


Long-term observation of patients with cancer – an entity-independent registry for healthcare and translational research at the University Medicine Dresden (Cancer-Reg-VT)

Langzeitbeobachtung von Patienten mit Krebserkrankungen – ein entitätsunabhängiges Register für Versorgungsforschung und translationale Forschung an der Universitätsmedizin Dresden (Cancer-Reg-VT)



Authors

Thomas Datzmann^{1, 2} , Olaf Schoffer¹, Jochen Schmitt^{1, 2}, Heiko Böhme², Johannes Fritzmann^{3, 2}, Marius Distler^{3, 2}, Ulrike Ubbelohde², Esther Giehl-Brown^{3, 2}, Tjark Henke^{3, 2}, Mechthild Krause^{4, 5}, Hanno Glimm^{6, 5}, Martin Bornhäuser^{7, 2}, Jürgen Weitz^{3, 2}

Affiliations

- 1 Zentrum für Evidenzbasierte Gesundheitsversorgung (ZEGV), Medizinische Fakultät und Universitätsklinikum Carl Gustav Carus, TU Dresden, Dresden, Germany
- 2 Nationales Centrum für Tumorerkrankungen Dresden (NCT/UCC), Medizinische Fakultät und Universitätsklinikum Carl Gustav Carus, Technische Universität Dresden; Helmholtz-Zentrum Dresden-Rossendorf (HZDR), DKFZ, Dresden, Germany
- 3 Klinik und Poliklinik für Viszeral-, Thorax- und Gefäßchirurgie (VTG), Medizinische Fakultät und Universitätsklinikum Carl Gustav Carus, TU Dresden, Dresden, Germany
- 4 Abteilung für Radioonkologie, Nationales Centrum für Tumorerkrankungen Dresden (NCT/UCC), Medizinische Fakultät und Universitätsklinikum Carl Gustav Carus, Technische Universität Dresden; Helmholtz-Zentrum Dresden-Rossendorf (HZDR), DKFZ, Dresden, Germany
- 5 Deutsches Konsortium für Translationale Krebsforschung, DKTK, Dresden, Germany
- 6 Abteilung für Translationale Medizinische Onkologie, Nationales Centrum für Tumorerkrankungen Dresden (NCT/UCC), Medizinische Fakultät und Universitätsklinikum Carl Gustav Carus, Technische Universität Dresden; Helmholtz-Zentrum Dresden-Rossendorf (HZDR), DKFZ, Dresden, Germany
- 7 Medizinische Klinik I, Medizinische Fakultät und Universitätsklinikum Carl Gustav Carus, TU Dresden, Dresden, Germany

Key words

Cancer diseases, Data linkage, SHI data, HRQOL

Schlüsselwörter

Krebserkrankungen, Datenverknüpfung, GKV-Daten, HRQOL

Bibliography

Gesundheitswesen 2023; 85 (Suppl. 3): S226–S234

DOI 10.1055/a-2129-7651

ISSN 0949-7013

© 2023. The Author(s).

This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial-License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Georg Thieme Verlag KG, Rüdigerstraße 14,
70469 Stuttgart, Germany

Correspondence

Dr. Thomas Datzmann
TU Dresden, Zentrum für Evidenzbasierte Gesundheitsversorgung (ZEGV)
Fetscherstraße 74
01307 Dresden
Germany
thomas.datzmann@ukdd.de

Deutsche Version und Zusätzliches Material unter:

<https://doi.org/10.1055/a-2129-7651>

ABSTRACT

Introduction Translational research is important, especially in medicine where decisions affect people's lives. Clinical reg-

istries and the studies embedded in them allow the depiction of actual care practice under routine conditions. Translating the findings of health services research back into clinical research through prospective cohort studies has the potential to drive medical innovations faster, more effectively and, above all, in a more targeted manner. These must therefore be a central component of cutting-edge oncological research.

Objective The aim of the registry is the establishment of clinical cohorts and the provision of a comprehensive, high-quality data set for oncological diseases.

Methods/Design The registry will prospectively record all patients treated for cancer at Dresden University Hospital (UKD). In addition to the data from the hospital information systems (ORBIS, TDS, GEPADO, etc.), monitoring of health-related quality of life (HRQOL) is to be carried out at regular intervals at the beginning and during the course of treatment. In addition, individual linkage with data from clinical cancer registries and health insurance companies (including AOK PLUS) is planned for a period of five years before and after inclusion. All these data will be merged in a registry database. The selection of variables and measurement time points is closely based on the guidelines for colorectal carcinoma of the international initiative ICHOM (International Consortium for Health Outcomes Measurement). The study management software (STeVe) separates personal identification characteristics (IDAT) and medical data (MDAT) at an early stage. The independent trust centre of the TU Dresden (Treuhandstelle) ensures that no personal data enter the registry database. It is thereby also ensured that the data owners involved (UKD, biobank, health insurance company, cancer registry, patient) only receive the personal data they need for allocation. The MOSAIC software tools recommended by the TMF (Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V.) are used to manage the pseudonyms.

