

# Spezifikation für Krebsregisterdaten

~~2024 V02~~

2025 V01

Technische Dokumentation

# Informationen zum Bericht

## BERICHTSDATEN

---

Spezifikation für Krebsregisterdaten. Technische Dokumentation. ~~2024-V02~~2025 V01

Datum der Veröffentlichung ~~15. November 2023~~1. Juli 2024

## AUFTRAGSDATEN

---

Auftraggeber                      Gemeinsamer Bundesausschuss (G-BA)

**Hinweis:**

Aus Gründen der leichteren Lesbarkeit wird im Folgenden auf eine geschlechtsspezifische Differenzierung verzichtet. Entsprechende Begriffe gelten im Sinne der Gleichbehandlung für beide Geschlechter.

# Inhaltsverzeichnis

Tabellenverzeichnis .....	7
Abbildungsverzeichnis .....	10
Abkürzungsverzeichnis .....	12
Glossar .....	13
1 — Einleitung .....	14
1.1 — Zielsetzung der Technischen Dokumentation .....	14
1.2 — Spezifikationsbegriff .....	14
1.3 — Gültigkeit von Spezifikationspaketen .....	15
2 — Prozesse zur Datenselektion und zur Datenübermittlung .....	17
2.1 — Datenselektion und Filterbeschreibung .....	18
2.1.1 — Selektionsstufen .....	20
2.1.2 — Filterkriterien des QS-Filters .....	22
2.1.3 — Externe Parameter des QS-Filters .....	24
2.1.4 — Ergebnis der Datenselektion .....	26
2.2 — Datenexport .....	27
2.2.1 — Exportdateien .....	29
2.2.2 — XML-Struktur .....	31
2.2.3 — Segmentierung .....	31
2.2.4 — Nullmeldung .....	31
2.3 — Datenprüfung und Regeln .....	32
2.4 — Lieferzeiträume und Datenübermittlung .....	35
2.4.1 — Lieferzeiträume .....	35
2.4.2 — Vorgaben für die Pseudonymisierung der Krankenversichertennummer .....	36
2.4.3 — Vorgaben für die Dateiverschlüsselung und Übermittlung .....	37
2.4.4 — Dateiverschlüsselung .....	38
2.5 — Rückprotokollierung .....	38
2.5.1 — Empfangsbestätigung der Vertrauensstelle .....	38
2.5.2 — Datenflussprotokoll .....	39
3 — Prozesse bei den Krebsregistern .....	41

4	Prozesse bei der Vertrauensstelle .....	43
5	Prozesse bei der Auswertungsstelle .....	45
6	XML-Schemata .....	46
6.1.1	Kompositionsmodell .....	46
6.1.2	Schnittstellen .....	47
6.1.3	XML-Struktur des Datencontainers (KKR-Schnittstelle) .....	47
6.1.4	XML-Struktur der Exportdaten .....	54
6.1.5	XML-Struktur der Rückprotokollierung (Datenflussprotokoll) .....	55
6.1.6	Versionierung der XML-Schemata .....	62
6.2	XML-Verschlüsselung .....	62
6.3	Die Rückprotokollierung .....	63
6.3.1	Empfangsbestätigung .....	63
6.3.2	Datenflussprotokoll .....	64
6.3.3	Fehlermeldungen .....	64
6.3.4	Prüfergebnisse und Ergebnisprotokollierung .....	66
7	Releaseplanung .....	70
7.1	Termine für die Releaseplanung .....	70
7.2	Fehler- und Changemanagement .....	71
	Impressum .....	73
	Tabellenverzeichnis .....	7
	Abbildungsverzeichnis .....	10
	Abkürzungsverzeichnis .....	12
	Glossar .....	13
1	Einleitung .....	14
1.1	Zielsetzung der Technischen Dokumentation .....	14
1.2	Spezifikationsbegriff .....	14
1.3	Gültigkeit von Spezifikationspaketen .....	15
2	Prozesse zur Datenselektion und zur Datenübermittlung .....	17
2.1	Datenselektion und Filterbeschreibung .....	18
2.1.1	Selektionsstufen .....	20
2.1.2	Filterkriterien des QS-Filters .....	22

2.1.3	Externe Parameter des QS-Filters.....	24
2.1.4	Ergebnis der Datenselektion .....	26
2.2	Datenexport.....	27
2.2.1	Exportdateien.....	29
2.2.2	XML-Struktur.....	31
2.2.3	Segmentierung.....	31
2.2.4	Nullmeldung .....	31
2.3	Datenprüfung und Regeln .....	32
2.4	Lieferzeiträume und Datenübermittlung.....	35
2.4.1	Lieferzeiträume.....	35
2.4.2	Vorgaben für die Pseudonymisierung der Krankenversichertennummer.....	36
2.4.3	Vorgaben für die Dateiverschlüsselung und Übermittlung .....	37
2.4.4	Dateiverschlüsselung .....	38
2.5	Rückprotokollierung .....	38
2.5.1	Empfangsbestätigung der Vertrauensstelle.....	38
2.5.2	Datenflussprotokoll .....	39
3	Prozesse bei den Krebsregistern.....	41
4	Prozesse bei der Vertrauensstelle .....	43
5	Prozesse bei der Auswertungsstelle .....	45
6	XML-Schemata .....	46
6.1.1	Kompositionsmodell.....	46
6.1.2	Schnittstellen .....	47
6.1.3	XML-Struktur des Datencontainers (KKR-Schnittstelle) .....	47
6.1.4	XML-Struktur der Exportdaten.....	54
6.1.5	XML-Struktur der Rückprotokollierung (Datenflussprotokoll) .....	55
6.1.6	Umgang mit Umlaut-Domains und E-Mail-Adressen .....	61
6.1.7	Versionierung der XML-Schemata .....	62
6.2	XML-Verschlüsselung .....	62
6.3	Die Rückprotokollierung.....	63
6.3.1	Empfangsbestätigung .....	63
6.3.2	Datenflussprotokoll .....	64

6.3.3 Fehlermeldungen ..... 64

6.3.4 Prüfergebnisse und Ergebnisprotokollierung ..... 66

7 Releaseplanung ..... 70

7.1 Termine für die Releaseplanung ..... 70

7.2 Fehler- und Changemanagement ..... 71

Impressum ..... 73

# Tabellenverzeichnis

Tabelle 1: Komponenten für das Spezifikationspaket.....	15
Tabelle 2: Struktur der Tabelle <code>Version</code> .....	16
Tabelle 3: Tabellen in der Spezifikationsdatenbank für die Datenselektion.....	18
Tabelle 4: Ausschnitt aus der Tabelle <code>ModulSpezifikationsjahr</code> .....	22
Tabelle 5: Zu den Operatoren des Pseudocodes analoge SQL-Operatoren.....	23
Tabelle 6: Struktur der Tabelle <code>FilterListeWert</code> .....	25
Tabelle 7: Inhalt der Tabelle <code>Funktion</code> .....	25
Tabelle 8: Struktur der Tabelle <code>Wertebereich</code> .....	26
Tabelle 9: Tabellen in der Spezifikationsdatenbank für den Datenexport.....	28
Tabelle 10: Struktur der Tabelle <code>Schlüsselwert</code> .....	29
Tabelle 11: Datenkategorien im Datenfluss.....	30
Tabelle 12: Liste der IK-Nummern der Krebsregister für die Exportdatei.....	30
Tabelle 13: Tabellen in der Spezifikationsdatenbank für die Datenprüfung und Regeln.....	32
Tabelle 14: Struktur der Abfrage <code>FeldDefinition</code> .....	32
Tabelle 15: Struktur der Abfrage <code>PruefschrittPruefung</code> .....	34
Tabelle 16: Status des Dokuments.....	39
Tabelle 17: Status des Datensatzes.....	39
Tabelle 18: XML-Schemata für die VST.....	43
Tabelle 19: Verwendbare Schemata, Ablageort und Verwendungszweck.....	47
Tabelle 20: Symbole in den XML-Schema-Diagrammen.....	47
Tabelle 21: Attribute: <code>root</code> -Element.....	49
Tabelle 22: Kindelemente: <code>document</code> -Element.....	50
Tabelle 23: Attribute: <code>body</code> -Element.....	52
Tabelle 24: Attribute: <code>sent</code> -Element.....	54
Tabelle 25: Attribute: <code>patientenstammdaten</code> -Element.....	55
Tabelle 26: Attribute: <code>protocol</code> -Element.....	56
Tabelle 27: Attribute: <code>validation_provider</code> -Element.....	57
Tabelle 28: Attribute: <code>validation_item</code> -Element.....	58
Tabelle 29: Attribute: <code>status_document</code> -Element.....	59
Tabelle 30: Attribute: <code>Status</code> -Element.....	60
Tabelle 31: Attribute: <code>Error</code> -Element.....	60
Tabelle 32: Attribute: <code>Error</code> -Element.....	61
Tabelle 33: Rückmeldungen der Vertrauensstelle.....	64
Tabelle 34: Aufbau und Attribute des Elements <code>&lt;error&gt;</code> .....	64
Tabelle 35: Kind-Elemente des <code>&lt;error&gt;</code> -Elements.....	65
Tabelle 36: Verwendete XPath-Syntax.....	66

Tabelle 37: Meilensteine der Releaseplanung der Spezifikation 2024 für das erste Jahr .....	70
Tabelle 38: Meilensteine der Releaseplanung ab der Spezifikation 2025.....	71
Tabelle 1: Komponenten für das Spezifikationspaket.....	15
Tabelle 2: Struktur der Tabelle <code>Version</code> .....	16
Tabelle 3: Tabellen in der Spezifikationsdatenbank für die Datenselektion.....	18
Tabelle 4: Ausschnitt aus der Tabelle <code>ModulSpezifikationsjahr</code> .....	22
Tabelle 5: Zu den Operatoren des Pseudocodes analoge SQL-Operatoren .....	23
Tabelle 6: Struktur der Tabelle <code>FilterListeWert</code> .....	25
Tabelle 7: Inhalt der Tabelle <code>Funktion</code> .....	25
Tabelle 8: Struktur der Tabelle <code>wertebereich</code> .....	26
Tabelle 9: Tabellen in der Spezifikationsdatenbank für den Datenexport .....	28
Tabelle 10: Struktur der Tabelle <code>SchlüsselWert</code> .....	29
Tabelle 11: Datenkategorien im Datenfluss.....	30
Tabelle 12: Liste der IK-Nummern der Krebsregister für die Exportdatei .....	30
Tabelle 13: Tabellen in der Spezifikationsdatenbank für die Datenprüfung und Regeln .....	32
Tabelle 14: Struktur der Abfrage <code>FeldDefinition</code> .....	32
Tabelle 15: Struktur der Abfrage <code>PruefschrittPruefung</code> .....	34
Tabelle 16: Status des Dokuments .....	39
Tabelle 17: Status des Datensatzes.....	39
Tabelle 18: XML-Schemata für die VST .....	43
Tabelle 19: Verwendbare Schemata, Ablageort und Verwendungszweck.....	47
Tabelle 20: Symbole in den XML-Schema-Diagrammen .....	47
Tabelle 21: Attribute: root-Element.....	49
Tabelle 22: Kindelemente: document-Element .....	50
Tabelle 23: Attribute: body-Element.....	52
Tabelle 24: Attribute: sent-Element.....	54
Tabelle 25: Attribute: patientenstammdaten--Element.....	55
Tabelle 26: Attribute: protocol-Element.....	56
Tabelle 27: Attribute: validation_provider-Element.....	57
Tabelle 28: Attribute: validation_item-Element.....	58
Tabelle 29: Attribute: status_document-Element .....	59
Tabelle 30: Attribute: Status-Element.....	60
Tabelle 31: Attribute Error-Element.....	60
Tabelle 32: Attribute: Error-Element.....	61
Tabelle 33: Rückmeldungen der Vertrauensstelle.....	64
Tabelle 34: Aufbau und Attribute des Elements <code>&lt;error&gt;</code> .....	64
Tabelle 35: Kind-Elemente des <code>&lt;error&gt;</code> -Elements .....	65
Tabelle 36: Verwendete XPath-Syntax .....	66
Tabelle 37: Meilensteine der Releaseplanung ab der Spezifikation 2025.....	71





# Abbildungsverzeichnis

Abbildung 1: Workflow anhand der Spezifikationsdatenbank.....	17
Abbildung 2: Tabelle DatenfeldAttribut .....	19
Abbildung 3: Beispiel QS-Filter- Ausgangssituation .....	21
Abbildung 4: Beispiel QS-Filter- Patientenfilter .....	22
Abbildung 5: Mapping.....	28
Abbildung 6: Umgang mit Widersprüchen gemäß § 8 oKFE-RL .....	44
Abbildung 7: root-Element und Kind-Elemente header und body .....	49
Abbildung 8: Header-Element und Kind-Element document .....	50
Abbildung 9: body-Element und Kind-Elemente case und statistic .....	52
Abbildung 10: case-Element und seine Kind-Elemente patient und qs_data .....	53
Abbildung 11: Element statistic und Kind-Element sent .....	53
Abbildung 12: Element sent .....	54
Abbildung 13: Element patientenstammdaten .....	54
Abbildung 14: Schema des Datenflussprotokolls „response_BAS_KKR.xsd“ .....	56
Abbildung 15: Element protocol und Kind-Elemente validation_provider und status_document .....	56
Abbildung 16: Element validation_provider und Kind-Element validation_item ....	57
Abbildung 17: Element validation_item und Kind-Element status .....	58
Abbildung 18: Element status_document .....	59
Abbildung 19: Element status und Kind-Element error .....	60
Abbildung 20: Element error und Kind-Elemente rule_id, rule_type, list und error_message .....	60
Abbildung 21: Element protocol und Kind-Elemente validation_item und status_case .....	61
Abbildung 22: Beispiel einer Empfangsbestätigung der VST an ein Krebsregister .....	63
Abbildung 23: Beispiel eines <error>-Elements .....	65
Abbildung 24: Beziehungen zwischen <validation_item> im header und <validation_item> im body über die id .....	67
Abbildung 1: Workflow anhand der Spezifikationsdatenbank.....	17
Abbildung 2: Tabelle DatenfeldAttribut .....	19
Abbildung 3: Beispiel QS-Filter- Ausgangssituation .....	21
Abbildung 4: Beispiel QS-Filter - Patientenfilter .....	22
Abbildung 5: Mapping.....	28
Abbildung 6: Umgang mit Widersprüchen gemäß § 8 oKFE-RL .....	44
Abbildung 7: root-Element und Kind-Elemente header und body .....	49
Abbildung 8: Header-Element und Kind-Element document .....	50

Abbildung 9: <code>body</code> -Element und Kind-Elemente <code>case</code> und <code>statistic</code> .....	52
Abbildung 10: <code>case</code> -Element und seine Kind-Elemente <code>patient</code> und <code>qs_data</code> .....	53
Abbildung 11: Element <code>statistic</code> und Kind-Element <code>sent</code> .....	53
Abbildung 12: Element <code>sent</code> .....	54
Abbildung 13: Element <code>patientenstammdaten</code> .....	54
Abbildung 14: Schema des Datenflussprotokolls „ <code>response_BAS_KKR.xsd</code> “ .....	56
Abbildung 15: Element <code>protocol</code> und Kind-Elemente <code>validation_provider</code> und <code>status_document</code> .....	56
Abbildung 16: Element <code>validation_provider</code> und Kind-Element <code>validation_item</code> .....	57
Abbildung 17: Element <code>validation_item</code> und Kind-Element <code>status</code> .....	58
Abbildung 18: Element <code>status_document</code> .....	59
Abbildung 19: Element <code>status</code> und Kind-Element <code>error</code> .....	60
Abbildung 20: Element <code>error</code> und Kind-Elemente <code>rule_id</code> , <code>rule_type</code> , <code>list</code> und <code>error_message</code> .....	60
Abbildung 21: Element <code>protocol</code> und Kind-Elemente <code>validation_item</code> und <code>status_case</code> .....	61
Abbildung 22: Beispiel einer Empfangsbestätigung der VST an ein Krebsregister .....	63
Abbildung 23: Beispiel eines <code>&lt;error&gt;</code> -Elements .....	65
Abbildung 24: Beziehungen zwischen <code>&lt;validation_item&gt;</code> im header und <code>&lt;validation_item&gt;</code> im <code>body</code> über die <code>id</code> .....	67

## Abkürzungsverzeichnis

Abkürzung	Bedeutung
AES	Advanced Encryption Standard
BAS	(Bundes-)Auswertungsstelle
DFP	Datenflussprotokoll
G-BA	Gemeinsamer Bundesausschuss
GFL	Gesundheitsforen Leipzig GmbH
GUID	Globally Unique Identifier
ICD	International Statistical Classification of Diseases and Related Health Problems
IQTIG	Institut für Qualitätssicherung und Transparenz im Gesundheitswesen
KFRG	Krebsfrüherkennungs- und registergesetz
KKR	Klinisches Krebsregister
LID	Leistungserbringeridentifizierende Daten
oBDS	Einheitlicher onkologischer Basisdatensatz
oKFE-RL	Richtlinie für organisierte Krebsfrüherkennungsprogramme
OPS	Operationen- und Prozedurenschlüssel
PID	Patientenidentifizierende Daten
PKCS	Public Key Cryptography Standard
PKI	Public-Key-Infrastruktur
QS	Qualitätssicherung
RL	Richtlinie
SECON	Security-Schnittstelle für den Datenaustausch im Gesundheits- und Sozialwesen
SFTP	SSH File Transfer Protocol
SGB	Sozialgesetzbuch
VST	Vertrauensstelle
W3C	World Wide Web Consortium
XML	Extensible Markup Language
XSD	Extensible Schema Definition
XSLT	Extensible Stylesheet Language Transformation

## Glossar

Begriff	Bedeutung
Spezifikationsjahr	Ein Spezifikationsjahr bezieht sich auf das Jahr, in dem die Spezifikation gültig ist und Daten auf Grundlage dieser Spezifikation selektiert und exportiert werden. Das Spezifikationsjahr deckt sich mit einem Kalenderjahr.
Gemeinsamer Bundesausschuss (G-BA)	Oberstes Beschlussgremium der gemeinsamen Selbstverwaltung der Ärzte, Zahnärzte, Psychotherapeuten, Krankenhäuser und Krankenkassen in Deutschland. Er bestimmt in Form von Richtlinien den Leistungskatalog der Gesetzlichen Krankenversicherung (GKV) für mehr als 70 Millionen Versicherte und legt damit fest, welche Leistungen der medizinischen Versorgung von der GKV erstattet werden. Darüber hinaus beschließt der G-BA Maßnahmen der Qualitätssicherung für den ambulanten und stationären Bereich des Gesundheitswesens.
ICD-GM	engl.: International Classification of Diseases – German Modification. Auf deutsche Verhältnisse angepasste Version der ICD, die vom Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) erstellt wird.
Indexleistung	Medizinische Leistung, die den Startpunkt für ein Verfahren (die Auslösung) genau definiert (im Regelfall OPS-Kodes).
Matching	Eindeutiger Verweis von Ausgangs- auf Zieldaten (1:1-Beziehung)
Operationen- und Prozedurenschlüssel (OPS)	Kodierungssystem für medizinische Leistungen bzw. verbindlicher Abrechnungsschlüssel zwischen medizinischen Leistungserbringern und Kostenträgern.
Patientenidentifizierende Daten (PID)	Daten, die eindeutig einen bestimmten Versicherten identifizieren (z. B. Versichertennummer).
Plausibilitätsprüfung	Statistisches Verfahren, mit dem die Dokumentationsdaten auf erlaubte und/oder fehlende Werte, Widerspruchsfreiheit, Werteverteilung und bekannte Korrelationen geprüft werden.
Pseudocode	Programmcode, der das zugrundeliegende Prinzip eines Algorithmus beschreibt, selbst aber nicht lauffähig ist. Er dient zur Veranschaulichung, unabhängig von der konkret zu verwendenden Programmiersprache.
QS-Filter	Der QS-Filter ist ein Algorithmus, der auf Grundlage festgelegter Kriterien die für die Qualitätssicherung zu dokumentierenden Patienten und deren Daten „filtert“. Die Kriterien hierzu werden in einer Spezifikation definiert.
Selektion	Auswahl bestimmter Daten aus einem Datenpool anhand zuvor definierter Kriterien.
XSLT	Extensible Stylesheet Language Transformation. Programmiersprache zur Transformation von XML-Dokumenten in andere XML-Dokumente oder andere Dokumentformate wie HTML. Im QS-Kontext kann es auch für Datenprüfung und Protokollerstellung verwendet werden.

