

STIPENDIENPROGRAMM

Hamburger Krebsgesellschaft



PROJEKTBERICHT

Epidemiologie des Mammakarzinoms mit Hirnmetastasen in Deutschland

Projektlaufzeit

01.03.2025 – 28.02.2026 (12 Monate)

Bericht-Stand

08. Juni 2026

Antragstellende

PD Dr. Elena Laakmann — UKE Hamburg
Dr. Annemarie Schultz — Hamburgisches Krebsregister
PD Dr. Frederik Peters — Hamburgisches Krebsregister
Leonie Rosenberg — HAW Hamburg

1. Zusammenfassung

Das Projekt verfolgte zwei parallele Stränge, die in einem Proof-of-Concept zusammengeführt werden. Strang 1 umfasste den Aufbau einer datenschutzkonformen Container-Lösung, mit der Auswertungsskripte an die Daten der 15 Landeskrebsregister reisen können — ohne dass Patientendaten die jeweilige Registerumgebung verlassen. Strang 2 umfasste zwei epidemiologische Publikationen zur Hirnmetastasierung beim Mammakarzinom auf Basis des bundesweiten ZfKD-Datensatzes: eine deskriptive Analyse (Paper 1, eingereicht) und eine Competing-Risks-Risikofaktorenanalyse (Paper 2, in Vorbereitung). Als Proof-of-Concept werden die aus beiden Publikationen hervorgegangenen Analyseskripte über KIKA verteilt in allen 15 Registern ausgeführt und die aggregierten Ergebnisse am HKR konsolidiert — der Nachweis, dass Container-Lösung und Analysemethodik gemeinsam funktionieren (Paper 3).

Beide Projektstränge wurden planmäßig umgesetzt. Das entwickelte Verfahren — die Container-basierte Lösung KIKA (KI in der Krebsregister-Datenanalyse) — ist einsatzbereit, mit Forscher- und IT-Handbuch sowie Beispielskripten ausgestattet und steht den 15 Landeskrebsregistern als fertiges Air-Gap-Installationspaket zur Verfügung. Mit der Vorstellung beim DKFZ am 01.06.2026 beginnt die Phase der bundesweiten Implementierung; die Zusammenführung beider Stränge im Proof-of-Concept ist dabei der zentrale nächste Schritt. Das Projekt hat damit nicht nur das ursprünglich versprochene Faktenblatt zur Hirnmetastasierung greifbar nahegebracht, sondern eine Infrastruktur geschaffen, die für jede künftige seltene Konstellation wiederverwendbar ist.

2. Ausgangssituation und Projektziel

In den letzten Jahrzehnten ist durch die flächendeckende Erfassung klinischer Krebsregisterdaten auf Basis des bundeseinheitlichen onkologischen Basisdatensatzes (oBDS) ein Datenbestand entstanden, der für die Versorgungsforschung und die Generierung evidenzbasierter S3-Leitlinien-Empfehlungen einen unschätzbaren Wert hat. Die externe Nutzung dieser Daten ist jedoch durch die hohen — und berechtigten — Schutzerfordernisse sowie die zugehörigen Antragsverfahren erheblich verlangsamt. Zwischen Antragstellung und Datenverfügbarkeit liegen typischerweise mehrere Monate.

Bei seltenen Konstellationen wie der Hirnmetastasierung beim Mammakarzinom potenziert sich dieses Problem: Allein im Hamburgischen Krebsregister waren zu 2023 nur 20 dokumentierte Fälle mit Hirnmetastasen (BRA-Lokalisation) bei C50-Diagnose vorhanden — eine Datenmenge, die für die im Antrag genannten differenzierten Auswertungen (Histologische Subtypen, TNM-Stadium, Therapielinien, Überleben) bei weitem nicht ausreicht. Damit ist es nötig, Daten aller 15 Landeskrebsregister einzubeziehen.

Das Projekt zielte darauf ab, diesen Engpass strukturell zu überwinden: Statt Daten zentral zusammenzuführen, soll der Code „zu den Daten reisen“. Jedes Landeskrebsregister behält die volle Hoheit über die eigenen Patientinnendaten; nur aggregierte, nicht-personenbezogene Ergebnisse werden geteilt.

