	int	yes	EORTC - QLQ - PR25 (score 1 – not at all, to 4 – very much): Q54. Did you have ejaculation problems (e.g. dry)?
EORTC_Q55	int	yes	EORTC - QLQ - PR25 (score 1 – not at all, to 4 – very much): Q55. Have you felt uncomfortable about being sexually intimate?
biochemicalRelapseFUp	bool	No	Biochemical Relapse on Follow-Up. true, false
psasFUp	List<PSA>	No	List of PSA objects (see PSA sheet)
prospective	bool	No	Prospective or Retrospective data

Field	Type	Optional	Description
useCaseType	string	No	Use case identifier. Mandatory value: "UC1+2+5+7b+9"

Example:


```
{
 "age": "59",
 "dre": "Positive",
 "biopsyBeforeMRI": true,
 "previousAdenomectomy": false,
 "psas": [
 {
 "total": 18,
 "free": 0,
 "ratio": 0,
 "date": "2020-03-15T00:00:00",
 "baseline": true
 }
 ],
 "relativePCa": "True",
 "relativePCaAge": "65",
 "relativeBCa": "False",
 "relativeBCaAge": "",
 "relativeAggressiveCancer": "True",
 "mriDate": "2020-06-15T00:00:00",
 "prostatectomyMethod": "Robotic-assisted laparoscopic",
 "prostatectomyNerveSparing": true,
 "carcinomaClass": "Urothelial carcinoma",
 "lesions": [
 {
 "volume": 651,
 "gleason1": 5,
 "gleason2": 4,
 "diam1": 31,
 "diam2": 13,
 "location": [
      "BLTZa",
      "BRTZa",
```

```
      "MLTZa",
      "MRTZa",
      "ARTZa"
 ],
 "x": 1,
 "y": 2,
 "z": 3,
 "pi_rads": 5,
 "index_lesion": true
 }
 ],
 "lesionsCoordinateSystem": "Physical",
 "T_Prostatectomy": "pT1a",
 "N_Prostatectomy": "pN0",
 "resectionMarginsStatus": "Positive",
 "extraProstaticExtension": "Present",
 "perineuralInvasion": "Present",
 "seminalVesicalInvasion": "Present",
 "wheelerRanking": "L3F",
 "epic26": 54,
 "EORTC_Q52": 2,
 "EORTC_Q53": 3,
 "EORTC_Q54": 3,
 "EORTC_Q55": 3,
 "biochemicalRelapseFUp": false,
 "psasFUp": [
 {
 "total": 12,
 "free": 0,
 "ratio": 0,
 "date": "2020-12-14T00:00:00",
 "baseline": true
 }
 ],
 "prospective": false,
 "useCaseType": "UC1+2+5+7b+9"
}
```

Table 15 Form 1+2+8+9 JSON objects description

Field	Type	Optional	Description
age	string	no	Integer number converted to string
dre	string	yes	Digital Rectal Examination. Possible Values: "Positive", "Negative", "Not Assessed"
biopsyBeforeMRI	bool	no	A biopsy was performed before MRI. true, false
previousAdenomectomy	bool	no	An Adenomectomy was performed before MRI. true, false
psas	List<PSA>	no	List of PSA objects. There must be one object with the 'baseline' field set to 'true' (see PSA sheet)
relativePCa	string/bool	yes	First degree relative with Prostate Cancer. One possible value from: "True", "False", "Unknown"
relativePCaAge	string	yes	Age of first degree relative when diagnosed with Prostate Cancer.
relativeBCa	string/bool	yes	First degree relative with Breast Cancer. One possible value from: "True", "False", "Unknown"
relativeBCaAge	string	yes	Age of first degree relative when diagnosed with Breast Cancer.
relativeAggressiveCancer	string/bool	yes	First degree relative with Aggressive Cancer. One possible value from: "True", "False", "Unknown"

Field	Type	Optional	Description
mriDate	string/DateTIme	no	Date of the MRI (e.g., "2021-12-21T00:00:00Z")
mriPositive	bool	no	MRI confirmed cancer. true, false
biopsyPositive	string	no	Biopsy confirmed cancer. true, false
biopsyType	string	no	One possible value from: "Systematic", "Fusion", "Systematic+Fusion", "In-Bore"
carcinomaClass	string	no	One possible value from: "Acinar adenocarcinoma", "Intraductal carcinoma", "Ductal adenocarcinoma", "Urothelial carcinoma", "Adenosquamous carcinoma", "Squamous cell carcinoma", "Basal cell carcinoma", "Adenocarcinoma with neuroendocrine differentiation", "Small cell neuroendocrine carcinoma", "Large cell neuroendocrine carcinoma", "Not Assessed"
lesions	List<Lesion>	no	List of Lesion objects. There must be one object with the 'index_lesion' field set to 'true' (see Lesion sheet)
lesionsCoordinateSystem	string	yes	One possible value from: "Physical", "Pixel"
T_MRI	string	no	One possible value from: "cTX", "cT2a", "cT2b", "cT2c", "cT3a", "cT3b", "cT4"
N_MRI	string	yes	One possible value from: "cNX", "cN0", "cN1"

Field	Type	Optional	Description
followUps	List<FollowUp>	no	List of FollowUp objects (see FollowUp object)
prospective	bool	no	Prospective or Retrospective data
useCaseType	string	no	Use case identifier. Mandatory value: "UC1+2+5+9"

Example:


```
{
 "age": "55",
 "dre": "Positive",
 "biopsyBeforeMRI": true,
 "previousAdenomectomy": false,
 "psas": [
 {
 "total": 6.2,
 "free": 0,
 "ratio": 0,
 "date": "2020-03-20T00:00:00",
 "baseline": true
 }
 ],
 "relativePCa": "False",
 "relativePCaAge": "0",
 "relativeBCa": "False",
 "relativeBCaAge": "0",
 "relativeAggressiveCancer": "False",
 "mriDate": "2020-03-30T00:00:00",
 "mriPositive": false,
 "biopsyType": "Systematic+Fusion",
 "carcinomaClass": "Not Assessed",
 "biopsyPositive": false,
 "lesions": [
 {
 "volume": 32,
 "gleason1": 2,
 "gleason2": 3,
```

```
"diam1": 4,
"diam2": 4,
"location": [
     "MLTZa"
],
"x": 11,
"y": 12,
"z": 13,
"pi_rads": 2,
"index_lesion": true
}
],
"lesionsCoordinateSystem": "Physical",
"T_MRI": "cT2a",
"N_MRI": "cN0",
"followUps": [
{
"fupPSAConfirmedNormal": false,
"fupMRIConfirmedNegative": true,
"fupBiopsyConfirmedNegative": true,
"imagingReport": true,
"psasFUp": [
     {
     "total": 6.3,
     "free": 0,
     "ratio": 0,
     "date": "2020-07-31T00:00:00",
     "baseline": true
     }
],
"mriDateFUp": "2020-10-16T00:00:00",
"TFUp_MRI": "cT2a",
"NFUp_MRI": "cN0",
"TFUp_Biopsy": "pT1",
"lesionsFUp": [
     {
     "volume": 32,
     "gleason1": 2,
     "gleason2": 3,
     "diam1": 4,
     "diam2": 4,
```

```
      "location": [
      "MLTZa"
      ],
      "x": 11,
      "y": 12,
      "z": 13,
      "pi_rads": 2,
      "index_lesion": true
      }
],
"biopsyPerformedFUp": true,
"biopsyTypeFUp": "Fusion",
"biopsyDateFUp": "2020-11-12T00:00:00",
"likelihoodProgression": 1,
"carcinomaClassFUp": "Not Assessed"
},
{
"fupPSAConfirmedNormal": false,
"fupMRIConfirmedNegative": true,
"fupBiopsyConfirmedNegative": true,
"imagingReport": true,
"psasFUp": [
      {
      "total": 6.5,
      "free": 0,
      "ratio": 0,
      "date": "2021-07-31T00:00:00",
      "baseline": true
      }
],
"mriDateFUp": "2021-10-20T00:00:00",
"TFUp_MRI": "cT2a",
"NFUp_MRI": "cN0",
"TFUp_Biopsy": "pTX",
"lesionsFUp": [
      {
      "volume": 32,
      "gleason1": 2,
      "gleason2": 3,
      "diam1": 4,
      "diam2": 4,
```

```
      "location": [
      "MLTZa"
      ],
      "x": 11,
      "y": 12,
      "z": 13,
      "pi_rads": 2,
      "index_lesion": true
      }
 ],
 "biopsyPerformedFUp": true,
 "biopsyTypeFUp": "Systematic+Fusion",
 "biopsyDateFUp": "2021-11-10T00:00:00",
 "likelihoodProgression": 1,
 "carcinomaClassFUp": "Not Assessed"
 }
 ],
 "prospective": false,
 "useCaseType": "UC1+2+8+9"
}
```


FollowUp Object

Table 16 FollowUp JSON objects description

Field	Type	Optional	Description
fupPSAConfirmedNormal	bool	no	Follow-up PSA confirmed normal. true, false
fupMRIConfirmedNegative	bool	no	Follow-up MRI confirmed negative. true, false
fupBiopsyConfirmedNegative	bool	no	Follow-up biopsy confirmed negative. true, false
imagingReport	bool	no	There is an Imaging report. true, false
psasFUp	List<PSA>	no	List of PSA objects (see PSA sheet)
mriDateFUp	string/DateTime	no	Date of the follow-up MRI (e.g., "2021-12-21T00:00:00Z")

