Data Explained

Radiotherapy Dataset (RTDS)

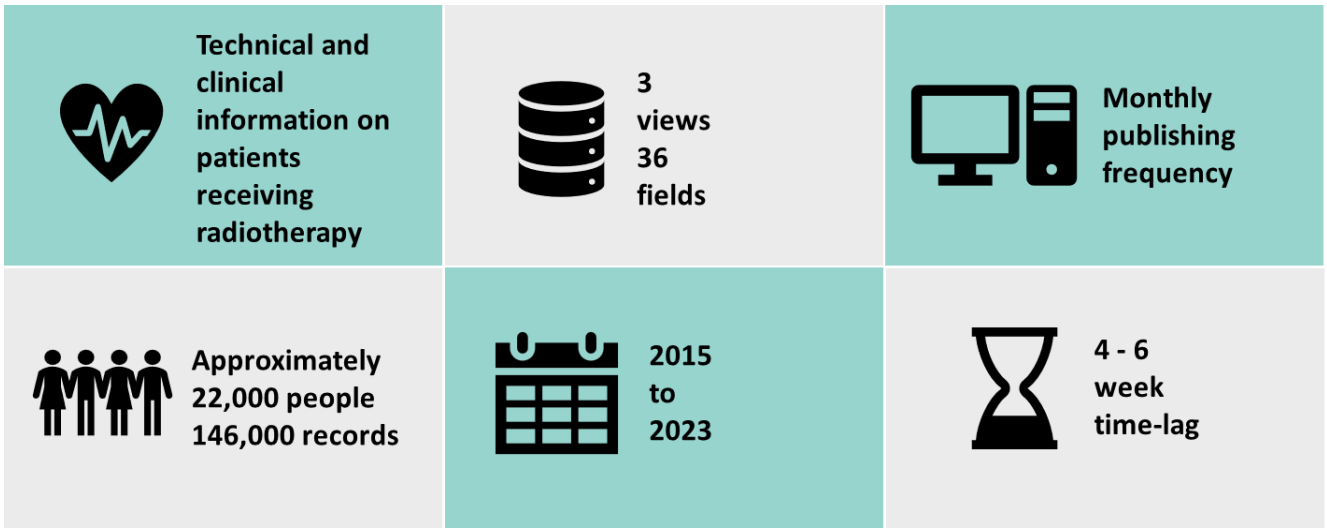
Authors: Dr Laura Baker, Dr Arron Lacey, Professor Ashley Akbari
Date: February 2025

This Data Explained summarises the Radiotherapy Dataset (RTDS). This output is intended to help guide users interested in this data for the first time, and future use of the data within the SAIL Databank towards research delivery and to provide feedback into future data source development and documentation.

The data discussed in this Data Explained was made securely available through an approved SAIL project via the independent Information Governance Review Panel (IGRP) – project 1598. The data used in this approved SAIL project comes from Digital Health and Care Wales (DHCW) and was accessed through the Secure Anonymised Information Linkage (SAIL) Databank. The data was not originally collected for research and it is expected that there are gaps and inconsistencies in its recording, a number of which are detailed in the following. The work presented in this Data Explained is correct at the point of publication. Views expressed in this Data Explained are those of the researchers and not necessarily those of ADR Wales partner organisations.

Overview

The Radiotherapy Dataset (RTDS), accessible through the SAIL Databank, includes detailed information on radiotherapy treatments in Wales. The infographic below summarises the data sources' features at the time of this publication.



Introduction

Cancer is a complex and significant public health issue, being a leading cause of death globally (Siegel et al., 2023; WHO, 2020). The increasing incidence of cancer, combined with an ageing population, places greater demand on healthcare services and highlights the need for effective detection, treatment, and intervention strategies.

Radiotherapy is a widely used cancer treatment that employs radiation (e.g., x-rays) to target and destroy cancer cells. It is a crucial component of many cancer treatment plans, particularly for localised cancers (CRUK, 2023). However, despite its effectiveness, radiotherapy can cause undesirable side effects and health complications. Understanding these effects is essential for improving treatment strategies and patient outcomes.

The RTDS was initially known as the 'National Radiotherapy Data Set' and was developed and implemented by Public Health England (PHE) in 2009 (Dawson, 2015). It aimed to gather technical and clinical information on patients receiving radiotherapy. The primary objectives were to:

"*To improve availability and use of data already held about treatment given to patients by radiotherapy facilities and to direct the development of systems supporting data collection and use in the small number of facilities that do not yet hold such data*" (Wells, 2015). In addition, it also aims to: "*Collect consistent and comparable data across all NHS providers of radiotherapy services in England and Wales in order to provide intelligence for service planning, commissioning, clinical practice and research and the operational provision of radiotherapy services across England and Wales*" (Cancer Implementation Group, 2022). Furthermore, as a standard, it requires all NHS radiotherapy providers to: "*Collect and submit standardised data monthly against a nationally defined dataset*" (Cancer Implementation Group, 2022).

How is the data collected?

In 2015, the NHS Wales Informatics Service (NWIS), which is now known as Digital Health Care Wales (DHCW), facilitated the introduction of RTDS in Wales. This was achieved by adopting the RTDS Management Service from England, in collaboration with the National Clinical Analysis and Specialised Applications Team (NATCANSAT), to deliver the service in Wales (Wells, 2015).

The implementation of RTDS affected the three radiotherapy service centres in Wales:

- Velindre Cancer Centre, Cardiff.
- Singleton Oncology Centre, Swansea.
- North Wales Cancer Treatment Centre, Rhyl.