3. Umgesetzte Arbeitspakete

3.1 Arbeitspaket 1 — Datenschutzkonformer Ansatz: KIKA-Container

Konzeptionell entschieden sich die Projekt-Beteiligten gegen aufwendige Spezialsoftware (vantage6) und für die Eigenentwicklung einer Container-basierten, niedrighwelligen Lösung. Die rechtlich und technisch heterogene Krebsregisterlandschaft in Deutschland erfordert ein Werkzeug, das ohne vertiefte technische Kenntnisse einsetzbar ist und dezentral — auch ohne Internetverbindung — in den Registern vor Ort funktioniert.

Die Anschubfinanzierung der Hamburger Krebsgesellschaft ermöglichte es, weitere Fördermittel einzuwerben: Das Vorhaben setzte sich bei der Ausschreibung „Call for Transfer“ des C4T PIER PLUS durch (Projektname KIKA — KI in der Krebsregister-Datenanalyse), woraufhin die BWFG zusätzlich eine wissenschaftliche Dienstleistung an die Hochschule für Angewandte Wissenschaften Hamburg vergab. So konnte über eine Kooperation mit dem Fachbereich Informatik der HAW der so genannte „KIKA-Container“ programmiert werden.

Architektur:

- KIKA besteht aus 8 Docker-Containern (Docker Compose): PostgreSQL-Datenbank, Next.js-Frontend (Port 8090), FastAPI-Backend, Import-Worker mit XSD-Prüfung, Redis-Job-Warteschlange, nginx Reverse-Proxy, Schema-Migrations sowie ein vollständig konfigurierter RStudio-Server (Port 8787) mit vorinstallierten R-Paketen (data.table, survival, cmprsk, ggplot2, DBI/RPostgres). Gesamtpaket ca. 1,8 GB.
- Vollständige Air-Gap-Fähigkeit: kein Internetzugang im laufenden Betrieb erforderlich.
- Einheitliches Datenbankschema, das exakt der oBDS-Spezifikation folgt und auf alle Register-Implementierungen passt.
- Desktop-Launcher (KIKA.bat / KIKA.ps1) für nicht-technische Nutzer:innen — Doppelklick-Start ohne PowerShell-Kenntnisse. Systemvoraussetzungen: Windows 10/11 Pro 64-Bit, Docker Desktop ≥ 4.30 mit WSL2-Backend, 16 GB RAM, 50 GB SSD.

Lieferumfang für die Register:

- Komplettes Air-Gap-Installationspaket (kika-air-gap.tar.zst, ca. 2,0 GB)
- IT-Handbuch mit Schritt-für-Schritt-Installation, Backup-Anleitung und Troubleshooting
- Forscher-Handbuch mit Erklärungen zu Datenmodell, Skript-Konvention und Output-Erstellung
- Datenbank-Schema-Dokumentation mit allen 17 Tabellen und Feldern

3.2 Arbeitspaket 2 — Harmonisierung und Pilotanalyse

Parallel zur Containerentwicklung wurden die inhaltlichen Analysen zur Hirnmetastasierung auf Basis des bundesweiten ZfKD-Datensatzes durchgeführt. Auf Basis des Variablenkatalogs aus dem Antrag entstanden zwei Analyseskripte: eines für die deskriptive Epidemiologie (Paper 1) und eines für die Competing-Risks-Risikofaktorenanalyse (Paper 2). Beide Skripte sind modular aufgebaut und können in jedem Landeskrebsregister über KIKA identisch ausgeführt werden — der Einsatz in der verteilten Analyse (Paper 3) ist damit direkt vorbereitet.

Die Identifikation der Hirnmetastasen erfolgt über den oBDS-Code BRA in der JSON-strukturierten Spalte `distant_metastasis` der Tabelle `tumor_follow_up`.