Discussion/Conclusion With the registry, previously missing evidence on the effectiveness, safety and costs of diagnostic and therapeutic measures can be made, taking into account long-term and patient-reported outcomes of routine care. The data potentially allow for the identification of barriers to and facilitators of innovative promising cancer diagnostics and therapies. They also enable generation of scientifically relevant hypotheses in the field of translational and outcomes research.

ZUSAMMENFASSUNG

Einleitung Translationale Forschung ist wichtig, insbesondere in der Medizin, wo Entscheidungen das Leben von Menschen beeinflussen. Klinische Register und die darin eingebetteten Studien ermöglichen die Abbildung der tatsächlichen Versorgungspraxis unter Routinebedingungen. Die Rückführung

der Erkenntnisse aus der Versorgungsforschung in die klinische Forschung durch prospektive Kohortenstudien hat das Potenzial, medizinische Innovationen schneller, effektiver und vor allem zielgerichteter voranzutreiben. Dies muss daher ein zentraler Bestandteil der onkologischen Spitzenforschung sein.

Zielsetzung Ziel des Registers ist der Aufbau von klinischen Kohorten und die Bereitstellung eines umfassenden, qualitativ hochwertigen Datensatzes für onkologische Erkrankungen.

Methoden/Design Das Register wird prospektiv alle Patient:innen erfassen, die am Universitätsklinikum Dresden (UKD) wegen Krebs behandelt werden. Zusätzlich zu den Daten aus den Krankenhausinformationssystemen (ORBIS, TDS, GEPADO, etc.) soll in regelmäßigen Abständen zu Beginn und im Verlauf der Behandlung ein Monitoring der gesundheitsbezogenen Lebensqualität (HRQOL) durchgeführt werden. Darüber hinaus ist eine individuelle Verknüpfung mit Daten aus klinischen Krebsregistern und Krankenkassen (u. a. AOK PLUS) für einen Zeitraum von fünf Jahren vor und nach dem Einschluss geplant. Alle diese Daten werden in einer Registerdatenbank zusammengeführt. Die Auswahl der Variablen und Messzeitpunkte orientiert sich eng an den Leitlinien für das kolorektale Karzinom der internationalen Initiative ICHOM (International Consortium for Health Outcomes Measurement). Die Studienmanagementsoftware (STeVe) trennt frühzeitig persönliche Identifikationsmerkmale (IDAT) und medizinische Daten (MDAT). Die unabhängige Treuhandstelle der TU Dresden stellt sicher, dass keine personenbezogenen Daten in die Registerdatenbank gelangen. Damit ist auch sichergestellt, dass die beteiligten Dateneigentümer (UKD, Biobank, Krankenkasse, Krebsregister, Patient) nur die personenbezogenen Daten erhalten, die sie für die Zuordnung benötigen. Zur Verwaltung der Pseudonyme werden die von der TMF (Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V.) empfohlenen MOSAIC-Softwaretools eingesetzt.

Diskussion/Schlussfolgerung Mit dem Register können bisher fehlende Erkenntnisse über die Wirksamkeit, Sicherheit und Kosten von diagnostischen und therapeutischen Maßnahmen unter Berücksichtigung von Langzeit- und patientenberichteten Outcomes aus der Routineversorgung gewonnen werden. Die Daten erlauben potenziell die Identifizierung von Barrieren und Förderfaktoren für innovative, vielversprechende Krebsdiagnostik und -therapien. Sie bieten auch die Möglichkeit, wissenschaftlich relevante Hypothesen im Bereich der Translations- und Outcome-Forschung zu generieren.

LIST OF ABBREVIATIONS

AOK PLUS	General Local Health Insurance Company in Saxony/Thuringia
EBM	Uniform assessment framework for the billing of outpatient medical services
GEPADO	Patient Information System at the UKD (contains genetic information)
GKV	Statutory health insurance
ICD10-GM	International Statistical Classification of Diseases and Related Health Problems 10th Revision German Modification
ICHOM	International Consortium for Health Outcomes Measurement
IDAT	Person-identifying data
KDAT	Data from health insurance companies and cancer registries
KV	Health insurance number
MDAT	Medical data
NCT	National Centre for Tumour Diseases Dresden
OPS	Operation and Procedure Code for coding operations, procedures and general medical measures
ORBIS	Hospital Information System at the UKD
TDS	Tumour Documentation System at the UKD
UCC	University Cancer Center Dresden
UKD	University Hospital Dresden