In 2016, PHE took over the management of RTDS Version 5 and partnered with the Welsh Cancer Intelligence and Surveillance Unit (WCISU) to fully implement RTDS in Wales. Together, they developed a system that enabled radiotherapy centres to submit data on a monthly basis, starting with patients treated from the previous month (Cancer Implementation Group, 2022).

RTDS data submissions from each centre began on the following dates (Cancer Implementation Group, 2022):

- South West Wales Cancer Centre: April 2016.
- North West Wales Cancer Centre: March 2018.
- Velindre Cancer Centre: November 2018.

The Welsh data is a subset of the 'National Radiotherapy Data Set,' focusing specifically on radiotherapy data items, with fewer demographic details. It covers all patients receiving the following types of radiotherapy:

- Teletherapy: External beam radiotherapy.
- Brachytherapy: Internal radiotherapy delivered using automated remote after loading machines.
- All other Brachytherapy: Internal radiotherapy for malignant disease (Wells, 2015).

Data from all radiotherapy-providing healthcare centres in Wales is collected and submitted monthly to DHCW. These submissions include patient records from the preceding month. The initial Welsh Government data standard specified that RTDS data is generated using toolkits provided by the RTDS Management Service (Dawson, 2015). These toolkits extract local data from Oncology Management Systems and the Patient Administration System (PAS), converting it into the required RTDS format (Dawson, 2015). The resulting files are then sent via the NATCATSAT data upload facility (N3) as encrypted Microsoft Access files (Dawson, 2015).

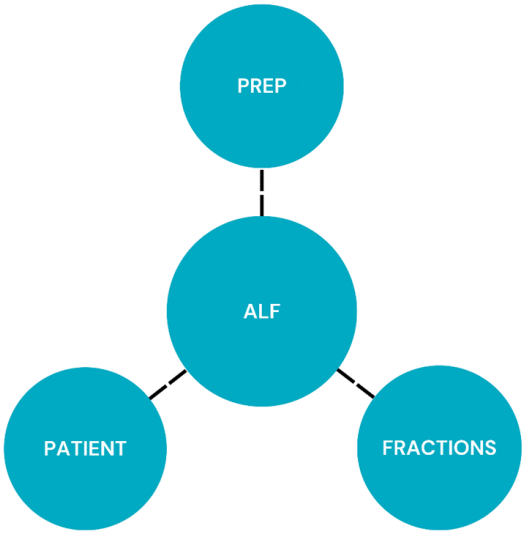
NHS England has produced detailed documentation, including a user guide (National Disease Registration Service, 2021) and a technical guide (Public Health England, 2021), which provide extensive background information on the data fields and the process radiotherapy data undergoes

before being finalised in RTDS. Additionally, a cohort profile focused on NHS England (NHSE) has been published, offering further context for working with the dataset (Sandhu et al., 2023).

Data Linkage

RTDS holds detailed individual-level records that document each individual through their radiotherapy treatment. Each record is linked via an encrypted NHS number (ALF within the SAIL Databank).

RTDS data in SAIL contains 36 fields and is organised into three views (**Figure 1**).
Further details and meta-data on the RTDS data source can be found on the online meta-data gateway (HDR UK, 2024).



- **RTDS.Patient**:
  Demographic and diagnostic data, made up of 11 fields.
- **RTDS.Prep:**
  A mixture of fields with some procedure information presented. This view provides eight fields.
- **RTDS.Fractions:**
  Radiotherapy treatment data, including 17 fields.

Figure 1. Overview of data views and common linkage fields in the radiotherapy dataset within the SAIL Databank.
Three views currently exist for RTDS, with the *alf_pe* (a double encrypted NHS patient identifier) used to facilitate data linkage across the views. Metadata relating to the fields can be found here.

Accessing RTDS through the SAIL Databank allows radiotherapy treatment data to be linked with other data sources, including routinely-collected electronic health record (EHR), administrative, survey, registry and other data sources. A full list of data sources available via the Databank is provided here:
https://healthdatagateway.org/en/search?type=datasets&sampleAvailability=&publisherName=SAIL.

This data linkage supports the development of more specific and targeted research questions. For example, it enables studies on how radiotherapy treatment parameters might influence the onset of other comorbidities, as well as many other ways in which linkage could be used to evaluate trends at the population level.

What can the data be used for?

At the time of the release of this Data Explained output, the RTDS in the SAIL Databank contains data coverage between 2010 to 2023, with records containing a decision-to-treat date spanning from 2010 to 2023 and treatment end dates ranging from 2018 to 2023. The data source contains approximately 146,000 records, representing over 22,000 unique individuals (**Table 1**).

Table 1. Overview of the number of records, unique identifiers and null values in the radiotherapy dataset views. The number of total records and unique encrypted NHS numbers (ALFs) within each view of the RTDS data source, providing insight into the scope and distribution of data across different individual-related views.

Counts	RTDS View		
	Fraction	Patient	Prep
Total Records	145,238	24,687	87,828
Number of Patients (ALFs)	22,076	22,679	22,896
Null Values	2,219	26	1,758

As a standalone data source, RTDS data can be used to identify trends in radiotherapy treatment. Several examples have been carried out to demonstrate the capability of the data, providing insight into the broad potential of the RTDS data source when designing population-level cancer studies.

Example 1. A longitudinal overview of the number of planned radiotherapy treatments and the number of individuals receiving radiotherapy treatment in Wales.
Using RTDS, it is possibly to identify trends in the number of individuals undergoing radiotherapy treatment (**Figure 2**).