# 1 Einleitung

## 1.1 Zielsetzung der Technischen Dokumentation

Nach § 65c Absatz 8 SGB V sind die Krebsregister befugt und verpflichtet, einen zweckgebundenen Ausschnitt aus den von ihnen erhobenen Daten für die gesetzliche Qualitätssicherung (QS), also die Durchführung von Verfahren gemäß §§ 136 ff., und für die Methodenbewertung gemäß §§ 135, 137c und 137e SGB V zur Verfügung zu stellen. Aus datenschutzrechtlichen Gründen und den allgemeinen Grundsätzen der Datensparsamkeit und der Datenvermeidung (§ 3a BDSG) folgend, müssen die für die Zwecke der Qualitätssicherung und der Methodenbewertung erforderlichen Daten und ihre jeweiligen Empfänger spezifiziert und in Richtlinien des G-BA festgelegt werden.

Das vorliegende Dokument ist Teil der Allgemeinen Spezifikation zur Nutzung von Daten gemäß § 65c Absatz 8 SGB V. Es beschreibt die Auswahl von Daten, die Erzeugung von XML-Exportdateien, deren Aufbereitung für eine Übertragung und die Interpretation des Rückprotokolls.

Die XML-Exportdatei verwendet die Daten, die nach den Vorgaben des G-BA aus den Leistungserbringerdaten bei den klinischen Krebsregistern für die Methodenbewertung an die Gesundheitsforen Leipzig GmbH (GFL) als die durch den G-BA beauftragte Auswertungsstelle zu übermitteln sind.

Im vorliegenden Dokument werden die Begriffe QS-Daten und QS-Datei stellvertretend für Inhalte im Kontext der vorgenannten Methodenbewertung (konkret der Programmbeurteilung gem. oKFE-RL) verwendet.

## 1.2 Spezifikationsbegriff

Die Spezifikation ist die Gesamtheit aller Vorgaben für die Selektion, den Export und die Übermittlung der Daten für ein spezifisches Krebsfrüherkennungsprogramm nach oKFE-RL. Eine Spezifikation bezieht sich auf das Jahr, in dem die Spezifikation gültig ist und Daten auf Grundlage dieser Spezifikation selektiert und exportiert werden (sog. Spezifikationsjahr). Das Spezifikationsjahr deckt sich mit einem Kalenderjahr.

Die Spezifikation wird als Spezifikationspaket auf den Webseiten des IQTIG zum Download zur Verfügung gestellt. Ein Spezifikationspaket besteht aus einzelnen Komponenten. Komponenten von IQTIG-Spezifikationen sind im Regelfall die Spezifikationsdatenbank, die Technische Dokumentation und XML-Schemata. Die Komponenten der Spezifikation für Krebsregisterdaten sind in Tabelle 1 dargestellt.

Tabelle 1: Komponenten für das Spezifikationspaket

Komponente	Inhalt
Spezifikationsdatenbank	Definitionen zur Datenselektion, Datenexport, Datenprüfung, Lieferzeiträume und Rückprotokollierung
Technische Dokumentation	Dokumentation der Inhalte der Datenbank und Beschreibung der Selektion von Daten, der Erzeugung von XML-Exportdateien, deren Aufbereitung für eine Übertragung und der Interpretation des Rückprotokolls.
XML-Schemata	Definitionen für Exportdateien und Protokolle im Datenfluss
Komponentenübersicht	Jedem Paket liegt eine Auflistung der enthaltenen Komponenten bei.
Übersicht Änderungen	Übersicht über die Änderungen zur vorhergehenden Version. Die Übersicht wird als Text-Datei gepflegt.
Benennungsschema	Jedem Paket liegt ein Dokument bei, welches über das Benennungsschema der Spezifikation informiert. Das Benennungsschema ist für alle Spezifikationen des IQTIG gültig und wird zentral gepflegt.

Die technische Umsetzung der Anforderungen an die Verfahren zur Datenselektion, Datenexport, Datenprüfung, Lieferzeiträume und Rückprotokollierung werden in der Spezifikationsdatenbank definiert. Sie enthält zusammen mit der Technischen Dokumentation alle Informationen, die die Krebsregister für die Selektion und den Export der Krebsregisterdaten brauchen.

Sowohl die Spezifikationspakete als auch die einzelnen Komponenten werden nach einem einheitlichen Muster benannt, das bereits im Namen die relevanten Informationen wie Betriebsart (Regel-/Testbetrieb), Spezifikationssjahr und Versionierung enthält (siehe Komponente Benennungsschema).

Jede Komponente eines Spezifikationspakets kann also unabhängig von den anderen Komponenten aktualisiert werden. Durch die Versionierung sowohl auf der Ebene der Pakete als auch auf der Ebene der Komponenten ist gewährleistet, dass der aktuelle Stand leicht ersichtlich ist. Zudem wird die Kommunikation über die anzuwendenden Bestandteile der Spezifikation erleichtert.

Jedem Paket liegt eine Auflistung der enthaltenen Komponenten und eine Übersicht über die Änderungen zur vorhergehenden Version bei.

1.3 Gültigkeit von Spezifikationspaketen

Die wichtigsten Eigenschaften einer Version sind der Versionsname (Attribut name) und der Gültigkeitszeitraum (Attribute ab und bis). Der Gültigkeitszeitraum einer Version ist in der Regel ein Kalenderjahr (z. B. 01.01.2024 bis 31.12.2024). In der Spezifikationsdatenbank für Krebsregisterdaten sind in der Tabelle version Informationen hinterlegt, welche Klarheit bezüglich der Gültigkeit der jeweiligen Spezifikationsdatenbank bringen.

Versionen können den Status „in Entwicklung“ oder „final“ haben. Diese Zustände werden in der Nachschlagetabelle `VersStatus` verwaltet. Das Attribut `gueltig` zeigt die gültige Version der Datenbank an. Nur eine einzige Version darf als gültig markiert sein. Darüber hinaus verwaltet die Tabelle `Version` die Historie der Versionen: Welche Vorgängerversion vorher gültig war, kann über das Attribut `fkVersion` ermittelt werden. Mit jedem Release erfolgt ein neuer Eintrag in diese Tabelle. Beschreibende Erläuterungen der einzelnen Feldinhalte können der nachstehenden Tabelle 2 entnommen werden.

Tabelle 2: Struktur der Tabelle `Version`

Datenfeld	Beschreibende Erläuterung
<code>id</code>	Eindeutige Nummer
<code>name</code>	Bezeichnung der Spezifikationsdatenbank entsprechend dem Benennungsschema für Spezifikationskomponenten
<code>bezeichnung</code>	Kurzbeschreibung der Spezifikationsdatenbank
<code>ab</code>	Beschreibt den Zeitpunkt, ab wann dieses Release Gültigkeit erlangt (und somit das vorige Release abgelöst wird)
<code>bis</code>	Entspricht der Angabe im Feld <code>ab</code> eines nachfolgenden Releases. Bei neuen Einträgen bleibt diese Angabe zunächst leer und wird erst im Rahmen eines Folge-releases befüllt. Eine Spezifikationsdatenbank bleibt demzufolge gültig bis es durch ein aktuelleres Release zu einem späteren Zeitpunkt abgelöst wird.
<code>pub</code>	Datum der Veröffentlichung. Diese Angabe liegt zeitlich stets mindestens 3 Monate vor der Angabe im Feld <code>ab</code>
<code>fkVorversion</code>	Referenz auf das vormals gültige Release
<code>fkVersStatus</code>	Einordnung der Spezifikationsdatenbank (Entwicklung, final, Service-Release)



## 2 Prozesse zur Datenselektion und zur Datenübermittlung

In diesem Kapitel werden die Spezifikationsdatenbank und die daraus folgenden grundlegenden Prozessschritte beschrieben, welche für **alle** Verfahrensteilnehmer im Rahmen der Datenübermittlung der selektierten Daten an die Auswertungsstelle Gesundheitsforen Leipzig (GFL) notwendig sind. Zusätzliche und verfahrensteilnehmerspezifische Prozesse werden in den darauffolgenden Kapiteln erläutert:

- Kapitel 3 (Prozesse bei den Krebsregistern),
- Kapitel 4 (Prozesse bei der Vertrauensstelle) und
- Kapitel 5 (Prozesse bei den Auswertungsstellen).

### Themenspezifischer Aufbau der Spezifikationsdatenbank

Die Nutzung der aktuellen Spezifikationsdatenbank kann anhand des Spezifikationspakets und der darin enthaltenen Spezifikationskomponente „Komponentenübersicht“ sichergestellt werden. Wie weiter oben beschrieben, bringt ebenfalls die Tabelle `version` in der Spezifikationsdatenbank Klarheit bezüglich der Gültigkeit (siehe Abschnitt 1.3).

Inhaltlich kann die Access-Datenbank in fünf Themenbereiche unterteilt werden (siehe Abbildung 1). Die Unterteilung stellt zugleich einen Workflow von Datenselektion bis zur Datenübermittlung und Rückprotokollierung dar.

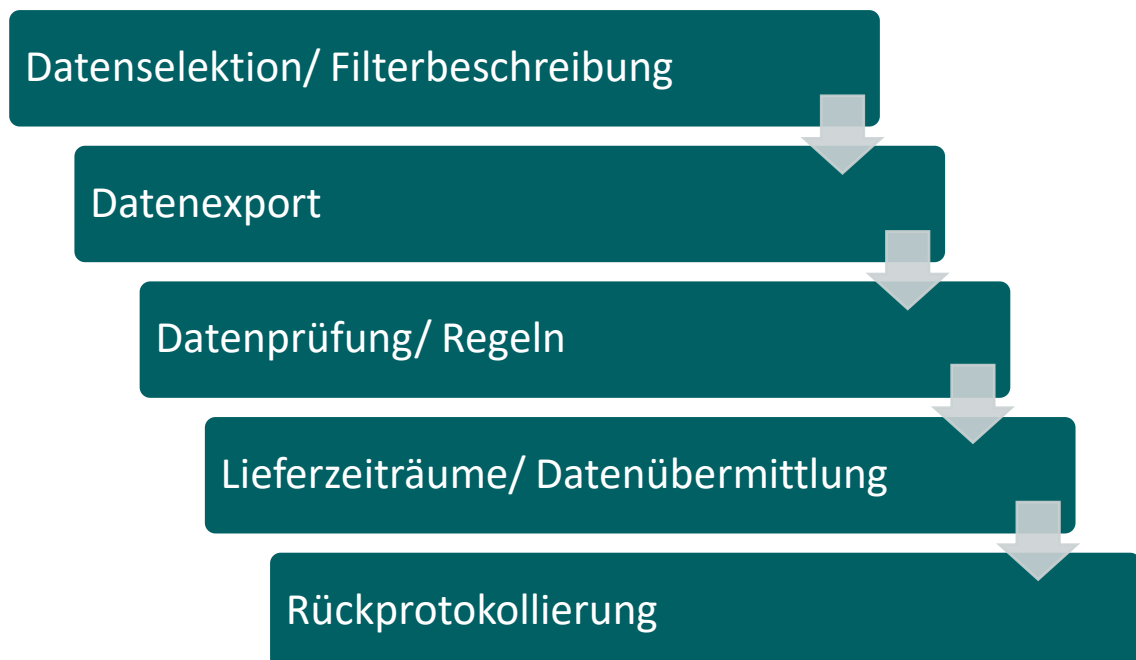


Abbildung 1: Workflow anhand der Spezifikationsdatenbank

## 2.1 Datenselektion und Filterbeschreibung

Bei der Nutzung von Leistungserbringerdaten bei den klinischen Krebsregistern für Zwecke der Methodenbewertung ist die Datensparsamkeit von höchster Bedeutung. Daher gilt grundsätzlich, dass nur Daten der Patienten selektiert und exportiert werden dürfen, die für das jeweilige Verfahren notwendig sind. Die Selektion von Daten geschieht immer für ein konkretes Verfahren (ein konkretes Modul).<sup>1</sup>

Welche Daten durch das klinische Krebsregister selektiert werden, ist anhand festgelegter Filterkriterien spezifiziert. Die Kriterien werden als QS-Filter bezeichnet.

Die Filterkriterien sind in den in Tabelle 3 dargestellten Datenbanktabellen der Spezifikationsdatenbank hinterlegt:

Tabelle 3: Tabellen in der Spezifikationsdatenbank für die Datenselektion

Tabelle in der Spezifikationsdatenbank	Beschreibung
DatenfeldAttribut	Zuordnung der Eingangsdaten zu den Exportattributen
Datenquelle	Datenquelle (Unterteilung nach oBDS und Administrative Daten)
FilterListe	Kodelisten zur Verwendung in Selektionsfiltern
FilterListeModul	Relationentabelle zur Verknüpfung von Filterlisten mit mehreren Modulen
FilterListeWert	Inhalt der Kodelisten aus der Tabelle „FilterListe“
Funktion	Funktionen zur Ableitung von Werten aus Datenfeldern
ModulSpezifikationssjahr	Erweiterte Angaben zu den Modulen je Spezifikationsjahr, Darstellung des QS-Filters als Pseudocode
Wertebereich	Wertebereiche (z. B. Zeiträume/ Zeitfilter) zur Verwendung in Selektionsfiltern

In der Spezifikationsdatenbank wird der sog. Eingangsdatensatz definiert (siehe Abbildung 2). Der Eingangsdatensatz ist in der Tabelle *DatenfeldAttribut* festgelegt. Er beschreibt ein Datenmodell, welches eine eindeutige Kennung und Herkunft für jedes Datenfeld der Spezifikation bereithält. Als Basis für die Eingangskennung dient der einheitliche onkologische Basisdatensatz (oBDS).

<sup>1</sup> Definitionen in der Spezifikationsdatenbank sind modulbezogen. In der Regel entspricht ein Modul einem bestimmten Verfahren.

DatenfeldAttribut												
	id	ber.	fkFunktion	fkDatenquelle	nummer	gruppe	zV	feldbezeichnung	definition	fkElement	fkAttribut	fkMussKann
+	330			oBDS	3.1	Patienten Stammdaten		Krankenversicherungsnummer	Eindeutige GK-Versicherten-Nummer des Patienten nach § 290 SGB V in gehäushter Monat und Jahr der Geburt einer Person.	Patienten-eGK-Hash	V	M
+	331			oBDS	3.10	Patienten Stammdaten		Patienten Geburtsdatum		patientenstammdaten	gebdatum	M
+	332			oBDS	10.1	Residualstatus		Beurteilung des lokalen Residualstatus nach Abschluss der Operation	Lokale Beurteilung der Residualklassifikation nach Resektion, bezieht sich auf das, was	residualstatus	residualstatusop	K
+	333			oBDS	10.2	Residualstatus		Gesamtbeurteilung des Residualstatus	Gesamtbeurteilung der Residualklassifikation der Erkrankung einschließlich etwaiger	residualstatus	residualstatusgesamt	K
+	334			oBDS	8.6	TNM-Klassifikation		TNM c_p_Praefix_T	Gibt an, ob die Klassifikation klinisch oder pathologisch erfolgte.	tnm	c_p_praefix_t	M
+	335			oBDS	8.9	TNM-Klassifikation		TNM T-Kategorie	Ausbreitung des Primärtumors, erfolgt gemäß Tumorentität nach TNM.	tnm	tnm_t-kategorie	M
+	337			oBDS	20.1	Tod		Sterbedatum	Datum des Todes.	tod	sterbedatum	K
+	339			oBDS	20.3	Tod		Todesursache ICD	Todesursache im Sinne des Grundeidens.	todesursache	icd	M
+	340			oBDS	20.2	Tod		Tod tumorbedingt	Krebs-Tod-Relation.	tod	tod_tumorbedingt	K
+	341			oBDS	5.6	Diagnose		Primärtumor Diagnosedatum	Datum, angegeben in Tag, Monat und Jahr, an dem die meldepflichtige Diagnose	diagnose	diagdatum	M
+	342			oBDS	5.1	Diagnose		Primärtumor Tumordiagnose ICD Code	Kodierung einer meldepflichtigen Erkrankung nach der aktuellen ICD-GM Version.	diagnose	icd	M
+	343			oBDS	13.2	OP		OP Datum	Datum der OP.	op	datum	M
+	344			oBDS	8.11	TNM-Klassifikation		TNM N-Kategorie	Ausbreitung von regionären Lymphknotenmetastasen, erfolgt gemäß Fehlen oder Vorhandensein von	tnm	tnm_n-kategorie	M
+	345			oBDS	8.12	TNM-Klassifikation		TNM M-Kategorie	Fernmetastasen, gemäß Tumorentität nach	tnm	tnm_m-kategorie	M
+	346			oBDS	13.3	OP		OPS	Gibt an, welche Operation durchgeführt wurde.	op	ops	M
+	348			oBDS	5.4	Diagnose		Primärtumor Topographie ICD-O	Bezeichnung der Topographie einer Erkrankung nach der aktuellen ICD-O Version.	diagnose	icd_o	M

Eingangsdatensatz

Abbildung 2: Tabelle DatenfeldAttribut

Das IQTIG verweist in der jeweiligen Eingangskennung (Referenzierung auf den oBDS, Eingangsdatensatz) der Datenfelder stets auf die aktuelle Version des oBDS, welche zum Zeitpunkt der Veröffentlichung des Spezifikationspakets vorliegt. Die klinischen Krebsregister können so den Bezug aus den referenzierten Datenfeldern der Spezifikationsdatenbanktabelle DatenfeldAttribut zu ihren eigenen Datenbeständen herstellen.



