FDZ-Gesundheit

# Access to health data via the FDZ-Gesundheit: Structures and processes explained in brief

**Since October 2025, the Research Data Centre for Health (FDZ-Gesundheit) has provided Germany with a consolidated data source that makes pseudonymised claims data from statutory health insurance accessible in a secure environment for defined purposes. The aim is to strengthen the basis for research and health policy decisions and, furthermore, to prepare for connectivity with the EHDS. The key questions addressed in this article are what opportunities and requirements are associated with the use of this data source.**

## What types of data analysis are typically possible

>> Despite all its limitations, statutory health insurance claims data is particularly suitable for questions that can be addressed through diagnoses, services/procedures, dispensing of medicines and utilisation. Typical analyses include:

- Epidemiology and care provision: prevalence/incidence according to a defined case definition, regional differences, trends over the years.
- Care pathways and patient journeys: cross-sector pathways (outpatient and inpatient), courses of care before and after index events.
- Mapping of healthcare realities as a basis for various price negotiations.
- Treatment and medication analyses: prescription/dispensing patterns, persistence/adherence (as a proxy), treatment changes, concomitant medication.
- Utilisation and costs: service volumes, resource consumption, cost trends by subgroup or event.
- Evaluation of care models: observational designs (e.g. before-and-after or controlled cohorts), provided that exposure and endpoints can be reliably operationalised within the billing data.

It is important to have realistic expectations, as statutory health insurance (GKV) claims data are not clinical or registry data. The documentation primarily serves billing purposes and often reflects clinical measurements, severity levels or patient-reported outcomes only indirectly. The validity and robustness of statutory health insurance claims data therefore require a transparent and well-founded,and in some cases relatively complex,operationalisation and are often underpinned by sensitivity analyses.

## How is access regulated?

Anyone can access the FDZ-Gesundheit, including the commercial healthcare sector. There is no restriction to specific institutions. It is

important that the data may only be used for one of the legally defined purposes. This is mandatorily checked as part of the application process. The law lists 10 permissible purposes of use (Section 303e(2) of the German Social Code, Book V). These include, for example, improving the quality of care, scientific research, analyses of healthcare provision, analyses of the effectiveness of cross-sectoral forms of care, and the development, further development and monitoring of the

safety of medicinal products, medical devices, examination and treatment methods, aids and remedies, digital health and care applications, and artificial intelligence systems in the healthcare sector.

The use of the data for other purposes is not permitted. In particular, the data made available must not be used for market research, advertising or sales activities. To ensure this, Legal, Compliance and Data Protection should be consulted in case of doubt before submitting an application. This is important as breaches can have far-reaching consequences.

The fees for an FDZ application are governed by the Data Transparency Fees Regulation (DaTraGebV) and are payable by the applicant. The following components are particularly relevant in practice:

| Gebühren eines FDZ-Antrags | | |
|---|---|---|
| **Gebührenbestandteil** | **Typischer Ansatz** | **Worauf zu achten ist** |
| Grundgebühr (Antragsbearbeitung) | 4.000 Euro | unabhängig vom Umfang der Auswertung |
| Auswertung/Bereitstellung je Berichtsjahr | 1.000 Euro pro Berichtsjahr | je mehr Jahre beantragt, desto höher die Zusatzgebühr |
| Nutzung gesicherte virtuelle Umgebung | 1.000 Euro pro Nutzungstag | straffe Arbeitsplanung reduziert Nutzungstage |
| Zwischenergebnisse | 109 Euro pro Zwischenergebnis | budgetär berücksichtigen |
| Beratung (optional) | 100 bis 3.000 Euro | abhängig vom Umfang und budgetär berücksichtigen |
| Schulung (optional) | 1.000 Euro (Basis) + 600 Euro je Zusatzmodul | für neue Teams |
| Sonstige Auslagen | Lizenzgebühren für Nutzung bestimmter angebotener Services | vorab im Antragsverfahren abklären und budgetär berücksichtigen |

***Table 1:*** **Fees for an FDZ application. Source: own representation based on information from FDZ-Gesundheit.**

**Calculation example (without cost reduction): Assumption:** Project with 5 reporting years (e.g. 2019–2023), 14 days of analysis work in the virtual environment, plus consultancy fees of 1,000 euros.

