

Meta-Analysis of Radiotherapy Audit Data

Elizabeth A. COOKE^{a,1}, Ellie SMYTH^a, Spencer A. THOMAS^a, Catharine H. CLARK^{a,b,c} and Mohammad HUSSEIN^a

^a National Physical Laboratory, Hampton Road, Teddington, TW11 0LW, UK

^b University College London NHS Foundation Trust, London, NW1 2PG, UK

^c University College London, London, WC1E 6BT, UK

ORCID ID: Elizabeth A. Cooke <https://orcid.org/0000-0003-3843-8393>, Spencer A.

Thomas <https://orcid.org/0000-0002-9260-1561>, Catharine H. Clark

<https://orcid.org/0000-0003-2287-8212>, Mohammad Hussein <https://orcid.org/0000-0002-1597-740X>

Abstract. Standardised digital records can improve access and quality of health services and enable meta-analysis of data which may reveal unknown patterns, which improve our understanding of the data. In this study, we report on meta-analysis of data curated from advanced radiotherapy dosimetry audits conducted at hospitals across the UK using test objects developed at the National Physical Laboratory. This meta-analysis highlights hospitals which are performing within expectation, or outside of expected intervals and would benefit from measurement support. Anonymised hospitals with low precision or accuracy treatment plans are identified, enabling support and improvements where appropriate. The results presented may be used to provide insight to hospitals and inform areas of focus for improved predictions of radiation dose that are tailored to a given hospital. Moreover, this analysis can enable auditors and regulators to provide additional services or recommendations, and potentially identify previously unknown patterns or dependencies in the data.

Keywords. Radiotherapy audits; meta-analysis; data visualisation; standardisation

1. Introduction

In the current digital age, data from measurements are collected routinely for a wide range of applications. Curating these data in structured tables can greatly enhance data access, searchability and reuse through data queries [1, 2]. In addition, the use of digital forms to input data can lower the incidence of human error. As use of digital technologies increases, storing digital information in standardised and searchable formats is crucial for complex areas such as healthcare [3, 4], and in particular in those areas related to clinical decision making, such as in radiotherapy. Standardised digital records can improve access and quality of health services and enable meta-analysis of data which may reveal previously unknown patterns, clusters or correlations improving our understanding of the data, measurements, and their application domains [5, 6].

In this study, we use retrospective data curated from a range of advanced radiotherapy dosimetry audits conducted at hospitals across the UK to show how meta-

¹Corresponding Author: Dr Elizabeth Cooke, National Physical Laboratory, Hampton Road, TW11 0LW, UK; E-mail: elizabeth.cooke@npl.co.uk.

analysis can be used to improve operational processes and patient care. We describe a meta-analysis of the audit data to obtain insights into audit performance. In particular, we highlight the use of this analysis in comparing the performance of different hospital treatment planning systems (TPS). This analysis can identify hospitals that would benefit from measurement support, such as improved use, optimisation, and refinement of their TPS model parameters to predict the dose distribution, or improved local measurements for quality assurance of radiotherapy treatment. This may lead to higher accuracy and precision in radiotherapy, ultimately leading to reduced radiation outside the target volume, improved treatment and patient outcomes.

2. Methods

The dosimetry audits analysed here assess the positional and dosimetric accuracy of Stereotactic Body Radiotherapy (SBRT; also known as Stereotactic Ablative Body Radiotherapy, SABR) lung and spine treatment delivery. This is achieved using alanine pellets (solid state dosimeters) [7] and GafChromic film [8] dosimetry, placed in a test object developed at the National Physical Laboratory (NPL) and designed to imitate the thorax. The audits contain several predictions and measurements, and information regarding the type of equipment used both at the hospitals and by NPL in the audit.

Here, we focus on retrospective data from lung ($n=49$) SBRT dosimetry audits with an average of 9 ± 2 dosimetry measurements per audit. We compare the dose predicted by each hospital's TPS and the measured dose in the dosimeters in the test object. Comparing the predicted and measured doses, we can assess the quality of a hospital's radiotherapy planning as part of the audit process. The predicted value is calculated by each respective hospital using their radiotherapy TPS, which is specialised software that simulates the radiotherapy dose deposition. The measured value is calculated by irradiating alanine pellets in the test object and correcting the measurements using a reference measurement.

We extract, standardise, and analyse the data from each audit conducted by NPL, ensuring all units are consistent and data is recorded in the same format. We generate visualisations of the data to allow hospitals and auditors to be able to easily visualise and compare audit data from across multiple hospitals and provide recommendations for where improvements are required for individual sites or for new protocols.

Since early 2024, NPL has developed the Dosimetry Audits for Advanced Radiotherapy (DAART) database [9], see Figure 1, which will streamline the future process of gathering and analysing audit data, generating audit reports for the hospitals, and reduce errors in data transfer. The DAART database also allows auditors to visualise and compare different audit data conducted at the same, or different (anonymised) hospitals.