Field	Type	Optional	Description
TFUp_MRI	string	no	One possible value from: "cTX", "cT2a", "cT2b", "cT2c", "cT3a", "cT3b", "cT4"
NFUp_MRI	string	yes	One possible value from:"cNX", "cN0", "cN1"
TFUp_Biopsy	string	yes	One possible value from:"pTX", "pT1", "pT2"
lesionsFUp	List<Lesion>	no	List of Lesion objects (see Lesion sheet)
biopsyPerformedFUp	bool	no	A biopsy was performed. true, false
biopsyTypeFUp	string	no	One possible value from: "Systematic", "Fusion", "Systematic+Fusion", "In-Bore"
biopsyDateFUp	string/DateTime	no	Date of the follow-up Biopsy (e.g., "2021-12-21T00:00:00Z")
likelihoodProgression	int	no	Likelihood of progression. Values from 1 to 5 according to the scale
carcinomaClassFUp	string	no	Carcinoma classification on follow-up. One possible value from: "Acinar adenocarcinoma", "Intraductal carcinoma", "Ductal adenocarcinoma", "Urothelial carcinoma", "Adenosquamous carcinoma", "Squamous cell carcinoma", "Basal cell carcinoma", "Adenocarcinoma with neuroendocrine differentiation", "Small cell neuroendocrine carcinoma", "Large cell neuroendocrine carcinoma", "Not Assessed"

Likelihood of progression. Values from 1 to 5 according to the scale:

Table 17 Table of likelihood of radiological progression

Likert	Assessment of the likelihood of radiologic progression	Example
1	Resolution of previous features suspicious on MRI	Previously enhancing area no longer enhances
2	Reduction in volume and/or conspicuity of previous features suspicious on MRI	Reduction in size of previously seen lesion that remains suspicious for clinically significant disease
3	Stable MRI appearance: no new focal/diffuse lesions	Either no suspicious features or all lesions stable in size and appearance
4	Significant increase in size and/or conspicuity of features suspicious for prostate cancer	The lesion becomes visible on diffusion-weighted imaging; a significant increase in the size of the previously seen lesion
5	Definitive radiologic stage progression	The appearance of extracapsular extension, seminal vesicle involvement, lymph node involvement, or bone metastasis

A.3 Curation Tools

The motion-correction application, which is instantiated by clicking the Motion-correct button, performs *inter-volume motion-correction of a DWI or DCE series*.

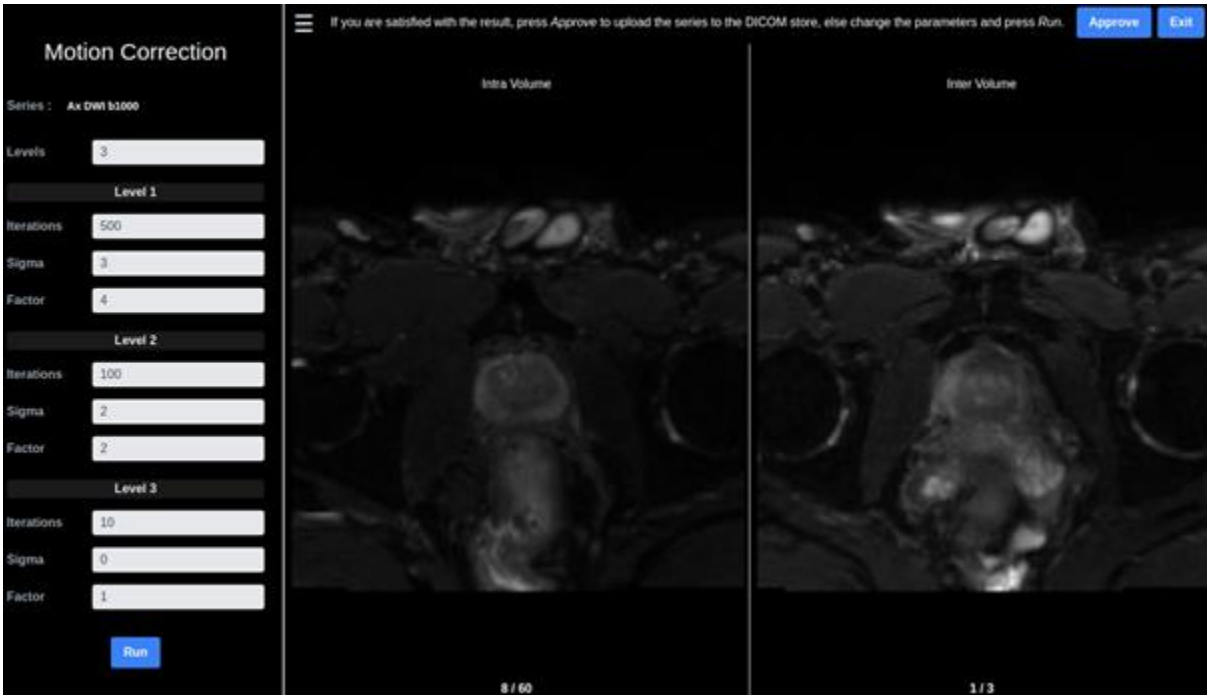


Figure 43: The motion-corrected series can be concurrently reviewed for intra- and inter-volume motion in two side-by-side viewers.

The first volume of the 4D series is automatically selected as the reference volume. The motion-corrected series can be concurrently reviewed for intra- and inter-volume motion in two side-by-side viewers, as shown in the next figure. The user may scroll through all the slices of all volumes and the middle slice per volume, respectively, in order to identify excess motion. For instance, on the left-hand side of the split-screen one can scroll through all total 60 slices of all volumes (20 slices x 3 volumes). On the right-hand side, one can scroll through the middle slice of each of the 3 volumes.

If the result is satisfactory, the *Approve* button at the top-right of the screen accepts the result by sending it to the Cloud Staging Area and advances the user to the co-registration application taking as input the result of the motion-correction phase and the already selected static series.

The co-registration application co-registers the motion-corrected series to a T2w image. The moving (motion-corrected) and static (T2) series are also colour-coded in green and red, respectively, as illustrated in the next figure.

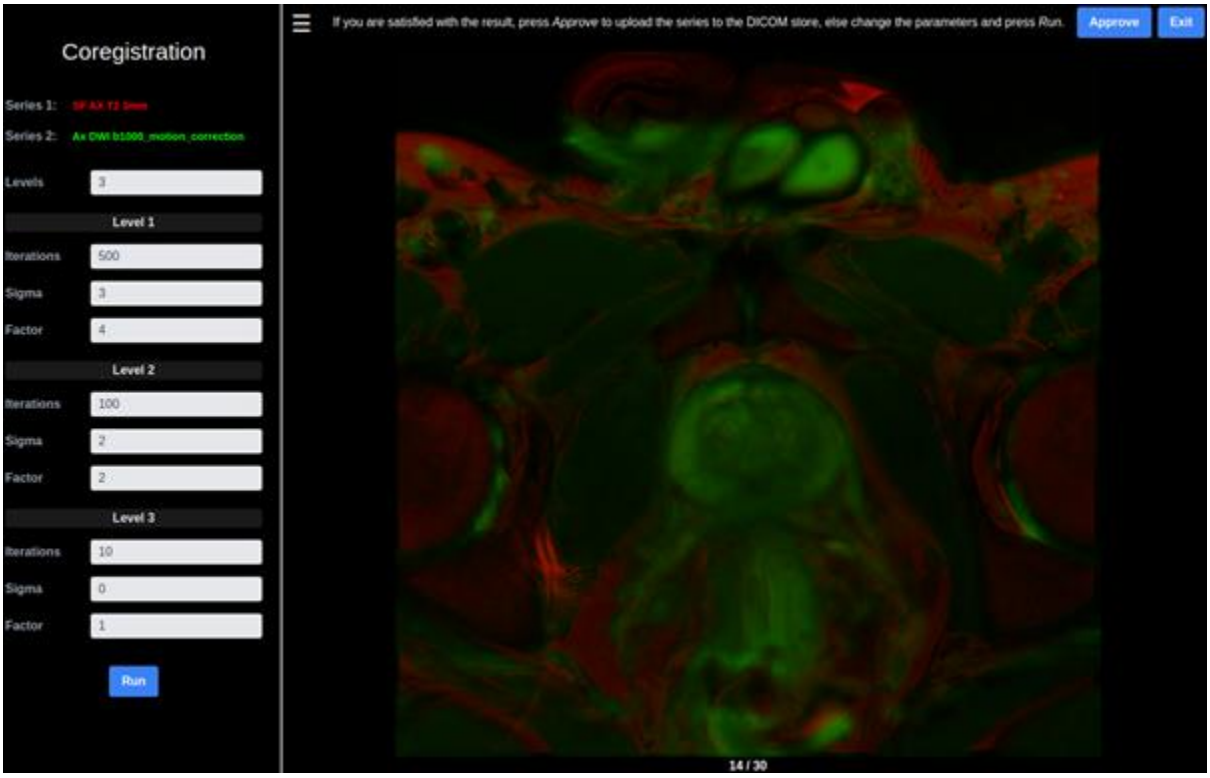


Figure 44: The moving (motion-corrected) and static (T2) series as depicted upon the T2w image, colour-coded in green and red, respectively.

In this case, the 1st volume of the previously motion-corrected series has been laid over the T2w series and the result for slice 14/30 is displayed. Again, if the result is satisfactory, the *Approve* button at the top-right of the screen accepts the result by sending it to the Cloud Staging Area and directs the user back at the beginning in order to select a new series to curate.

In either curation phase, if the result is unsatisfactory, the registration hyper-parameters on the left-hand side may be refined in order to re-run the curation function by clicking the *Run* button.

As already shown on the left-hand side of the figures above, the inputs for configuring the motion-correction and co-registration algorithms are the same. The underlying registration algorithm uses a combination of transformation steps: translation, rigid, and affine to progressively compute an affine transformation to register two 3D volumes via a similarity metric. Each step is performed at multiple levels to align two volumes using a multi-resolution strategy. The number of levels, iterations per level, smoothing factory per level, and down sampling factor per level are configurable via the settings on the left-hand side. By default, 3 levels are used. Level 1 uses 500 iterations, sigma 3, and factor 4, which means that, if the original shape of the image was (X, Y, Z) voxels, then the shape of down-sampled image would be (X/4,

Y/4, Z/4). Level 2 uses 100 iterations, sigma 2, and factor 2. Level 3 uses 10 iterations, sigma 0, factor 1, which means that the image is neither smoothed nor down-sampled in this last step.

