In-house Developed Application to Collect DICOM Data for Radiotherapy CT Dose Audits

Christopher Hamill-Taylor, Emma McIntosh, Sarah Butt
Ninewells Hospital, Dundee, NHS Tayside

Background: In accordance with the Ionisation Radiation (Medical Exposure) Regulations 2017 [1], all radiation exposures must be optimised. To comply with this, annual dose audits are carried out in radiotherapy, comparing CT dose statistics from radiotherapy planning CT scans to national dose reference levels (NDRLs) for all major treatment sites [2]. This ensures local practice is appropriate and highlights opportunities for further patient dose optimisation.

In NHS Tayside, the local PACS database is manually interrogated to collect the required CT dose statistics (CTDIvol, DLP and scan length). Manual data collection is time intensive and subject to errors due to the large sample size, with 838 patients audited in 2023. A CT dose audit app has therefore been developed locally, which uses the DICOM metadata to collect the dose statistics, replacing the manual process.

Methods: Key DICOM tags were identified for extracting the dose information from CT sets. This was carried out by using Python 3.10 [3] with the Pydicom [4] and Pandas [5]. A data workflow was then identified, in which the app would be able access the collected dose information. A test script was set up to verify the script was able to obtain the correct information. Finally, a user interface was developed using the PyQT5 library [6]. The app was validated, any differences were investigated for root causes, and modifications made as required.

Results: The CT dose audit app was successfully developed. This was validated by comparing the output data from the app to the manually collected data from the 2023 audit. This comparison demonstrated the app was able to reproduce the manually captured data for all treatment sites except Lung 4DCT, where a there was a discrepancy noted in the results. In some 4DCT scans, the number of images can vary between phase sets; the app reviews only the 0% phase which may not always have the maximum number of images. The discrepancy resulted in only a 0.6 mm difference in the mean scan length between the manual audit and app. The accuracy of the app was therefore deemed acceptable, and this was approved for use in the 2024 audit.

Discussion: The introduction of the CT dose audit app has reduced the time to perform this task, allowing annual audit results to be available more quickly. The app can also be used to efficiently assess the effectiveness of any changes in practice, supporting the ongoing work in radiation dose optimisation required under IR(ME)R. The automated nature of data collection avoids transcription errors, as demonstrated through the validation where the app was able to highlight errors in the manual audit data from 2023. At present, the small discrepancy seen in Lung 4DCT remains a known issue and is under investigation.

Conclusion: The CT dose audit app is now in routine use in the department for annual audits, and the 2024 audit has been successfully completed using this app. This has reduced the time burden and number of possible errors compared to the manual audit process. Future updates to the dose audit app will consider integration of CBCT and kV imaging to support the optimisation of on-treatment verification imaging in radiotherapy.

Key references:

[1] 'The Ionising Radiation (Medical Exposure) Regulations 2017'.
[2] T. J. Wood *et al.*, 'IPEM topical report: the first UK survey of dose indices from radiotherapy treatment planning computed tomography scans for adult patients', *Phys. Med. Biol.*, vol. 63, no. 18, p. 185008, Sep. 2018, doi: 10.1088/1361-6560/aacc87.
[3] 'Python', Python.org. Accessed: Mar. 26, 2025. [Online]. Available: https://www.python.org/
[4] Darcy Mason *et al.*, *pydicom/pydicom: pydicom 3.0.1*. (Sep. 22, 2024). Zenodo. doi: 10.5281/ZENODO.13824606.
[5] The pandas development team, *pandas-dev/pandas: Pandas*. (Sep. 20, 2024). Zenodo. doi: 10.5281/ZENODO.13819579.