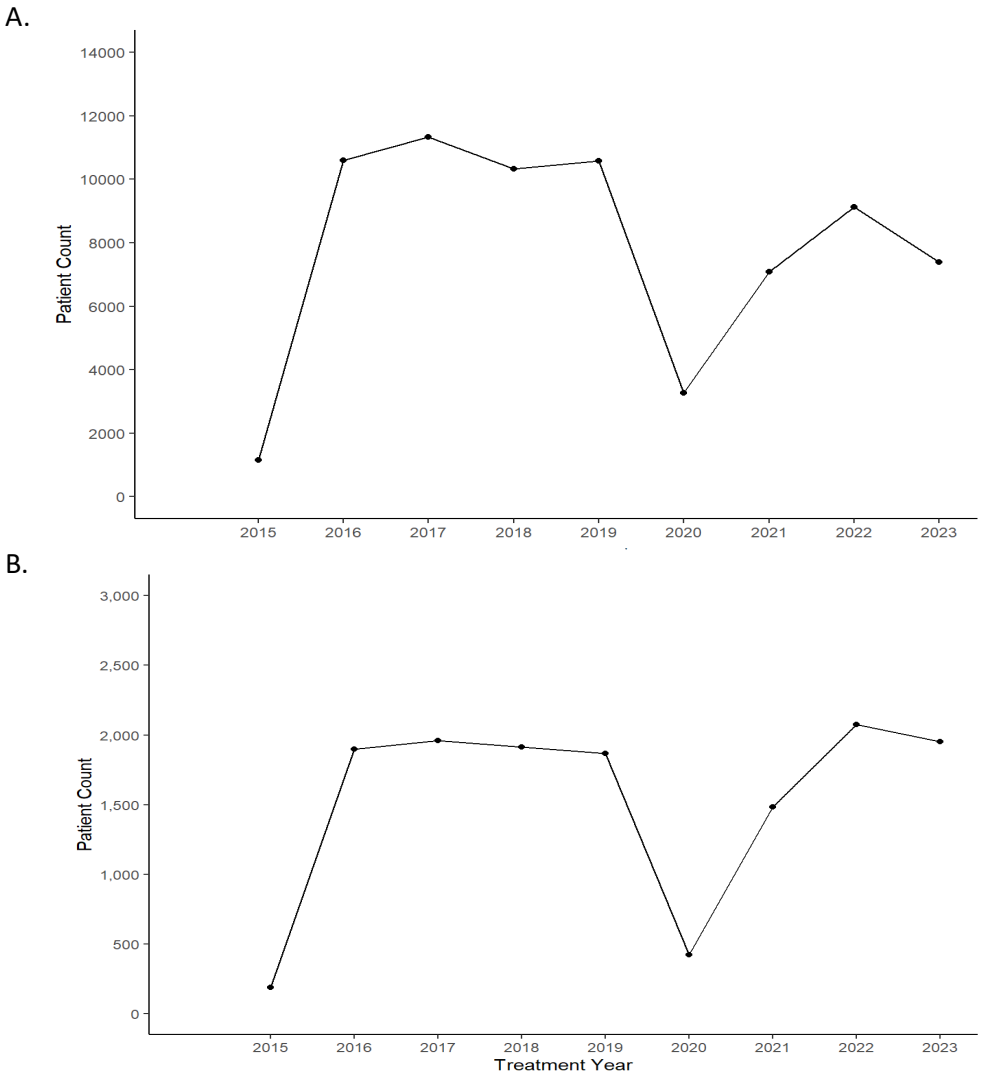


Figure 2. Annual count of radiotherapy treatment plans and individuals receiving radiotherapy in Wales.
A. **Longitudinal view of the number of radiotherapy treatment plans in Wales**. Distinct patient identifiers (ALFs) were counted per year based on the treatment start date.
B. **Longitudinal view of the number of individuals receiving radiotherapy treatment in Wales**. Distinct ALFs were counted per year based on the documented year of diagnosis.
Counts of fewer than 10 individuals and cases with no recorded decision-to-treat date were excluded. The analysis was limited to the years 2015-2024 due to the available data coverage.

Example 2. Frequency of radiotherapy treatments issued per healthcare provider in Wales.
RTDS allows for the identification of the number of individuals who have received radiotherapy treatment per healthcare provider by using the *organisation_name* field and counting ALFs. Additional parameters, such as decision to treat date, are also captured within the RTDS, enabling further refinement of incidence analysis (**Figure 4**).

This example demonstrates the available individual counts when designing a study using RTDS. It's important to note that many of the "decision to treat" dates occurred before the 1900s, requiring a data cleaning process to ensure accurate yearly counts.