- Basic fee: €4,000
- Additional fee per reporting year: 5 years at €1,000 each = €5,000
- Use of virtual working environment (SPE): 14 × €1,000 = €14,000
- Consultancy and two interim reports: €1,000 plus €218
- Total: €24,218 (plus other expenses)

This hypothetical example shows that the number of reporting years, the need for consultancy and, in particular, the number of days spent using the secure analysis environment (SPE) can significantly influence the total

cost of an application. It therefore makes sense to develop the scripts and test logs as far as possible before submitting the application for access and to concentrate the time spent in the SPE to a few intensive working days. Furthermore, some programming skills are required for the translation of the analysis script.

The DaTraGebV provides for a reduction ‘to one-tenth’ of the fees, subject to application, if the application is submitted by certain institutions. These include, amongst others, universities, university hospitals, publicly funded non-university research institutions, and public

healthcare research institutions. A potential reduction does not apply if data access is sought jointly with other institutions that do not belong to the eligible institutions, or if data access is granted as part of a contract research project for a non-eligible institution.

In practice, this means that for industry-related projects, it is worth clarifying at an early stage whether the industry partner requires access to the secure environment or whether collaboration via a project advisory board, coordination of methods and discussion of results without data access may suffice. In cases of uncertainty, it is advisable to seek prior clarification with the FDZ-Gesundheit so that roles, application format and fee implications are clear and predictable.

**What does a typical timeline look like: from application to result**

The actual processing times depend on the complexity and completeness of the application in question. The FDZ-Gesundheit expressly emphasises that thorough preparation of the application is a prerequisite for swift processing. If you have any questions, you can consult the helpdesk, where you can quickly obtain clarification.

A phased approach to project management has proven effective:

**Phase 1 – Concept and data logic (internal)**

Specify the research question, intended use, case definitions, endpoints, observation window, required variables and reporting years. Define early on which results are to be exported later (tables) and what minimum cell sizes are required, as the volume of results is limited. Plan to test the script in advance using the Public Use File.

**Phase B – Application preparation following registration of the institution and roles**

Formulate the data cut and methodology plan in such a way that it is verifiable. The FAQs from the FDZ-Gesundheit emphasise that, when assessing research questions, the intended use is particularly relevant, as is the suitability of the data cut. In addition, the smallest possible data set and the maximum number of cells available in the results are assessed.

**3. Phase C – Review cycles and queries**

Respond promptly to any queries from the FDZ-Gesundheit. Common points include: data minimisation, unambiguous variable lists, plausibility of endpoints, planned aggregation/output formats or data volume.

**4. Phase D – Working in the virtual working environment (SPE) and work preparation**

Define the software and

script strategy, and, where appropriate, request interim results. The aim is to ensure that productive work in the SPE can be completed within a few days of use.

**5. Phase E – Analysis and request for results**

Evaluation in the virtual working environment/analysis room; followed by a request for the results set. The final results or result tables must be requested; these will only be provided after review by the FDZ Health, particularly with regard to re-identification risks. Adjustments may be necessary for small sample sizes. Graphs cannot currently be exported.

**6. Phase F – Publication and report preparation**

Interpret results, clarify the limitations of routine statutory health insurance data (coding logic, proxies), and provide documentation of operationalisation to ensure reproducibility. The FDZ-Gesundheit recommends publishing a paper with results in the database; accepted applications are listed in the application register.

To optimise cost and time planning and management, the aforementioned preparations should be carefully considered before accessing the analysis room. Any additional steps will delay the project or may incur usage days in the virtual analysis environment.

**Conclusion and checklist: What are the stumbling blocks for applications?**

The FDZ-Health offers a new means of accessing data and will, in the coming years, become a key resource for the analysis of statutory health insurance (GKV) billing data and for addressing issues relating to healthcare provision and research. With regard to the scope of data, it should be noted that for rare diseases, the minimum number per cell is 30 according to the FDZ, though this can be reduced to 5 upon justified request. It is not yet possible to assess how the data minimisation requirement will affect the scope of accessible data beyond this.