[6] Riverbank Computing, ‘PyQt5 Reference Guide — PyQt Documentation v5.15.7’. Accessed: Mar. 26, 2025. [Online]. Available: https://www.riverbankcomputing.com/static/Docs/PyQt5/

Using open-source NLP to derive actionable insight from free text patient comments

[1]Rebecca E Woodrow, [1]Neide Simões-Capela, [1]Timothy Cross, [1]Daria Cosentino
Clinical Informatics – Clinical Services Group, HCA Healthcare UK

Key Words: natural language processing, patient experience, R, clinical informatics

Background. Patient feedback drives healthcare improvement **[1]** with free-text comments, providing unique insights around the patient's journey beyond that available from structured questions. However, reviewing vast amounts of patient comments is challenging for clinical staff to process and retain. Advances in natural language processing (NLP), such as sentiment analysis and topic modelling help **[2]** but often rely on complex, data-heavy black-box methods that require secure environments to ensure data protection, or can be difficult to interpret for the end user **[3].** Here, we present how out-of-the-box NLP techniques can analyse free-text patient experience comments within a secure local environment, using modest datasets, enabling non-specialist staff to take meaningful action.

Methods. We analysed free-text comments received via our Patient Experience Survey platform to the question; "Is there anything else you would like to tell us for example things that went well, where we could have done better or specific staff you'd like to mention?". Data were included from 1st April 2024 – 1st March 2025 inpatient visits, after cleaning and de-duplication, resulting in 10,120 unique responses. All analyses were performed locally via R-Studio (v4.2.1). Text data were cleaned and the pre-trained pipeline UDPipe **[4]** was used to run tokenization, Part-Of-Speech (POS) tagging, and lemmatization. Tagged noun frequency was calculated across the period, and per month to track frequent topic evolution. Secondly, an in-house adaptation of SentimentR **[5]** was used to calculate mean unbounded sentiment per sentence (<0 negative, >0 positive), and combined with POS-tagged data to link noun presence with sentence sentiment. Impact **[6]** was estimated as noun frequency x mean sentiment, to identify areas for celebration (frequency > Q3, sentiment > Q3) action (frequency > Q3, sentiment < 0), and monitoring (frequency > Q3 or sentiment < 0). Finally, nouns within each target area were analysed for cooccurrence with other nouns and verbs within sentence, seeking contextual information.

Results. POS tagging identified n=5804 unique lemmatized nouns across the dataset. Most frequent nouns were 'room' (n=1381), 'time' (n=1300), 'experience' (n=1099), 'day' (n=1066), and 'food' (n=839), and were consistently frequent topics per month. 26,381 sentences were included for sentiment analysis, whereby 45.5% included at least one of the top 50 nouns. Examples of positively-associated nouns were attention (0.54, SD = 0.39), and service (0.44, SD = 0.39) and examples of negatively-associated nouns were hour (-0.03, SD = 0.26) and medication (-0.01, SD = 0.32). Three nouns were highlighted for celebration; 'stay', 'service', and 'experience', with high cooccurrence additionally present with other nouns such as 'nurse' and 'staff', and descriptors such as 'excellent' and 'positive'. Two nouns were highlighted for action; 'room' and 'hour', however showed little cooccurrence. Instead, 'room' presented with other nouns such as 'food', 'tv', and 'floor', and descriptors such as 'cold'. A further n=10 nouns were identified for monitoring with low sentiment but infrequent, and n=7 identified as frequent but within indiscriminate sentiment.

Conclusion. These explorations used out-of-the-box NLP tools to identify key themes in patient comments, guiding celebration, monitoring, and action. They require no model training, work in a secure local environment, and importantly remain interpretable for non-specialist, enabling staff to focus on and drive hospital-wide improvements.

Key references

[1] Doyle C, Lennox L, Bell D. *A systematic review of evidence on the links between patient experience and clinical safety and effectiveness*. BMJ Open. 2013 Jan 3;3(1):e001570.
[2] Khanbhai M, Anyadi P, Symons J, Flott K, Darzi A, Mayer E. *Applying natural language processing and machine learning techniques to patient experience feedback: a systematic review*. BMJ Health Care Inform. 2021 Mar;28(1):e100262.
[3] Velupillai S, Suominen H, Liakata M, Roberts A, Shah AD, Morley K, Osborn D, Hayes J, Stewart R, Downs J, Chapman W, Dutta R. *Using clinical Natural Language Processing for health outcomes research: Overview and actionable suggestions for future advances*. J Biomed Inform. 2018 Dec;88:11-19.
[4] Straka M, Straková J. *Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe*. Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies, pages 88–99, Vancouver, Canada. Association for Computational Linguistics.
[5] Rinker, T. W. *sentimentr: Calculate Text Polarity Sentiment*". 2021. version 2.9.1., https://github.com/trinker/sentimentr