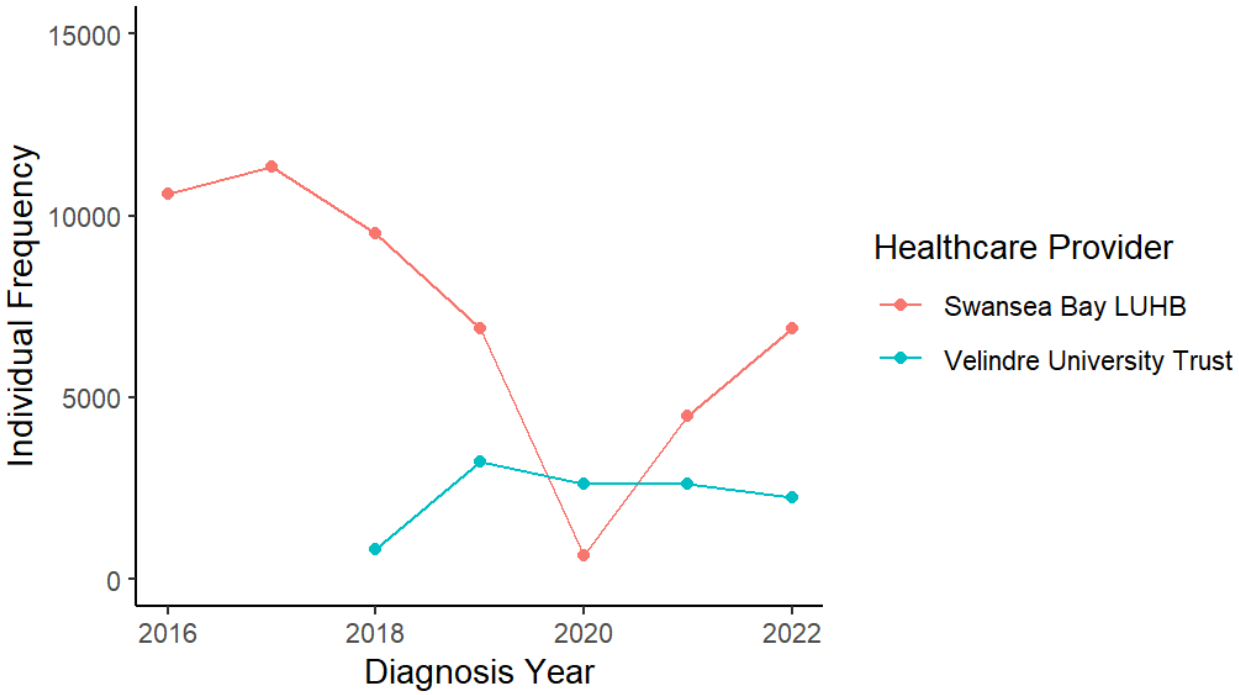


Figure 3. A longitudinal overview of the number of radiotherapy treatments per healthcare provider in Wales.
The number of individuals (ALFs) was counted annually for each healthcare provider. Data points with values less than or equal to 10 were excluded from the analysis. Additionally, any data where the decision to treat occurred before the 1900s was also excluded (As such, all data from North Wales was excluded from this Data Explained output).

Linking different views within the RTDS data source offers the potential for more refined analysis of radiotherapy incidence. For example, joining *RTDS.patient* and *RTDS.fractions* allows for the exploration of sex-specific differences in radiotherapy treatment intentions in Wales (**Figure 4**).

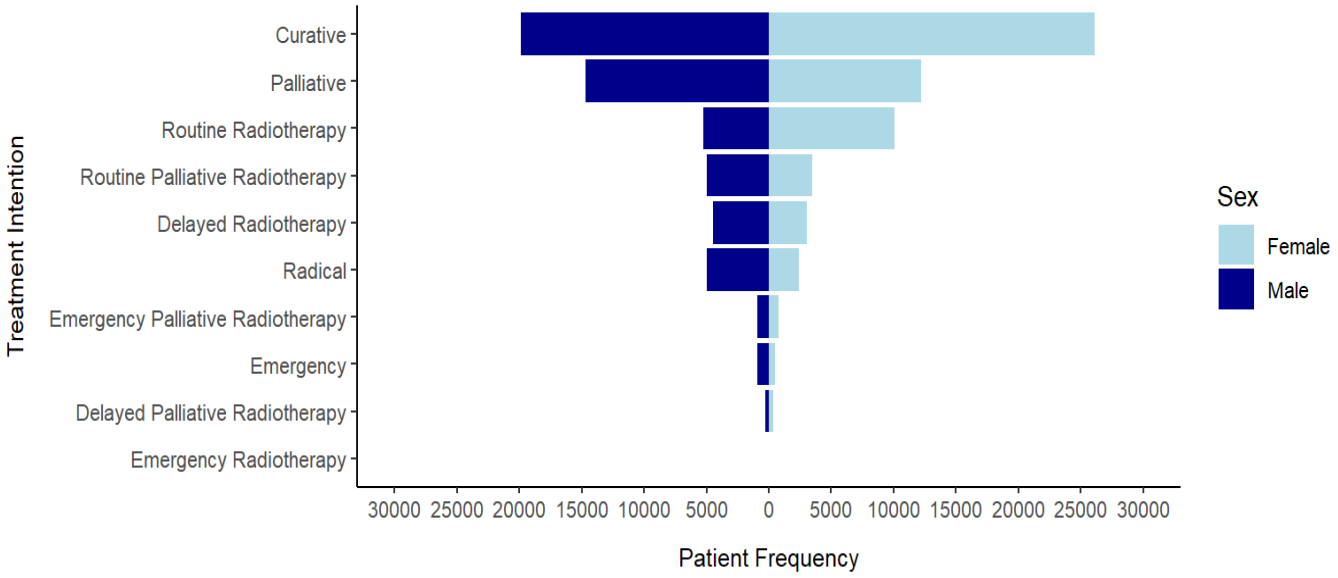


Figure 4. Frequency of radiotherapy treatment intentions by sex category.
Linkage between ***RTDS.patient*** and ***RTDS.fractions*** views was performed to count individual records by sex and treatment intention. Counts greater than 10 and null values for treatment intention were excluded from the analysis.