More specifically, the number of levels specifies the number of resolutions. For each level, the number of iterations (to solve the optimization problem) has to be specified along with a smoothing (Gaussian kernel sigma) and a scaling (down sampling) factor. It is recommended to exploit a higher smoothing and more down sampling for more iterations in early levels in order to achieve an early, good-enough transformation before more fined-grained transformations towards the end. The pre-filled parameters are considered a good starting point and should work for most cases.

Finally, when the *Approve* button is clicked, the derived image is uploaded to the Cloud Staging Area as a new DICOM series with a new Series Instance UID, and any related series, like the original or static series, is added to the Referenced Series Sequence public DICOM tag. Note that if a study is selected for curation again, the already approved and stored results are not modified or deleted. Curation results are stored in an append-only fashion.

A.4 Annotation Tools

To perform segmentation, click on the **"Q" button** in the Study Viewer to launch Annotation Tools developed by QUIBIM. It will open a new browser tab displaying the images and the tools for annotation.

Figure 45: Annotation environment.

Table 6 shows the main buttons found in the annotation environment. The tool provides two modes of operation:

- Manual image segmentation
- Automatic image segmentation (based on an AI based DL model developed by QUIBIM)

Table 18: Icons present in the annotation environment and their main functionality.

Icon	Description
	Add notes to the study (not used for the moment)
	Browse through the slices with the mouse wheel.
	Activate/deactivate zoom and pan.
	Center the image by resetting zoom and pan.
	Adjust the image contrast
	Set the initial windowing
	Select a predefined colormap
	Show/hide the information overlaid on the viewer area
	Show/hide the DICOM header tags
	Show MPR view
	Open a new view to load an additional series
	Save changes (overwrites previous DICOM Seg file)
	Save changes creating a new DICOM Seg file

Icon	Description
	Download the DICOM Seg file locally
	Undo changes
	Redo changes
	Remove all ROIs
	Change label
	Activate brush tool
	Activate eraser tool
	Launch automatic prostate segmentation

Details about the functionalities of the tool are reported in a dedicated manual (IFU-instructions for use) which is provided to the users and available in the ProstateNet. For complementarity purposes part of the abovementioned is incorporated in the ANNEX section of the current deliverable.

Zoom-in / zoom-out and pan

The button  activates the zoom and pan tools. Once they are activated:

- Zoom-in / zoom-out: scroll up and down to zoom-in and zoom-out respectively. The mouse pointer is used as the center to zoom the image.
- Pan: click and move the mouse to the desired direction to move the image.

The button  resets the zoom and pan to center the image in the viewer.

Change image contrast

The button  activates the tool to change image windowing. Once it is activated:

- Change the window center: click and hold down the left mouse button, and move the mouse up or down to increase or decrease respectively the window centre.
- Change window width: click and hold down the left mouse button, and move the mouse left or right to increase or decrease respectively the window width.

The button  resets window center and window width values to their original values.

Multiplanar Reconstruction (MPR) view

Imaging studies are acquired in a concrete plane, commonly axial. However, it is possible to reconstruct the other two orthogonal planes, coronal and sagittal, from the initially acquired one.

The user can trigger this reconstruction view by clicking the  button located in each of the cells. Clicking it will open a new view where the user will be able to see the axial, coronal, and sagittal views. In addition, the brush tool (see section 6) is compatible with this view, this means that the three different planes can be used to delineate ROIs.

Open a secondary series

A secondary series can be opened. For that, click on the button  . The viewer will be split in two. Select the series to load on the series list (left bar) and drag and drop it to the new view.

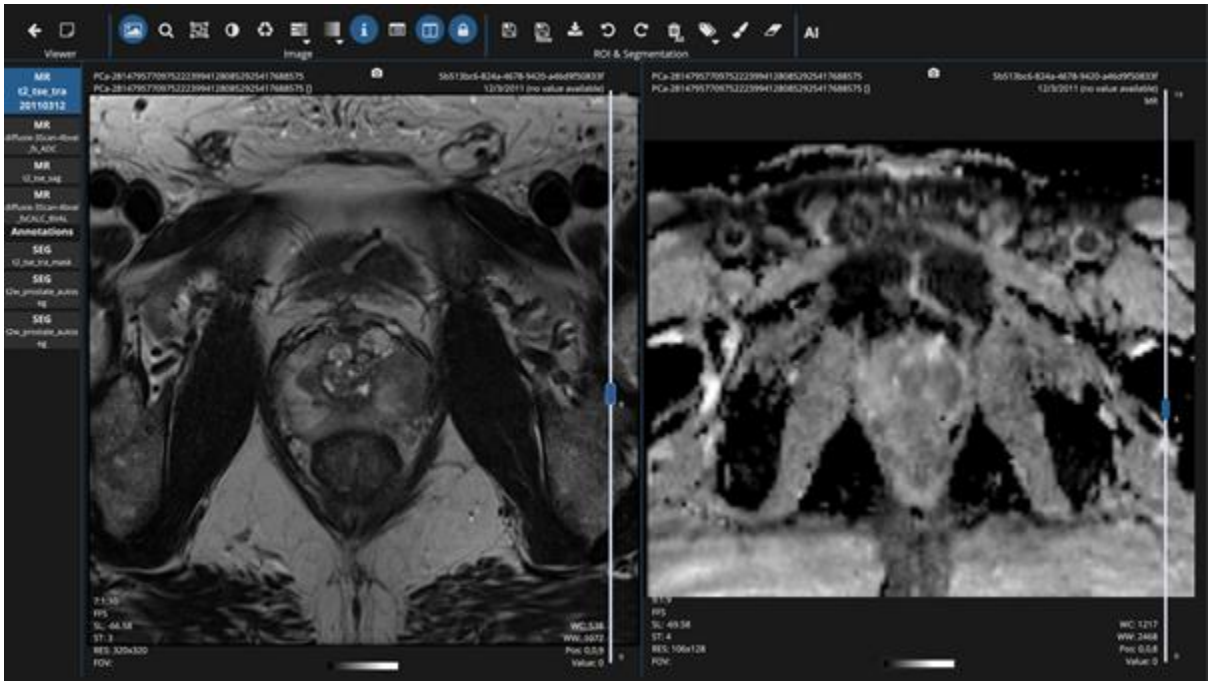


Figure 46. Screenshot from the annotation tool integrated in the ProstateNet.

The button  allows for synchronizing both series on the z-axis. If it is activated (lock closed), when changing the image slice, the other series is automatically positioned. If you want to de-synchronize the views, deactivate the button (opened lock - ).

Select a label and create new ones

The button  opens a dropdown to select the label of the region to delineate. By default, the labels shown in Figure 10 are initialized. These belong to the regions to delineate when doing the prostate gland segmentation:

- TZ+CZ: Transition Zone and Central Zone.
- PZ: Peripheral Zone
- SV: Seminal Vesicles

Or the lesions segmentation. By default, three different labels are initialized (Lesion 1 – Lesion 3), however, new labels can be created of new lesions must be segmented by clicking the button "New Label". When clicking on "New label" a colour picker is opened to select the desired colour for the new label. New labels will be automatically named "Label 4", "Label 5"…

Figure 47: Default labels are initialized in the annotation environment.

By clicking on the trash button  , a complete region can be deleted.

By clicking on the eye button  , a complete region can be hidden/shown.

To change the name or the colour of a specific label, the button "Configuration" must be clicked (Figure 10). The new modal shows, per row, each defined label. Per label the following actions can be conducted:

- Change label colour: Click over the coloured square to open a colour picker to select the new colour.
- Change label name: Change the label name by changing the value of the "Label" field.
- Defined drawing intensity range: A minimum and/or maximum intensity value can be introduced to define a specific range over which the brush tool draws per each specific label.
    - Min. value: voxels with image intensity values below this limit will not be drawn.
    - Max. value: voxels with image intensity values over this limit will not be drawn.

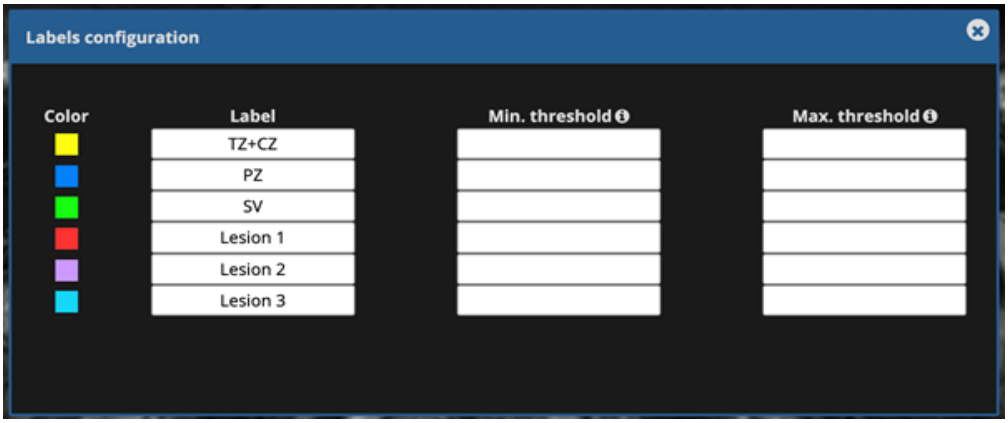


Figure 48: Labels configuration modal.

Execute automatic segmentation tool

When annotating the prostate gland, an automatic segmentation tool can be used as a pre-segmentation to ease the annotation workflow. It requires a T2w MRI sequence in the axial plane.

To execute the analysis, click on the AI button. When the segmentation finishes with no errors, the following message will be shown (Figure 22). If there is an incident during the execution, an error message will be shown.