It is clear, however, that the FDZ-Gesundheit will continue to evolve even after its launch and will enable the inclusion of additional health data and sources in the future, such as from the electronic patient record (ePA) or links to register data.

With regard to the structure of the available data, the publication of Data Model 4 has been announced for 2026, which, among other things, provides data for the years 2024 and 2025 on a quarterly basis. From the perspective of a competitive environment within the industrial healthcare sector, it may be advisable to continue seeking analyses based on the previously accessible pools of statutory health insurance (GKV) billing data. In the case of FDZ applications, the complexity of an analysis may be limited, as the maximum number of results is restricted or may necessitate a modification of the applications. In principle, prior to any analysis, it should be assessed in advance which data source is best suited and most appropriately usable for the specific type of query.

The following aspects relating to an application to the FDZ-Gesundheit are important:

- A permissible purpose of use must be stated.

• The intended use and the research question must be compatible; the assessment is based on whether the research question is compatible with the stated purpose.

• Successful applications are listed in the application register once the results have been provided.

• The data subset should be as small as possible and as large as necessary, i.e., the list of variables, reporting years and sub-populations should be closely aligned with the analysis plan and justified.

• Data set description (structure, tables and variables): The published description shows the data structure and the available data models by reporting year (Data Model 1: 2009–2015; Data Model 2: 2016–2018; Data Model 3: from 2019).

• Public Use File (PUF): A fully anonymised data product intended specifically for technical development and script testing – not for robust substantive findings.

• Document operationalisation: case definitions, index date, endpoints, time window, exclusion criteria.

• Carefully plan result formats and volume in advance, i.e. table size, minimum cell sizes and the level of aggregation, to ensure that result verification runs smoothly.

• Cost control can be scaled, amongst other things, via the number of days of use required, the number of reporting years, and the need for consultancy/training; in other words, the scripts, plausibility checks and results tables should be clearly prepared in advance of access; consolidate work within the virtual working environment, in particular by specifying the necessary technical requirements at the time of application.

• Clarify cooperation arrangements (particularly with industry) in advance from both a legal and organisational perspective, as they may preclude cost reductions.

One thing is certain: even with the introduction of the European Health Data Space (EHDS), the FDZ-Gesundheit will remain a key player in health research in Germany. Exploring its potential therefore presents an opportunity for all and strengthens data-driven decision-making within the German healthcare system. <<

**by: Dr. rer. oec. Katja[Gehrke1] ORCID: 0000-0001-5516-457X Dr. iur. Friederike[Goeters2] ORCID: 0009-0007-9592-4001**

1: Institute for Descriptive Health Data Analysis OHG (DGDA), Hamburg

2: Solicitor, Möhrle Happ Luther, Hamburg

**Abstract**

**Introduction:** Since October 2025, the Research Data Centre for Health (FDZ-Gesundheit) has provided Germany with a consolidated data source that makes pseudonymised statutory health insurance claims data accessible in a secure environment for defined purposes. The aim is to strengthen the basis for research and health policy decisions and, furthermore, to prepare for connectivity with the EHDS. The key questions addressed in this article are what opportunities and requirements are

associated with the use of this data source. **Methods:** This overview provides a descriptive classification of the new data offering. Factors considered include suitability for research projects as well as practical requirements regarding application procedures, data extraction and analysis logic. **Results:** The FDZ data offer high case numbers and cross-sectoral analyses of healthcare provision issues; however, the volume of results and the complexity of analyses are currently limited. At the same time, its use requires thorough and technical preparation, as well as a sound understanding of statutory health insurance (GKV) data. **Discussion:** Since mid-February 2026, 82 institutions have registered, 53 applications have been submitted and 18 analysis spaces have been utilised, with initial results already in the review phase. By the end of the year, at least 300 research projects are expected to have been initiated. The results will be published in the public application register. The first entry in the register comes from the GKV-Spitzenverband. The FDZ-Gesundheit can enrich healthcare research as a data source and is expected to offer additional potential in future through the continuous expansion of data content. However, initial experiences also reveal limitations. **Keywords:** Research Data Centre for Health, FDZ-Gesundheit, access to statutory health insurance claims data, application fees, checklist