Introduction

Frequently used translational pathways in the field of oncology research make use of Randomised Controlled Trials (RCTs), especially for efficacy analyses of therapeutic interventions. However, due to their selective inclusion and exclusion criteria, the specific setting and the restriction to a concrete/specific question in a strictly defined study population, RCTs may inadequately reflect clinical reality [1]. Clinical registries and the studies embedded in them, on the other hand, allow the depiction of actual health care practice under routine conditions. They can reveal, for example, weak points in therapy and thus have the potential to optimise therapy processes. The general credo of science at present is that RCTs alone are not sufficient to advance clinical research and decisively improve patient care [1]. As a method of translational research, the analysis of prospective registry data combines controlled, prospective observations with routine patient care. Furthermore, clinical cohorts and registries serve as an excellent basis for embedding clinical trials. A back-translation of findings from health services research through prospective cohort studies into clinical research has the potential to drive medical innovation faster, more effectively and, above all, in a more targeted manner [1]. They must therefore be a central component of cutting-edge oncological research [2].

In addition, in recent years there have been increased international efforts to standardise routine care in order to create the basis for more effective care, data harmonisation and research into outcomes under routine conditions. An outstanding international ini-

tiative, especially for the oncological field, is ICHOM (International Consortium for Health Outcomes Measurement) [3]. In so-called standard sets, it was specified in detail which data should be collected from which patient at which time using which method (measurement) and which outcomes should be used (minimum set).

Ensuring high-quality, patient-centred care is a major challenge for the health care system. A precise understanding of the risk factors, the benefits and effectiveness of therapeutic measures under routine conditions, as well as the interplay between outpatient and inpatient care are prerequisites for monitoring health care and guiding future health development. The treatment costs as well as the costs for psychological and other secondary health demands caused by cancer must also be considered.

Objectives of the register

The primary goals of the research project can practically be assigned to two areas, health services research and clinical prognosis research. In the area of health services research, questions in the field of translational research are to be answered in addition to effectiveness and outcomes analyses. Specifically, the following questions of outcome research, translational research and clinical research will be addressed and a suitable data basis will be created for answering these questions::

- How effective and safe are oncological treatments and care concepts, taking into account long-term and patient-reported outcomes?
- Is the standard medical care of patients with cancer adequate and of high quality?
- What are the barriers or facilitators to the implementation of innovative, promising cancer therapies?
- Which individual factors are prognostically relevant or influence the response to therapy and the course of the disease?

Based on this, it will be possible to generate science-relevant hypotheses in the field of translational and outcomes research. Recent developments in personalised medicine (targeted therapy) and novel treatment approaches such as immunotherapy have raised hopes of significantly improve cancer survival in the future [4, 5] and at least transforming cancer from a fatal to a chronic disease. The registry is ideally suited to address these questions through the establishment of clinical cohorts and the provision of comprehensive, high-quality data sets for oncological diseases.

Methods/Design

Data basis and management

The data basis of the registry is formed by clinical data (routine data of the hospital information system as well as data from the clinical cancer registries), quality-of-life data (questionnaires), biodata (exclusively results of biomarker analyses from standard care or with separate consent of the patient), histology (exclusively data from standard care or with separate consent of the patient), data of imaging procedures and statutory health insurance (SHI) data (diagnoses, procedures, prescriptions). These data must be collected

individually for each patient and transferred to the register at the NCT/UCC Dresden. However, within the framework of the registry, we mainly bring together data from clinical information systems, cancer registries, and SHI data that has already been collected and do not re-collect data that has already been documented. Therefore, there is no additional burden on the treating physicians due to duplicate documentation within the registry. Data collection is closely based on the ICHOM standard set for colorectal adenocarcinoma [3], but is individually adapted for each cancer entity. Currently, internationally harmonised standard sets are available in oncology for lung, prostate, breast and colorectal cancer. The timing of the follow-up surveys in particular can vary depending on the entity and is based on the usual follow-up of these patients at the UKD in order to ensure the most economical integration of the registry into routine procedures (see Figure S1 in the appendix). In addition, biomarkers, possible complications and other aspects of the cancer entity must be adapted. A detailed breakdown of the variables to be used, the timing and the origin of the data can be found in the appendix as an example for pancreatic cancer (Table S2 & S3).

The aim is to establish a prospective patient cohort. Prospective here, however, refers to the time of recruitment. Data should also be collected retrospectively over a period of 5 years before the inclusion of the patient. In the case of purely post-treatment patients, data from the last few years must be retrospectively searched. The follow-up should also be at least 5 years. A patient-specific end-of-follow-up is not planned. However, the follow-up times vary between the respective disease status. For example, in the case of non-progressive pancreatic cancer, a quarterly follow-up is planned in the first year after treatment and then every six months. In the event of disease progression and renewed intervention (surgery, chemotherapy, etc.), follow-up visits should again take place at 3-month intervals in the first year and then every six months (see Figure S1 in the appendix). Interim visits during a stay at the UKD, e. g. in case of complications, are documented and the data on these visits are also transferred to the registry database.