Hinweis

Datenexport aus dem wohnortbezogenen klinischen Krebsregister

Das für die Datenselektion und Datenübermittlung verantwortliche klinische Krebsregister ist das wohnortbezogene klinische Krebsregister zum Zeitpunkt der Diagnosestellung. Die klinischen Krebsregister stellen über den registerübergreifenden Datenabgleich (RÜD) mit den behandlungsortbezogenen klinischen Krebsregistern sicher, dass vor dem Datenexport alle Informationen im wohnortbezogenen Krebsregister vorliegen. Auf diesem Weg sollen doppelte Datenexporte aus den wohnort- und behandlungsortbezogenen Krebsregistern vermieden werden. Gehen dennoch Doppelmeldungen bei der Auswertungsstelle (GFL) ein, werden diese die ersten zwei Jahre nicht abgelehnt, sondern verarbeitet. Es erfolgt anschließend eine Evaluation über die Doppelmeldungen und ggf. eine Anpassung der technischen Spezifikation.

Zusätzlich zur Eingangskennung wird in der Tabelle DatenfeldAttribut das Format des Datenfeldes für den Export spezifiziert, das für die weitere Verarbeitung erforderlich ist (siehe Spalten fkElement, fkAttribut, fkMussKann). Die in den klinischen Krebsregistern vorliegenden und zum Export vorgesehenen Daten müssen in das durch das IQTIG spezifizierte Exportformat konvertiert werden. Weitere Informationen zum Datenexport sind in Abschnitt 2.2 hinterlegt.

### 2.1.1 Selektionsstufen

Die Selektionskriterien werden durch den G-BA verfahrensspezifisch festgelegt. Es handelt sich in der Regel um eine Liste von Codes (z. B. ICD-Kodes), die als Ein- und Ausschlusskriterien auf spezifizierten Datenfeldern zur Anwendung kommen. Die verfahrensspezifisch festgelegten Selektionskriterien sind in der Spezifikation als jeweils ein Modul festgelegt. Es kann ein bis zu zweistufiger QS-Filter verwendet werden.

#### 1. Patientenfilter

In der ersten Filterstufe wird zunächst über den Patientenfilter sichergestellt, dass nur Patienten eingeschlossen werden, die in einem genau definierten Zeitraum mindestens eine bestimmte Diagnose erhalten haben. Über diese Aufgreifkriterien wird die Relevanz des Falles für ein bestimmtes Modul begründet.

Nur diejenigen Patienten, die in der ersten Filterstufe erfasst wurden, werden für die zweite Filterstufe (Leistungs- und Medikationsfilter) berücksichtigt (insofern eine zweite Selektionsstufe festgelegt wurde).

#### 2. Leistungs- und Medikationsfilter

Derzeit sind keine Verfahren in der Spezifikation enthalten, für die eine zweite Selektionsstufe erforderlich ist. Dieser Abschnitt wird entsprechend überarbeitet, insofern ein Verfahren mit zwei Selektionsstufen spezifiziert wird.

#### Beispiel

Drei weibliche und drei männliche Patienten (senkrechte Achse, siehe Abbildung 3) besuchen diverse Ärzte und erhalten entsprechende medizinische Leistungen (Kästen):

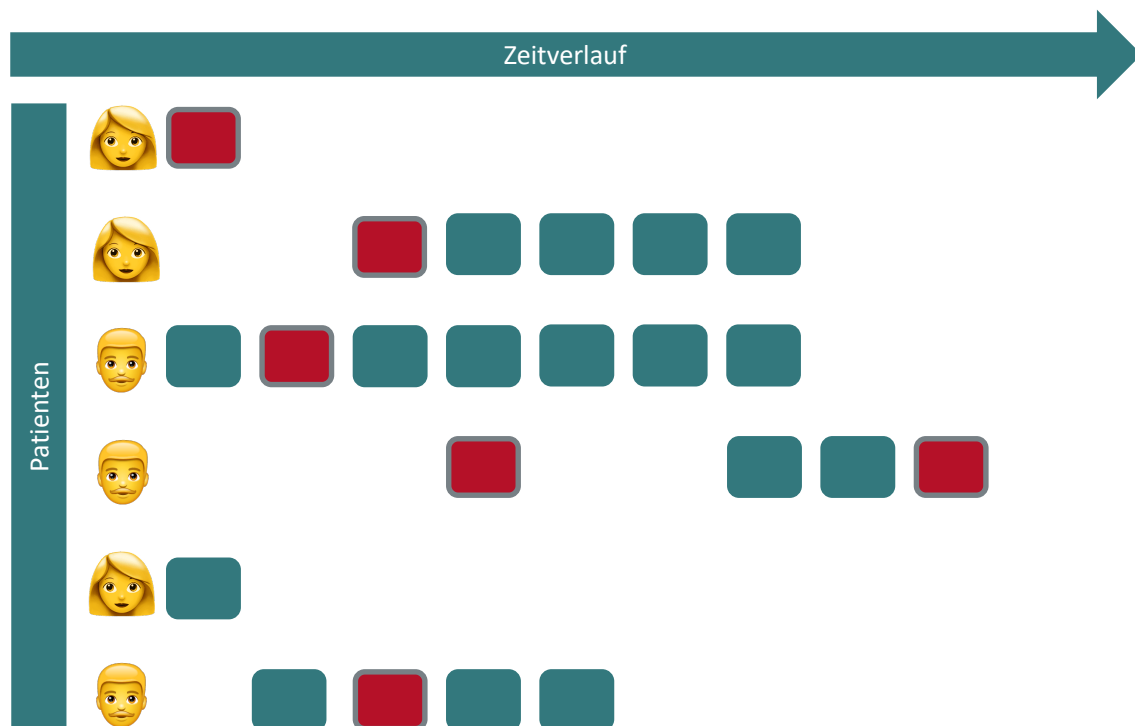


Abbildung 3: Beispiel QS-Filter- Ausgangssituation

Durch den Patientenfilter werden ausschließlich Patienten ausgewählt, die in einem bestimmten Zeitraum eine das Modul betreffende Indexleistung (z. B. Diagnose, rote Kästen) erhalten haben. Diese ist in Abbildung 4 mit einem „X“ gekennzeichnet. Nur vier Patienten entsprechen den Kriterien des Patientenfilters:

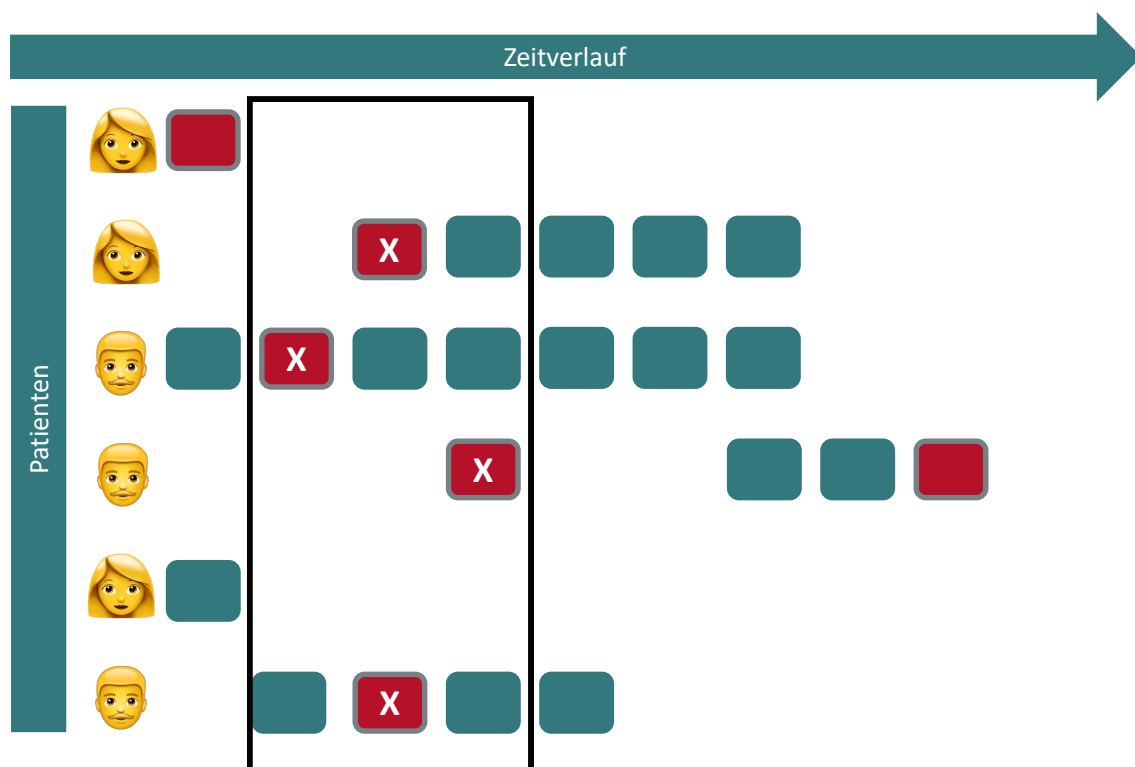


Abbildung 4: Beispiel QS-Filter – Patientenfilter

Die roten Kästen, die nicht im schwarzen Rahmen in Abbildung 4 enthaltenen sind, entsprechen zwar der geforderten Indexleistung, werden aber aufgrund des Indexfilters (Zeitraum des Patientenfilters) nicht in die Auswahl eingeschlossen.

### 2.1.2 Filterkriterien des QS-Filters

Die in Abschnitt 2.1.1 beschriebenen Filterstufen bilden gemeinsam den QS-Filter und sind für jedes Verfahren anhand eines jeweiligen Moduls in der Spezifikationsdatenbank abgebildet.

Die Definition des QS-Filters ist als Pseudocode und als Beschreibung in der Tabelle `ModulSpezifikationsjahr` hinterlegt. Darin befinden sich folgende Felder:

Tabelle 4: Ausschnitt aus der Tabelle `ModulSpezifikationsjahr`

Datenfeld	Bedeutung	Verwendung
<code>fkModul</code>	Name des Moduls	Eingrenzung auf das Exportmodul
<code>spezifikationsjahr</code>	Spezifikationsjahr	Eingrenzung auf das Spezifikationsjahr, in dem die Daten selektiert und exportiert werden
<code>patientenFilter</code>	Pseudocode des Patientenfilters	Logischer Aufbau des QS-Filters mit dem die Patientenauswahl stattfindet
<code>patientenFilterText</code>	Textbeschreibung des Patientenfilters	Textuelle Erläuterung zu dem Pseudocode des QS-Filters
<code>leistungsMedikationsFilter</code>	Pseudocode des Leistungs- und Medikationsfilters	Derzeit keine Verwendung
<code>leistungsMedikationsFilterText</code>	Textbeschreibung des Leistungs- und Medikationsfilters	Derzeit keine Verwendung

Da die genaue Anwendung des QS-Filters auf die Datenbestände der Krebsregister durch diese individuell umgesetzt werden muss, sind die Filter in Form eines Pseudocodes angegeben, der im Folgenden beschrieben wird.

#### Pseudocode der QS-Filter

Der Patientenfilter liegt pro Modul vor und wird als Pseudocode dargestellt.

Die Syntax des Pseudocodes lehnt sich an die logischen Ausdrücke in bekannten Programmier- und Skriptsprachen an. Jedoch haben die Operatoren deutsche Namen, z. B. `UND` statt `AND` oder `ODER` statt `OR` (weitere Beispiele sind in der Tabelle 5 aufgelistet).

Der Pseudocode verwendet Referenzen auf Datenfelder sowie externe Parameter (siehe Abschnitt 2.1.3) und verknüpft diese mittels Operatoren. Jede Bedingung (Filterkriterium) hat folgenden Aufbau:

<Operand> <Operator> <Operand>

- Der linke Operand ist das Referenzfeld.
- Als Operator kann jeder der Operatoren aus Tabelle 5 verwendet werden.
- Die auf der rechten Seite erlaubten Operanden sind:
  - Literale (z. B. Ganzzahlen, Zeichenketten)
  - Kodelisten, in denen auch die Codes eines Schlüssels referenziert werden können
  - ICD-Listen (Alle Codes sind terminal)

Tabelle 5: Zu den Operatoren des Pseudocodes analoge SQL-Operatoren

Präzedenz*	Operator	SQL-Analogie	Erläuterung
0	IN	IN	Die Variable und die Feldelemente müssen gleichen Typs sein. Prüft, ob sich das genannte Element der linken Seite in der Liste auf der rechten Seite befindet.
0	EINSIN	IN	Mengenoperator. Prüft, ob sich mindestens eines der Elemente der linken Seite in der Liste auf der rechten Seite befindet.
0	INNERHALB	BETWEEN <Minimum> AND <Maximum>	Wertebereich. Vergleich eines Wertes bzw. eines Datums gegen eine untere und eine obere Grenze, wobei die Grenzen eingeschlossen sind.
1	<	<	Vergleichsoperator „kleiner“
1	>	>	Vergleichsoperator „größer“
1	<=	<=	Vergleichsoperator „kleiner gleich“
1	>=	>=	Vergleichsoperator „größer gleich“
2	=	=	Vergleichsoperator „gleich“
2	<>	<>	Vergleichsoperator „ungleich“
3	UND	UND	Logische UND-Operation: Beide Seiten müssen erfüllt sein.

Präzedenz*	Operator	SQL-Analogie	Erläuterung
4	ODER	ODER	Logische ODER-Operation: Eine oder beide Seiten müssen erfüllt sein.

### Beispiel

In einem Pseudocode wird der Mengenoperator EINSIN verwendet:

```
oBDS.5.1.Diagnose.Primärtumor Tumordiagnose ICD Code@icd EINSIN Filterliste.ZERVIXK_ICD_INDEX
```

Die linke Seite des Operators repräsentiert die konkreten Daten im klinischen Krebsregister. Für eine konkrete Patientin soll geprüft werden, ob eine ihrer Tumordiagnosen den Vorgaben der Spezifikation zum Einschluss der Patientin in den Datenexport entspricht. Die Einschlusskriterien entsprechen der rechten Seite des Operators EINSIN in Form einer Liste von ICD-Kodes.

Die ICD-Kodeliste „ZERVIXK\_ICD\_INDEX“ ist in der Spezifikationsdatenbank in der Tabelle `FilterListeWert` mit folgenden ICD-Kodes hinterlegt: C53.0, C53.8, C53.9, D06.0, D06.1, D06.9.

Befindet sich das Element der linken Seite (Operand: icd) in der Liste auf der rechten Seite (Operand: Filterliste.ZERVIXK\_ICD\_INDEX), wird die Patientin für den Datenexport vorgesehen.

### String-Vergleich bei Leistungen, Diagnosen und Prozeduren:

Die Mengenoperatoren basieren für Leistungen, Diagnosen oder Prozeduren auf einem Vergleich von Zeichenketten (String-Vergleich). Hierfür müssen in den vorliegenden ICD-Kodes die Suffixe „#“, „†“, „+“, „\*“ und „!“ ignoriert werden.

### Beispiel

Folgende Operationen führen zum gleichen Ergebnis:

- ('Z37.9') EINSIN ('Z37.9'; 'Z37.0'; 'Z37.1'; 'Z37.2'; 'Z37.3')
- ('Z37.9!') EINSIN ('Z37.9'; 'Z37.0'; 'Z37.1'; 'Z37.2'; 'Z37.3')

### NULL-Werte im Pseudocode:

Das Vorhandensein eines Attributs wird mit der Schreibweise `attributname <> NULL` gekennzeichnet. Analog bezeichnet `attributname = NULL` das Fehlen eines Attributs.

### 2.1.3 Externe Parameter des QS-Filters

Im Pseudocode der QS-Filter wird auf drei Arten von externen Parametern Bezug genommen: Kodelisten, Funktionen und Wertebereiche. Diese Entkopplung der Parameter aus dem Pseudocode dient dazu, den Pseudocode übersichtlicher zu halten und eine Pflege der QS-Filter und der zugehörigen Datenbankabfragen bei Änderungen von Kodelisten und Betrachtungszeiträumen zu erleichtern.



## Kodelisten

Die in den Ein- und Ausschlusskriterien verwendeten Kodelisten (Filterlisten) sind in der Tabelle `FilterListeWert` enthalten.