[6] Cammel, S.A., De Vos, M.S., van Soest, D. et al. How to automatically turn patient experience free-text responses into actionable insights: a natural language programming (NLP) approach. BMC Med Inform Decis Mak 20, 97 (2020).

When IT says no: going it alone (with AI help) to make a proof of concept and a call for change

Dom Withers - Radiotherapy Physics - Barking, Havering & Redbridge University Hospitals NHS Trust

Aims: In September 2024, an amendment to IR(ME)R 2017 introduced a requirement to establish Diagnostic Reference Levels (DRLs) for verification images acquired on linacs. Our department generates approximately 60 CBCT DICOM Radiation Dose Structured Reports (RDSRs) per day, which include CBCT protocol parameters such as Dose Length Product (DLP), used to define DRLs. Manual extraction via DICOM viewers or text editors is possible but inefficient: a colleague in Diagnostic Radiology spent over 80 hours extracting data for a year's data from a single room. A simple script is the obvious solution. To date, most departmental computing projects have not involved the use of patient-identifiable data. In October 2024, the Trust's IT department was asked to install Python and MySQL on a Trust PC to enable development of a data extraction tool and database. The request was denied, citing security concerns and the absence of a Trust software development policy. This work describes how development progressed, using non-patient data, to show proof-of-concept but also to show Trust and IT management that such projects need to be brought "in from the cold", and that high-level support of scientific computing is essential in modern healthcare provision.

Method: This work was carried out on a personal, i.e. non-Trust, dual-boot Windows 11/Linux Mint PC, with no prior experience in Python, some MySQL knowledge, above-average Linux skills, and using ChatGPT Plus (self-funded) to guide the process. Initial prompts generated scripts extracting simple text outputs from RDSRs and progressed to more complex functionality. Phantom CBCT RDSRs were used to avoid using patient data—initially 4 for basic testing, later expanded to 740 for batch processing.

Results: The first generated Python script output the full set of DICOM file structure, tags, descriptions and values. In successive refinements, relevant data were identified and isolated. Further prompts guided installation and configuration of MySQL and development of a Python script to populate a database. Another prompt generated a Python Flask web interface, displaying selectable fields with an "Export to CSV" option. Another version of the script was developed to search directories deeper than one level and write extracted data directly to CSV files. The final versions extracted dose data from 740 RDSRs into both database and CSV formats. Tests on a selection of RDSR files showed that the data had been extracted correctly. Working versions were created in Linux and Windows. The total time spent creating these tools was around 8 hours.

Discussion: It is freely acknowledged that a little knowledge and the use of ChatGPT is absolutely not an optimal way to go about developing software, but a lot was learned and it was completely effective in obtaining the required results for this simple task. Checking ChatGPT output is important, and testing is essential: whilst it makes mistakes, finding and correcting them actually adds to the learning experience. Creating a bespoke tool for extracting useful dose data automatically into an analysable format took a short amount of time. Even in their present forms, these tools could be useful for analysing and auditing doses on the Trust's three linacs and radiology equipment. Further development could include tools to analyse results and display them graphically. Support for installation on Trust PCs is essential if this is to be used for patient data.

Conclusion: With accessible tools, little experience and some curiosity, basic data extraction tools can be created. The resulting workflow facilitates auditing of CBCT verification doses in compliance with IR(ME)R 2017. It is hoped this example will act as both a practical solution and a catalyst for organisational change in the Trust, to encourage it to implement policies that support scientific computing and in-house software development in MPCE and other departments - without having to resort to external resources or un-safe/insecure practices.

Key Words: DICOM, DRLs, IR(ME)R, Python, MySQL