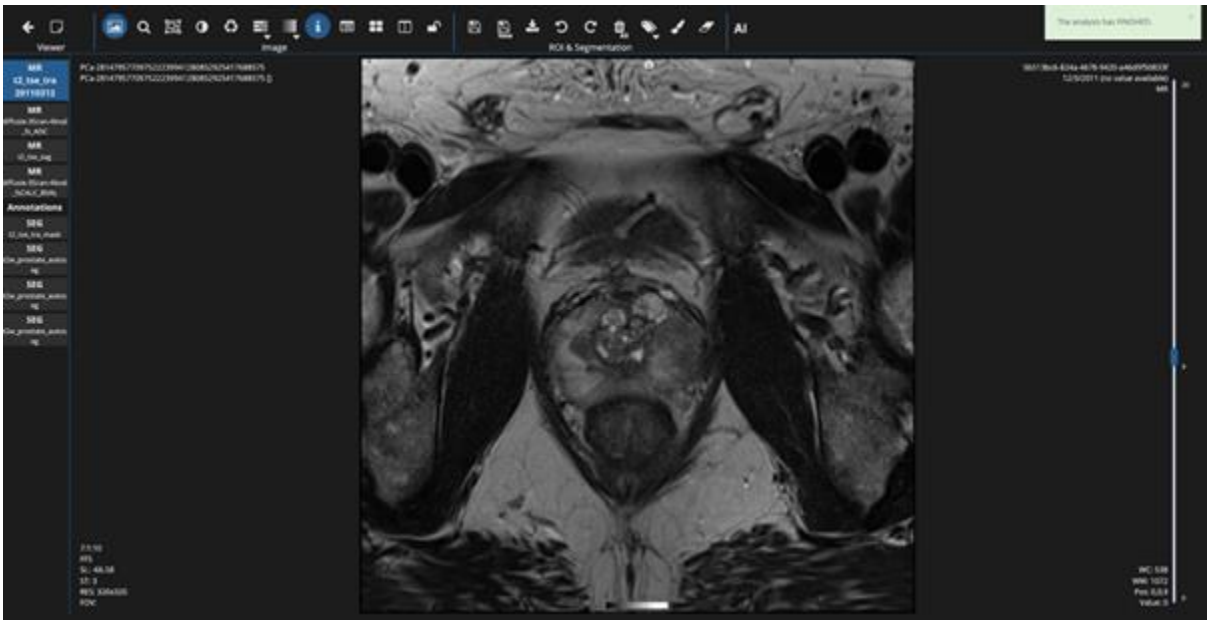


Figure 49: Prostate automatic segmentation finished successfully.

Once the segmentation finishes, a new segmentation series will appear in the bar on the left named “t2w_prostate_autoseg”. By clicking on it, it will be loaded over the T2w series (Figure 23).

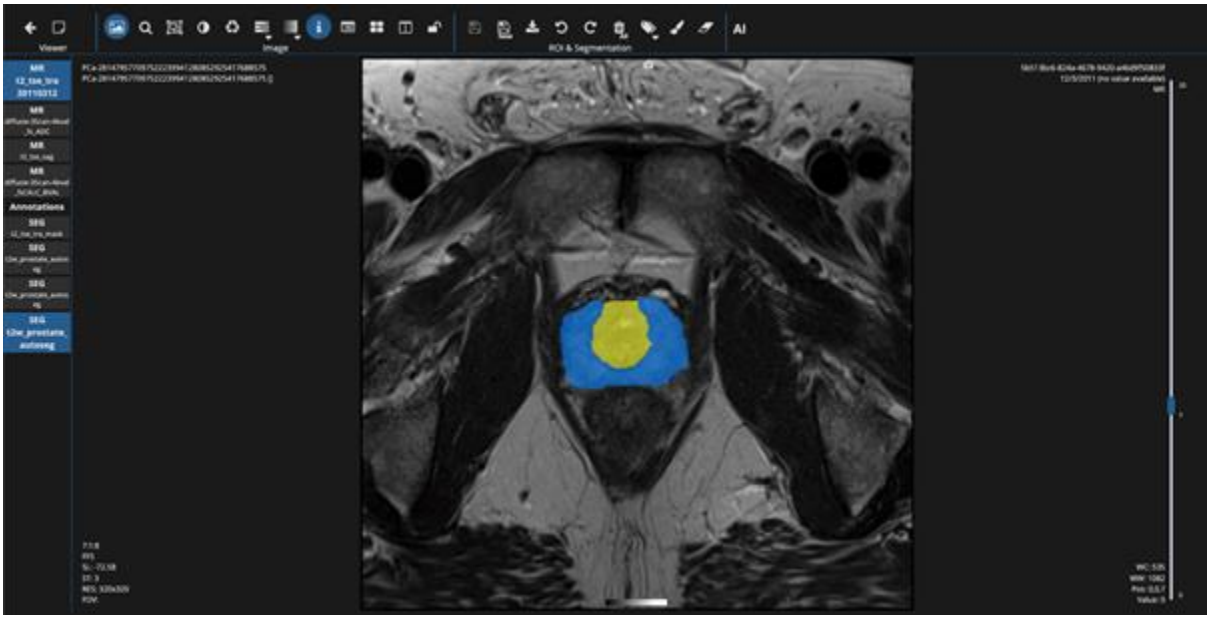


Figure 50: Automatic prostate segmentation overlaid over the T2w series.

To edit the segmentation, activate the brush tool  , select the label to edit from the label list,  and follow section 7. Once the correction is finished, store the annotation as a new segmentation file (section 8 – Save new) with the name “prostate”.

Segment a region

To segment, a specific region, make sure the appropriate label has been selected (see section 5).

Once selected, the brush tool can be activated by clicking on the button  .

When activated, a circular cursor appears, this defines the region that is going to be segmented. To delineate a region, click the left mouse button and without ceasing the click, move the mouse through the region to segment.

The size of the brush and the opacity of the mask can be edited either through the keyboard or through the panel (Figure 51).

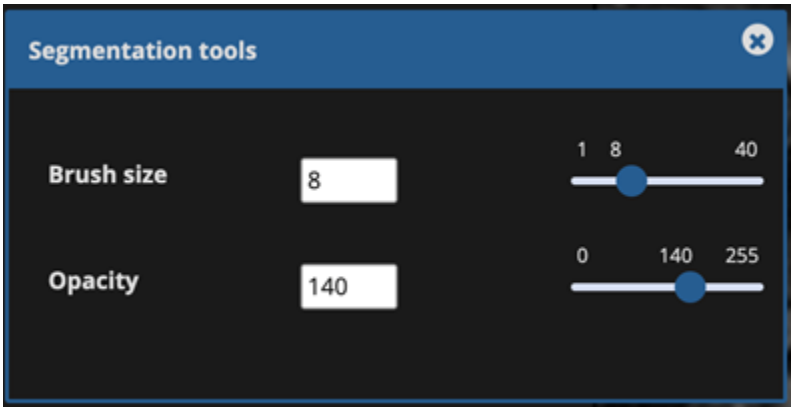


Figure 51: Panel to control brush size and mask opacity.

To increase/decrease the brush size one of the following options is possible:

- Define the specific size by introducing a whole number in the panel (min: 1, max: 40).
- Change the brush size dynamically through the slider in the panel.
- Press the buttons "+" and "-"on the keyboard, respectively.

To increase/decrease the mask opacity one of the following options is possible:

- Define the specific opacity by introducing a whole number in the panel (min:0, max:255).
- Change the opacity dynamically through the slider in the panel.
- Use the keyboard controls:
  - "A": decrease opacity
  - "D": increase opacity
  - "S": hide/show mask

To remove a region from the mask there are two different options:

- Use the right button from the mouse.
- Activate the eraser tool by clicking the button 

Save segmentation mask

To save the drawn segmentation mask, there are three main options:

- Save new: 

To save a new segmentation file. When clicked, a modal is opened to add a description of the mask to save. This description will be used, together with the original series description to define

the series description of the new DICOM file as <series_description_original_series>_<mask_description>.

Type the mask description and click on “Save” to finally save the mask.

- Save: 

To override the previously loaded segmentation file.

- Save local: 

This button downloads locally a DICOM Seg file with the drawn mask. This button does NOT save the mask in the platform database.

A.5 Metadata Catalogue

Welcome page

After successfully signing in, you will be redirected to the ProCAncer-I metadata catalogue welcome page (**Error! Reference source not found.**). The welcome page features a menu bar (top-left) and options for signing out of the application (top-right).

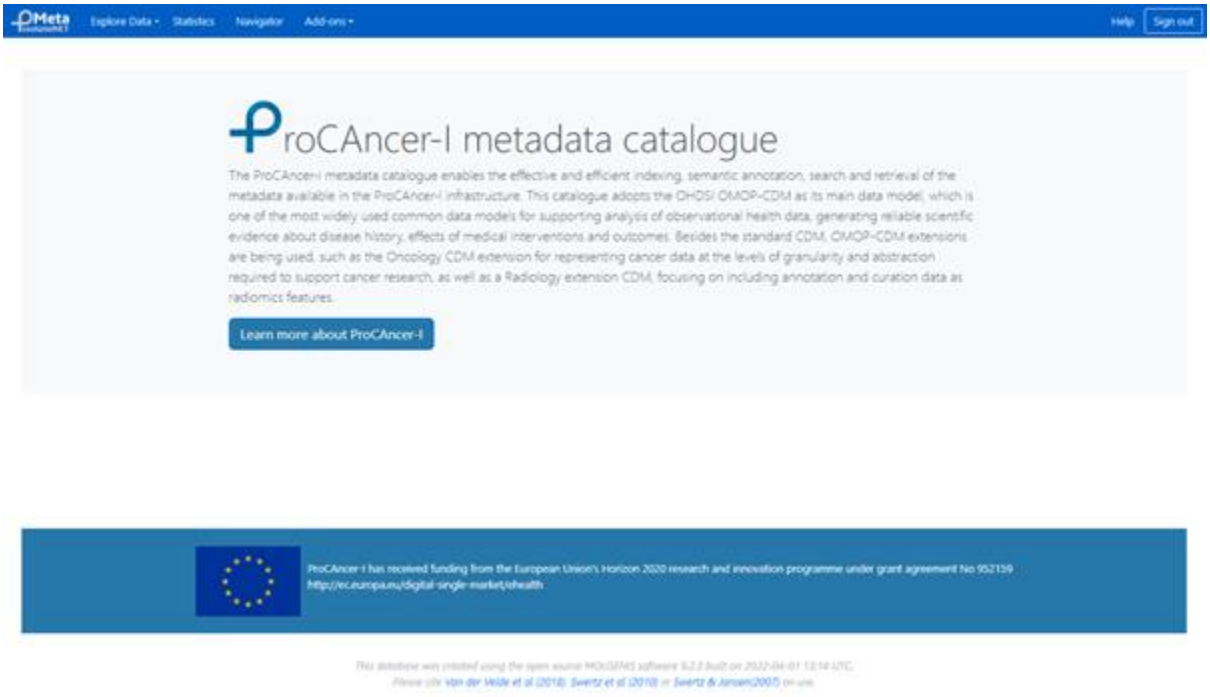


Figure 52: ProCAncer-I metadata catalogue welcome page.

