# RT-HaND\_I: Updating from PACS- Summary

September 2024

# Introduction

Diagnostic and follow-up imaging for Head and Neck Cancer patients available up until the implementation of EPIC (4th October 2023) were ingested into XNAT. Only scans considered pertinent by H&N clinicians to the diagnosis, staging, treatment and follow-up of head and neck cancer patients were ingested, a summary can be seen in Appendix 1 or the full inclusion/exclusion criteria are in document 12a RT-HaND\_I Imaging Study Descriptions, known Include vs Exclude.

# Background

PACS (Sectra) was queried for the study descriptions and accession numbers of all diagnostic imaging sessions contained for each patient. The resultant study descriptions returned were used to filter the accession numbers to enable the selection of the imaging sessions to ensure only relevant data was ingested using the REST-API into the data lake. Study descriptions considered relevant were any orders pertinent to the diagnosis, staging, progression or ongoing monitoring of treatment side effects of HNC. This included all staging FDG nuclear medicine scans (NM), all staging (e.g. CT chest/abdomen and pelvis) and relevant anatomy (head and neck area) CTs and MRIs and all dental x-ray (XR) and video fluoroscopy (VF) studies. Unspecified CTs were inspected and included if relevant. A brief summary is shown in Appendix 1. The data flow diagram for this process is shown below.

# 

# Ingestion:

## Existing patients imaging data update (aka prospective imaging data collection of existing patients)

To update the lake for previous retrospective patients, PACS should be queried for new studies since the last ingestion. The descriptions should be used to select the studies for ingestion. There is a list of agreed *include* and *exclude* studies to use to initially filter the list. This can be found in document “12a RT-HaND\_I Imaging Study Descriptions, known Include vs Exclude”. It is likely that not all study descriptions will be included on either the *include/exclude* list. This is due to updates in scanning protocols and new machines. A clinician should be consulted regarding the new study descriptions and the *include/exclude* lists updated accordingly. With imaging studies to be ingested filtered out, CSC can then ingest via REST-API with access to the PACS-XNAT pipeline given with respect to clinical priorities.

Patients known to be deceased should not be included in this query. A data extract from EDW can confirm known deceased patients.

## New patients image data upload (prospective imaging data collection of new patients)

Separately, PACS should be queried for new patients. There should be no date limitation on when these studies were performed. The descriptions should be used to select the studies for ingestion as above. After the initial data extraction of these patients, they should be transferred to the “existing” patient list and follow the above patient update pathway after the initial harvest.

# Updating the data availability summary spreadsheet

A spreadsheet summary of the data available in XNAT exists as a .csv appendage to the data lake in XNAT. This enables the XNAT administrator to know easily what imaging is available per patient.

## Ingestions

After all the PACS images have been successfully ingested, a summary spreadsheet of each type (e.g. MRI, CT, planar x-ray) of imaging session can be downloaded. A python script takes these files and produces a summary of the images available per patient with the difference in date of the images from the start date of the radiotherapy. This summary is then uploaded to the data lake as a .csv summary. This gives a good idea of the data available in XNAT.

## Failures to ingest via XNAT

The CSC team will return a list of failures via .csv file of json (JavaScript Object Notation) code. This list contains the patient ID, accession number, study description and unique identifier (UID) of the studies that were not able to be ingested via XNAT. This can be probed to check that the current failure rates are in line with historic upload failure rates and that there is not a pipeline issue.