Within the scope of the study, the consent of the study participants to the utilisation of their SHI routine data as well as the data of the clinical cancer registries will be documented via the registry-specific consent form (see appendix). In the consent form, patients can agree to have their personal data transmitted to the registry by the respective health insurance company and cancer registry for a period of 5 years before and after the date of registry inclusion. This data allows the individual patient histories to be described and presented without gaps. This is an important supplement to the parameters obtained in the course of treatment at Dresden University Hospital. The administration of the declarations of consent, including patient rights (objection, information, etc.), is carried out in the independent trust office of the TU Dresden (THS). Access to the data, which is specifically extracted from the register for scientific research, is only permitted within the framework of Dresden University Medicine for the specific, above-mentioned questions of healthcare and clinical research. For questions that go beyond the original purpose described here, an application in accordance with §75 Section 2 SGB X must be submitted to the Saxon State Ministry for Social Affairs and Social Cohesion (SMS) or to the re-

spective competent supervisory authority and needs to be approved.

Registry design

The registry is a prospective observational clinical cohort study on patients with cancer.

Participants, interventions and outcomes

Consecutively, all patients with at least one inpatient hospital stay due to a cancer diagnosis (ICD10-GM C00–97) who are treated at the UKD or are at tumour aftercare are initially included. Patients must be at least 18 years old at the time of inclusion. A control group is not planned. There will be no study-related intervention, only the observation of patients in routine care. Additional case definitions are needed for additional determinants. These include concomitant diseases or metastases, which are defined according to individual ICD-10-GM codes, specialists involved in the treatment, who are defined and selected according to the physician group code, or diagnostic and therapeutic services, which are defined according to OPS codes, EBM numbers and central pharmaceutical numbers (PZN), respectively. The recommendations of the Good Practice Secondary Data Analysis (GPS) [6] as well as the Good Practice Data Linkage (GPD) [7] are followed for the various case definitions in order to carry out adequate internal diagnosis and further code validations. Potential endpoints of this study are 30-day and 1 to 5-year mortality, relative survival, survival after Brenner (period approach), progression- and relapse-free survival, overall survival, overall clinical outcome, complications of treatment (short-term), quality-of-life outcomes (general well-being, physical functioning, emotional functioning, social functioning, mental functioning), direct costs, end-of-life hospitalisation rate, and proportion of study participants dying in hospice. Differentiation between primary and secondary outcomes is not necessary. In addition, the analysis of potential influencing factors for remission and relapse, as well as treatment-specific complications and description of patient pathways and treatment sequences will be possible. A summary of the outcomes, measurement instruments and measurement times can be found in the appendix (exemplary for pancreatic carcinoma).

The registry is already in the development phase. Internal hospital structures, processes and data sources, including IT interfaces, are being established. The integration of the project's internal patient management software (STeVe) and the THS, including the automatic management of consent forms, has also been established. Currently, a functioning registry for 2 cancer entities (colorectal adenocarcinoma, pancreatic carcinoma) is operated in the follow-up, including the monitoring of quality of life (first patient in: second half of 2020). However, the registry still needs to be enriched with data from health insurers and cancer registries and thus needs access to a.o. outpatient treatments and diagnoses of the participants. However, extensive discussions have already taken place with health insurance companies and clinical cancer registries, as well as preliminary work on this, so that the aim is to integrate these data sources quickly, at least for some of the participants. As of December 2022, 141 people with adenocarcinoma of the colon (C18), 181 people with adenocarcinoma of the rectum (C20) and 148 people with pancreatic carcinoma (C25) are in the

► **Table 1** Patient characteristics at baseline: Socio-demographic data, data on disease severity and data on quality of life (QOL) are given. Based on our inclusion criteria, the registry, which started in 2020, includes patients with newly discovered cancers as well as patients undergoing follow-up care or with current treatment for a recurrence/metastasis. Patients whose primary diagnosis and/or treatment took place outside the University Hospital Dresden can therefore also be included in the registry if they are in contact with the UKD. There are occasional cases in which the primary diagnosis and the inclusion of the patient in the register are many years apart.