Tabelle 6: Struktur der Tabelle `FilterListeWert`

Datenfeld	Bedeutung	Verwendung
<code>id</code>	ID	Eindeutige Zuordnung des Datenbankeintrags
<code>fkFilterListe</code>	Bezeichnung der Liste	Ist im Pseudocode der QS-Filter referenziert, ist Fremdschlüssel zur Tabelle <code>FilterListe</code> . Hieraus ist auch das jeweilige Modul erkennbar, welche in den Tabellen <code>FilterListe</code> und <code>Modul</code> hinterlegt ist.
<code>code</code>	Wert des Kodes	Dieser Wert wird auf Gleichheit geprüft

## Funktionen

Wie oben kurz beschrieben entsprechen einige Datenfelder des Exportformats nicht unmittelbar einem Datenfeld der Eingangsdaten, sondern werden mittels einer Funktion von diesem abgeleitet oder auch unabhängig von den Eingangsdaten bestimmt. Dies ist der Fall, wenn in der Tabelle `DatenfeldAttribut` das Feld `berechnet` den Wert `true` enthält. In diesem Fall wird auf das referenzierte Datenfeld eine `Funktion` angewandt. Beschreibung und Berechnungsvorschrift der Funktion können der Tabelle `Funktion` in der Spezifikationsdatenbank entnommen werden.

Tabelle 7: Inhalt der Tabelle `Funktion`

Funktion	Formel	Beschreibung
<code>year</code>	Stelle 1 bis 4 aus einem Datum im Format JJJJ-MM-TT	Jahr aus einem Datum
<code>source</code>	Herkunft der Daten entsprechend dem oBDS und dem organspezifischen Modul „Prostatakrebs“, aus dem Datenbestand der Krebsregister zu entnehmen, siehe Schlüssel 'quelle'	Kennzeichen der genauen Datenquelle des Falles
<code>format</code>	Ist WAHR, wenn das Format des ersten Parameters dem zweiten Parameter entspricht	Prüfung der gesetzlichen Krankenversicherungsnummer auf korrektes Format
<code>LENGTH</code>	Länge eines übergebenen Parameters	Länge der gesetzlichen Krankenversicherungsnummer
<code>LEFT</code>	Nimmt zwei Parameter: Ein Datum und eine Zahl x und gibt die ersten x Stellen des Datums zurück.	Die ersten 2 Stellen der gesetzlichen Krankenversicherungsnummer

## Wertebereiche

Die in den Ein- und Ausschlusskriterien verwendeten Wertebereiche (entspricht den Zeiträumen des Patientenfilters und des Leistungs- und Medikationsfilters) sind in der Tabelle *Wertebereich* enthalten.

Tabelle 8: Struktur der Tabelle *Wertebereich*

Datenfeld	Bedeutung	Verwendung
modul	Bezeichnung des Moduls	Eingrenzung auf ein bestimmtes Modul
spezifikations-jahr	Angabe des Spezifikationsjahres	Eingrenzung auf ein bestimmtes Spezifikationsjahr
name	Bezeichnung des Wertebereichs	Ist im Pseudocode der QS-Filter referenziert, mögliche Unterscheidung in <i>Index-jahr</i> (Zeitraum des Patientenfilters) und <i>Zeitfilter</i> (Zeitraum des Leistungs- und Medikationsfilters)
min	Kleinster Wert (inklusive)	Wertebereich, gegen den auf Einschluss geprüft wird
max	Größter Wert (inklusive)	
version	Version des Wertebereichs	Überprüfung der Aktualität
datum	Änderungsdatum	Intern
fkVersionRelease	Release-Version der aktuellen Liste	Versionskontrolle

Die Minimum- und Maximumwerte eines Wertebereichs können im Pseudocode mit der Schreibweise *NAME.min* bzw. *NAME.max* referenziert werden.

### 2.1.4 Ergebnis der Datenselektion

Die Anzahl der zu exportierenden Datensätze pro Modul und Spezifikationsjahr ergibt sich aus der Anwendung des Patientenfilters. Jeder Datensatz entspricht dabei einem Patienten und wird als *case* im XML-Datensatz angegeben. Der Umfang der zu exportierenden Daten je Datensatz ergibt sich aus dem Vorhandensein entsprechender Daten bei den Krebsregistern, die im Patientenfilter definiert sind.



#### Hinweis

Die Selektion durch die Filterkriterien erfolgt dabei jeweils auf Ebene des Best-of-Datensatzes entsprechend des oBDS.

Wenn ein Patient die Filterbedingungen erfüllt, ist er zu exportieren. Unterelemente (z. B. TNM-Angaben, Diagnosen usw.) werden nur dann mit exportiert, wenn sie:

- als Pflichtfelder in der Struktur der Exportdaten vorgesehen sind (Tabelle DatenfeldVerwendung),
- durch einen Filter selektiert wurden und
- Sekundärkodierungen zu einer durch den Filter selektierten Angabe sind

Fälle, die durch den Patientenfilter selektiert wurden, sind beim Export mit dem zusätzlichen Attribut `patientenfilter="1"` zu versehen.

Da die Basis für den Datenexport der Best-of-Datensatz des oBDS genutzt wird, wird davon ausgegangen, dass ein Patient bei der Selektion nur einmal ausgelöst werden kann. Wird ein Patient mit mehr als einem positiven Prüfergebnis aufgegriffen, so ist dieser nicht doppelt in den Exportdaten aufzuführen. Vielmehr ist eine Zusammenführung der selektierten Daten vorzunehmen und der Patient als ein Fall (`case` im XML) zu exportieren. Das zusätzliche Attribut `patientenfilter="1"` ist für den zusammengesetzten Fall anzugeben.

## 2.2 Datenexport

Der Datenexport hat zum Ziel, die im vorherigen Schritt selektierten Daten der Krebsregister im von der Bundesauswertungsstelle (BAS) spezifizierten Format in Exportdateien für die weitere Verarbeitung und Übermittlung abzulegen.



### Hinweis

Für jedes Modul mit Krebsregisterdaten ist von jedem (wohnortsbezogenen) Krebsregister eine separate Datenlieferung je Spezifikationsjahr zu exportieren.

---

### Dateien und Dateiformate

Um die Daten für die weitere Verarbeitung exportieren zu können, müssen Art, Format und Struktur der zu erzeugenden Dateien klar definiert sein.

### Mapping

Wie bereits in Abschnitt 2.1 beschrieben bezieht sich die Spezifikation lediglich auf die Eingangsdaten bei den klinischen Krebsregistern und den oBDS anstatt auf die individuell verschiedenen Datenspeichersysteme bei den Krebsregistern. Die Umsetzung des Datenexports erfordert deshalb, dass das einzelne Krebsregister in der Lage ist zu identifizieren, wo und in welcher Form die in ihrem System selektierten Daten innerhalb der Datenstruktur der Exportdateien abzulegen sind.

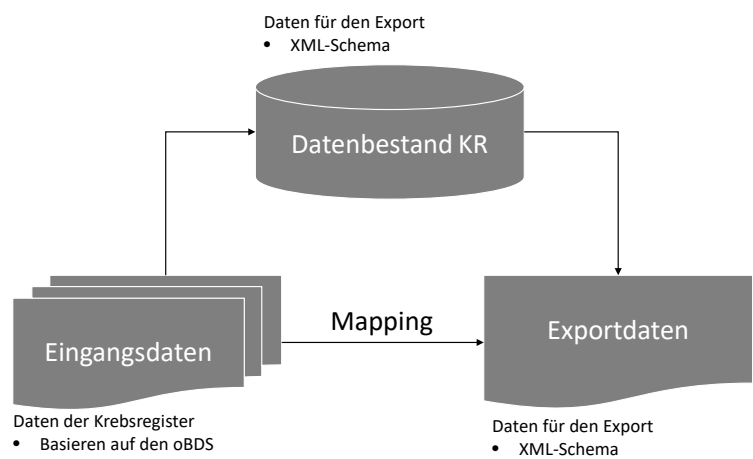


Abbildung 5: Mapping

Dazu sind in der Spezifikationsdatenbank verschiedene Tabellen angelegt, welche die Überleitung der Eingangsdaten zu den Exportdaten beschreiben und gleichzeitig Hinweise auf die Struktur der Exportdatei als XML-Struktur liefern:

Tabelle 9: Tabellen in der Spezifikationsdatenbank für den Datenexport

Tabelle in der Spezifikationsdatenbank	Beschreibung
Attribut	Liste der verwendeten Exportattribute
BasisTyp	Datentypen der verwendeten Exportattribute
DatenfeldAttribut	Zuordnung der Eingangsdaten zu den Exportattributen
DatenfeldVerwendung	Angabe der Verwendung der Eingangsdaten je Modul
Datenquelle	Datenquelle (Unterteilung nach oBDS und Administrative Daten)
Element	Liste der Exportelemente
ElementDatentyp	Angabe der modulabhängigen XML-Elementtypen der Exportelemente
ElementRelation	Hierarchie der Exportelemente je Modul
MussKann	Label „Muss-/Kann-Feld“
RelationTyp	Hierarchie der Exportelemente je Modul
Schluessel	Schlüssel zur Verwendung in Attributwerten
SchluesselWert	Schlüsselwerte zur Tabelle „Schluessel“

Mit Hilfe dieser Tabellen können die einzelnen Datenfelder der Eingangsdaten eindeutig einem entsprechenden Datenfeld in den Exportdateien zugeordnet werden.

## Beispiel Mapping

Das Ergebnis der Datenselektion gibt an, dass das Eingangsdatenfeld „TNM T-Kategorie“ für das Modul ZERVIXK exportiert werden muss. In der Spezifikationsdatenbank in der Tabelle `DatenfeldAttribut` ist ersichtlich, dass dieses Datenfeld im Element `<tnm>` als Attribut „tnm\_t-kategorie“ exportiert werden muss. Ebenso ist angegeben, dass es sich bei diesem Datenfeld um ein Muss-Datenfeld handelt, d. h. es muss in jedem Fall im Exportdatensatz enthalten sein (siehe Spalte `fkMussKann` bzw. Tabelle `DatenfeldVerwendung`).

Die Tabelle `DatenfeldAttribut` wird zudem in der Tabelle `DatenfeldVerwendung` über eine Fremdschlüsselbeziehung referenziert. Dort werden neben allen notwendigen Exportdatenfeldern für das Modul ZERVIXK auch die Datenfelder angegeben, welche in einem der Filterstufen (Pseudocode) auftreten (Reminder: Alle Datenfelder, die in den Filterstufen benannt sind und für die ein positives Prüfergebnis ausfällt, sind zu exportieren, siehe Abschnitt 2.1.1 Selektionsstufen). Das Beispielattribut „tnm\_t-kategorie“ ist nicht in einem der Filterstufen als Filterkriterium enthalten, muss laut der Tabellen `DatenfeldAttribut` und `DatenfeldVerwendung` aber exportiert werden.

Bei dem zu exportierenden Attribut „tnm\_t-kategorie“ handelt es sich zudem um einen Schlüssel (siehe Spalte `fkSchluessel` in Tabelle `Attribut`, Tabelle `SchluesselWert`).



### Hinweis

#### Schlüsselwerte

Für einige zu exportierenden Datenfelder sind nur bestimmte Schlüsselwerte zulässig (siehe Tabelle `Attribut` in den Spalten `fkBasisTyp` und `fkSchluessel`). In diesem Fall ist im Feld `schluessel` ein Verweis auf die entsprechende Werteliste angegeben. Die zulässigen Werte des jeweiligen Schlüssels sind der Tabelle `SchluesselWert` zu entnehmen.

Tabelle 10: Struktur der Tabelle `SchluesselWert`

Datenfeld	Bedeutung
id	Feldkennung des Schlüsselwertes
fkSchluessel	Bezeichnung des Schlüssels
code	Schlüsselwert
bezeichnung	Bezeichnung des Schlüsselwertes

### 2.2.1 Exportdateien

Art und Anzahl der Exportdateien in einer Datenlieferung ergeben sich aus den verschiedenen Datenkategorien und ihrem Empfänger im Datenfluss.

In der Datenlieferung sind fünf Kategorien zu unterscheiden:

Tabelle 11: Datenkategorien im Datenfluss

Datenkategorie	XML-Element	Beschreibung
AdministrativeDaten	header	Enthält Meta-Informationen zum Datenlieferanten, über das Datenpaket, Informationen zur Organisation des Datenflusses und zur Identifizierung der Transaktion. Pro Datei ist der Header nur einmal vorhanden. Jede Stelle im Datenfluss muss diese Daten einsehen können.
Krebsregisteridentifizierende Daten (KRID)	body	Enthält die Pseudo-IKNR des Krebsregisters. Hierbei handelt es sich um die IKNR, die von den Krebsregistern im Rahmen des registerübergreifenden Datenaustausches genutzt wird. Bei den krebregisteridentifizierenden Daten wird die IK Nummer des jeweiligen Krebsregisters exportiert. Die entsprechenden IK Nummern sind der Tabelle 12 zu entnehmen.  Tabelle 12
Patientenidentifizierende Daten (PID)	patient	Die patientenidentifizierenden Daten (PID) sind die eGK-Versichertennummern der Patienten. Die PID werden in gehashter Form (siehe Abschnitt 2.4.2) an die VST mit dem öffentlichen Schlüssel der VST verschlüsselt. Die VST ersetzt die PID durch ein verfahrensspezifisches Patientenpseudonym.
QS-Daten	qs_data	In diesem Element werden die medizinischen Daten des Patienten gespeichert. Diese Daten dürfen vom Krebsregister und der BAS eingesehen werden. Sie müssen von dem absendenden Krebsregister mit dem öffentlichen Schlüssel der Auswertungsstelle verschlüsselt werden.

Bei den krebregisteridentifizierenden Daten wird die IK Nummer des jeweiligen Krebsregisters exportiert. Die entsprechenden IK Nummern sind der Tabelle 12 zu entnehmen.

Tabelle 12: Liste der IK-Nummern der Krebsregister für die Exportdatei

Krebsregister/ Bundesland	IK Nummer
Baden-Württemberg	678000018
Bayern	678700026
Brandenburg/ Berlin	670700018
Bremen	673100018
Hamburg	671500018

Hessen	675300015
Mecklenburg-Vorpommern	670300014
Nordrhein-Westfalen	673700014
Rheinland-Pfalz	676500019
Saarland	679300013
Sachsen-Anhalt	671000024
Schleswig-Holstein	671300016
Thüringen	676000014
Sachsen Chemnitz	677700013
Sachsen Zwickau	677700024
Sachsen Dresden	677900015
Sachsen Leipzig	678200010

### 2.2.2 XML-Struktur

Als Exportformat für den Datenexport aus dem klinischen Krebsregister über die Vertrauensstelle zur Auswertungsstelle wird XML genutzt. Die Struktur der XML-Dateien wird jeweils durch ein XML-Schema eindeutig definiert.

Die grundlegende XML-Struktur ist im Abschnitt 6 beschrieben. Die spezifische XML-Struktur ergibt sich erst im Kontext eines bestimmten Moduls.



#### Hinweis

XML-Beispiele befinden sich im Paket der XML-Schemata. Nähere Informationen sind im Abschnitt 6 zu finden.

### 2.2.3 Segmentierung

Um die Größe der zu übertragenden Dateien zu begrenzen, dürfen für Verfahren der Methodenbewertung nicht mehr als 50.000 Datensätze in einer Exportdatei enthalten sein. Diese Vorgaben sind modulspezifisch in der Spezifikationsdatenbank in Tabelle Modul in der Spalte *segmentierung* hinterlegt. Die Richtlinienzugehörigkeit der einzelnen Module ist in Spalte *richtlinie* hinterlegt. Als Datensatz für die Zählung gilt ein Patient mit allen dazugehörigen Falldaten pro Exportmodul und je Spezifikationsjahr.

### 2.2.4 Nullmeldung

Sollte für die Kombination aus einem Modul und einem Spezifikationsjahr vom Patientenfilter kein Patient aufgegriffen worden sein, so ist dennoch eine Transaktion im Sinne einer Nullmeldung durchzuführen. Im Unterschied zu einer Transaktion deren Ergebnismenge nach Anwendung des

Patientenfilters nicht leer ist, enthält eine Nullmeldung keine case-Elemente. Weiterhin ist das Attribut `total_count` im header auf „0“ zu setzen.

Dadurch wird sichergestellt, dass zu jedem Lieferzeitfenster von jedem Krebsregister eine Transaktion durchgeführt werden kann.

## 2.3 Datenprüfung und Regeln

In der Spezifikationsdatenbank sind die in Tabelle 13 dargestellten Tabellen für die Datenprüfung relevant:

Tabelle 13: Tabellen in der Spezifikationsdatenbank für die Datenprüfung und Regeln

Tabelle in der Spezifikationsdatenbank	Beschreibung
Strenge	Label „Strenge der Prüfregel“
Verarbeitungsschritt	Liste von Verarbeitungsschritten für Prüfungen
VerarbeitungsschrittPruefung	Zugehörigkeit der Prüfungen zu den Verarbeitungsschritten
Pruefung	Beschreibung der Prüfung
Fehlermeldung	Konkrete Fehlermeldung
Pruefbereich	Kategorisierung der Prüfung

Vor dem Export der XML-Dokumente wird als Strukturprüfung eine Schemaprüfung durchgeführt. Die Felddefinitionen sind entweder dem Schema (`./interface_KKR/Interface_KKR.xsd`) oder alternativ der Abfrage `FeldDefinition` der Spezifikationsdatenbank zu entnehmen.

Die Abfrage `FeldDefinition` enthält dazu folgende Angaben:

Tabelle 14: Struktur der Abfrage `FeldDefinition`

Spalte	Bedeutung
<code>modul</code>	Verfahren
<code>eingangskennung</code>	Eindeutige Identifikation des Datenfeldes
<code>attribut</code>	XML-Attribut eines Elements
<code>basisTyp</code>	Datentyp2 (ZAHL, GANZEZAHL, DATUM, SCHLUESSEL, etc.)
<code>schluessel</code>	Falls <code>basisTyp</code> dem Wert „SCHLUESSEL“ entspricht, wird der konkrete Schlüsselname angegeben. Die einzelnen Schlüsselcodes befinden sich in der Spezifikationsdatenbank in der Tabelle <code>SchluesselWert</code> .
<code>basisTypBezeichnung</code>	Bezeichnung des Datentyps

<sup>2</sup> Details siehe Tabelle `BasisTyp` in der Spezifikationsdatenbank



Spalte	Bedeutung
format	Format
formatRegExp	Regulärer Ausdruck
laenge	Feldlänge: x = genaue Länge ..x = maximale Länge
nachKommaLaenge	Maximale Nachkommastellenlänge
min	Minimalwert eines Gültigkeitsbereichs
max	Maximalwert eines Gültigkeitsbereichs
mussKann	Angabe, ob Pflichtfeld oder optional: K = Kann-Feld (optional) M = Muss-Feld (Pflichtfeld)

Aus dieser Abfrage ergeben sich folgende Prüfungen:

- Überprüfung des Formats (Feldinhalt)
- Überprüfung der Feldlänge
- Überprüfung der Schlüsselcodes
- Überprüfung der numerischen Wertebereiche
- Überprüfung der Muss-Kann-Eigenschaft

### Administrative Prüfungen

Aktuell sind in der Spezifikation ausschließlich administrative Prüfungen hinterlegt. Diese sind verfahrensunabhängig und definieren, welche Prüfungen im Rahmen des Exports (vom Krebsregister bis zur Auswertungsstelle) innerhalb einer Einrichtung und in welcher zeitlichen Reihenfolge durchzuführen sind. Administrative Prüfungen beziehen sich nicht auf Datenfelder innerhalb der Exportdaten, sondern auf übergeordnete Prüfungen, wie die Einhaltung des Schemas, Prüfung auf fehlende Segmente, etc.