Using the "Explore data" menu

On the top-left part of the screen resides the "Explore Data" menu item (Figure 25), where you can view, filter, and search the metadata that is currently available in the metadata repository. This includes metadata about 1) the use case clinical data currently uploaded through the eCRF Data Upload Tool (either from single or batch mode); 2) the imaging study metadata that get uploaded along with the clinical data; 3) metadata about the annotation tasks, and finally 4) metadata about the curation tasks.

In the following sections, we will describe how you can find, search and filter data depending on various scenarios.

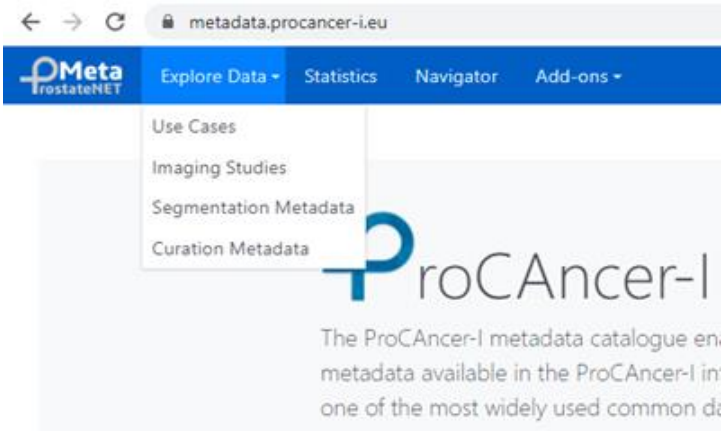


Figure 53: Explore data menu.

Selecting the "Use Cases" menu item

By selecting the "Explore data" -> "Use Cases" menu item at the top left area of the metadata catalogue user interface, you are able to explore the use case clinical data, as filled through the eCRF, along with the imaging metadata (Figure 26). In addition, apart from this information you can explore some metadata concerning the upload process itself, such as the institution uploading the current content, the time of the uploading, the method that the data was uploaded (single-mode or batch-mode) if it was validated by someone, and finally if the data refers to prospective or retrospective information. This information is under the attribute "Dataset Item", which you can expand and select the attributes you would like to see in the table view.

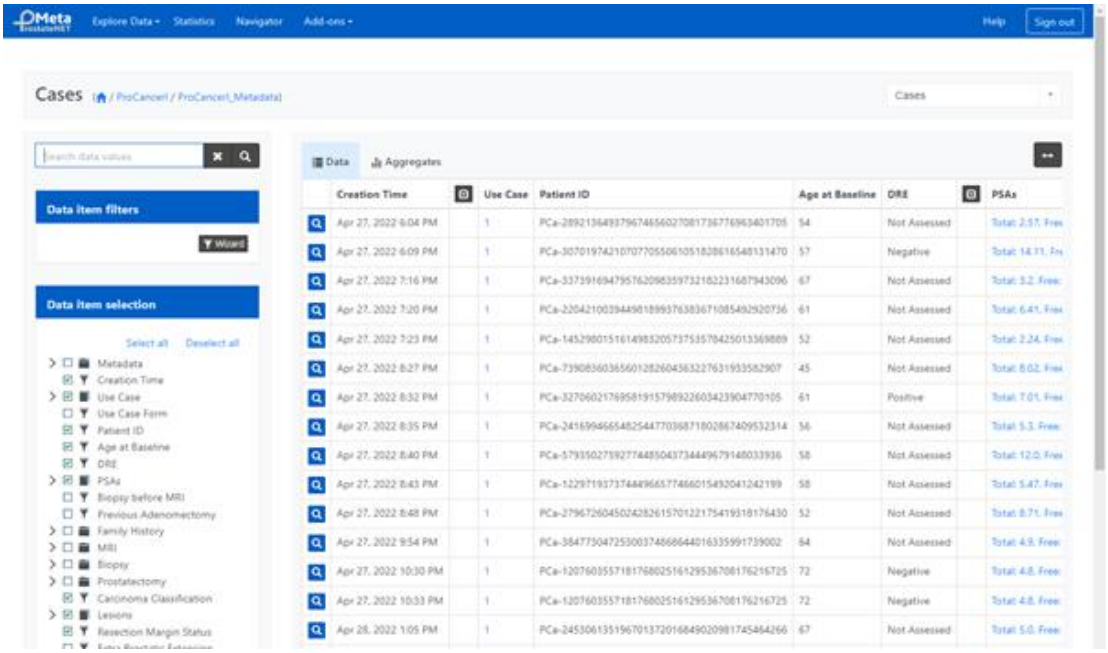


Figure 54: Explore Use Case Clinical and Imaging Metadata

The main area of the explorer shows all the data available in the selected table. The left part of the screen shows the options for searching, filtering, and projecting options which will be described below. In case, you want the main data area to be shown on the entire screen, you can select the  button on the top right part of the screen. Then, by clicking on the  button, you can return to the previous state where the filtering/searching options are available.

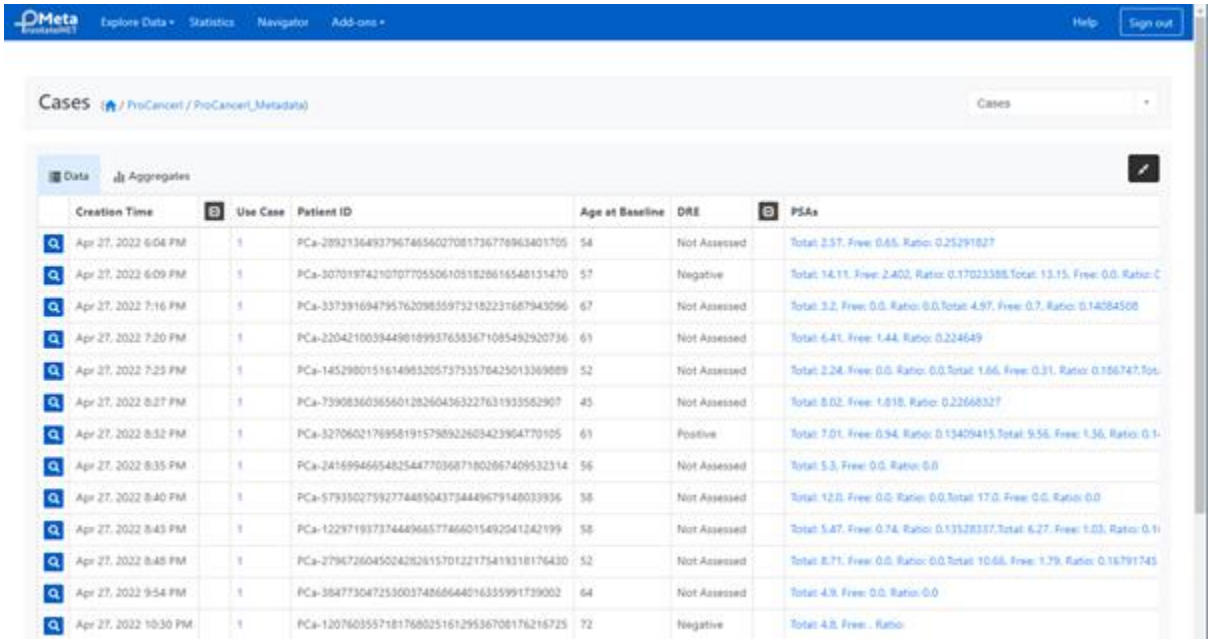
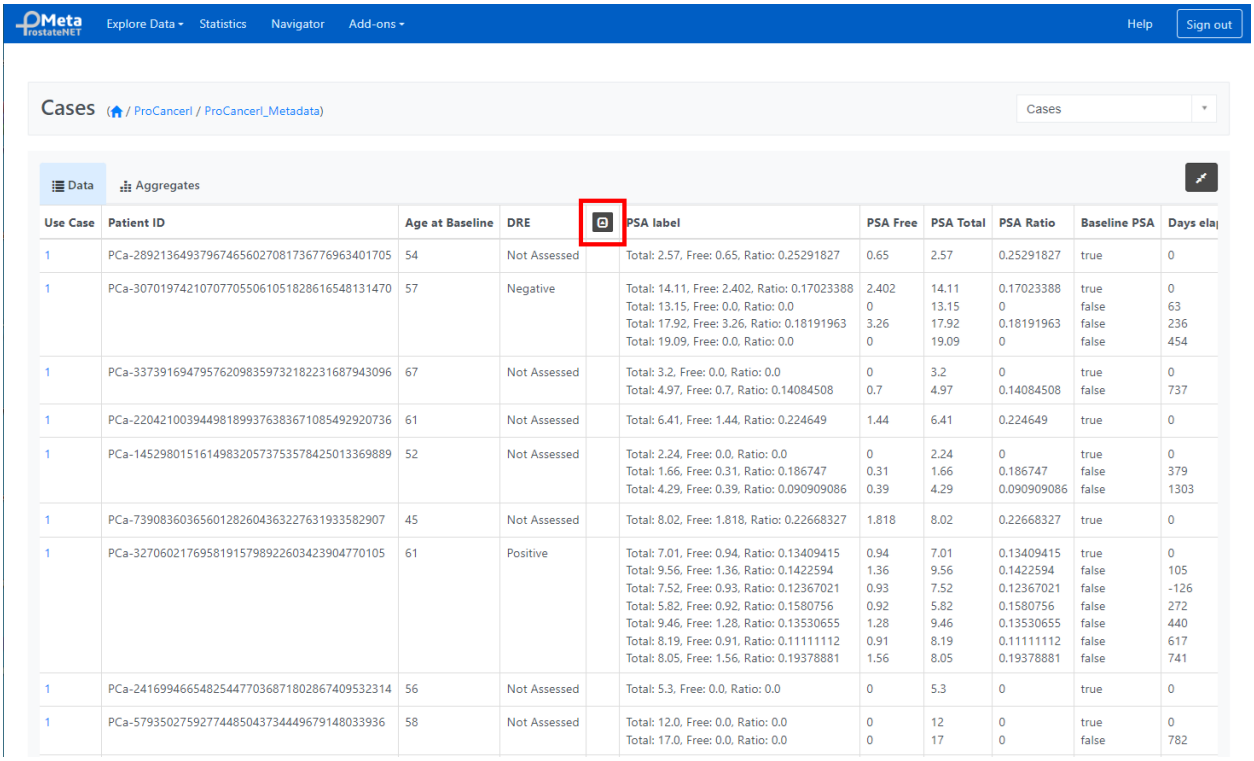