Entity	Adenocarcinoma colon C18		Adenocarcinoma rectum C20		Pancreatic carcinomas C25	
	n	%	n	%	n	%
	141	100	181	100	148	100
Age at diagnosis in years (mean ± SD)	59.8 ± 14.0		60.4 ± 11.5		61.4 ± 12.4	
Gender (female)	54	38.3	50	27.6	69	46.6
Family status						
married	84	59.6	126	69.6	94	63.5
widowed	13	9.2	8	4.4	14	9.5
divorced	15	10.6	17	9.4	11	7.4
single	19	13.5	14	7.7	14	9.5
separated	2	1.4	4	2.2	2	1.4
unknown	8	5.7	12	6.6	13	8.8
UICC stage at initial diagnosis						
UICC I	8	5.7	20	11.0	21	14.2
UICC II	26	18.4	16	8.8	41	27.7
UICC III	32	22.7	61	33.7	18	12.2
UICC IV	41	29.1	46	25.4	37	25.0
Not available	34	24.1	38	21.0	31	20.9
Grading at initial diagnosis						
L (1,2)	69	48.9	102	56.4	70	47.3
H (> 2)	45	31.9	40	22.1	30	20.3
Indefinable	9	6.4	1	0.6	0	0.0
Not available	18	12.8	19	10.5	16	10.8
L/V category at initial diagnosis						
L0	67	47.5	114	63.0	74	50.0
L1	54	38.3	39	21.5	36	24.3
Missing	20	14.2	28	15.5	38	25.7
V0	83	58.9	115	63.5	71	47.9
V1	37	26.2	37	20.5	35	23.6
V2	1	0.7	0	0.0	2	1.4
Missing	20	14.2	29	16.1	40	27.0
Mutation status						
KRAS mutated/wildtype	19/26	13.5/18.4	21/39	11.6/21.5	---	---
NRAS mutated/wildtype	1/28	0.7/19.9	3/39	1.7/21.5	---	---
BRAF mutated/wildtype	7/37	5.0/26.2	3/50	1.7/27.6	---	---
MSI stable/unstable	28/6	19.9/4.3	62/1	34.3/0.6	---	---
Comorbidities						
Other tumor C00-C97 (without C44, C18-21 & C77-79)	37	26.2	39	21.5	37	25.0
Chronic kidney disease (CKD) - N18	10	7.1	11	6.1	9	6.1
Liver disease K71-76	35	24.8	37	20.4	25	16.9
Coronary heart disease (CHD) - I20-25	12	8.5	12	6.6	11	7.4
Heart failure - I50-52	14	9.9	24	13.3	12	8.1
Hypertensive diseases - I10-15	76	53.9	107	59.1	96	64.9
Diabetes (Typ2) - E11-14	27	19.1	26	14.4	76	51.4
Metastasis						
Total	116	82.3	154	85.1	118	79.7
Synchrone metastasis within 100 days	46	32.6	48	26.5	42	28.4

► **Table 2** Overview of the treatments of the registry patients: Resection of the primary cancer is the initial therapy for both colorectal adenocarcinoma and pancreatic carcinoma. In the case of rectal carcinoma, approx. 66 % of the therapies are carried out before resection (neoadjuvant). In colon and pancreatic carcinoma, these neoadjuvant therapies play a subordinate role. Chemotherapies carried out after primary resection (adjuvant), on the other hand, are used particularly frequently for pancreatic carcinoma (approx. 65 %). In order to show the actual burden of all patients in the registry and not only those who have undergone resection, the therapies for recurrences and metastases have been added to the therapies for the primary cancer in the lower part of table. Radio- and radiochemotherapy are mainly used in the treatment of adenocarcinoma of the rectum (C20). Radiochemotherapy has a subordinate role in the treatment of adenocarcinoma of the colon and pancreatic carcinoma.

Entity	Adenocarcinoma colon C18		Adenocarcinoma rectum C20		Pancreatic carcinomas C25	
	n = 141	%	n = 181	%	n = 148	%
Resection of the primary tumour	119	84.4	146	80.7	102	68.9
Of which with R0	109	91.6	139	95.2	95	93.1
Complete remission after resection (ypT0)	2	1.7	10	6.8	1	1.0
No complete remission after resection (ypT>0)	10	8.4	87	59.6	18	17.6
Repeated surgery within 30 days	13	10.9	34	23.3	17	16.7
Resection + CTx for primus	42	40.8	77	57.9	66	64.7
Therapies (primus, recurrence, metastasis)						
OPx (any surgical intervention)	134	95.0	170	93.9	147	99.3
CTx	108	76.6	141	77.9	116	78.4
RTx	27	19.1	73	40.3	29	19.6
RCTx	2	1.4	56	30.9	7	4.7
CTx: chemotherapy; OPx: surgery; RTx: radiotherapy; RCTx radiochemotherapy						

► **Table 3** Quality of life data at the time of study inclusion based on patient-reported EORTC-C30 questionnaires (cross-sectional): Baseline information on the patients' general quality of life based on the generic questionnaire instrument C30 of the "European Organization for Research and Treatment of Cancer" (EORTC [8]). Mean values and their standard deviation for six separate QOL domains and the general state of health are shown. In comparison, the values for pancreatic carcinoma are somewhat lower than those for colon and rectal carcinoma.