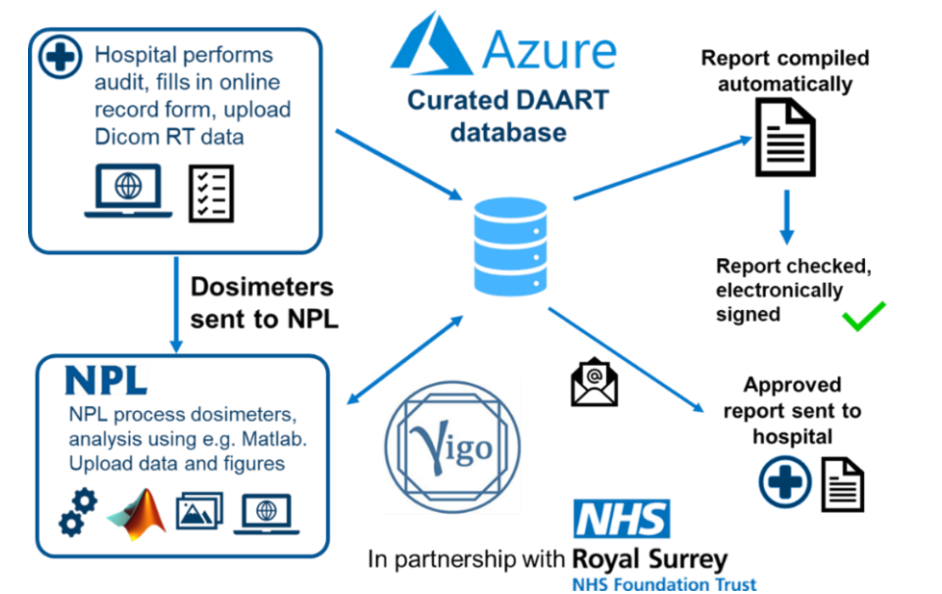


Figure 1: Schematic of the DAART database. Hospitals are able to upload the results from an audit directly to the database, where NPL can then process and analyse the data. The data is stored in a structured, standardised format, allowing additional analysis and meta-analysis across audits.

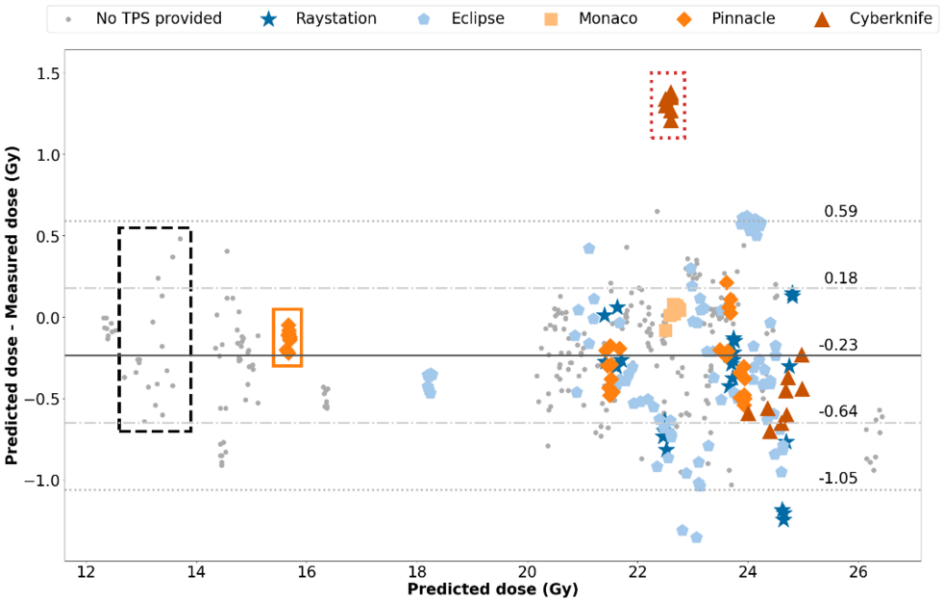


Figure 2: Comparison plot of the lung SBRT measured and predicted dose values. Each point is one of the measurements from an audit of a hospital (49 audits, 435 measurements). The mean (solid black line), and one (dot-dashed grey lines) and two standard deviations (dotted grey lines) either side of the mean are shown, allowing the easy identification of outliers. Coloured symbols highlight different TPS used, allowing comparison with similar set-ups. Boxes highlight measurements from specific hospitals showing low precision (black dashed), low accuracy (red dotted) and high precision and accuracy (orange solid) measurements.

3. Results

Figure 2 shows a comparison of the lung SBRT measured and predicted dose values for retrospective audits conducted by NPL. The difference between predicted and measured dosages from all audit data is observed to be not normally distributed using the Kolmogorov-Smirnov test [10] ($p = 10^{-44}$), therefore all statistics are computed using the bootstrap method [11] with 200 samples. Figure 2, along with output summary statistics, allows hospitals and auditors to be able to easily visualise and compare audit data.

In Figure 2, each datapoint is highlighted according to the TPS used at the hospital. This allows a comparison of audited TPS to examine comparability between systems and hospitals. We observe that different TPS are comparable, showing similar variability across all systems. Being able to compare like-for-like systems is critical to understanding the source of any deviations and ultimately improve measurements and patient care.