Zu finden sind diese Prüfungen in der Abfrage `PruefschrittPruefung`. Über die Spalte `einrichtung` können sie je nach Einrichtung (KKR, VST und BAS) gefiltert werden. Bestandteil der Abfrage sind auch die Fehlermeldungen, die als Ergebnis einer Prüfung mit negativem Ergebnis angezeigt werden.

In Tabelle 15 wird die Struktur der Abfrage `PruefschrittPruefung` beschrieben. Die Fehlermeldungen und weitere Informationen im Falle eines negativen Prüfergebnisses werden nicht durch das Krebsregister, sondern durch die die Daten entgegennehmende Stelle in das Datenflussprotokoll eingetragen. Das Krebsregister ist jedoch verpflichtet, durch eigene Prüfungen sicherzustellen, dass keine fehlerhaften Dokumente exportiert werden.

Tabelle 15: Struktur der Abfrage *PruefschrittPruefung*

Datenfeld	Bedeutung	Anwendungsrelevanz
prozessschritt	Ordnet die Prüfung in den zeitlichen Kontext des Prozessschrittes ein	Implementierung der Prüfprozesse in einem Datenservice
pruefung	Beschreibung der Prüfung	Datenservice
einrichtung	Einrichtung, die die Prüfung durchzuführen hat	Datenservice und Anwender
pruefbereich	Kategorisierung der Prüfung nach: Spezifikation, Transaktion, Dechiffrierung, Schema, PID_Pseudonym, Sonstige Prüfung	Rückprotokollierung
prozessschrittSort	Legt die Reihenfolge der Prozessschritte fest	Datenservice
pruefungSort	Stellt die Prüfung innerhalb des Prozessschrittes in zeitliche Reihenfolge	Datenservice
fehlermeldungId	Id der Fehlermeldung	Rückprotokollierung
fehlermeldung	Konkrete Fehlermeldung	Rückprotokollierung
fehlermeldungParameter	Fehlermeldungen sind parametrisierbar. Parameter befinden sich in <spitzen Klammern> innerhalb der Fehlermeldung. Im Feld fehlermeldungParameter werden alle gültigen Parameter aufgelistet und erläutert. Beispiel:  <params details="Technische Details" /> beschreibt den Parameter „details“ in der Fehlermeldung „Die QS-Datei konnte nicht entschlüsselt werden: <details>“	Rückprotokollierung
strenge	Strenge einer negativ verlaufenen Prüfung: Mögliche Werte sind „hart“ und „weich“	Rückprotokollierung
verursacher_wenn_vst	Gibt den Verursacher einer negativ verlaufenen Prüfung an, wenn die prüfende Einrichtung die VST ist und der Verursacher des Fehlers eindeutig feststeht	Datenservice, Rückprotokollierung
verursacher_wenn_bas	Gibt den Verursacher einer negativ verlaufenen Prüfung an, wenn die prüfende Einrichtung die BAS ist und der Verursacher des Fehlers eindeutig feststeht	Datenservice, Rückprotokollierung

Durch die Spalten `prozessschrittSort` und `pruefungSort` wird die Reihenfolge der Prüfungen innerhalb einer Einrichtung festgelegt.

Eine Filterung der Abfrage in der Spalte `einrichtung` nach „KKR“ ergibt alle für die Krebsregister vor dem Export durchzuführenden Prüfungen. Schlägt hierbei eine Prüfung mit der Strenge „hart“ fehl, ist die Ursache dafür zu beheben und der Export neu durchzuführen.

**Hinweis**

Prüfungen mit der Strenge „hart“ (siehe Angabe in der Datenbanktabelle `Pruefung`) sind bis auf Weiteres als „weich“ zu interpretieren. Das bedeutet, im Fall eines negativen Prüfungsergebnisses durch die VST bzw. die Auswertungsstelle wird die Datenlieferung angenommen. Im Datenflussprotokoll muss die jeweilige Datenannahmestelle aber die entsprechenden Regelverletzungen mit dem `rule_type = „D“` kenntlich machen (siehe XML-Schema).

Das Krebsregister überarbeitet daraufhin die Datenlieferung und nimmt eine Korrekturlieferung im Rahmen der Korrekturlieferungsfrist vor.

---

## 2.4 Lieferzeiträume und Datenübermittlung

### 2.4.1 Lieferzeiträume

Die Datenübermittlung der Daten von den Krebsregistern an die Auswertungsstellen der okFE erfolgen zu festgelegten Lieferzeitfenstern. Diese sind der Tabelle `Lieferfrist` in der Spezifikationsdatenbank zu entnehmen. Bei der Festlegung der Lieferzeitfenster, insbesondere für die spezifikationsjahresgebundenen Datenexporte, wurde als Grundannahme davon ausgegangen, dass die für einen Export notwendigen Informationen zu einem Patienten zwischen 9 Monate und einem Jahr nach der Erstdiagnose vollständig im klinischen Krebsregister vorliegen.

Neben der ersten Datenlieferung pro Modul und Spezifikationsjahr ist eine Korrekturlieferung möglich. Diese ist ebenfalls zeitlich befristet und kann der Tabelle `Lieferfrist` entnommen werden.

---

**Beispiel 1**

Für das Modul ZERVIXK ist eine jährliche Datenlieferung pro Modul und Spezifikationsjahr spezifiziert. Diese Datenlieferung liegt für das Spezifikationsjahr 2024 im Zeitraum vom 01.05.2024 bis zum 31.05.2024. Eine Korrekturlieferung ist bei Bedarf möglich und liegt zwischen dem 01.06.2024 und 15.06.2024.

---

### 2.4.2 Vorgaben für die Pseudonymisierung der Krankenversichertennummer

Der Export der Krankenversichertennummer (eGK-Versichertennummer) aus den klinischen Krebsregistern an die unabhängige Vertrauensstelle des G-BA erfolgt in pseudonymisierter Form. Die entsprechenden rechtlichen Grundlagen wurden in § 25a Absatz 4 Satz 6 SGB V getroffen:

*Ein Abgleich der Daten nach Satz 4 und der Daten, die nach § 299 zum Zwecke der Qualitätssicherung an eine vom Gemeinsamen Bundesausschuss bestimmte Stelle übermittelt werden, mit Daten der epidemiologischen oder klinischen Krebsregister ist unter Verwendung eines aus dem unveränderbaren Teil der Krankenversichertennummer des Versicherten abgeleiteten Pseudonyms zulässig, sofern der Versicherte nicht schriftlich oder elektronisch widersprochen hat.*

Die folgende Gesetzesbegründung wurde vom Gesetzgeber textiert:

*In der Praxis stand dem Abgleich allerdings entgegen, dass die meisten Landesgesetze die Übermittlung von Krebsregisterdaten an die Vertrauensstelle des Gemeinsamen Bundesausschusses (G-BA) auf Basis der technischen Umsetzungslösung des G-BA bisher nicht gestatten. Die Änderungen in § 25a berücksichtigen dies und machen den Datenabgleich gangbarer. Sie setzen im Wesentlichen eine von den Ländern vorgeschlagene Lösung um, wonach die Krebsregisterdaten an die Vertrauensstelle des G-BA zusammen mit einem Pseudonym übermittelt werden, das aus dem unveränderbaren Teil der Krankenversichertennummer gebildet wird. Teil dieser Gesamtlösung ist zudem, dass die Länder auf dieser Grundlage die erforderlichen technischen, organisatorischen und teilweise rechtlichen Anpassungen zur Umsetzung dieser Lösung vornehmen. In einigen Ländern sind lediglich kurzfristig umsetzbare technische Anpassungen erforderlich.*

Die konkrete technische Umsetzung der Gesetzesvorgaben wird durch die Formulierungen der „Umsetzungsskizze Einbindung Krebsregister“ in Version 1.6 vom 7. Dezember 2021 der GFL vorgegeben. Hierzu wird ein zweistufiges Hashverfahren auf die eGK-Versichertennummer in den klinischen Krebsregistern angewendet, bevor die Daten an die Vertrauensstelle exportiert werden. Die eGK-Versichertennummer wird als erstes mittels Message-Digest Algorithm 5 (MD5) und anschließend mittels Secure Hash Algorithm (SHA256) pseudonymisiert. Die Vertrauensstelle überführt anschließend den ihnen übermittelten eGK-Hash final mittels des für die QS-Verfahren und die oKFE-Programme genutzten Pseudonymisierungsalgorithmus in das finale Patientenpseudonym.

### Technische Voraussetzungen<sup>3</sup>

Es ist wichtig, dass alle Hash Berechnungen auf Byte Arrays ausgeführt werden und keine Strings im Hex-Format verwendet werden. Liegt ein MD5 Hash schon vor, so ist dieser in ein Byte Array umzuwandeln. Dabei muss der Hex-Wert in den Dezimal-Wert konvertiert werden (Beispiel: 'A5' => 0xA5 => 165). Falsch wäre einen MD5 Wert, der als String im Hex-Format vorliegt, direkt in ein Byte Array umzuwandeln (Beispiel: 'A5' => [65, 53] bzw. 'a5' => [97, 53]).

1. eGK Nummer (String, erster Buchstabe ist großgeschrieben) in ein Byte Array umwandeln. Eine Zeichenkodierung muss hier nicht beachtet werden, da in einer eGK Nummer keine Sonderzeichen enthalten sind.
2. MD5 Hash des "eGK Byte Arrays" generieren. Es wird ein MD5 Byte Array erzeugt.
3. SHA256 Hash des "MD5 Byte Arrays" generieren. Es wird wieder ein Byte Array erzeugt.
4. "SHA256 Byte Array" in einen String (Hex-Format) umwandeln. Groß-Klein-Schreibung spielt hier keine Rolle, da in der Vertrauensstelle der Hash wieder in ein Byte Array umgewandelt wird.

### Pseudocode

```
String eGK = "A123456789";
byte[] eGK_bytes = eGK.getBytes();
byte[] md5 = md5(eGK_bytes);
byte[] sha256 = sha256(md5);
String sha256_string = toHexString(sha256);
print(sha256_string) =>
'22A5E613A1B9084A6B52B09CD30D7D59CED2DFC29E4BC77BBDCC3A8E5C1F7773'
```

### Prüfung der Implementierung

Die Umsetzung der Vorgaben der Verschlüsselung können auf der folgenden Internetseite geprüft werden. Die URL leitet den Nutzer direkt zu der korrekten Kombination der Hashing-Algorithmen weiter.

[https://cyberchef.io/#recipe=MD5\(\)From\\_Hex\('Auto'\)SHA2\('256',64,160\)&input=OTeyMzQ1Njc4OQ](https://cyberchef.io/#recipe=MD5()From_Hex('Auto')SHA2('256',64,160)&input=OTeyMzQ1Njc4OQ)

### 2.4.3 Vorgaben für die Dateiverschlüsselung und Übermittlung

Die Übermittlung der Daten vom klinischen Krebsregister an die Vertrauensstelle erfolgt mittels REST-Schnittstelle. Hierfür wird von der Vertrauensstelle ein Java Softwareprogramm (REST-Client) bereitgestellt. Der REST-Client ist auf Windows-PCs lauffähig und übernimmt die Komprimierung.

<sup>3</sup> Mit der folgenden URL wird exakt der Hashing-Algorithmus abgebildet, mit dem die KKR die eGK hashen müssen, bevor sie diese exportieren. URL: [https://cyberchef.io/#recipe=MD5\(\)From\\_Hex\('Auto'\)SHA2\('256',64,160\)&input=OTeyMzQ1Njc4OQ](https://cyberchef.io/#recipe=MD5()From_Hex('Auto')SHA2('256',64,160)&input=OTeyMzQ1Njc4OQ)

Der Link kann auch dafür genutzt werden, um gegen zu prüfen, ob die eGK-Versichertennummern nach dem korrekten Algorithmus erstellt werden.

mierung und Verschlüsselung der Exportdatei und die sichere Übermittlung der Daten an die Vertrauensstelle. Der REST-Client inklusive des Nutzerhandbuchs liegt dem Spezifikationspaket bei. Bei technischen Fragen wenden Sie sich bitte per E-Mail ([helpdesk-gbavst@nortal.com](mailto:helpdesk-gbavst@nortal.com)) oder per Telefon (+49 30-318 05 09-57) an den Helpdesk der Vertrauensstelle.

#### 2.4.4 Dateiverschlüsselung

~~Die Datenblöcke patient und ge~~Der Datenblock~~gs~~\_data ~~sind~~ist mit einer asymmetrischen XML-Verschlüsselung anhand des öffentlichen Schlüssels des jeweiligen Zieladressaten zu versehen. Da einige Krebsregister z. B. Drittdienstleister für die Erstellung der Exportdateien nutzen, kann optional auch der Datenblock patient mit der gehashten eGK-Nummer vor der Übermittlung an die VST XML-verschlüsselt werden.

Die programmspezifischen öffentlichen Schlüssel zur Verschlüsselung werden den Krebsregistern zur Verfügung gestellt. Die personenidentifizierenden Daten sind immer mit dem jeweils aktuellen öffentlichen Schlüssel der Vertrauensstelle und die medizinischen Daten sind immer mit dem jeweils aktuellen öffentlichen Schlüssel der oKFE-Auswertungsstelle zu verschlüsseln. Somit ist datenschutzrechtlich sichergestellt, dass an keiner Stelle die Zusammenführung von nicht pseudonymisierten, personenidentifizierenden Daten und medizinischen Daten stattfinden kann.

Die Verschlüsselung baut auf dem W3C-XML-EncryptionStandard auf. Zur Verschlüsselung stellt das IQTIG im Rahmen der bisherigen Verfahren ein Tool (GPacker-/XPacker-Tool; abrufbar unter <https://iqtig.org/datenerfassung/servicedateien/>) zur Verfügung. Dieses kann von den Verfahrensteilnehmern genutzt werden, um die Implementierung der Verschlüsselung nicht selbst vornehmen zu müssen.



#### Hinweis

Für die Datenübertragung per REST wird das Transportprotokoll HTTPS genutzt. SFTP stellt die Transportverschlüsselung über das Protokoll sicher.

---

## 2.5 Rückprotokollierung

### 2.5.1 Empfangsbestätigung der Vertrauensstelle

Nach erfolgreicher Eingangsverarbeitung erhält das Krebsregister von der Vertrauensstelle eine Empfangsbestätigung, die den Erhalt einer Datenlieferung bestätigt. Ist eine Datenlieferung nicht bearbeitbar, wird anstelle einer Empfangsbestätigung ein Datenflussprotokoll an das Krebsregister übermittelt.

Die genaue Struktur und ein Beispiel sind detailliert in Abschnitt 6.3 beschrieben.

### 2.5.2 Datenflussprotokoll

Ein Datenflussprotokoll (DFP) wird durch folgende Stellen und in folgenden Fällen erstellt und an die Krebsregister übermittelt:

- Vertrauensstelle: Wenn eine Weiterverarbeitung aufgrund von Fehlern nicht möglich ist (der Dokumentenstatus wurde auf „ERROR“ gesetzt). Diese DFP werden je Lieferung unter Verwendung des Header-Attributs `id` erstellt
- Auswertungsstelle: Wenn die Auswertungsstelle oKFE die Eingangsdatei schemakonform von der VST erhält. In diesem Fall enthält das Protokoll die Prüfergebnisse aller Prüfinstanzen (VST, Auswertungsstelle). Diese DFP werden ebenfalls transaktionsbezogen unter Verwendung des Header-Attributs `set_id` erstellt.

Dem Datenflussprotokoll kann anhand des Attributs `v` des Elements `<status_document>` entnommen werden, ob die Übermittlung der Transaktion (Auswertungsstelle) bzw. eines Segments (VST) erfolgreich war:

Tabelle 16: Status des Dokuments

Ergebnis	Bedeutung	Konsequenz
OK	Transaktion erfolgreich: Keine Auffälligkeiten	Keine
WARNING	Transaktion erfolgreich: Es sind Verstöße gegen weiche, aber nicht gegen harte Prüfungen aufgetreten.	Abhängig von den aufgetretenen Fehlern ist von den KKR zu entscheiden, ob die Lieferung nach Beheben der Fehler erneut durchzuführen ist.
ERROR	Transaktion nicht erfolgreich: Es sind Verstöße gegen harte Prüfungen aufgetreten, die die Weiterverarbeitung verhindert.	Die Fehler sind durch die KKR zu beheben und die Datensätze sind anschließend erneut zu übermitteln.  Hinweis: Wenn der Status „D“ gesetzt ist, gibt es auf Datensatzebene keine ERROR-Meldung, sondern WARNING.