**Access to health data via the FDZ-Gesundheit: a concise explanation of structures and processes**

**Introduction:** Since October 2025, the Research Data Centre for Health (FDZ-Gesundheit) has provided Germany with a consolidated data source that makes pseudonymised claims data from statutory health insurance accessible in a secure environment for defined purposes. The aim is to strengthen the foundation for research and health policy decisions and to prepare for connectivity with the EHDS. The central questions of this article are what opportunities and requirements are associated with the use of this data source. **Methods:** This overview provides a descriptive analysis of the new data offering. Factors considered include suitability for research projects as well as practical requirements for data requests, data extraction, and analysis logic. **Results:** The FDZ data offer a high number of cases and cross-sectoral analyses of healthcare issues; however, the volume of results and the complexity of analyses are currently limited. At the same time, using the data requires thorough and technical preparation as well as a solid understanding of statutory health insurance (GKV) data. **Discussion:** Since mid-February 2026, 82 institutions have registered, 53 applications have been submitted, and 18 analysis spaces have been utilised, with the first results already in the review phase. By the end of the year, at least 300 research projects are expected to have been initiated. The results will be published in a public application registry. The first entry in the registry comes from the GKV-Spitzenverband. The FDZ-Gesundheit can enrich health services research as a data source and is expected to offer additional potential in the future through the continuous expansion of its data content. However, initial experiences also reveal limitations. **Keywords:** Research Data Centre for Health, FDZ-Gesundheit, access to GKV billing data, application fees, checklist

**Authors' declaration**

The authors declare that there are no conflicts of interest.

**Citation**

Gehrke, K., Goeters, F.: “Access to health data via the FDZ-Gesundheit: Structures and processes explained in brief”, in “Monitor Versorgungsforschung” (03/26), pp. 53–56. http://doi.org/10.24945/MVF.03.26.1866-0533.2834

**REGINT project: Funding phase completed**

>> The REGINT project successfully completed its funding phase at the end of March 2026. The project results are now being systematically disseminated to the specialist public and incorporated into continuing professional development. At the DKVF 2026, REGINT is organising a symposium with external experts on the topic “REGINT Symposium: Where do we stand? Health services research using register data and the way forward”. The focus is on the use of registry data for evidence-based health services research. In addition, presentations of the case study results from Phase 2 are planned at further scientific events, including the 20th Annual Conference of the DGEpi in Ulm. In the context of continuing professional development, the project will be featured in a dedicated module within the REGIBA “Healthcare-related data” module (2026–2027). In parallel, a publication on Phase 1 and two further manuscripts based on the case studies are being finalised. <<

**DNVF Registry Working Group Expert Meeting**

>> At an online expert meeting of the DNVF Registry Working Group in March 2026, over 30 experts discussed the potential and limitations of Target Trial Emulation (TTE), which is not a specific statistical method but a conceptual framework for making assumptions about causal effects explicit and transparent. TTE is increasingly being used to evaluate the benefits of healthcare interventions such as cancer screening. However, challenges remain in the form of confounding (e.g. due to indication) and incomplete data, as Vanessa Didelez and Jochen Schmitt emphasised. Specific projects such as the ZEBra study (combining health insurance and cancer registry data) or the NANA project (feasibility of non-randomised analyses) show that TTE helps to identify unsuitable research questions at an early stage – yet data quality is crucial. <<

**TMF Working Group on Medical Registries**

>> On 27 May 2026, the first meeting of the EFMI working group on European medical patient registries took place as part of Medical Informatics Europe (MIE) 2026 in Genoa. Initiated by TMF e.V., the group aims to create a permanent European platform for knowledge exchange and harmonisation in order to successfully integrate registries into the European Health Data Space (EHDS). The focus was on promoting interoperability standards such as HL7 FHIR and OMOP, the development of data stewardship frameworks, and sustainable governance models. In addition, a workshop on the regulatory and technical challenges of the EHDS for registry operators will take place on the same day from 11:00–13:00. <<