3.3 Arbeitspaket 5 — Faktenblatt und methodische Publikation

- Am 01.06.2026 wurde KIKA bei der DKR-Jahrestagung am DKFZ Heidelberg vor den 15 Landeskrebsregistern vorgestellt (inkl. Live-Demo). Das Air-Gap-Paket steht nun über die dDatabox des Hamburgischen Krebsregisters zum Download bereit.
- Die aus Paper 1 und Paper 2 entwickelten Skripte werden via KIKA an die 15 Landeskrebsregister verteilt. Jedes Register führt die Skripte lokal aus und liefert ausschließlich aggregierte Ergebnisdateien zurück (kein Personenbezug).
- Die Konsolidierung der Aggregate am HKR, Erstellung des bundesweiten Faktenblattes und Einreichung als Paper 3 schließen den Kreis: Die ZfKD-basierten Befunde aus Paper 1 und 2 werden mit den aktuelleren, vollständigeren Registerdaten validiert und erweitert.
- Publikation der methodischen Arbeit (Container-Lösung, Architektur, Übertragbarkeit) ist parallel geplant.

3.4 Publikationen

Aus dem Projekt gehen drei Publikationen hervor:

Paper 1 — Epidemiologie der Hirnmetastasierung beim Mammakarzinom in Deutschland

Titel: „Cerebrale Metastasen beim Mammakarzinom: Eine bevölkerungsbasierte Kohortenstudie zu Epidemiologie, Therapiemustern und Überleben 2020–2023 in Deutschland“

Inhalt: Retrospektive bevölkerungsbasierte Analyse von 296 246 Mammakarzinom-Patientinnen aus dem ZfKD-Datensatz (2020–2023). 2.531 Patientinnen (0,85 %) entwickelten zerebrale Metastasen. Schwerpunkte: Inzidenz, Altersverteilung, molekulare Subtypen, Therapiemodalitäten und Gesamtüberleben (Median 12 Monate, subtyp-spezifisch 6–26 Monate).

Status: Manuskript eingereicht; Überarbeitung nach Reviewer-Feedback in Vorbereitung

Paper 2 — Risikofaktoren für Hirnmetastasen bei extrakraniell metastasierten Patientinnen

Titel: „Risk Factors for Brain Metastases in Extracranially Metastatic Breast Cancer — A Population-Based Competing Risks Analysis from German Cancer Registry Data“ (vorläufig)

Inhalt: Prädiktoren für die Hirnmetastasierung bei extrakraniell metastasierten Mammakarzinom-Patientinnen. Kohorte: 46.485 Patientinnen aus 14 deutschen Bundesländern (ZfKD-Datensatz, Diagnosejahre 2014–2023), 2.083 Hirnmetastasen-Ereignisse. Methodik: Fine-Gray Competing Risks. Zentrale Befunde: HER2+/HR–: SHR 4,95; TNBC: SHR 3,50; HER2+/HR+: SHR 2,74 vs. HR+/HER2– als Referenz. Lungen- und Lebermetastasen als unabhängige Prädiktoren.

Status: In Vorbereitung.

Paper 3 — KIKA: Datenschutzkonforme Container-Lösung für verteilte Krebsregister-Analysen

Inhalt: Paper 3 bringt die beiden parallelen Projektstränge zusammen. Die in Paper 1 und Paper 2 entwickelten Analyseskripte werden über KIKA verteilt in den 15 Landeskrebsregistern ausgeführt. Die Rückgabe ausschließlich aggregierter Ergebnisdateien erbringt den Proof-of-Concept, dass die datenschutzkonforme verteilte Auswertung über alle deutschen Register in der Praxis funktioniert. Zugleich liefert Paper 3 die bundesweit repräsentativsten Zahlen zur Hirnmetastasierung beim Mammakarzinom — auf deutlich größerer Fallbasis als die ZfKD-Analyse aus Paper 1.

Status: Vorstellung beim DKFZ Heidelberg am 01.06.2026 erfolgt. Publikation wird auf Basis des Feedbacks der Register vorbereitet und eingereicht.

Kooperationspartner: Hamburgisches Krebsregister (HKR), Universitätsklinikum Hamburg-Eppendorf (UKE), HAW Hamburg — Fachbereich Informatik (Prof. Kai von Luck)

4. Wesentliche Ergebnisse

4.1 Methodisch

KIKA ist die erste produktionsreife, datenschutzkonforme Container-Lösung für verteilte Auswertungen über alle deutschen Landeskrebsregister, die ohne komplexe Spezialsoftware auskommt. Die gesamte Lösung ist Open Source (4 GitHub-Repositorien).