Wenn der Dokumentenstatus „WARNING“ oder „ERROR“ ist, bedeutet dies, dass entweder auf Hebderebene oder auf Datensatzebene Fehler aufgetreten sind. Analog zum Status des Dokuments werden die Prüfergebnisse auf Datensatzebene im Element `<status_case>` festgehalten:

Tabelle 17: Status des Datensatzes

Ergebnis	Bedeutung	Konsequenz
OK	Datensatz angenommen: Keine Auffälligkeiten	Keine
WARNING	Datensatz angenommen: Es sind weiche Fehler aufgetreten.	Abhängig von den aufgetretenen Fehlern ist von den KR zu entscheiden, ob die Lieferung

		nach Beheben der Fehler erneut durchzuführen ist.
ERROR	<p>Es sind harte Fehler aufgetreten, die eine Aufnahme in den Datenpool verhindert haben.</p> <p>Es sind harte Fehler aufgetreten. Die Daten werden mit entsprechender FehlerEinstufung in den GFL-Importdatenpool aufgenommen. Es erfolgt eine Rückmeldung der Fehler im Datenflussprotokoll. Die Daten müssen korrigiert und erneut vom KKR geliefert werden.</p>	<p>Die Fehler sind durch die KKR zu beheben und die Datensätze sind anschließend erneut zu übermitteln.</p> <p>Hinweis: Wenn der Status „D“ gesetzt ist, gibt es auf Datensatzebene keine ERROR-Meldung, sondern WARNING.</p>

Ist ein Fehler aufgetreten, ist zu überprüfen, ob das Krebsregister dafür verantwortlich war. Den Hinweis hierzu gibt das optionale Attribut `originator` im `<error>`-Tag, falls dieses gesetzt ist und den Wert „Krebsregister“ enthält. Da dieses Attribut jedoch nur gesetzt werden darf, wenn ein Verursacher eindeutig feststeht, ist auch bei Fehlern ohne `originator`-Attribut zu überprüfen, ob der Fehler durch das Krebsregister verursacht wurde.

Einzelheiten zur Rückprotokollierung können dem Abschnitt 6.3, die XML-Struktur dem Abschnitt 6 entnommen werden.



### 3 Prozesse bei den Krebsregistern

Die klinischen Krebsregister haben in Bezug auf die Übermittlung der Krebsregisterdaten folgende Punkte zu berücksichtigen:

1. Die Datengrundlage für die Selektion und den Export der Daten bildet der Best-of-Datensatz
2. Das für die Datenselektion und Datenübermittlung verantwortliche klinische Krebsregister ist das wohnortbezogene klinische Krebsregister zum Zeitpunkt der Diagnosestellung. Die klinischen Krebsregister stellen über den registerübergreifenden Datenabgleich (RÜD) mit den behandlungsortbezogenen klinischen Krebsregistern sicher, dass vor dem Datenexport alle Informationen im wohnortbezogenen Krebsregister vorliegen. Auf diesem Weg sollen doppelte Datenexporte aus den wohnort- und behandlungsortbezogenen Krebsregistern vermieden werden.
3. In der XML-Exportdatei ist die RÜD-IK-Nummer des Krebsregisters aus Tabelle 12 als „Absenderadresse“ der Datenlieferung anzugeben, um eine Zuordnung der Lieferung zum Lieferanten vornehmen zu können.
4. Die eGK-Versichertennummern sind vor dem Export gemäß der Vorgaben in Abschnitt 2.4.2 zu hashen.
5. Liegen mehrere eGK-Nummern bei den Krebsregistern vor, ist die eGK-Nummer zum Zeitpunkt der Diagnosestellung zu übermitteln.
6. Bezüglich der zeitlichen Zuordnung des TNM wird das TNM zum Zeitpunkt der Diagnose übermittelt. Liegen zum Diagnosedatum ebenfalls mehrere TNM vor, ist das schwerwiegendste TNM zu übermitteln.
7. Bei fehlender Angabe des Attributs „c\_p-praefix\_t“ (siehe XML-Schema) in den Datenbeständen der Klinischen Krebsregister bzw. bei klinischer Feststellung der T-Kategorie ist der Wert „c“ einzutragen, insofern eine Angabe für die T-Kategorie vorliegt.

#### **Modulspezifische Hinweise für DARMK und ZERVIXK:**

1. Falls die UICC-Klassifikation für das Modul DARMK nicht vorliegt, jedoch vom klinischen Krebsregister auf Basis vorliegender Informationen selbst gebildet werden kann, ist diese durch das Krebsregister zu bilden und zu übermitteln.
2. Es wird immer das älteste OP-Datum, was im Zusammenhang mit dem Primärtumor steht, exportiert.
3. Falls für ein Modul mehrere Tumore (verschiedene Abschnitte, oder einmal C20, einmal C18) im Best-of-Datensatz des Patienten vorliegen, ist nur eine Übermittlung pro Patient vorzunehmen. Hierbei ist der schwerwiegendste Primärtumor zu wählen (siehe Abschnitt 2.1.4). Sollte eine Abschätzung zum schwerwiegendsten Primärtumor bei zeitlicher Synchronität nicht möglich sein, muss das Krebsregister abwägen, welcher Primärtumor übermittelt wird.

4. Morphologie: Für den Datenexport ist der diagnostisch-relevante Morphologie-Kode anzugeben.
5. In dem Fall, dass mehrere Todesursachen dem Krebsregister bekannt sind, ist die Todesursache zu übermitteln, welche im Zusammenhang mit der Krebsart entsprechend des Früherkennungsprogramms steht (entspricht dem Grundleiden des Patienten). Kann kein Zusammenhang mit der Krebsart hergestellt werden, ist die erste Angabe aus dem Best-of-Datensatz zu übermitteln.

## 4 Prozesse bei der Vertrauensstelle

In diesem Kapitel werden die technischen Hintergründe und durchzuführenden Prozesse beschrieben, die neben den in Kapitel 2 genannten Vorgaben zusätzlich bei der Vertrauensstelle relevant sind.

### Eingangsprüfung/Empfangsbestätigung

Die von den klinischen Krebsregistern übermittelten Datensätze werden entsprechend der vorgegebenen Prüfungen in der Spezifikationsdatenbank (z. B. auf Wohlgeformtheit, Schemakonformität) geprüft. Anschließend wird eine Empfangsbestätigung (entsprechend der Spezifikation) erzeugt und an die Krebsregister übermittelt.

Das Übertragungsverfahren für die Datenübermittlung von den klinischen Krebsregistern an die VST ist REST.

### Pseudonymisierung der PID und eGK-Hash-Werte der KKR

Die Vertrauensstelle pseudonymisiert die durch die klinischen Krebsregister gehashte eGK-Versichertennummer verfahrensspezifisch gemäß der Vorgaben des G-BA. Detaillierte Informationen hierzu befinden sich in Abschnitt 2.4.2.

Tabelle 18 zeigt, nach welchem XML-Schema die VST die Exportdateien aufbereiten muss:

Tabelle 18: XML-Schemata für die VST

Ablageort	XML-Schema	Verwendungszwecke
interface_ VST_BAS	interface_ VST_BAS.xsd	<ul style="list-style-type: none"> <li>▪ Zusammenführung der QS- und der PID Datensätze</li> <li>▪ Integration der Patientenpseudonyme in die entsprechenden QS-Datensätze</li> <li>▪ Datenübermittlung von der VST an die BAS</li> </ul>

### Umgang mit Widersprüchen gemäß § 8 oKFE-RL

Wenn Versicherte einer Verarbeitung Ihrer Daten zum Zweck der Programmbeurteilung nach § 6 oKFE-RL wirksam widersprochen haben, müssen gem. § 8 Abs. 4 Satz 4 in der VST die dazugehörigen Programmbeurteilungsdaten gelöscht und durch eine Kennung, dass ein Widerspruch vorliegt, ersetzt werden.

Bei Vorliegen entsprechender Widersprüche, erfolgt das „Einfügen der patient-Elemente“ sowie die „Pseudonymisierung der PID“ wie gehabt. Zusätzlich müssen jedoch zwei weitere Schritte vollzogen werden, deren strukturelle Zugehörigkeiten sich auf das Element `qs_data` beschränken, so wie in Abbildung 6 dargestellt.

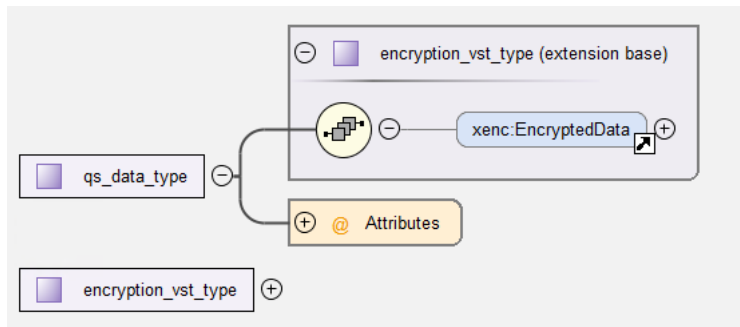


Abbildung 6: Umgang mit Widersprüchen gemäß § 8 oKFE-RL

In einem ersten Schritt erfolgt die Löschung der Programmbeurteilungsdaten. Hierfür muss der gesamte XML-verschlüsselte Inhalt des `qs_data`-Elementes entfernt werden. In einem zweiten Schritt ist die zuvor erfolgte Löschung der Programmbeurteilungsdaten im XML-Attribut `wider-spruch` des `qs_data`-Elementes anzuzeigen, so dass die Auswertungsstelle dies erkennen kann.

### Übermittlung an die Auswertungsstelle

Die Weiterleitung der Daten von der Vertrauensstelle an die Auswertungsstelle erfolgt mittels REST. Hierbei ist zu beachten, dass Daten für Module gemäß oKFE-RL an die GFL als Auswertungsstelle der oKFE zu übermitteln sind.

### Rückmeldung der Auswertungsstelle

Die von der VST empfangenen Daten werden von der Auswertungsstelle auf Schemakonformität geprüft. Nach Feststellen der Schemakonformität wird eine Empfangsbestätigung (entsprechend der Spezifikation) erzeugt und per REST an die VST übermittelt.

Bei einem Schemaverstoß wird anstelle der Eingangsbestätigung ein Protokoll mit entsprechenden Fehlermeldungen zurückgeschickt.

## 5 Prozesse bei der Auswertungsstelle

Die (Bundes-)Auswertungsstelle (BAS) hat in Bezug auf die Verarbeitung der Krebsregisterdaten folgende Punkte zu berücksichtigen:

1. Die Auswertungsstelle übermittelt eine Empfangsbestätigung an die VST, insofern in der Datenlieferung keine Fehler mit `originator` = „Vertrauensstelle“ enthält.
2. Andernfalls (es liegen Fehler vor), übermittelt die Auswertungsstelle ein Datenflussprotokoll anstatt der Empfangsbestätigung an die VST.
3. Eine Überprüfung der Daten auf Duplikate (Doppelmeldungen) für die oKFE-Module erfolgt durch die GFL ohne Ablehnung der Datenlieferung. Das heißt, insofern Doppelmeldungen eingehen, werden diese verarbeitet und nicht abgelehnt. Insofern eine Anpassung der Spezifikation erforderlich ist, ist dies durch die GFL an das IQTIG zu kommunizieren (siehe Abschnitt 7.2).
4. Jede Datenlieferung wird als komplett für den ausgewiesenen Zeitraum, das ausgewiesene Exportmodul, das Spezifikationsjahr und das einsendende Krebsregister betrachtet.
5. Die Zeichensatzkodierung des Rückprotokolls erfolgt in UTF-8.
6. Die Datenlieferungen sind für Verfahren gemäß oKFE-RL in Dateien zu jeweils maximal 50.000 Datensätzen segmentiert. Rückprotokolle beziehen sich immer auf die Gesamtdatenlieferung.
7. Die Annahme der Daten und der Versand der Rückprotokolle erfolgt per REST.

## 6 XML-Schemata

Das XML-Schema hat die Aufgabe, das aktuelle Datenflussmodell der G-BA-Richtlinie abzubilden:

- Abbildung der zu exportierenden Krebsregisterdaten
- Abbildung der Schnittstellen an den Institutionen-Übergängen (z. B. KKR-VST)
- Abbildung der Rückprotokollierung

Aus diesem Grund gibt es bei der Erstellung von Schemata, welche die Konformität von Richtlinie und Datenschutz sicherstellen sollen, kein „Allround-Schema“, welches alle Anforderungen an alle Beteiligten abdeckt, sondern eine „Schema-Familie“, aus der heraus gezielt für jede Schnittstelle („Interface“) eine passende Datenstruktur definiert wird.

### 6.1.1 Kompositionsmodell

Bei der Schema-Erstellung wurde auf ein Kompositionsmodell zurückgegriffen, in dem sich alle Teilschemata am Ende einen Namensraum teilen, um diese Schema-Familie besser warten zu können und gleiche Teilstrukturen nur einmal definieren zu müssen. Als Bezeichnung des Namensraums wurde „urn:gba:sqg“ gewählt. Zu diesem Namensraum werden die Bausteine je nach Bedarf über „includes“ zusammengestellt.

Das Kompositionsmodell macht es möglich, Konzepte aus der objektorientierten Programmierung – darunter fallen die Konzepte abstrakter Typ, Ersetzbarkeit von Typen, Wiederverwendung und Polymorphismus – zu nutzen.

Dadurch können Schemata erstellt oder genutzt werden, die generische Grundtypen definieren und diese Typen so erweitern, dass sie schnittstellenspezifisch sind, ohne das ursprüngliche Schema zu beeinflussen.

Dieses Kompositionsmodell soll hier näher erläutert werden.

---

#### Beispiel:

Die patientenidentifizierenden Daten existieren gemäß Datenflussmodell in drei Ausprägungen:

- im Klartext (Schnittstelle KKR)
- pseudonymisiert (Schnittstelle BAS)
- (pseudonymisiert und) verschlüsselt (Schnittstelle KKR-VST, VST-BAS)

Bei dieser Konzeption werden alle Ausprägungen vom selben Basisdatentyp geerbt, in einem zweiten Schritt die drei Ausprägungen konkretisiert und angepasst, und über „includes“ in die jeweilige Schnittstelle integriert.

---

### 6.1.2 Schnittstellen

In der folgenden Tabelle werden Schema-Dateien aufgeführt, die im Rahmen der Übermittlung von Sozialdaten Verwendung finden. Andere Dateien haben zwar ebenfalls die Dateiendung .xsd, sind aber keine vollständigen Schemata, sondern Bausteine für Schnittstellen.


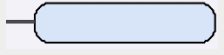
Tabelle 19: Verwendbare Schemata, Ablageort und Verwendungszweck

Schnittstelle <sup>4</sup> / Ablageort	Schema	Verwendungszweck
Krebsregister Interface_KKR	interface_KKR.xsd	Validierung der Exportdatei
KKR/VST Interface_KKR_VST	interface_KKR_VST.xsd	Datenübermittlung vom Krebsregister an die VST
	response_VST_KKR.xsd	Datenflussprotokoll von der VST an das Krebsregister
	response_receipt_VST_KKR.xsd	Empfangsbestätigung von der VST an das Krebsregister
VST/BAS Interface_VST_BAS	interface_VST_BAS.xsd	Datenübermittlung von der VST an die BAS
	response_BAS_VST.xsd	Datenflussprotokoll von der BAS an die VST
	response_receipt_BAS_.xsd	Empfangsbestätigung von der BAS an die VST
BAS/KKR Interface_BAS_KKR	response_BAS_KKR.xsd	Datenflussprotokoll von der BAS an das Krebsregister
BAS	interface_BAS.xsd	Datenstruktur der Datenübermittlung nach der Entschlüsselung der Exportdateien


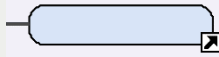


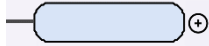

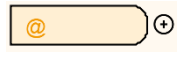
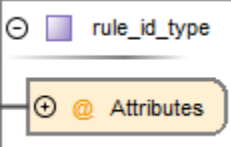
### 6.1.3 XML-Struktur des Datencontainers (KKR-Schnittstelle)

Zur Veranschaulichung der verwendeten XML-Schemata werden Diagramme verwendet, deren Symbole in der folgenden Tabelle kurz dargestellt und erläutert werden.

Tabelle 20: Symbole in den XML-Schema-Diagrammen

Symbol	Beschreibung
	<b>Optionales Element</b> Kardinalität 0..1 („0 oder 1“)
	<b>Obligatorisches Element</b> Kardinalität 1: das Element muss genau einmal vorkommen

<sup>4</sup> Die Zielgruppe ist aus der Schnittstelle zu entnehmen.

	<b>Mehrfach wiederholbares Element</b> Kardinalität: die erlaubte Anzahl der Elemente wird unter dem Symbol dargestellt (Beispiel: 1..n, n..m).
	<b>Referenzelement</b> Das referenzierte globale Element ist an anderer Stelle im Schema definiert.
	<b>Eine Folge von Elementen</b> Die Elemente müssen genau in der Reihenfolge vorkommen, in der sie im Schemadiagramm angezeigt sind.
	<b>Eine Auswahl von Elementen</b> Nur ein einziges Element aus der Liste kann ausgewählt werden.
	Ein Element mit Kind-Elementen
	<b>Optionales Attribut</b> Kardinalität 0..1 („0 oder 1“)
	<b>Obligatoisches Attribut</b> Kardinalität 1: das Attribut muss genau einmal vorkommen
	<b>Komplexer Datentyp</b> Der komplexe Datentyp wird mit einem Rahmen mit einem weißen bzw. lilafarbenen Hintergrund angezeigt.

Die wesentlichen Bestandteile der XML-Schemata werden in tabellarischer Form dargestellt. Die Darstellung umfasst folgende Eigenschaften des betrachteten Elements:

- Grafische Abbildung der Kind-Elemente und -Attribute
- Auflistung der Kind-Elemente
- Auflistung der Attribute sowie ihre Eigenschaften wie:
  - Name
  - XML-Datentyp (technische Bezeichnung: „Type“)
  - Muss-Kann-Feld (technische Bezeichnung: „Use“)
  - Konstante (technische Bezeichnung: „Fixed“)
  - Kurze Beschreibung (technische Bezeichnung: „Annotation“)

Grundsätzlich beginnt jede XML-Exportdatei mit einer Headerzeile (`<?xml version="1.0" encoding="utf-8" ?>`), gefolgt vom Wurzelement `<root>`, das den gesamten Inhalt einschließt. Als Zeichensatz wird „UTF-8“ verwendet.



Wurzelement <root>

Das Wurzelement ist eine Art Umschlag für alle XML-Typen in den Verfahren. Das Wurzelement besteht immer aus zwei Kind-Elementen (Zweigen) `<header>` und `<body>`.