In the preparation of the HN XNAT database, in a population of 2985 patients, upload failure rates were approximately 7% for CTs & MRIs with all investigated failures being due to imaging sessions being cancelled, images being found under different accession numbers or images being scheduled but not completed or no images available on PACS. This was an expected result due to clinical scientific computing having extensive experience with ingesting CT and MRI scans from PACS for other projects.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Modality | Total failures to ingest | Status |  | Checked |
| CT |  | Cancelled | 28 |  |
|  |  | No images found on PACS | 6 |  |
|  |  | Scheduled but not completed | 1 |  |
|  |  | Images under another study | 11 |  |
|  | 924 |  |  | 46 |
| MRI |  | Cancelled | 6 |  |
|  |  | Images under another study | 4 |  |
|  | 382 |  |  | 10 |
| NM (Nuclear Medicine scan) |  | Cancelled | 38 |  |
|  |  | No images found on PACS | 4 |  |
|  |  | Scheduled but not completed | 10 |  |
|  |  | Available on PACS | 20 |  |
|  | 313 |  |  | 72 |
| Video Fluoroscopy |  | Cancelled | 2 |  |
|  |  | No images found on PACS | 8 |  |
|  |  | Available on PACS | 2 |  |
|  | 628 |  |  | 12 |
| XR (Dental x-ray) |  | Cancelled | 7 |  |
|  |  | No images found on PACS | 4 |  |
|  |  | Available on PACS | 3 |  |
|  | 916 |  |  | 14 |
| Totals | **3163** |  |  | **154** |

Failure rates for NM, Video and XR were typically higher and the presented examples where images are available on PACS but are not ingested by XNAT. A table of investigated failures is shown above. Images available but not ingested included some PET-CTs which are extremely likely to be useful for future research. For this reason, we documented accession numbers that existed but that were not ingested so that individual researchers can decide if they want to manually extract this information to increase their study numbers. For many of the scans it will be obvious that the accession number was unused and therefore no images were ingested because there was a corresponding scan the week after or the images are attached to another study from the same day, but for cases such as the un-ingested PET-CTs we hope it will be useful.

Scans scheduled for 01-01-2099 are explicitly excluded from this list. This list also serves a record to try to re-ingest the cases should XNAT or PACS updates allow the flow of these images from one system to another. Currently XNAT cannot tell if a scan on PACS has been “cancelled” or if there are no images available.

A python script has been written which transforms the .json output into a file that can be appended to the EDW data record.

All python scripts are stored within the RTPHYS GitHub.

## Roles and responsibilities

|  |  |
| --- | --- |
| Role | Responsibility |
| Provide retrospective and prospective patient list | XNAT admin (with EDW assistance) |
| Query PACS for new patient data | CSC |
| Choose scans for ingestion | XNAT admin with clinician input |
| Ingest scans via REST-API | CSC |
| Document failures via .json | CSC |
| Investigate failures | XNAT admin |
| Update available data information & known failures | XNAT admin |
| Upload available data information to EDW | TY (As of Sept 2024) |

## Data stored in XNAT

A brief summary of the volume of data in XNAT (as of September 2024) is shown below;

|  |  |
| --- | --- |
| Diagnostic scan type | Number of scans |
| MRI | 5310 |
| CT | 11599 |
| PET | 1566 |
| Nuclear medicine | 325 |
| Video Fluoroscopy | 486 |
| Dental X-rays | 2885 |

On a per patient basis this corresponds to the below number of patients with at least the below diagnostic or follow-up scans available.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| > | 0 MR/PET-CT | 1 MR | 1 PET-CT | 1 MR and 1 PET-CT | 2 MR | 2 PET-CT | 2 MR and 1 PET-CT | 1 MR and 2 PET-CT | 2 MR and 2 PET-CT |
| 1 CT | 2638 | 1824 | 810 | 590 | 1162 | 351 | 388 | 284 | 202 |
| 2 CT | 1592 | 1206 | 480 | 367 | 825 | 212 | 263 | 179 | 138 |

# Appendix 1: Summary of relevant PACS imaging sessions

Brain-MRI

CTs – relevant categories only – such as staging CT scans, relevant anatomy areas (head and neck area)

Dental

Face-CT

Mandible PA – Dental

MRI – relevant categories only

Neck – CT

NM Whole Body FDG PET (also FDG half body, INGRES, Javelin, Figaro PETs)

OPG – X-ray

PET CT Scan

Research MRI

Unspecified CTs – relevant categories only

Video swallow

VFS

XR Dental categories

XR OPG categories