Example 4. Frequency of issued radiotherapy treatment.
By using the *opc_description* field, it is possible to identify the variations in radiotherapy treatment types provided to individuals in Wales. When combined with the *decision_to treat_date* field, this data allows for the evaluation of radiotherapy treatment trends over time (**Figure 5**).

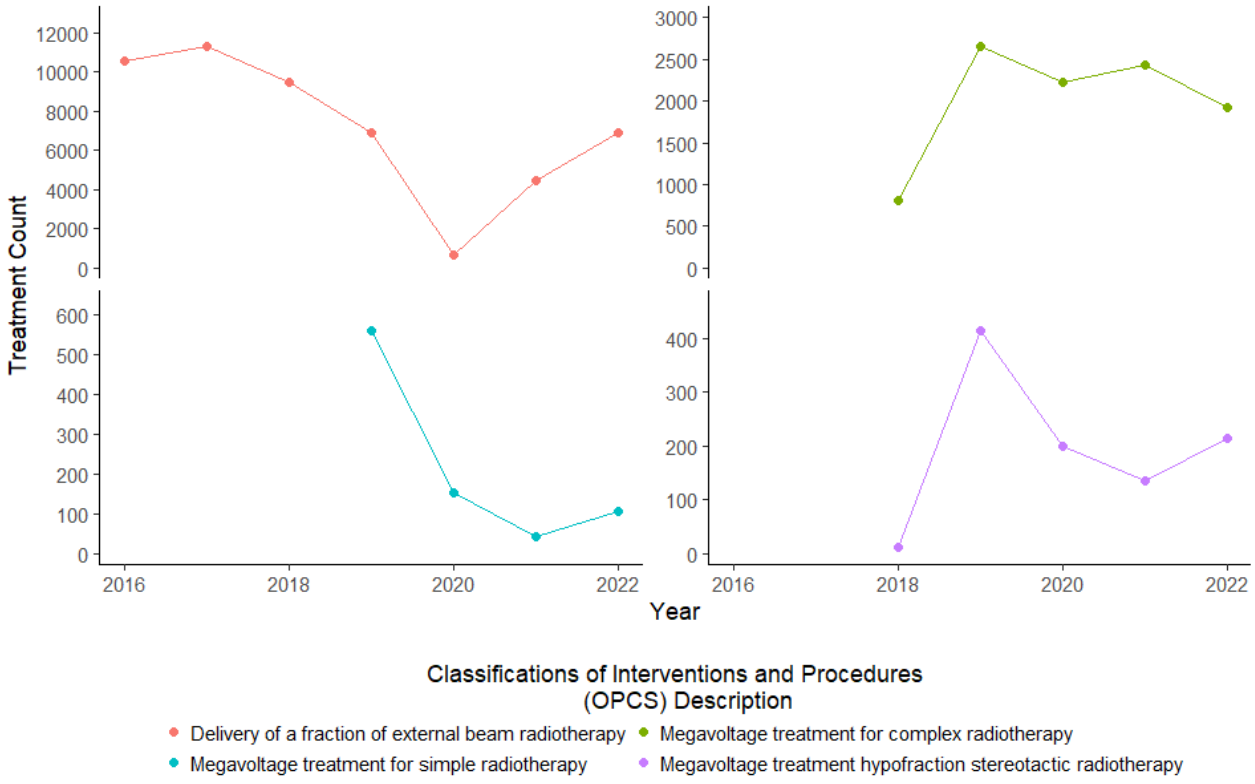


Figure 5. Longitudinal overview of prescribed radiotherapy treatment incidence.
Annual incidence of radiotherapy treatments. Individual counts were calculated per year for each treatment description that was provided within the RTDS. Entries with a decision-to-treat date of 1811 and counts less than 10 were excluded from the analysis. for specific counts, please see Appendix Table A4. This figure is presented in a faceted plot to better highlight the scales.

These insights may enable more targeted research to be conducted based on specific treatment types, further enhancing the understanding of radiotherapy practices and outcomes.

Example 5. Radiotherapy treatments issued per cancer type

Cancer-specific trends can also be identified using RTDS by counting individuals based on diagnosis ID (*ICD10_code*) at the time of their decision-to-treat date (**Figure 6**). These insights offer a foundation to understand treatment trends within different cancer groups.