Entity	Adenocarcinoma colon C18		Adenocarcinoma rectum C20		Pancreatic carcinomas C25	
Domain	n = 119		n = 150		n = 131	
	Mean *	± SD	Mean *	± SD	Mean *	± SD
Physical Functioning	71.77	22.62	72.7	22.47	64.99	25.36
Pain	67.68	29.31	69.03	30.55	62.09	30.17
Social Functioning	59.05	35.06	57.27	32.09	52.08	35.92
Cognitive Functioning	80.08	21.93	81.98	22.55	76.72	25.88
Role Functioning	65.41	35.2	61.19	31.26	52.69	32.8
Emotional Functioning	65.59	23.2	64.99	25.48	59.56	23.25
General health status	62	21.73	57.32	21.95	56.03	22.64
* QOL scores range from 0 to 100, with 100 representing the highest quality of life or the lowest symptom burden (pain).						

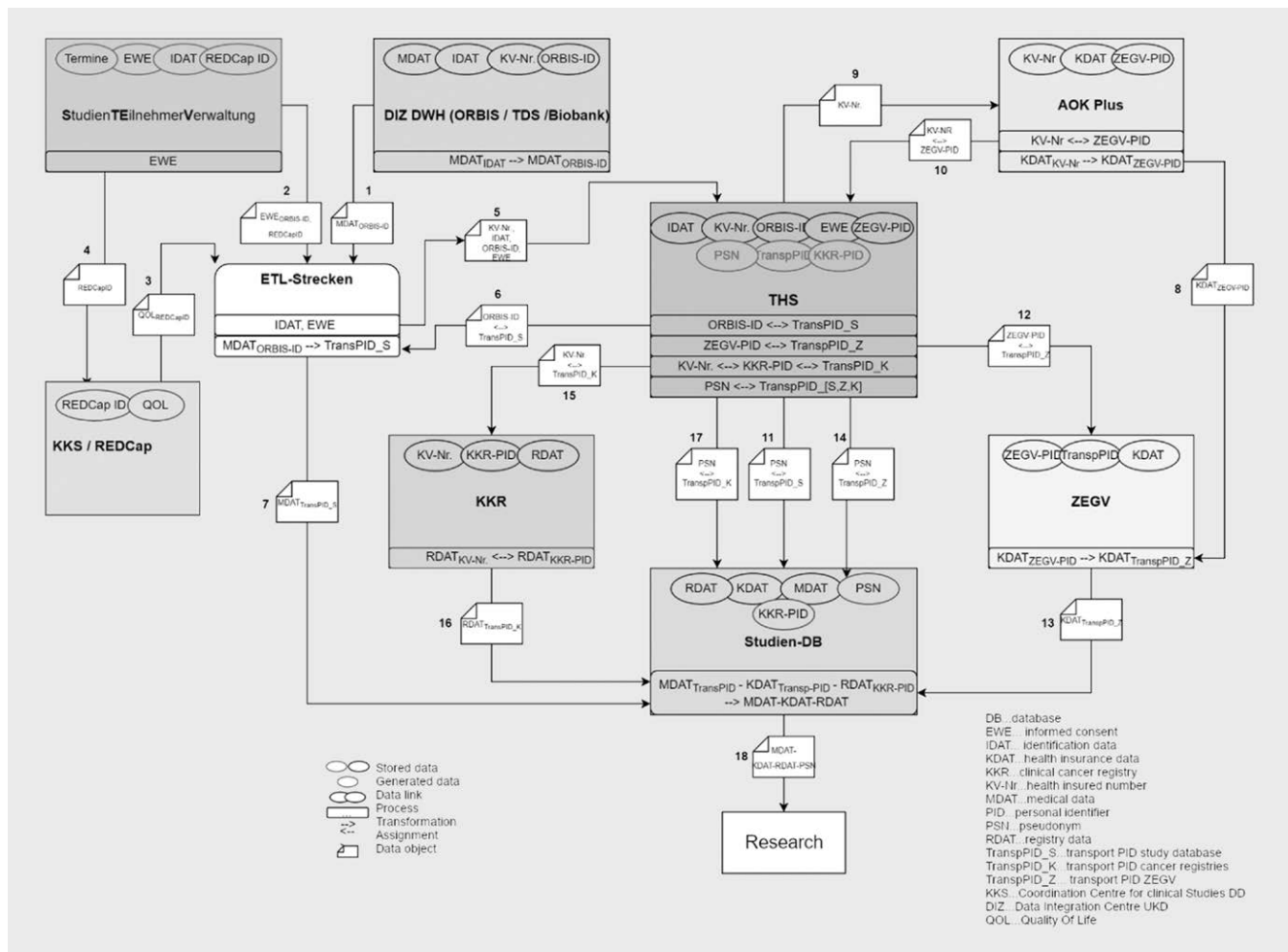
registry. A brief description of the present baseline can be found in ► **Tables 1–3**. The establishment phase of the registry also serves to estimate the effort and the material and personnel requirements for the planned roll-out of the registry to other cancer entities (melanoma, head and neck tumours, breast cancer, etc.).