Wesentliche Designentscheidungen:

- **Code zu Daten statt Daten zu Code:** Skripte werden geteilt, Patient:innendaten bleiben vor Ort.
- **Aggregat-basierter Austausch:** Die ausgetauschten Ergebnisdateien enthalten ausschließlich zusammengefasste Kennzahlen (kein:e Patient:in mit N<5 wird einzeln ausgewiesen).
- **Reproduzierbarkeit durch Containerisierung:** Jedes Register führt exakt das gleiche Skript auf der gleichen R-Umgebung mit den gleichen Paketen aus.

Die Lösung ist nicht auf das Pilotthema beschränkt. Jede künftige Forschungsfrage zu seltenen onkologischen Konstellationen kann nach demselben Muster bearbeitet werden.

4.2 Inhaltlich

Die inhaltlichen Analysen liefen parallel zur Containerentwicklung auf Basis des bundesweiten ZfKD-Datensatzes (296 246 Patientinnen, 2020–2023). Paper 1 liefert die erste bundesweite Epidemiologie der Hirnmetastasierung beim Mammakarzinom: 2.531 Patientinnen (0,85 %) entwickelten zerebrale Metastasen, mit subtypspezifischen Überlebensunterschieden (Median 6–26 Monate). Paper 2 stellt Prädiktoren der Hirnmetastasierung mittels Fine-Gray Competing Risks dar.

Die Datenauswertung hat wertvolle Hinweise auf Optimierungspotenziale in der Datenerfassung geliefert: Bei einem relevanten Teil der Patientinnen fehlten Angaben zum Mammakarzinom-Subtyp sowie zur Therapie nach Diagnosestellung der Hirnmetastasen. Zeitliche Trends lassen sich auf Basis des ZfKD-Datensatzes noch nicht belastbar auswerten, da die bundesweite Datenzusammenführung beim RKI erst kürzlich begonnen hat. Diese Erkenntnisse sind unmittelbar nutzbar: Sie geben den Krebsregistern konkrete Anhaltspunkte, an welchen Stellen die Erfassungsqualität gezielt gestärkt werden sollte — und damit eine Grundlage für robustere Analysen in der Zukunft.

4.3 Strukturell — Wirkung über das Pilotprojekt hinaus

Weitere strukturelle Effekte gehen über das ursprüngliche Projektziel hinaus:

1. **Zeitgewinn:** Die durchschnittliche Wartezeit auf Daten aus einem einzelnen Register sinkt durch KIKA von typischerweise drei Monaten auf wenige Tage.
2. **Wiederverwendbarkeit:** Die Investition in KIKA amortisiert sich mit jedem weiteren Forschungsprojekt — für seltene Tumore, regionale Versorgungsfragen und die Berechnung von Qualitätsindikatoren.

5. Ausblick

Kurzfristig (Q2 2026):

- Verteilung des Air-Gap-Pakets und des Hirnmetastasen-Skripts im Anschluss an die Vorstellung beim DKFZ Heidelberg (01.06.2026, abgeschlossen)

Mittelfristig (Q3–Q4 2026):

- Eingang der ersten Aggregate aus den Pilot-Registern (Ausführung der Paper-1/2-Skripte via KIKA)
- Konsolidierung der Aggregate, Validierung der ZfKD-Ergebnisse gegen Registerdaten
- Einreichung Paper 3 (Proof-of-Concept KIKA + bundesweite Hirnmetastasen-Epidemiologie)

Längerfristig möglich (2027+):

- Internationale Anbindung über Kooperation mit dem Netherlands Cancer Institute (Adaption für RARECAREnet-Projekte)
- Erweiterung um weitere klinische Module (z.B. Survival-Analyse, Subtyp-Analysen)

6. Anhang

6.1 Pilot-Outputs

Datei	Inhalt
01_anteil_hirnet.png	Anteil Hirnmetastasen an C50-Patientinnen
02_alter_bei_hirnet_C50.png	Altersverteilung bei Diagnose der Hirnmetastasen
03_zeit_diagnose_bis_hirnet.png	Zeit Erstdiagnose Mammakarzinom → Hirnmetastase
04_zeitverlauf_anteil.png	Anteil im zeitlichen Verlauf
Hirnmetastasen_Hamburg.xlsx	Alle Tabellen als Excel
agg_hirnet_hamburg.rds	Aggregat-Datei für Konsolidierung (keine Patientendaten)