Figure 55: Expanded data view

Sort

The data can be sorted by clicking on the headers of each column/attribute. If any column is a reference to other entities (such as the PSAs column), an  icon is shown to expand this reference to show the data from the referenced entity (Figure 56).

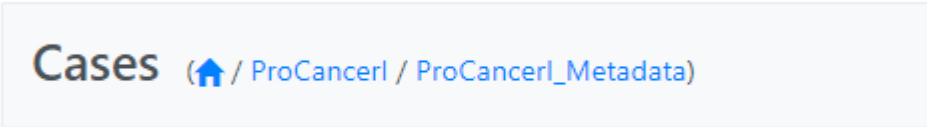


Use Case	Patient ID	Age at Baseline	DRE	PSA label	PSA Free	PSA Total	PSA Ratio	Baseline PSA	Days ela
1	PCa-289213649379674656027081736776963401705	54	Not Assessed	Total: 2.57, Free: 0.65, Ratio: 0.25291827	0.65	2.57	0.25291827	true	0
1	PCa-3070197421070770550610518286165481314 70	57	Negative	Total: 14.11, Free: 2.402, Ratio: 0.17023388 Total: 13.15, Free: 0.0, Ratio: 0.0 Total: 17.92, Free: 3.26, Ratio: 0.18191963 Total: 19.09, Free: 0.0, Ratio: 0.0	2.402 0 3.26 0	14.11 13.15 17.92 19.09	0.17023388 0 0.18191963 0	true false false false	0 63 236 454
1	PCa-337391694795762098359732182231687943096	67	Not Assessed	Total: 3.2, Free: 0.0, Ratio: 0.0 Total: 4.97, Free: 0.7, Ratio: 0.14084508	0 0.7	3.2 4.97	0 0.14084508	true false	0 737
1	PCa-220421003944981899376383671085492920736	61	Not Assessed	Total: 6.41, Free: 1.44, Ratio: 0.224649	1.44	6.41	0.224649	true	0
1	PCa-145298015161498320573753578425013369889	52	Not Assessed	Total: 2.24, Free: 0.0, Ratio: 0.0 Total: 1.66, Free: 0.31, Ratio: 0.186747 Total: 4.29, Free: 0.39, Ratio: 0.090909086	0 0.31 0.39	2.24 1.66 4.29	0 0.186747 0.090909086	true false false	0 379 1303
1	PCa-739083603656012826043632276319335829 07	45	Not Assessed	Total: 8.02, Free: 1.818, Ratio: 0.22668327	1.818	8.02	0.22668327	true	0
1	PCa-327060217695819157989226034239047701 05	61	Positive	Total: 7.01, Free: 0.94, Ratio: 0.13409415 Total: 9.56, Free: 1.36, Ratio: 0.1422594 Total: 7.52, Free: 0.93, Ratio: 0.12367021 Total: 5.82, Free: 0.92, Ratio: 0.1580756 Total: 9.46, Free: 1.28, Ratio: 0.13530655 Total: 8.19, Free: 0.91, Ratio: 0.11111112 Total: 8.05, Free: 1.56, Ratio: 0.19378881	0.94 1.36 0.93 0.92 1.28 0.91 1.56	7.01 9.56 7.52 5.82 9.46 8.19 8.05	0.13409415 0.1422594 0.12367021 0.1580756 0.13530655 0.11111112 0.19378881	true false false false false false false	0 105 -126 272 440 617 741
1	PCa-241699466548254477036871802867409532314	56	Not Assessed	Total: 5.3, Free: 0.0, Ratio: 0.0	0	5.3	0	true	0
1	PCa-579350275927744850437344496791480339 36	58	Not Assessed	Total: 12.0, Free: 0.0, Ratio: 0.0 Total: 17.0, Free: 0.0, Ratio: 0.0	0 0	12 17	0 0	true false	0 782

Figure 56: Expanding referenced entities.

Search, Filter

In the upper left corner of the “Explore Data” main page, the name of the currently selected table is displayed along with the links that redirect to the page showing all the available tables existing in the metadata repository (see Navigator section below for more information). Currently, the “Explore Data”à”Use Cases” option is selected, the “Cases” table is shown on the screen.



Below the name of the table, a search box is shown. This search box can be used to search all the data values for a certain search term (by searching the values of all the columns for the selected table).

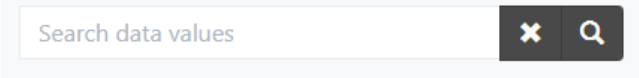


Figure 57: Data explorer search

Directly below the search box, the currently active attribute filters are shown. These can be edited by clicking on them. The cross trailing can be used to delete the currently used filter(s).

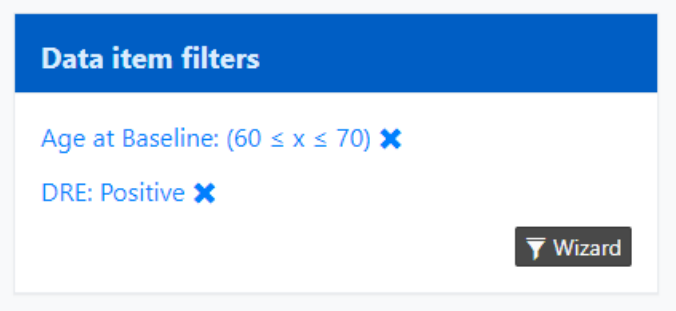


Figure 58: Data Explorer active filters.

Filters can be set from the attribute selection tree which is shown in Figure 59.

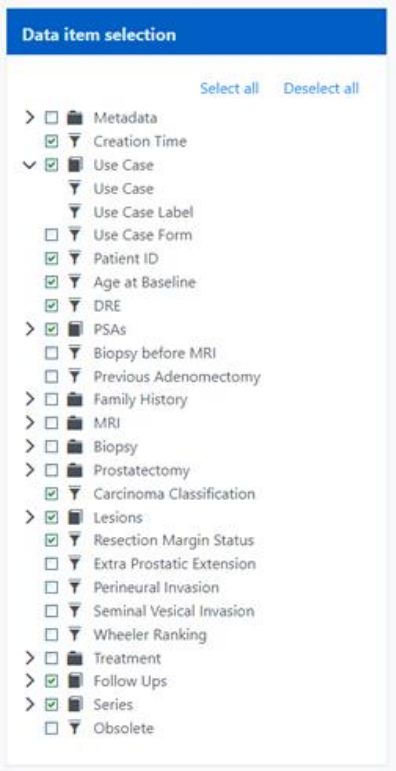
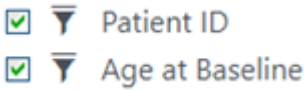


Figure 59: Attribute selection/filtering window.

Using the checkbox in front of each attribute, the visibility of this attribute in the table can be managed. The filter icon can be used to set filters for this specific attribute. There are three different types of attributes; atomic attributes, composite attributes, and attributes that refer to other tables in the repository. Atomic attributes refer to attributes that have a single value. For example, “PatientID” or “Age at baseline”, are both atomic attributes.



Atomic attributes are distinguished from the rest, by the filter icon in front of their name , which allows for filtering/searching for specific values. For example, by clicking on the filter icon, a popup window appears that shows a description of the selected attribute and options for filtering it, as shown in **Error! Reference source not found.**. You can click the plus button , for adding disjunctive criteria in your filtering, or removing one, by clicking on the minus button .

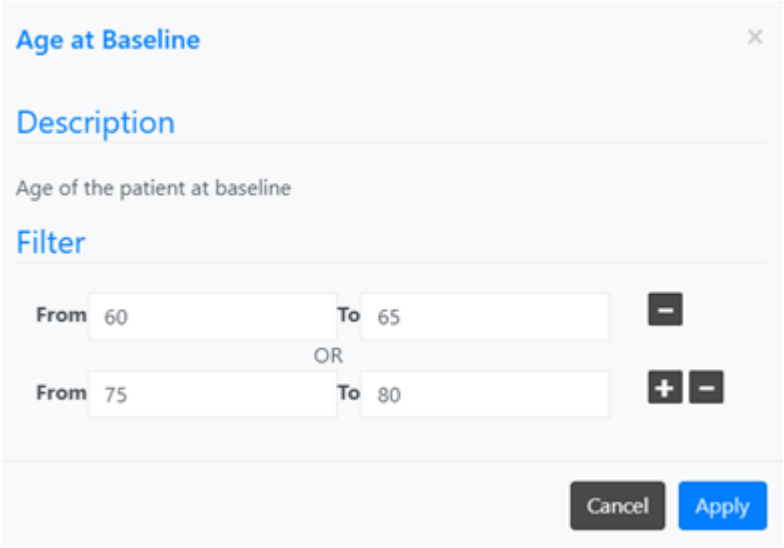


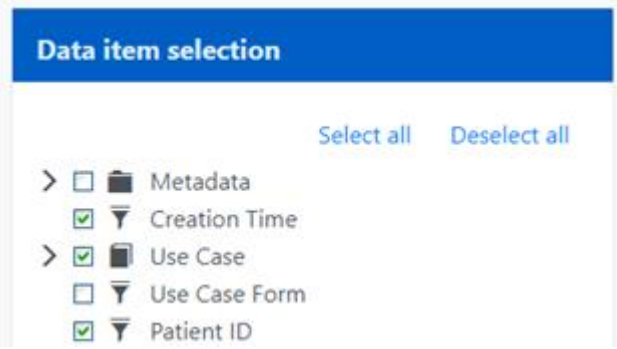
Figure 60: Filtering popup window.