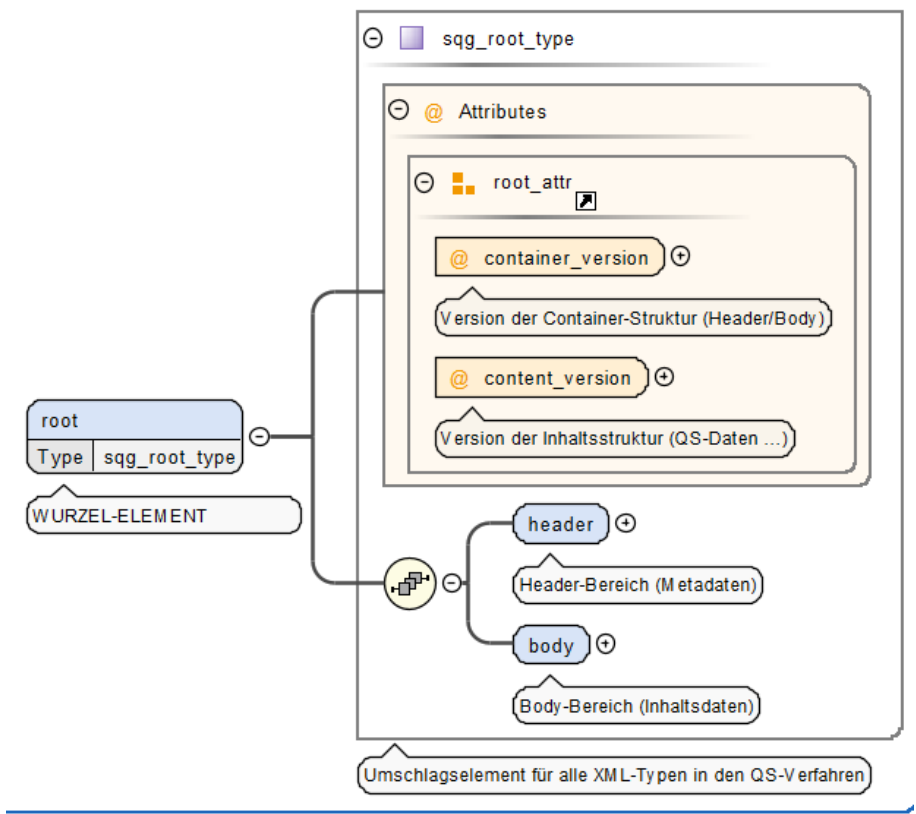


Abbildung 7: root-Element und Kind-Elemente header und body

Das `root`-Element hat zusätzlich zwei Attribute:

Tabelle 21: Attribute: root-Element

Attribut	Type	Use	Fixed	Beschreibung
container_version	xs:string	required	2.0	Ist ein fixer Wert und definiert die aktuell gültige Versionsnummer des Containers. Die Versionsnummer wird erhöht, wenn Änderungen am Schema des Containers (Umschlags) gemacht werden. Bei optionalen Änderungen wird die Versionsnummer beibehalten, um die Aufwärtskompatibilität zu gewährleisten.
content_version	content_version_type	required		Ist ein fixer Wert und definiert die aktuell gültige Versionsnummer des Inhalts.

Header-Bereich <header>

Das Element <header> besteht aus Metadaten (administrativen und meldebezogenen Daten) zu den eigentlichen Tumordaten, die im <body> enthalten sind.

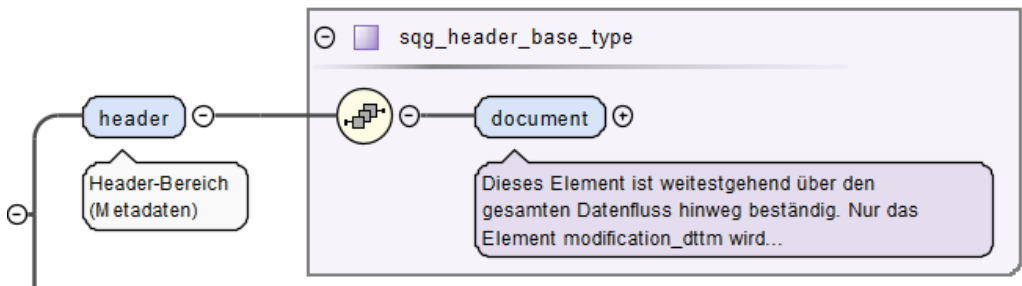


Abbildung 8: Header-Element und Kind-Element document

Das Element header hat ein Kindelement document welches wiederum diverse Kindelemente aufweist, die für den Datenfluss notwendigen Metainformationen beinhalten und in Tabelle 22 genauer beschrieben sind.

Tabelle 22: Kindelemente: document-Element

Element	Type	Kardinalität	Beschreibung
id	guid_type	1	Eindeutige ID einer Lieferung, wird vom Dokumentenersteller als GUID5 erzeugt. Nach Erstellung nicht mehr modifizierbar.
set_id	guid_type	1	Eindeutige ID für sämtliche Dokumente, einer Transaktion; wird vom KKR erzeugt. Nach Erstellung nicht mehr modifizierbar.  Für die Erstellung von transaktionsbezogenen Datenflussprotokollen für die KKR wird von der BAS diese set_id verwendet.
origination_dttm	xs:dateTime	1	Dieses Element ist der Zeitstempel der ursprünglichen Dokumenterzeugung.  Es darf nach seiner Erstellung nicht mehr modifiziert werden.
modification_dttm	xs:dateTime	1	Dieses Element ist ein Zeitstempel und muss bei jeder Modifikation des Dokuments aktualisiert werden.

<sup>5</sup> Ein Globally Unique Identifier oder kurz GUID ist eine global eindeutige Zahl mit 128 Bit (16 Bytes).

Element	Type	Kardinalität	Beschreibung
			Das Modifikationsdatum darf nicht vor dem Erstelldatum liegen.
validity_start_date	date_woTime_type	1	Beginn des Bezugszeitraums der übermittelten Daten (gemäß Vorgabe der jeweiligen verfahrensbezogenen Spezifikation).
validity_end_date	date_woTime_type	1	Ende des Bezugszeitraums der übermittelten Daten (gemäß Vorgabe der jeweiligen verfahrensbezogenen Spezifikation).
specification	enum_spez_type	1	Angabe der zum Export eingesetzten Version der Spezifikation. Hieraus ergeben sich insbesondere die eingesetzten Filterkriterien und Kodelisten.
data_flow	enum_flow_type	1	Dieses Attribut gibt an, für welchen Datenfluss dieses Dokument erzeugt wurde. Für die Übermittlung der Krebsregisterdaten ist der Wert „Krebsregisterdaten“ einzutragen.
data_target	enum_target_type	1	Dieses Element gibt an, welches Ziel der Datenfluss hat. Mögliche Werte sind:  Echtdatenpool: Bei Echtdaten für den Echtbetrieb (Regelbetrieb) Probedatenpool: Bei Echtdaten für vorläufige Auswertungen Testdatenpool: Bei Testdaten für Testzwecke  Die Kennzeichnung „Echtdatenpool“ ist dabei ausschließlich für den Echtbetrieb vorgesehen.
total_count	non_negative_integer_type	1	Gesamtzahl der exportierten case-Elemente in allen XML-Dateien einer Transaktion.

### Element header/protocol

Das Element `<protocol>` wird im Abschnitt 6.1.5 näher erläutert.

### Body-Bereich `<body>`

In dem Element `<body>` liegen die eigentlichen Nutzdaten:

- Patientenidentifizierende - bzw. Tumordaten (Stamm- und Behandlungsdaten)

■ krebsregisteridentifizierende Daten

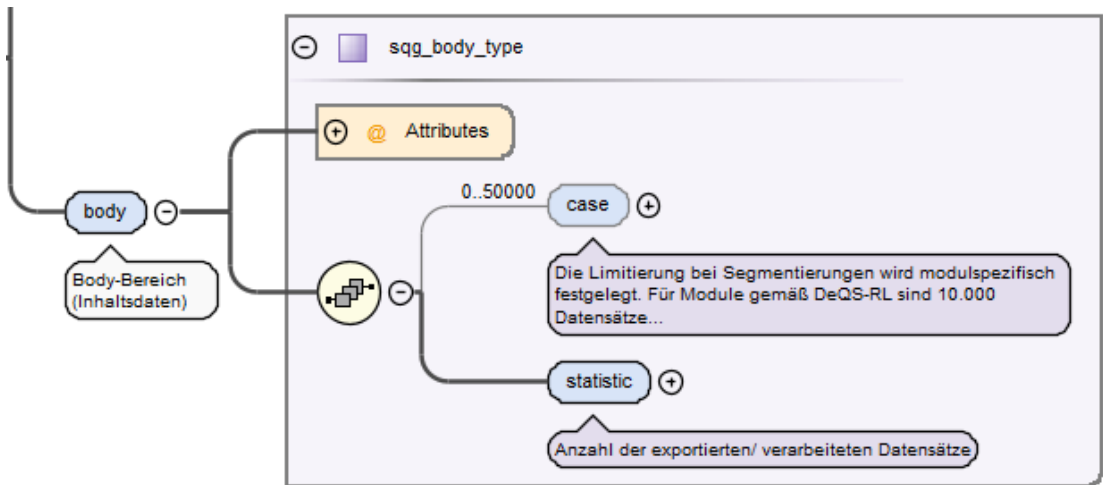


Abbildung 9: body-Element und Kind-Elemente case und statistic

Das body-Element hat folgende Attribute:

Tabelle 23: Attribute: body-Element

Attribut	Type	Use	Beschreibung
krebsregister_ik	kk_r_iknr_simple_type	required	Das Institutionskennzeichen des Krebsregisters
module	enum_modul_type	required	Enthält das Modul, aus dem die Datensätze stammen
pseud_procedure	enum_procedure_type	required	Enthält die Verfahrenskennung, die zusammengehörende Exportmodule vermerkt. Die PID-Felder eines Pseudonymisierungsverfahren müssen von der Vertrauensstelle mit demselben Schlüssel pseudonymisiert werden.

Element body/case

Das Element `<case>` entspricht einem Treffer des Patientenfilters und enthält genau einen Patienten-Datensatz.

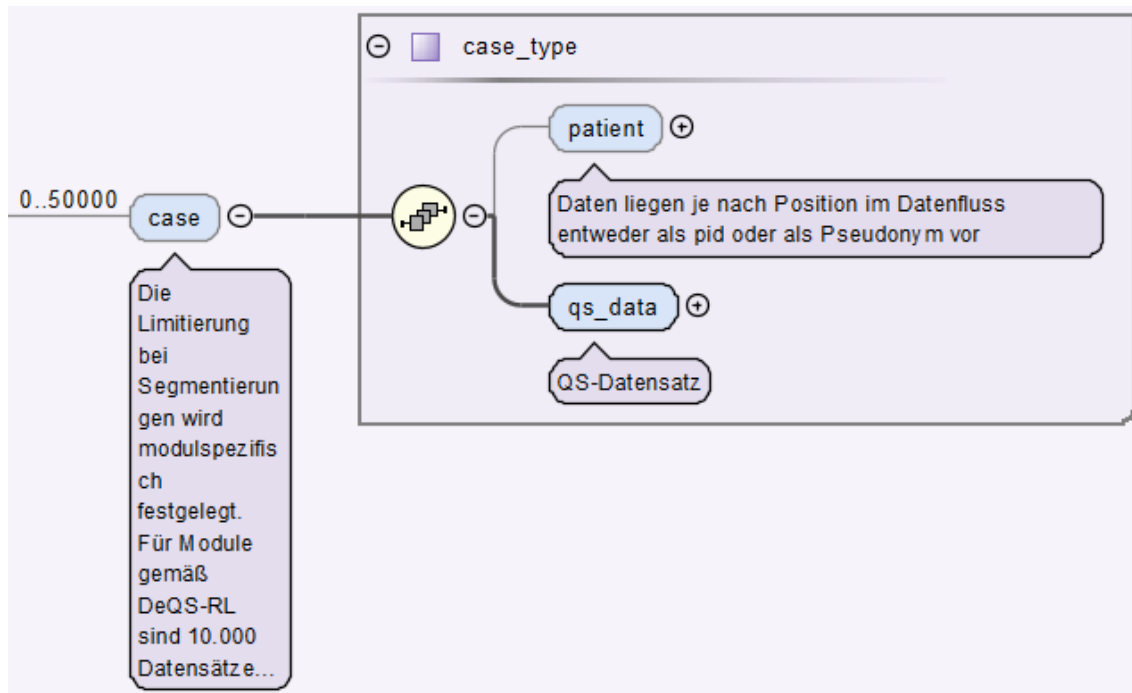


Abbildung 10: case-Element und seine Kind-Elemente patient und qs\_data

### Element body/statistic

Das Element dient dazu, Statistiken über die Datenlieferung des Absenders und über deren Verarbeitung durch die Auswertungsstelle aufzunehmen. Das Element **<statistic>** besteht aus einem **<sent>**- und einem **<processed>**-Block und ist nach der Prüfung bei der Auswertungsstelle Teil des Rückprotokolls.

Der **<sent>**-Block enthält dabei Informationen, wie viele Datensätze das Krebsregister in seiner Datenlieferung übermitteln möchte, während der **<processed>**-Block Informationen über die tatsächlich verarbeitbaren Datensätze enthält. Im Idealfall, also bei erfolgreicher Datenübermittlung stimmen die Angaben im **<processed>**-Block mit denen im **<sent>**-Block überein. Wurden hingegen fehlerhafte Datensätze von der Auswertungsstelle ausgemacht, werden diese im Element **<corrupted>** gezählt. Die Summe der Datensätze im processed-Block (corrupted + create) entspricht der Gesamtsumme der übermittelten Datensätze einer Lieferung durch ein Krebsregister.

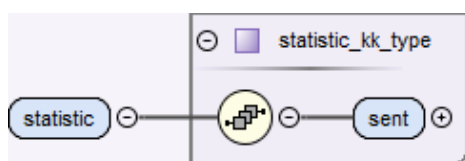


Abbildung 11: Element statistic und Kind-Element sent

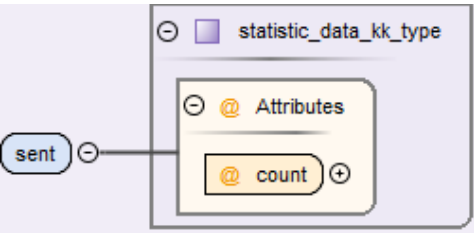


Abbildung 12: Element sent

Tabelle 24: Attribute: sent-Element

Attribut	Type	Use	Beschreibung
count	non_negative_integer_type	required	Gesamtzahl von Datensätzen in der XML-Datei, ist vom Datenlieferant auszufüllen

6.1.4 XML-Struktur der Exportdaten

Element body/case/qs\_data

Das Element `<qs_data>` ist ein Container für die verfahrensspezifischen Patientendaten. Dieser Datencontainer darf nur ein einziges Verfahren enthalten, es dürfen also innerhalb einer Datenlieferung keine Module gemischt bzw. kombiniert werden. Das Verfahren enthält einen weiteren Datencontainer, der dem Modulkürzel des Verfahrens entspricht (DARMK, ZERVIXK). Die Inhalte der einzelnen Verfahren können aus der Spezifikationsdatenbank bzw. ihrer begleitenden Dokumentation entnommen werden. Im Folgenden wird das Verfahren DARMK als Beispielverfahren für die Visualisierung der XML-Struktur verwendet.

Element qs\_data/[MODULKÜRZEL]/patientenstammdaten

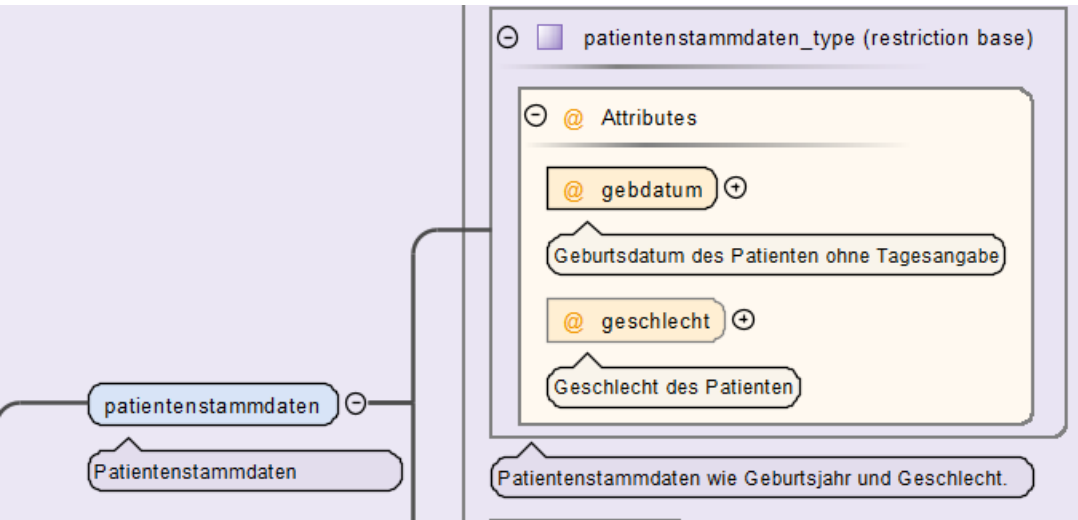


Abbildung 13: Element patientenstammdaten

Tabelle 25: Attribute: patientenstammdaten--Element

Attribut	Type	Use	Beschreibung
<u>geschlecht</u>	enum_geschlecht_type	optional	Geschlecht
<u>gebdatum</u>	mondatum_type	required	Geburtsjahr und Monat

Die Angabe des Geschlechts ist für das Modul ZERVIXK nicht vorgesehen. Deshalb ist das Attribut im Basistyp des patientenstammdaten\_type auf `optional` gesetzt. Über Erweiterungen wird dann die jeweilige use-Information (z. B. `prohibited` für ZERVIXK) festgelegt.

### 6.1.5 XML-Struktur der Rückprotokollierung (Datenflussprotokoll)

Für die Protokollierung der Prüfungen und deren Ergebnisse gibt es auf Dokumentenebene im Header und auf Fallebene im `<case_admin>` das Element `<protocol>`.

Auf Dokumentenebene sind alle Prüfungen zu dokumentieren, einschließlich der Prüfungen, die ausschließlich die Datensatzebene betreffen. Eine prüfende Einrichtung trägt sich als `<validation_provider>` in die entsprechende Auflistung ein und dokumentiert dann die von ihr durchgeführten Prüfungen in der Auflistung `<validation_item>`.