6.2 Verfügbare Materialien für die Krebsgesellschaft

- KIKA Air-Gap-Installationspaket (USB-Stick, ca. 1,8 GB)
- Vortragsfolien DKR-Jahrestagung Heidelberg (PowerPoint, 18 Folien)
- IT-Handbuch und Forscher-Handbuch (jeweils ca. 12 Seiten)
- Pilot-Skript Hirnmetastasen (R, modular dokumentiert)
- Quellcode aller Komponenten auf GitHub (4 Repositorien, MIT-Lizenz)

6.3 Screenshots KIKA-Applikation und Analyse-Outputs

Die folgenden Abbildungen zeigen die KIKA-Benutzeroberfläche sowie repräsentative Analyse-Grafiken aus dem integrierten RStudio.

Abbildung 1: XML-Upload — Startseite und Upload-Workflow

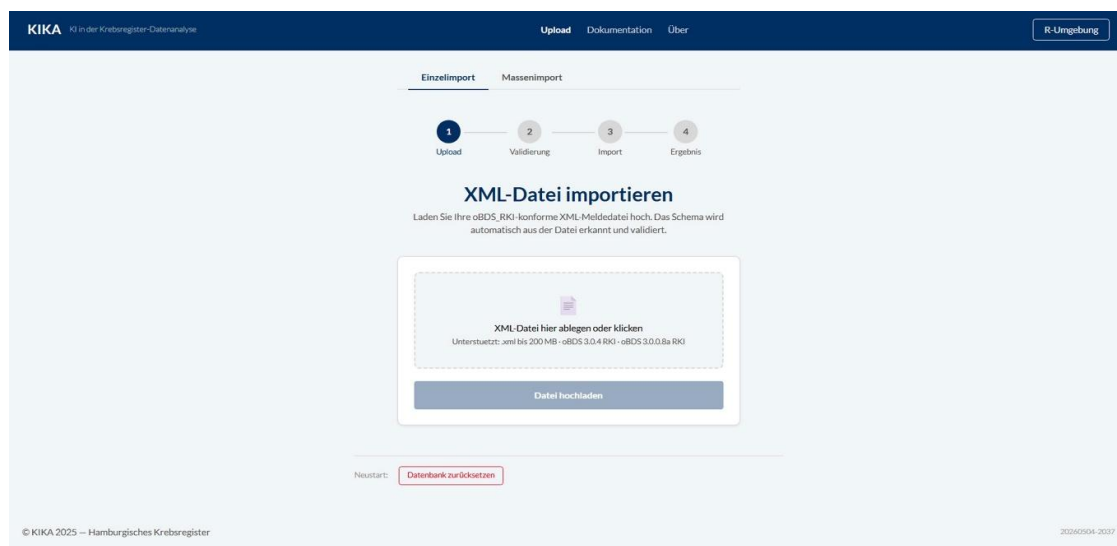


Abb. 1: KIKA Upload-Seite (Einzelimport). Vierstufiger Workflow (Upload → Validierung → Import → Ergebnis) als Fortschrittsleiste. Unterstützte Formate: oBDS 3.0.4 RKI und oBDS 3.0.0.8a RKI.

Abbildung 2: Importbericht nach erfolgreichem Import

KIKA KI in der Krebsregister-Datenanalyse Upload Dokumentation Über [R-Umgebung](#)

Einzelimport Massenimport

Upload Validierung Import Ergebnis

Fertig!
Import erfolgreich

Importbericht
Die Datei wurde erfolgreich validiert und importiert.

5.000 Patienten	2014-2023 Diagnosejahre
62 J. 22-99 Jahre Medianes Alter	5.000 Fälle

⚠ 501 Datenqualitätsprobleme — Datei wurde trotzdem vollständig importiert [I Qualitätsberichte CSV](#)

Die importierten Daten stehen jetzt in der R-Umgebung bereit.