Composite attributes are attributes that encompass several other attributes, acting as a group of the underlying attributes. For example, the attributes "MRI", and "Biopsy" are composite attributes that contain atomic attributes that pertain to information about the MRI and the Biopsy that has been performed. Composite attributes are distinguished from the rest by the folder icon , which when it gets expanded , reveals the underlying attributes. Similarly, you can filter these attributes the same way you can filter an atomic attribute.

Furthermore, some attributes refer to attributes existing in other tables in the repository. This occurs whenever a table has a many-many or one-many relationship with another table. For example, "PSAs" is such an attribute, because every patient can have more than one PSA associating with them. In this scenario, attributes have a square, document-like attribute as shown in **Error! Reference source not found.**.

Figure 61: Attributes that refer to other tables in the repository.

Finally, there is the option to "select all", or "deselect all" attributes to be shown on the page.



In the area with the active filters, you will also find the button Wizard at the bottom right to open the filter wizard. This is a popup screen that shows all the available attributes in one screen and allows to add filters for different attributes at the same time. Figure 62 shows the filter wizard popup window.

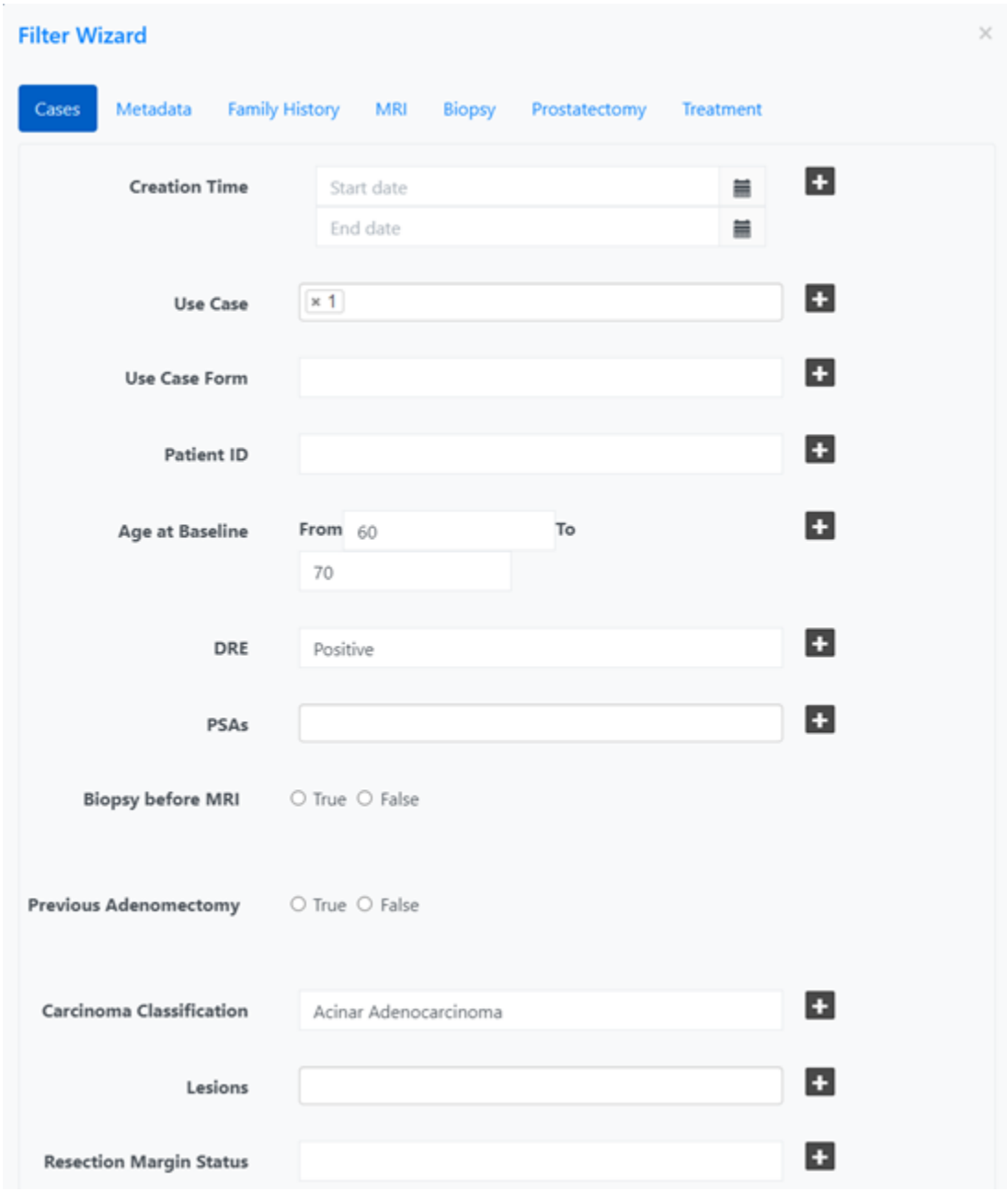


Figure 62: Filter Wizard popup window

Although filtering atomic and composite attributes is straightforward as described above, filtering attributes that refer to other attributes in referenced tables requires more attention in the filter wizard. This is due to the fact that in the filter wizard you can search for these attributes as a "whole entity". For example, searching for PSAs, you can search about the free, total, and ratio PSA, as well as if it is baseline PSA and the days elapsed from the baseline, all at once, as shown in Figure 29. This will match all the use cases in which *all* the PSA-related attributes have the selected values. If you want to filter attributes individually (e.g. filter based on the total PSA only), please filter them by using the attribute selection filtering window, as shown below.

Figure 63: PSA attribute filtering.

View tables

At the top right corner of the screen, a dropdown (table select) is shown which can be used to select any table you may wish to display.

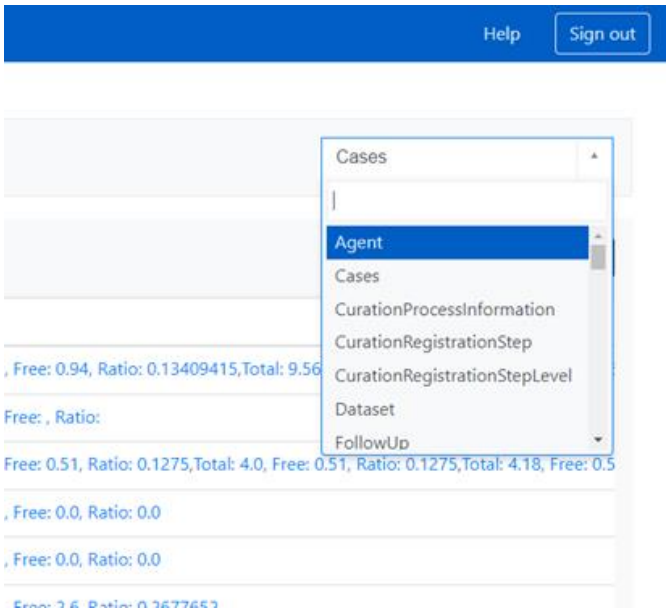


Figure 64: Data explorer table select.

Selecting the "Imaging Studies" menu item

By clicking on the "Explore Data" -→ "Imaging Studies" menu item (Figure 65), you can explore just the imaging study metadata for the cases that have been uploaded into the metadata repository.

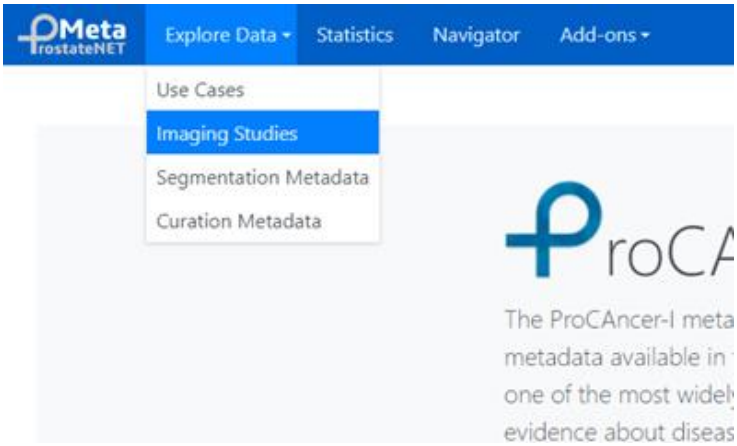


Figure 65: Explore Imaging Study metadata

This includes the Series UID and the Study UID of the MRI images that have been uploaded into the ProCAncer-I Image Repository along with various DICOM-related metadata. You can filter and search image-related metadata in the same way you can search/filter clinical data as described above.