Data protection, ethical and legal aspects

To ensure data protection, the data flow of the registry is embedded in the data flows of the NCT/UCC data warehouse. The study management software (STeVe) already separates personal identifying characteristics (IDAT) and medical data (MDAT) or secondary data of the health insurance companies and cancer registries

(KDAT) at an early stage. The THS (Treuhandstelle) intervenes to ensure that, on the one hand, no personally identifying information enters the registry database and, on the other hand, the data owners involved (UKD, biobank, health insurance company, cancer registry, patient) only receive the personally identifying information (ORBIS ID, KV number) that they need for the data management (see ► **Figure 1** and Table S1 in the appendix).

An application to link the routine data of the statutory health insurers with the data of the registry is being processed. On the basis of §75 of the SGB X, the transfer of social data for scientific research can also take place without individual consent if the interests of the person concerned worthy of protection are not impaired



► **Figure 1** Diagram of the data flow.

or the public interest in the research or planning significantly outweighs the confidentiality interest of the persons concerned and obtaining personal consent is impracticable. Although the patients' consent for the linkage of their SHI data within the registry has been obtained, the approval of the supervisory authority must still be obtained in order to obtain full legitimization.

The data collection is pseudonymised, as described in ► **Figure 1** and Table S1. The pseudonyms are administered by the THS. Without their involvement, the patient data made available for medical research (project-related partial data extract of the register data with coarsening of the data fields) cannot be traced back to a person, or only with disproportionately high technical effort. The identity comparison between the complementary data sources takes place exclusively in the THS. For each external data delivery, the THS generates unique transfer pseudonyms with which only the data of this transfer can be linked to the existing data in the study database. After the linkage of the data in the study database, these one-time transfer pseudonyms are deleted from the database. In the interest of data minimisation, complementary variables in external sources are only used if necessary for identity and plausibility checks in the THS and are not transferred to the study

database. The further linkage of these analysis data with other data sources is explicitly prohibited by binding contracts with data users' institutions. The names of patients and all other confidential information are subject to medical confidentiality, the provisions of the Federal Data Protection Act (BDSG), the European Data Protection Regulation (EU-DSGVO) and other state-specific regulations (e.g. Saxon Hospital Act). The data in the register are stored for a period of 10 years due to legal requirements based on recommendations of Good Epidemiological Practice (GEP) [9].

Documentation and quality assurance

The entire project is carried out in accordance with the Declaration of Helsinki [10] as well as the recommendations of the Good Practice Secondary Data Analysis (GPS) [6], the Good Practice Data Linkage (GPD) [7] and Good Epidemiological Practice (GEP) [9] as well as the professional code of conduct for physicians of the responsible state medical association in the respective current versions. Extensive plausibility checks are carried out for all variables in the data set over the entire observation period. For the comprehensibility and reproducibility of the results, the syntax for statistical analysis is commented on, stored and, if necessary, made available. The in-

vestigation is carried out in accordance with standardised operating procedures (SOPs).

Consent of the participants

Participation in the registry is based on the informed consent principle. The participation is voluntary. The patient information letter (Document S1) and the informed consent form (document S2) can be found in the appendix.

DISCUSSION/CONCLUSION

Due to the inclusion of patients during the entire course of the respective cancer disease, other patient cohorts (prevalent or incident) from the registry can/must be considered depending on the research question. For many scientific questions of outcomes and health care research and the methodological requirements of these, a uniform, defined starting point of observation is necessary (e.g. for survival time analyses). Here, the evaluations of the prevalent data would potentially distort the effects examined, since all patients who died between the initial diagnosis and the start of the study cannot be taken into account. The longer the period between initial diagnosis and study inclusion, the greater the bias in the results (survivor bias [11]). This pitfall should always be taken into account when evaluating registry data/observational studies. The overview of therapies (► **Table 2**) gives an aggregated insight into the treatment of the patients with colorectal adenocarcinoma and pancreatic carcinoma in the registry. In general, in the register it is possible to present and evaluate individual treatment histories. In particular, the inclusion of further outpatient treatments through the linkage with the data of the health insurers makes it possible to present extensive treatment histories of the patients. The registry data can further be used to describe the effectiveness, safety and cost-effectiveness of oncological treatments under routine conditions, the quality of care for cancer patients at the UKD and to derive forecasts of future care needs. In addition, disease and late effects can be identified and quantified with regard to possible tertiary prevention measures. The register can be used to map the care pathways and structure of patients with cancer in an interdisciplinary and cross-sectoral manner and to depict the utilisation of services. In addition, the quality as well as the effectiveness and safety of care can be analysed by analysing progression and survival, taking into account relevant patient-, disease-, therapy- and physician-specific factors. The results can be used to identify care deficits and possible intersection problems within the care system. The project offers the opportunity for evidence-based improvement of the care of insured persons with tumour diseases as well as the structural and organisational development of oncological care. Projects based on the use of secondary data from AOK PLUS as well as the cancer registries will contribute to the development of sustainable solution proposals for these important health policy problems.