[Daten in R-Umgebung analysieren](#)

[I Importbericht CSV](#) [Weitere Datei importieren](#)

Neuart: [Datenbank zurücksetzen](#)

© KIKA 2025 — Hamburgisches Krebsregister 2026/0504-2037

Abb. 2: Importbericht nach erfolgreichem Import (synthetischer Testdatensatz, 5.000 Patientinnen). 501 Datenqualitätsprobleme protokolliert, ohne den Import zu blockieren.

Abbildung 3: Dokumentationsseite — Handbücher und Datenbank-Schema

KIKA KI in der Krebsregister-Datenanalyse Upload Dokumentation Über [R-Umgebung](#)

Dokumentation

Alles, was du zur Arbeit mit KIKA brauchst — als Forscher, als Administrator oder beim Schreiben eigener Skripte.

Datenbank-Schema Alle Tabellen, Felder, Beziehungen und Enum-Werte der Krebs-DB	Forscher-Handbuch Daten erkunden, Projekte anlegen, R-Pakete, Skripte einbringen	IT-Handbuch Installation, Anpassungen, Backup, Update, Troubleshooting
---	--	--

© KIKA 2025 — Hamburgisches Krebsregister 2026/0504-2037

Abb. 3: Integrierte Dokumentationsseite mit drei Einstiegspunkten: Datenbank-Schema, Forscher-Handbuch und IT-Handbuch. Vollständig offline — kein Internetzugang nötig.

Abbildung 4: Über-Seite — Kernfunktionen und Datenschutzhinweis

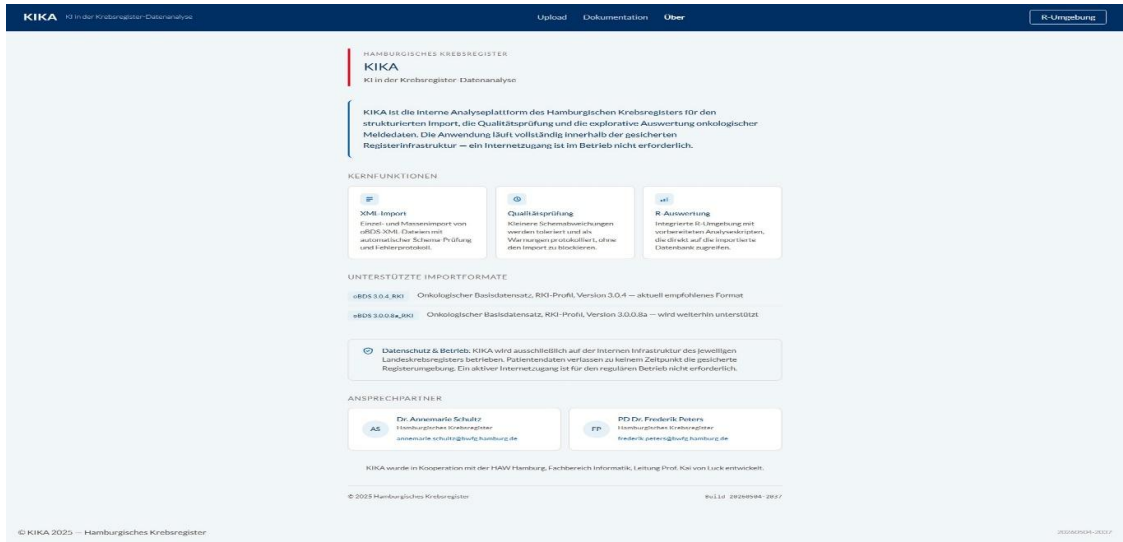


Abb. 4: Kernfunktionen (XML-Import, Qualitätsprüfung, R-Auswertung) und Datenschutzhinweis: Patientendaten verlassen zu keinem Zeitpunkt die gesicherte Registerumgebung.