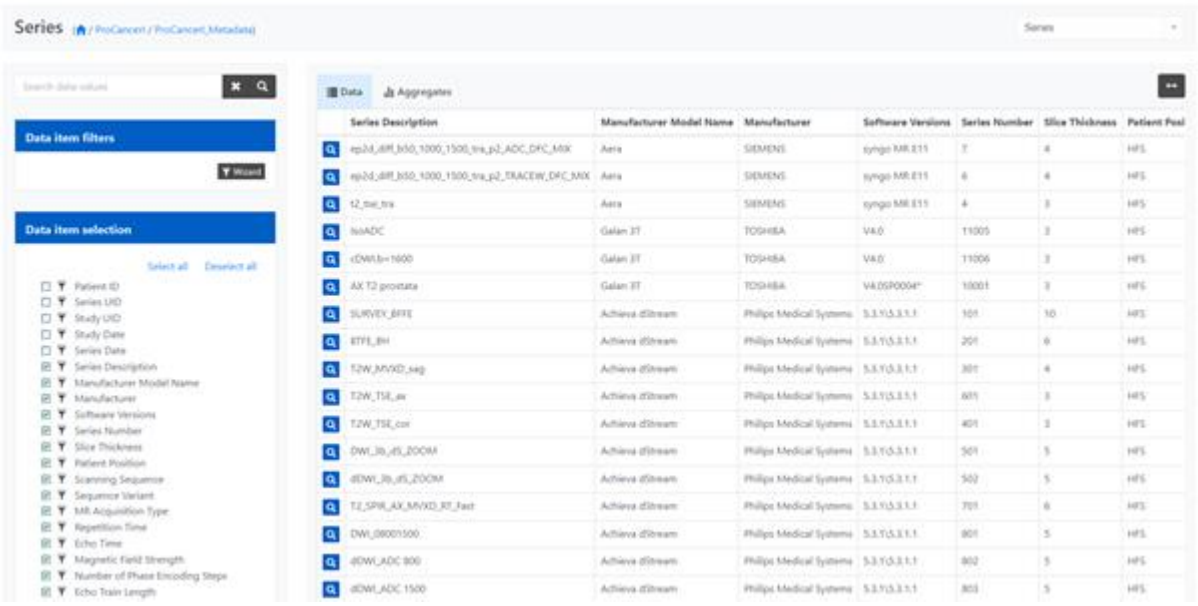


Figure 66: MRI Series metadata view

Selecting the "Segmentation Metadata" menu item

By clicking on the "Explore Data" à "Segmentation Metadata" menu item (Figure 67), you are able to explore the segmentation-related metadata. Currently, the metadata include: the segmentation date, who did the segmentation, the name of the segmented series, the source series id and the study id, the derived series id (i.e. the segmented series), the numbers and names of the segments, the software and version of the segmentation tool used, some algorithmic related metadata (such as the name of the algorithm, method, etc.) and finally the slice thickness, spacing between slices and pixel spacing.

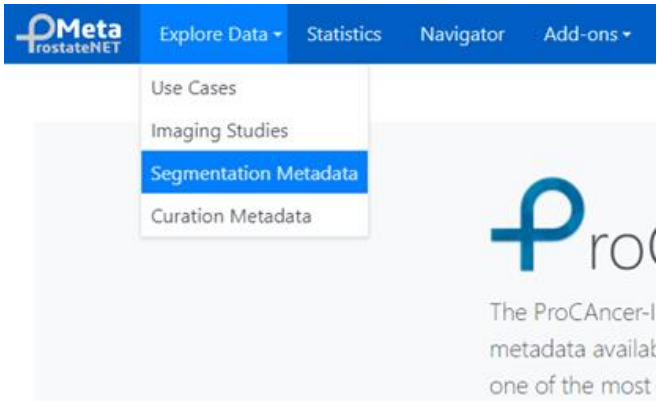


Figure 67: Segmentation Metadata menu item`

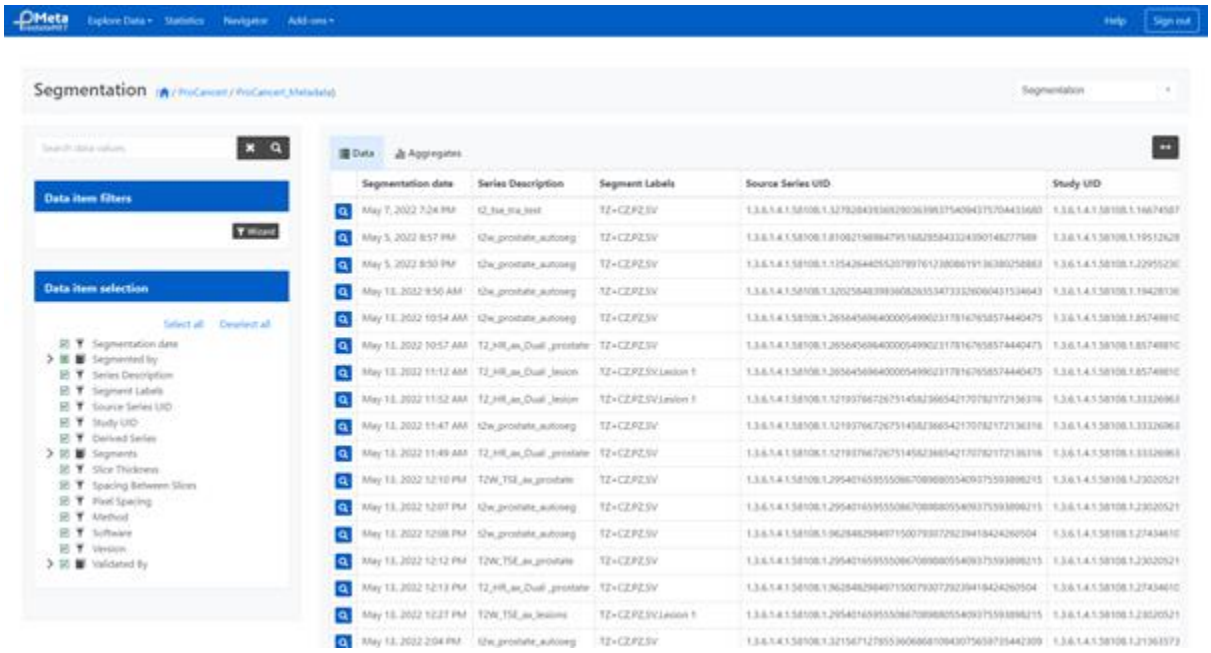


Figure 68: Segmentation metadata

Selecting the “Curation Metadata” menu item

Similarly, by clicking on the “Explore Data” à “Curation Metadata” menu item, you can explore the metadata related to the data curation tasks, co-registration, and motion correction. Currently, the metadata include the date of the curation task, who did the task, the name of the task (co-registration or motion correction), what are the source ids of the images that were curated (which in the case of co-registration there is an additional static series id), the derived series id, and some additional algorithmic related metadata (such as registration steps, parameters, etc.).

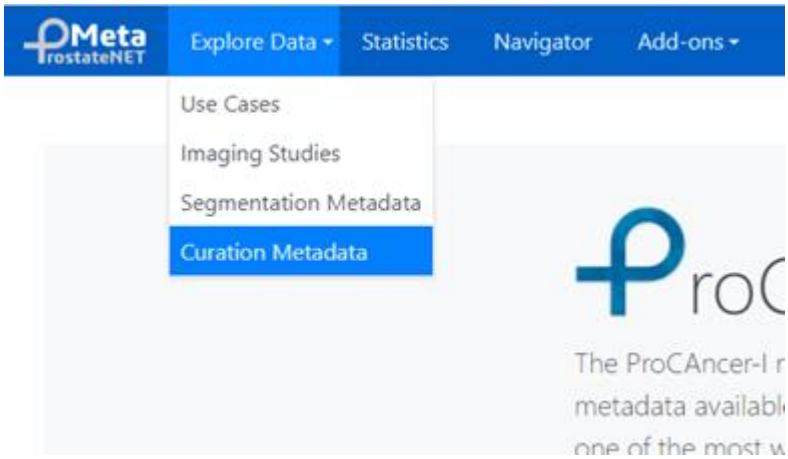


Figure 69: Explore Curation related metadata

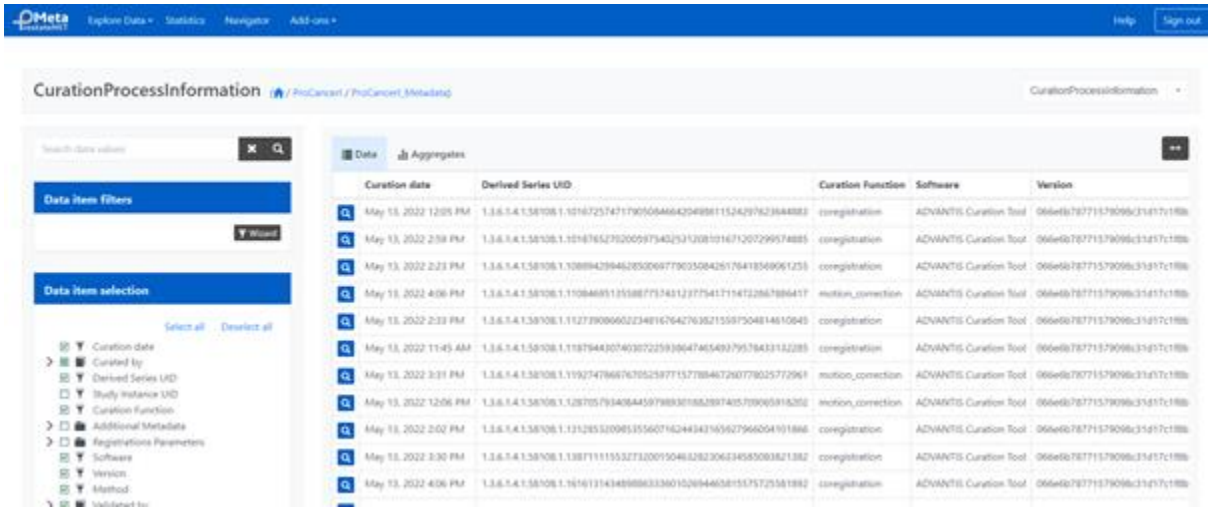


Figure 70: Curation metadata view