Statement of Ethics

Study approval statement: The study protocol for colorectal carcinoma and pancreatic carcinoma received approval from the Ethics Committee of the Medical Faculty of the TU Dresden (EK487102019).

Consent to participate statement: Participation in the registry is based on the informed consent principle. The participation is voluntary.

Author Contributions

All authors were involved in the conception and construction of the registry. OS, JS, TD, HB, and MD prepared the draft manuscript. All authors agreed to the final version of the manuscript.

Data Availability Statement

The data of the study are confidential. They can only be used for the oncological questions mentioned in the study protocol by the participating physicians of the Dresden University Medicine. Use within Dresden University Medicine beyond this purpose is only possible with the consent of the patients and the responsible supervisory authorities after application §75, SGB X. External third parties are excluded from using the data.

Listed in the BQS directory of medical registers in Germany
(https://registergutachten.bqs.de/reg_db/detail.php?pid=NCT%2FUCC+Registerstudienplattform)

Contributor's Statement

None

Acknowledgements

Besonderer Dank gilt Andrea Kloß und Thoralf Stange für ihre äußerst wertvolle Mitarbeit beim Aufbau und Betrieb des Registers. Wir danken Frau Dr. Haufe für ihre Unterstützung bei Fragen zum Register. Dem Direktorium des NCT/UCC Dresden danken wir herzlich für die Unterstützung im Rahmen der Core Unit Registry Trial Plattform.

Conflict of Interest

Unrelated to this study, JS received institutional funding for investigator-initiated research from Sanofi, Pfizer, Novartis, ALK, and acted as a payed consultant for Sanofi, Novartis, ALK and Lilly. OS is a member of the certification commission "Skin Cancer Centers" of the German Cancer Society. MB stated the employment at UKD, member of NCT board (project financing), fees and travel expenses from companies Jazz, Alexion, Novartis and project financing from AbbVie, Roche, Ad-Board: Jazz. The other authors declare that they have no conflict of interest.

References

- [1] Westfall JM, Mold J, Fagnan L. Practice-Based Research—"Blue Highways" on the NIH Roadmap. *JAMA* 2007; 297: 403–406. DOI: 10.1001/jama.297.4.403
- [2] Domenghino A, Walbert C, Birrer DL et al. Consensus recommendations on how to assess the quality of surgical interventions. *Nature Medicine* 2023; 29: 811–822. DOI: 10.1038/s41591-023-02237-3
- [3] Zerillo JA, Schouwenburg MG, van Bommel ACM et al. An International Collaborative Standardizing a Comprehensive Patient-Centered Outcomes Measurement Set for Colorectal Cancer. *JAMA Oncol* 2017; 3: 686–694. DOI: 10.1001/jamaoncol.2017.0417
- [4] Hanahan D. Rethinking the war on cancer. *Lancet* (London, England) 2014; 383: 558–563. DOI: 10.1016/s0140-6736(13)62226-6
- [5] Lowy DR, Collins FS. Aiming High--Changing the Trajectory for Cancer. *The New England journal of medicine* 2016; 374: 1901–1904. DOI: 10.1056/NEJMp1600894
- [6] Swart E, Gothe H, Geyer S et al. Gute Praxis Sekundärdatenanalyse (GPS): Leitlinien und Empfehlungen. *Gesundheitswesen* 2015; 77: 120–126
- [7] March S, Andrich S, Drepper J et al. Gute Praxis Datenlinkage (GPD). *Gesundheitswesen* 2019; 81: 636–650
- [8] Aaronson NK, Ahmedzai S, Bergman B et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute* 1993; 85: 365–376. DOI: 10.1093/jnci/85.5.365
- [9] Deutsche Gesellschaft für Epidemiologie (DGEpi). Leitlinien und Empfehlungen zur Sicherung von Guter Epidemiologischer Praxis (GEP) 2004
- [10] World Medical Association. Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 2013; 310: 2191–2194. DOI: 10.1001/jama.2013.281053
- [11] Hu ZH, Connett JE, Yuan JM et al. Role of survivor bias in pancreatic cancer case-control studies. *Annals of epidemiology* 2016; 26: 50–56. DOI: 10.1016/j.annepidem.2015.11.001