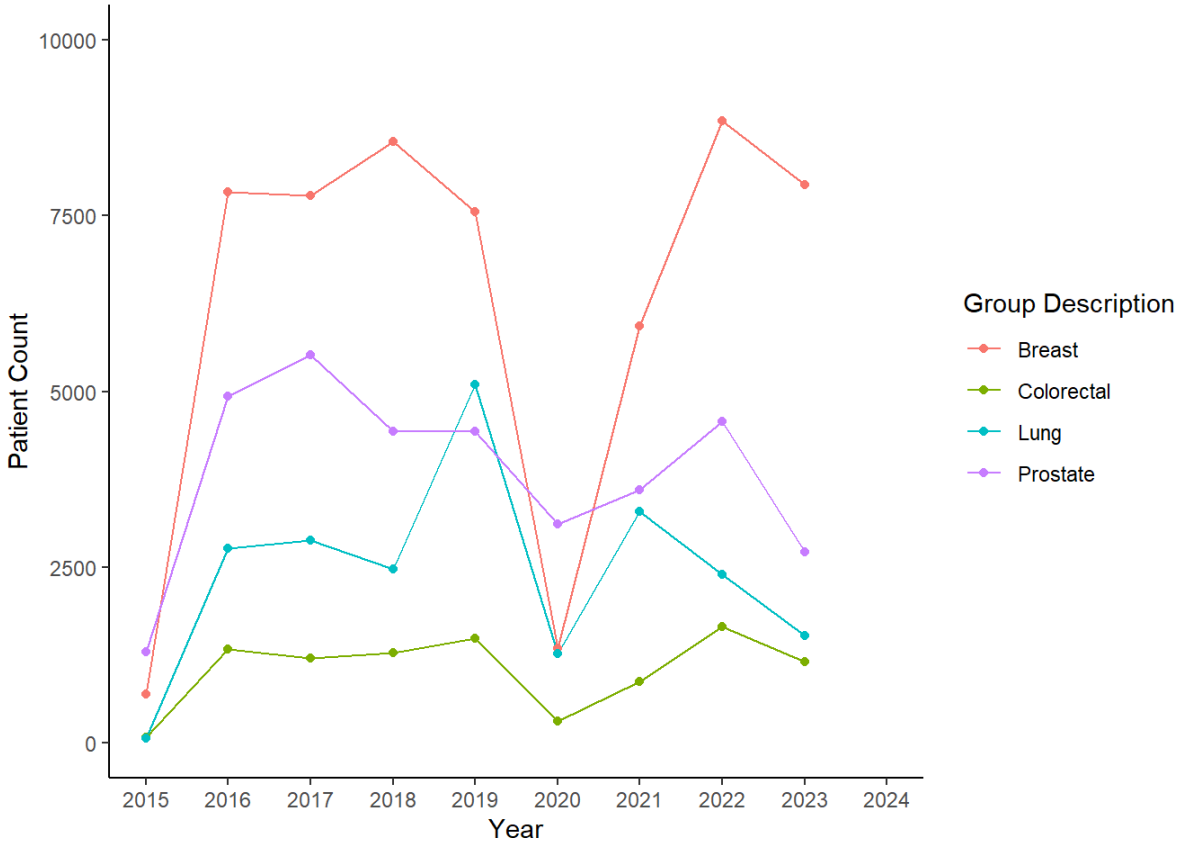


Figure 6. Incidence of radiotherapy treatment plans for the four most prevalent cancers in Wales. Distinct *alf_pe* were counted per International Classification of Disease (ICD-10) cancer diagnosis, per year. Values < 10 were excluded from the analysis. The analysis was further refined to only include data between 2014-2024, as this period exhibited the best coverage.

To enhance this analysis, parameters could be further refined to observe specific radiotherapy treatment trends for each cancer type, incorporating covariates such as sex or healthcare provider based on the research question.

Existing research or examples of previous research

At a national level, the RTDS has been used to study various population-level trends in radiotherapy, for example, in longitudinal studies that explore variations in radiotherapy service delivery, demand, and activity across different regions and periods (Hoskin et al., 2013; Round et al., 2013).
In addition, it has also been employed to examine how cancer treatment practices have evolved over time (Catto et al., 2023) and supported research into the outcomes of cancer patients undergoing radiotherapy and providing insights into the effectiveness of this treatment in managing the disease (Morris et al., 2016). Furthermore, RTDS has been used to investigate the impact of other pressures on healthcare services and how these strains affect the provision and timing of radiotherapy treatments (Lewis et al., 2021).
More recently, the data source has been applied in research that uses machine learning techniques to classify treatment categories based on patient data (Ahn et al., 2024). These studies demonstrate the significance of RTDS in enhancing the understanding and improvement of radiotherapy services and patient outcomes at a population level.

There is a notable lack of Wales-specific research studies using the RTDS, likely due to the relatively recent implementation of RTDS in Wales, with data submission to RTDS only beginning in 2016.
Regional population insights are crucial for addressing more specific and relevant research questions that can provide context and evidence to inform local policy decisions. Expanding the use of RTDS in Wales could lead to more targeted studies that reflect the unique healthcare needs and challenges within the region.

Data limitations encountered and suggested improvements

Although documentation indicating that Public Health Wales (PHW) worked closely with PHE to develop RTDS in Wales (Cancer Implementation Group, 2022), it is unclear whether the same documentation (National Disease Registration Service, 2021; Public Health England, 2021), applies to Welsh RTDS data. Furthermore, the data views available within the SAIL Databank do not align with the categories outlined in PHE documentation, such as Episode, Prescription, Exposure, and OPCDS (Sandhu et al., 2023). Additionally, the fields within SAIL's RTDS data source differ from those listed in the NHS data dictionary and cohort publications. For example, key fields such as Attendance Identifier, Appointment Date, Radiotherapy Priority, and Anatomical Treatment Site in the NHS data dictionary are not present in the SAIL data source (NHS, 2024; Sandhu et al., 2023).

The lack of comprehensive documentation detailing the data pathway that the RTDS in Wales follows, especially when compared to its English counterpart is a notable limitation of this data in Wales. This gap creates uncertainty about the processes involved in data collection, transformation, deposition and versioning within the RTDS in Wales. Additionally, while RTDS in Wales is a subset of the national standard, there are notable discrepancies between what is recorded in NHS data dictionaries and what is available in the SAIL Databank, as previously mentioned. These inconsistencies suggest that documentation for the English RTDS pathways cannot be fully applied in a Welsh context.