Abbildung 5: RStudio in KIKA — Kaplan-Meier Gesamtüberleben nach Subtyp

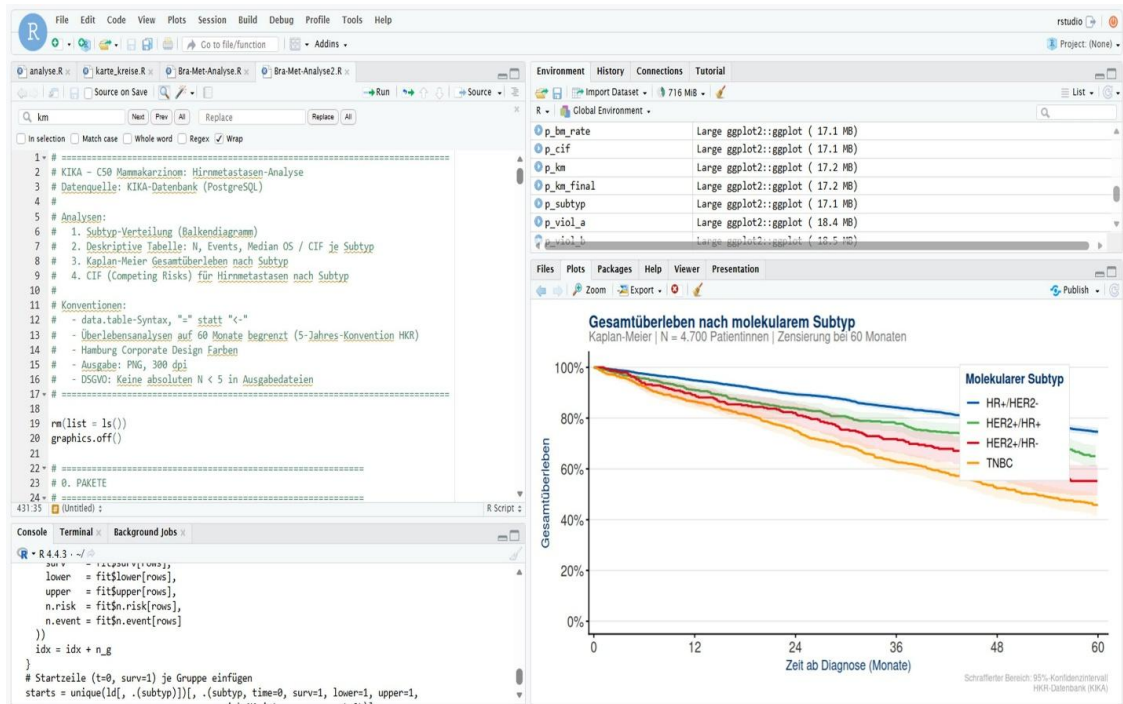


Abb. 5: RStudio-Viewer mit Kaplan-Meier-Kurven nach molekularem Subtyp (HR+/HER2-, HER2+/HR+, HER2+/HR-, TNBC). Synthetische Daten, N = 4.700 Patientinnen mit bekanntem Subtyp.

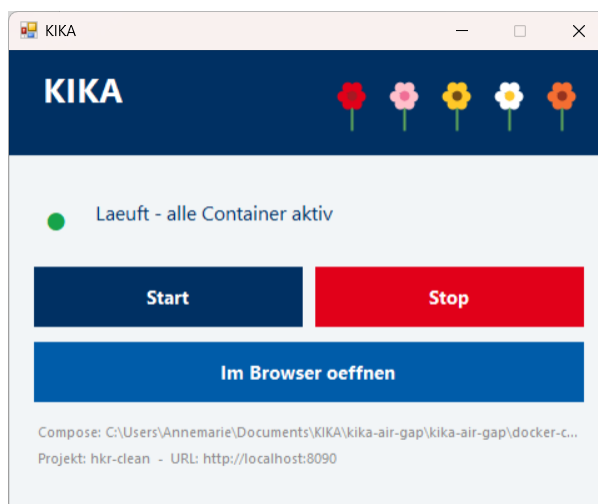
Abbildung 6: Desktop-Launcher — KIKA starten und stoppen ohne Kommandozeile

Abb. 6: KIKA Desktop-Launcher. Start/Stop per Doppelklick, ohne PowerShell oder Docker Desktop direkt bedienen zu müssen.

6.4 Kontakte

Hamburgisches Krebsregister:

Dr. Annemarie Schultz — annemarie.schultz@bwfg.hamburg.de

PD Dr. Frederik Peters — frederik.peters@bwfg.hamburg.de

Universitätsklinikum Hamburg-Eppendorf, Klinik für Gynäkologie:

PD Dr. Elena Laakmann — e.laakmann@uke.de