**Obituary: Peter Christian Scriba**

>> On 4 April 2026, Prof. Dr. med. Dr. h. c. Peter Christian Scriba passed away at the age of 90. With his passing, German medicine has lost one of its most influential scientists and a figure who inspired generations of

colleagues through his combination of intellectual clarity, human warmth and unwavering commitment to patients. <<

# References

Buzer.de (consolidated version of the DaTraGebV) as a means of accessing Section 11, (accessed 19 December 2025) Buzer.de: Regulation on the detailed regulation of procedures under the GDNG (amendments including those relating to fees). (accessed 19 December 2025) DaTraGebV Section 11 'Fee increases and reductions' (consolidated version). FDZ-Gesundheit: Frequently Asked Questions – FAQ, https://www.forschungsdatenzentrum-gesundheit.de/infoportal/faqs, accessed: 10 February 2026 FDZ-Gesundheit: Fee Schedule, https://www.forschungsdatenzentrum-gesundheit.de/infoportal/antragsvorbereitung/gebuehrenverordnung, accessed: 10 February 2026 Research Data Centre (FDZ-Health): Application Portal, https://www.forschungsdatenzentrum-gesundheit.de/infoportal/antragsportal, (accessed 19 December 2025) Research Data Centre-Health (2026) Research Data Centre: Application Register, accessed on 17 March 2026 Häussler, B., Hoffmann, W., Stegmaier, P. (2023) We need experimental spaces for healthcare research, in Monitor Versorgungsforschung (01/23), pp. 6–15. http://doi.org/10.24945/MVF.01.23.1866-0533.2470 Horenkamp-Sonntag et al. (2017) Validity of pharmaceutical data in statutory health insurance routine data in: Monitor Versorgungsforschung 03/17, pp. 53–59, doi: 10.24945/MVF.03.17.1866-0533.2019 Ihle P et al. (2023) The Health Research Data Centre – Developments to date and future steps in Gesundheitswesen 2023; 85 (Suppl. 2): pp. 99–100 Ludwig, M., Schneider, K., Heß, S., Broich, K. (2024) Establishment of the new 'Health Research Data Centre' for data provision to the scientific community in

Bundesgesundheitsbl 2024 · 67:131–138, https://doi.org/10.1007/s00103-023-03831-z

March S et al. (2023) Health Research Data Centre – Vision for further development from the perspective of healthcare research 2023; 85 (Suppl. 2): pp. 145–153, Thieme E-Journals – Das Gesundheitswesen / Abstract

Prasser, F., Riede, N., Wolter, S., Corr, D., Ludwig, M. (2024) Artificial Intelligence and Secure Health Data Use in the AI-FDZ Project: Anonymisation, Synthesis and Secure Processing for Real-World Data in Bundesgesundheitsbl 2024 · 67:171–179, https://doi.org/10.1007/s00103-023-03823-z

Schneider, K., Heß, S., Weber, S. (2024) Known challenges and new opportunities of real-world data in Bundesgesundheitsbl 2024 · 67:129–130, https://doi.org/10.1007/s00103-023-03826-w, 103_2023_Article_3826.pdf

Stegmaier, P. (2019) "Research Data Centre from 2022 at the earliest", in: "Monitor Versorgungsforschung" (06/19), pp. 12–13, doi: 10.24945/MVF.06.19.1866-0533.2182 Stegmaier, P. (2023) "GDNG and DigiG are on the way", in "Monitor Versorgungsforschung" (05/23), pp. 18, 20, http://doi.org/10.24945/MVF.05.23.1866-0533.2537 Swart Enno et al. (2021) Setting the course now, in Gesundheitswesen 2021; 83 (Suppl. 2): pp. 139–141, a-1537-9722.pdf

Wenzel M et al. (2025) State of play regarding research using secondary data at the Health Research Data Centre, Thieme, Open Access