NHS Wales set out Quality Performance Indicators (QPIs (*Radiotherapy, 2023*)), that aim to measure the quality of the service and patient care. Here, the time from the decision to treat to the start of radiotherapy treatment is evaluated and Monitoring Death Within 30- and 90-Days of Radiotherapy (*Radiotherapy, 2023*). As a standalone data source, RTDS lacks information such as the cancer diagnosis date, which is essential for calculating metrics like the average time from diagnosis to treatment. To address this, linking RTDS to other data sources is necessary. Furthermore, the RTDS does not include any fields related to death, which limits its ability to support analyses of key performance indicators related to patient outcomes.

Additionally, including fields related to disease progression (for example, difference in pre and post treatment staging and residual disease) would also enhance the data source, allowing for more comprehensive analyses of treatment efficacy and disease outcomes.

Summary

Assessing the full value of RTDS as a standalone data source presents an opportunity for improvement due to several challenges previously mentioned. Enhancing documentation, addressing inconsistencies in available fields, and linking to other data sources for crucial information such as diagnosis dates will significantly boost its utility. RTDS can become a resource for in-depth research. Additionally, more specific research avenues related to radiotherapy can be explored by fully leveraging the potential of the RTDS data source.

References

Ahn, S., Oh, E. J., Saleem, M. I., & Tham, T. (2024). Machine Learning Methods in Classification of Prolonged Radiation Therapy in Oropharyngeal Cancer: National Cancer Database. *Otolaryngology--Head and Neck Surgery: Official Journal of American Academy of Otolaryngology-Head and Neck Surgery*. https://doi.org/10.1002/ohn.926

Cancer Implementation Group. (2022). *Radiotherapy Dataset (RTDS) V6* (DATA STANDARD CHANGE NOTICE ISRN 2021 / 031; pp. 1–4). Welsh Government. https://dhcw.nhs.wales/information-services/information-standards/data-standards/data-standards-files/data-standard-change-notices-docs/dscns-2022/radiotherapy-dataset-rtds-v6/

Catto, J. W. F., Mandrik, O., Quayle, L. A., Hussain, S. A., McGrath, J., Cresswell, J., Birtle, A. J., Jones, R. J., Mariappan, P., Makaroff, L. E., Knight, A., Mostafid, H., Chilcott, J., Sasieni, P., & Cumberbatch, M. (2023). Diagnosis, treatment and survival from bladder, upper urinary tract, and urethral cancers: Real-world findings from NHS England between 2013 and 2019. *BJU International*, *131*(6), 734–744. https://doi.org/10.1111/bju.15970

CRUK. (2023). *What is radiotherapy?* https://www.cancerresearchuk.org/about-cancer/treatment/radiotherapy/what-is-radiotherapy

Dawson, C. (2015). *WELSH INFORMATION STANDARDS BOARD DATA SET CHANGE NOTICE RTDS* (Government : ISRN 2013 / 001; DATA SET CHANGE NOTICE, pp. 1–25). https://dhcw.nhs.wales/information-services/information-standards/data-standards/data-standards-files/data-standard-change-notices-docs/dscns-2015/dscn-2015-03-radiotherapy-data-set-final-pdf/

HDR UK. (2024). *Radiotherapy Dataset (RTDS)*. https://web.www.healthdatagateway.org/dataset/f102ca50-d1e4-44ec-924e-be89446652af

Hoskin, P. J., Forbes, H., Ball, C., Riley, D., & Cooper, T. (2013). Variations in Radiotherapy Delivery in England—Evidence from the National Radiotherapy Dataset. *Clinical Oncology*, *25*(9), 531–537. https://doi.org/10.1016/j.clon.2013.04.006

Lewis, P. J., Morris, E. J. A., Chan, C. S. K., Darley, K., Sebag-Montefiore, D., & Evans, M. (2021). COVID RT – Assessing the Impact of COVID-19 on Radiotherapy in the UK. A National Cancer Research Institute Clinical and Translational Radiotherapy Research Working Group Initiative in Partnership with the Royal College of Radiologists, the Society of Radiographers and the Institute of Physics and Engineering in Medicine. *Clinical Oncology*, *33*(1), e69–e72. https://doi.org/10.1016/j.clon.2020.08.008

Morris, E. J. A., Finan, P. J., Spencer, K., Geh, I., Crellin, A., Quirke, P., Thomas, J. D., Lawton, S., Adams, R., & Sebag-Montefiore, D. (2016). Wide Variation in the Use of Radiotherapy in the Management of Surgically Treated Rectal Cancer Across the English National Health Service. *Clinical Oncology*, *28*(8), 522–531. https://doi.org/10.1016/j.clon.2016.02.002

National Disease Registration Service. (2021). *Radiotherapy Data Set (RTDS) User guide V 6.1.1.* (NHS Digital DAPB0111 Amd 84/2020; pp. 1–110). PHE. https://digital.nhs.uk/ndrs/data/data-sets/rtds#key-updates

NHS. (2024). *Radiotherapy Data Set* [NHS]. NHS Data Models and Dictionary. https://www.datadictionary.nhs.uk/data_sets/clinical_data_sets/radiotherapy_data_set.html?hl=rtds

Public Health England. (2021). *Radiotherapy Data Set (RTDS) Technical Guide* (Public Health England https://digital.nhs.uk/ndrs/data/data-sets/rtds#key-updates). https://digital.nhs.uk/ndrs/data/data-sets/rtds#key-updates

Radiotherapy. (2023, July 13). NHS Wales Executive. https://executive.nhs.wales/functions/networks-and-planning/cancer/clinical-hub/radiotherapy/