4. Discussion

The results presented in Figure 2 may be used to provide insight to hospitals by giving a quantitative comparison between them and other sites using similar equipment. This information may inform areas of focus for improved predictions of dose, or measurements, that are tailored to a given hospital, such as focusing on accuracy or precision, or both. Moreover, this can enable auditors and regulators to provide additional services or recommendations, inform or motivate new standards, set suitable tolerances for future audits, and potentially identify previously unknown patterns or dependencies in the data.

Comparing all audits allows hospitals and auditors to highlight instances where audits have highly repeatable measurements, i.e., low variance; have highly variable measurements (black dashed box, Figure 2); have small differences between measured and predicted values (solid orange box, Figure 2); or are outside the 95 % confidence interval for dose difference (dotted red box, Figure 2).

Statistical analysis of intra-audit measurements can reveal outliers and underlying distributions in the data, which may be useful when analysing audit performance. For example, the audit measurements highlighted with the red dotted box may be easily flagged by an auditor and a follow-up audit scheduled.

In the data shown in Figure 2, 62 % of datapoints have no TPS provided. This highlights the need for standardised data formatting and collection. By using a standardised database, DAART in future will decrease the prevalence of missing or unusable data. For example, by using prescribed data formats and recording of measurement units in a consistent format which are checked upon submission.

5. Conclusions

Meta-analysis of radiotherapy audits may reveal previously unknown patterns or clusters which may improve treatment and patient outcomes. In this study, meta-analysis of radiotherapy audit data easily highlights hospitals which are performing within expectation, or outside of expected intervals. Where hospitals' measurements can be improved, NPL can assist with training in good measurement technique, as well as help

develop new and improved protocols. We have also identified a large proportion of missing data related to the TPS and the requirement for standardised data collection in order to perform comparisons. Future work will continue to analyse the audit data as the curated DAART database grows, allowing further meta-analysis of standardised data. In addition, we will investigate and compare other audit types, such as spine SBRT, to identify new trends and patterns within these audits.

Once the DAART database is fully operational and contains a larger sample of audit data, future work into machine learning models and feature importance will be useful to uncover any hidden trends or influences within the audit process. Any such trends will provide evidence that certain conditions of the audit may be influencing its outcome, allowing auditors and hospitals greater insight when planning future treatments [12].

Consistent standardised data is critical for metrologically-sound analysis, particularly in patient-critical care such as radiotherapy treatment planning. Storing digital information in standardised and searchable formats is crucial for future work in data analysis in healthcare.

6. Acknowledgements

This work was funded by the UK Government's Department for Science, Innovation & Technology through the UK's National Measurement System programmes.

References

- [1] Smith NAS., Sinden D, Thomas SA, Romanchikova M, Talbott JE, Adeogun M. Building confidence in digital health through metrology. *Br J Radiol.* 2020; 93(1109):20190574, doi: 10.1259/bjr.20190574.
- [2] Thomas SA, Brochu F. Curation at the point of measurement and traceability of measurement workflows. *Measurement: Sensors.* 2022; 23:100399, doi: 10.1016/j.measen.2022.100399.
- [3] Fadahunsi KP, O'Connor S, Akinlua JT, et al. Information Quality Frameworks for Digital Health Technologies: Systematic Review. *J Med Internet Res.* 2021; 23(5):e23479, doi: 10.2196/23479.
- [4] Adane K, Muluye D, Abebe M. Processing medical data: a systematic review. *Arch Public Health.* 2013; 71(1):27, doi: 10.1186/0778-7367-71-27.
- [5] Schulz S, Stegwee R, Chronaki C. Standards in Healthcare Data. In Kubben P, Dumontier M, Dekker A. *Fundamentals of Clinical Data Science.*: Springer; 2019. 19-36 p, doi: 10.1007/978-3-319-99713-1_3.
- [6] Wang X, Zhang X, Li Z, Zhong W, Yang P, Mao C. A brief introduction of meta-analyses in clinical practice and research. *J Gene Med.* 2021; 23(5):e3312, doi: 10.1002/jgm.3312.
- [7] Sharpe P, Sephton J. Therapy level alanine dosimetry at the NPL. In: Anton, M, editor. *Proceedings of the 216th PTB Seminar "Alanine Dosimetry for Clinical Application"*; 2006; Braunschweig.
- [8] Ashland. GAFCHROMIC EBT FILM. [Online].; 2023 [cited 2024 Oct 23]. Available from: <http://www.gafchromic.com/gafchromic-film/radiotherapy-films/EBT/index.asp>.
- [9] National Physical Laboratory. Dosimetry Audits for Advanced Radiotherapy. [Online].; 2024 [cited 2024 Oct 24]. Available from: <https://daart.npl.co.uk/>.
- [10] Massey Jr., FJ. The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association.* 1951; 46(253):68-78, doi: 10.2307/2280095.
- [11] Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*: Chapman & Hall; 1993.
- [12] Noblet C, Maunet M, Duthy M, Coste F, Moreau M. A TPS integrated machine learning tool for predicting patient-specific quality assurance outcomes in volumetric-modulated arc therapy. *Physica Medica.* 2024; 118:103208, doi: 10.1016/j.ejmp.2024.103208.