Round, C. E., Williams, M. V., Mee, T., Kirkby, N. F., Cooper, T., Hoskin, P., & Jena, R. (2013). Radiotherapy Demand and Activity in England 2006–2020. *Clinical Oncology*, *25*(9), 522–530. https://doi.org/10.1016/j.clon.2013.05.005

Sandhu, S., Sharpe, M., Findlay, Ú., Roe, C., Broggio, J., Spencer, K., & Thackray, K. (2023). Cohort profile: Radiotherapy dataset (RTDS) in England. *BMJ Open*, *13*(6), e070699. https://doi.org/10.1136/bmjopen-2022-070699

Siegel, R. L., Miller, K. D., Wagle, N. S., & Jemal, A. (2023). Cancer statistics, 2023. *CA: A Cancer Journal for Clinicians*, *73*(1), 17–48. https://doi.org/10.3322/caac.21763

Wells, r. (2015, February 4). *Welsh Information Standards Board WHC 2015 / 006*. https://dhcw.nhs.wales/information-services/information-standards/data-standards/data-standards-files/data-standard-change-notices-docs/dscns-2015/dscn-2015-03-radiotherapy-data-set-final-pdf/

WHO. (2020). *The top 10 causes of death*. World Health Organisation. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

Glossary

Abbreviation	Name
ADR	Administrative Data Research UK
ALF	Anonymised Linkage Field
CReSt	The Cancer Research Strategy for Wales
DHCW	Digital Health and Care Wales
ESRC	Economic and Social Research Council
ICD-10	International Classification of Disease
HDR UK	Health Data Research United Kingdom
IGRP	Information Governance Review Panel
ICNARC	Intensive Care National Audit & Research Centre
NATCANSAT	National Clinical Analysis and Specialised Applications Team
NHS	National Health Service
NHSE	National Health Service England
PAS	Patient Administration System
PHE	Public Health England
PHW	Public Health Wales
QPI	Quality Performance Indicators
RTDS	Radiotherapy Dataset
SAIL	Secure Anonymised Information Linkage (Databank)
SCP	Suspected Cancer Pathway

TRE	Trusted Research Environment
WCSU	WCISU - Welsh Cancer Intelligence and Surveillance Unit
WCRC	Wales Cancer Research Centre
WPAS	Welsh Patient Administration System

Disclaimer

This work was produced using administrative data accessed through the SAIL Databank. The use of the data in this work does not imply the endorsement of SAIL or data owners in relation to the interpretation or analysis.
This work uses research data source which may not exactly reproduce National Statistics aggregates. National Statistics follow consistent statistical conventions over time and cannot be compared to Data First linked data source.

Acknowledgements

This work is supported by ADR Wales, part of ADR UK (Administrative Data Research UK), a partnership transforming access to public sector data in the UK to inform policy decisions that enhance people's lives. ADR UK is funded by the Economic and Social Research Council (ESRC), part of UK Research and Innovation ES/W012227/1]. The research was facilitated by the SAIL Databank, with funding from Health and Care Research Wales via the Wales Cancer Research Centre's CReSt catalytic funding (researcher salary and data access costs) and funding from the Roche Bioresource Data Accelerator programme (data access costs).

We gratefully acknowledge the contributions and support of Professor Sunil Dolwani, Professor Mererid Evans, Dr Mark Davies, Liz Merrifield, Dr Peter Giles and Jenni Macdougall.

Contact

Name: Laura Baker
Email: l.m.baker@swansea.ac.uk

Appendix

Table A1. Counts associated radiotherapy treatment types given per year (in association with Figure 5).

Treatment Year	Individual Count	OCPs Description
2015	1151	Delivery of a fraction of external beam radiotherapy
2016	10595	Delivery of a fraction of external beam radiotherapy
2017	11335	Delivery of a fraction of external beam radiotherapy
2018	804	Megavoltage treatment for complex radiotherapy
2018	12	Megavoltage treatment hypo fract stereotactic radiotherapy
2018	9512	Delivery of a fraction of external beam radiotherapy
2019	413	Megavoltage treatment hypo fract stereotactic radiotherapy
2019	10	Delivery of a fraction of total body irradiation
2019	40	Megavoltage treatment for adaptive radiotherapy
2019	2659	Megavoltage treatment for complex radiotherapy
2019	6903	Delivery of a fraction of external beam radiotherapy nec
2019	559	Megavoltage treatment for simple radiotherapy
2020	659	Delivery of a fraction of external beam radiotherapy nec
2020	200	Megavoltage treatment hypofract stereotactic radiotherapy
2020	12	Megavoltage treatment for adaptive radiotherapy
2020	154	Megavoltage treatment for simple radiotherapy
2020	10	Delivery of a fraction of total body irradiation
2020	2232	Megavoltage treatment for complex radiotherapy
2021	4468	Delivery of a fraction of external beam radiotherapy nec
2021	136	Megavoltage treatment hypofract stereotactic radiotherapy
2021	2431	Megavoltage treatment for complex radiotherapy
2021	43	Megavoltage treatment for simple radiotherapy
2022	1929	Megavoltage treatment for complex radiotherapy
2022	106	Megavoltage treatment for simple radiotherapy
2022	6882	Delivery of a fraction of external beam radiotherapy nec
2022	214	Megavoltage treatment hypofract stereotactic radiotherapy
2023	239	Megavoltage treatment for complex radiotherapy
2023	7110	Delivery of a fraction of external beam radiotherapy nec
2023	41	Megavoltage treatment for simple radiotherapy