Sollte die ursprüngliche XML-Datei nicht lesbar (z. B. nicht entschlüsselbar oder nicht wohlgeformt) sein, ist ein reduziertes Protokoll („Miniprotokoll“) zu erstellen.

Das Miniprotokoll unterscheidet sich zum vollständigen Protokoll in folgenden Punkten:

- Es enthält kein Header- und kein body-Element (Abbildung 14).
- Die ursprüngliche ID (GUID) der Export-Datei wird nur angegeben, wenn sie zur Verfügung steht und lesbar ist.

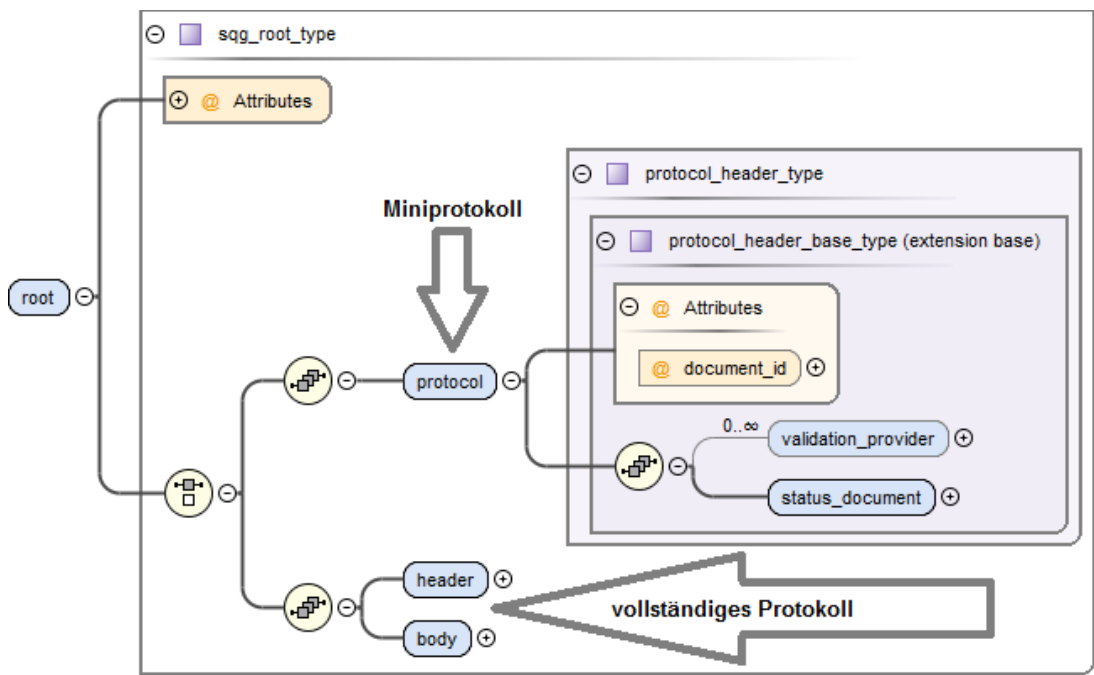


Abbildung 14: Schema des Datenflussprotokolls „response\_BAS\_KKR.xsd“

Nähere Informationen zur Rückprotokollierung können dem Abschnitt 6.3 entnommen werden. Die XML-Struktur der Rückprotokollierung für die Dokumenten- und die Fallebene ist identisch und wird hier näher erläutert:

Element header/protocol

Das Element `<protocol>` im ursprünglichen XML-Dokument mit den QS-Daten nimmt Informationen zu Prüfungen auf, die im Datenfluss durchgeführt wurden. Es ist Teil der Rückprotokollierung. Es soll gemeinsam mit dem Unterelement `<status_document>` vorhanden sein, um im nachfolgenden Datenfluss vorgenommene Prüfergebnisse aufzunehmen.

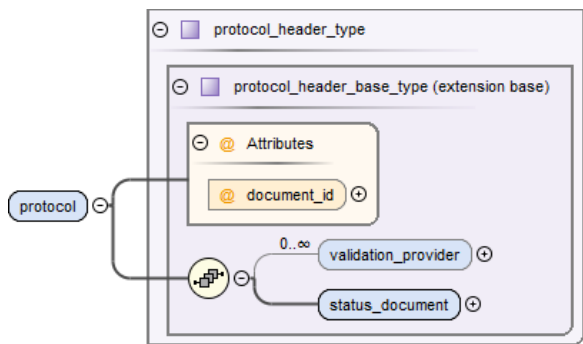


Abbildung 15: Element protocol und Kind-Elemente validation\_provider und status\_document

Tabelle 26: Attribute: protocol-Element

Attribut	Type	Use	Beschreibung
----------	------	-----	--------------



document_id	xs:string	optional	Wenn die ursprüngliche XML-Datei für die Rückprotokollierung nicht geeignet ist, muss die GUID des Exportdokuments in das Attribut document_id eingetragen werden, falls diese lesbar ist. Das Attribut ist also vornehmlich für Miniprotokolle notwendig.
-------------	-----------	----------	--

Element header/protocol/validation\_provider

Hier wird diejenige Stelle verzeichnet, die einen oder mehrere Prüfschritte durchgeführt hat. Die Ergebnisse der Prüfung werden in diesem Container abgelegt und Teil der Rückprotokollierung.

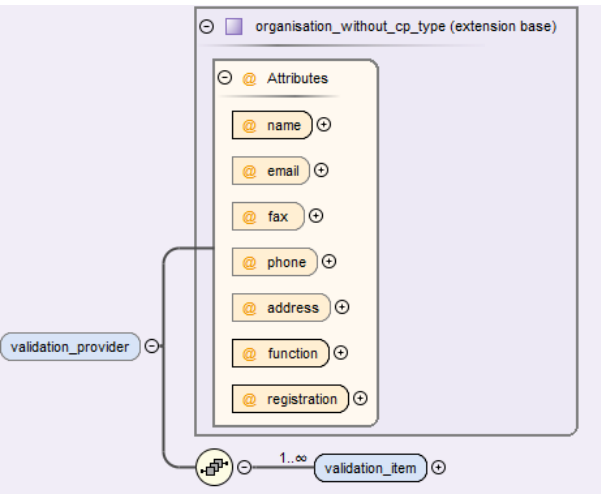


Abbildung 16: Element validation\_provider und Kind-Element validation\_item

Tabelle 27: Attribute: validation\_provider-Element

Attribut	Type	Use
name	xs:string	required
email	emailAddress_type	optional
fax	xs:string	optional
phone	xs:string	optional
address	xs:string	optional
function	enum_validation_provider_type	required
registration	registration_type	required

Element header/protocol/validation\_provider/validation\_item

Auf Dokumentenebene sind alle Prüfungen zu dokumentieren. Eine prüfende Einrichtung trägt sich als <validation\_provider> in die entsprechende Auflistung ein und dokumentiert dann die von ihr durchgeführten Prüfungen in der Auflistung <validation\_item>.

Als Ergebnis jeder Prüfung wird eine der folgenden Aussagen über das geprüfte Dokument getroffen:

- OK (Keine Auffälligkeiten)
- WARNING (Auffälligkeiten, die einer Weiterverarbeitung nicht im Weg stehen)
- ERROR (Auffälligkeiten bzw. Fehler, die eine Weiterverarbeitung des Dokuments ausschließen)

Das Ergebnis der Prüfung wird in das Attribut „V“ des Elements <status> im Element <validation\_item> eingetragen.

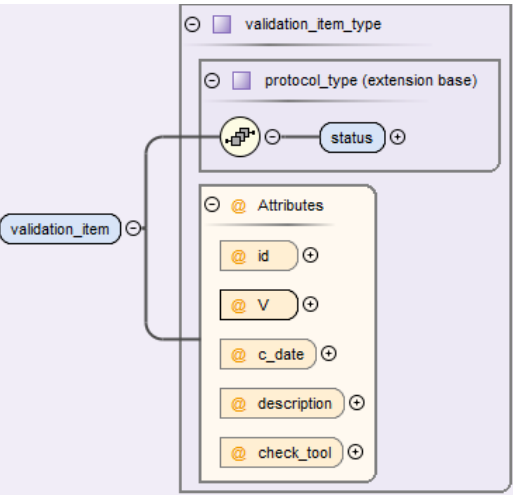


Abbildung 17: Element validation\_item und Kind-Element status

Das Element hat folgende Attribute:

Tabelle 28: Attribute: validation\_item-Element

Attribut	Type	Use	Beschreibung
id	xs:int	optional	Die ID muss nur dann vergeben werden, wenn eine Prüfung auf Datensatzebene stattfindet. Diese ID ist dokumentweit gültig und darf im Header nur einmal vorkommen. Prüfungen auf Datensatzebene (Element "case"), die zu dieser Prüfung gehören, werden über diese ID zugeordnet.
V	enum_validation_type	required	Dieser Wert bezeichnet die durchgeführte Prüfung anhand einer

Attribut	Type	Use	Beschreibung
			„enumeration“, die in sqg_header_protocol_base.xsd definiert wird.  Gültige Werte sind: Dechiffrierung, PID_Pseudonym, Schema, Spezifikation, Transaktion, sonstige Prüfung
c_date	smalldatetime_simple_type	optional	Hier kann ein Zeitstempel für die Verarbeitung angegeben werden.
description	xs:string	optional	Hier kann bei erklärungsbedürftigen Items eine Beschreibung hinterlassen werden (auf Dokumentenebene).
check_tool	xs:string	optional	Hier können beim Einsatz eines Tools für die Prüfung der XML-Dateien der Name und die Versionsnummer des Tools hinterlassen werden.

Element header/protocol/status\_document

Hier wird der Gesamtstatus des Dokuments angegeben. Das Attribut v kann also auf „OK“, „WARNING“ oder „ERROR“ stehen. Dieser Status kann nur geändert werden, wenn sich der Status des Dokuments im Laufe der Prüfung verschlechtert. „ERROR“ hat zur Folge, dass das Dokument komplett zurückgewiesen werden muss.

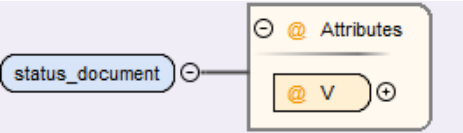


Abbildung 18: Element status\_document

Tabelle 29: Attribute: status\_document-Element

Attribut	Type	Use	Beschreibung
v	status_type	required	Mögliche Werte: OK/WARNING/ERROR

Element header/protocol/validation\_provider/validation\_item/status

Hier wird angegeben, ob die betroffene Prüfung ohne Fehler („OK“), mit Fehlern („WARNING“) oder mit fatalem Fehler („ERROR“) abgeschlossen wurde. Der Gesamtstatus des Dokuments entspricht jeweils dem schlechtesten Prüfergebnis. Bei der ersten Prüfung mit dem Ergebnis „ERROR“ muss die Weiterverarbeitung abgebrochen werden.

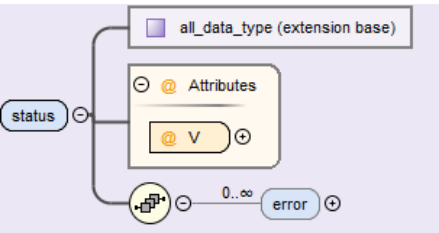


Abbildung 19: Element status und Kind-Element error

Tabelle 30: Attribute: Status-Element

Attribut	Type	Use
V	status_type	required

Darüber hinaus gibt es die Möglichkeit, eine beliebige Anzahl von `<error>`-Elementen mit einer `<error_message>` im `<status>`-Element unterzubringen.

Element header/protocol/validation\_provider/validation\_item/status/error

Ein `<error>`-Element nimmt Fehlerdaten auf. Als einziges verpflichtendes Unterelement gilt das `<error_message>`-Element. Die Elemente `<rule_id>` und `<rule_type>` sind spezifisch für die Anwendung der in der Spezifikationsdatenbank definierten Prüfungen (Tabelle Prüfung).

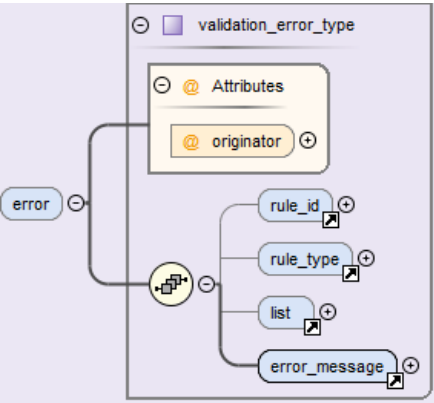


Abbildung 20: Element error und Kind-Elemente rule\_id, rule\_type, list und error\_message

Tabelle 31: Attribute Error-Element

Attribut	Type	Use	Beschreibung
originator	enum_organisation_type	optional	Mögliche Werte: Bundesauswertestelle, Vertrauensstelle, Krebsregister, undefined.

Das Element hat folgende Kind-Elemente:

Tabelle 32: Attribute: Error-Element

Element	Beschreibung
rule_id	Nummer der Prüfung bzw. der Fehlermeldung (Fehlerkode)
rule_type	Werte H (=hart), W (=weich) oder D (harte Prüfung, die zunächst weich interpretiert wird)
list	Liste der fehlerhaften Felder. Das Trennzeichen ist „ “ (siehe Abschnitt 6.3.3)
error_message	Fehlermeldung aus der Spezifikationsdatenbank als Freitext

### Element case/case\_admin/protocol

Dieses Element hat auf Datensatzebene eine ähnliche Struktur wie das oben beschriebene Element `<protocol>`.

Die Unterschiede sind:

- Während das `<protocol>`-Element im Header Ergebnisse der Prüfungen, die das Dokument insgesamt betreffen, aufnimmt, nimmt das `<protocol>`-Element im Body-Bereich die Ergebnisse der Prüfungen auf, die auf Datensatzebene erfolgen.

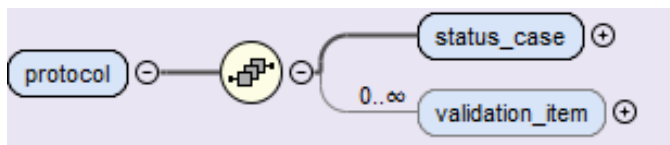


Abbildung 21: Element protocol und Kind-Elemente validation\_item und status\_case

- `<status_case>` beinhaltet jeweils das schlechteste Ergebnis aller auf Datensatzebene.
- Das Protokoll auf Datensatzebene hat kein Element `<validation_provider>` (Prüfstelle). Damit auch auf dieser Ebene die Ergebnisse der durchgeführten Prüfungen einer Prüfstelle zugeordnet werden können, müssen alle Ergebnisse einer Prüfung auf Fallebene mit einer gemeinsamen, dokumentweit eindeutigen ID im Attribut ID des Elements `<validation_item>` eingetragen werden.

### 6.1.6 Umgang mit Umlaut-Domains und E-Mail-Adressen

Die Nutzung von Umlauten ist durch die Internationalized Domain Names (IDN) möglich. Domainnamen wie z. B. <http://www.qualitätsklinikverbund.de/> sind damit nutzbar. Da die Verwendung von Zeichen aus einem Unicode-Zeichensatz in Domainnamen und E-Mail-Adressen immer noch Probleme in Anwendungsprogrammen bereiten, wurde das ASCII Compatible Encoding (ACE) entwickelt, in dem Regeln zur Überführung von Unicode-Zeichen in ASCII-Zeichen festgelegt wurden. Zu jeder Umlautdomain (IDN-Domain) wird parallel immer eine ACE-Domain registriert. Für das oben genannte Beispiel <http://www.qualitätsklinikverbund.de/> würde der ACE-Name [xn--qualittsklinikverbund-b2b.de](http://xn--qualittsklinikverbund-b2b.de) lauten.

E-Mail-Adressen dürfen im lokalen Teil (vor dem @-Zeichen) keine Umlaute enthalten. Die E-Mail-Adresse müller@muster.de wäre nicht möglich. Möglich ist lediglich mueller@muster.de. Es existieren experimentelle Request for Comments (RFCs)<sup>6</sup>, in denen UTF-8-codierte, internationale E-Mail-Adressen ermöglicht werden. Dies wird jedoch nicht flächendeckend im Produktiveinsatz genutzt, aus diesem Grund wird in dieser Spezifikation von der Annahme ausgegangen, dass der lokale Teil der E-Mail-Adresse immer durch ASCII-Zeichen repräsentiert wird.

In den E-Mail-Elementen des XML-Schemas sind ausschließlich Zeichen des ASCII-Zeichensatzes erlaubt. E-Mail-Adressen mit einem IDN Domainnamen sind entsprechend den ACE-Regeln durch ASCII-Zeichen darzustellen und zu übermitteln.

#### **6.1.66.1.7 Versionierung der XML-Schemata**

Das Schemamodell besteht im Wesentlichen aus zwei Teilstrukturen:

1. ein über die Zeit stabiler Umschlag (Datencontainer)
2. ein sich durch geänderte oder neu hinzugekommene Verfahren ändernder Inhalt

Da zu 1. eine spezifikationsübergreifende Konstanz erwartet wird – mit Versionswechseln unabhängig von der Spezifikation – und zu 2. ggf. eine jährliche mit der Spezifikation gekoppelte neue Version, soll die Versionskontrolle über zwei voneinander unabhängige Mechanismen erreicht werden:

- Für den Umschlag (Datencontainer) wird im Schema das Attribut `/root/@container_version` erzwungen und mit einem Fixwert besetzt. Eine Änderung im Schema, die nicht abwärtskompatibel ist, wäre ein Grund für einen Versionswechsel.
- Für die spezifikationsabhängige Struktur wird das Attribut `/root/@content_version` erzwungen und mit einem Fixwert besetzt. Hier gilt für den Versionswechsel im Prinzip dasselbe wie bei der „container\_version“. Die Versionsnummer wird jedoch nur erhöht, wenn der XML-Inhalt unabhängig von der zentralen Spezifikation geändert werden muss.

## **6.2 XML-Verschlüsselung**

Der Datentransport erfolgt über eine verschlüsselte HTTPS-Verbindung. Die Autorisierung erfolgt dabei über ein von der VST bereitgestelltes Client-Zertifikat. Der `qs_data`-Block, der die sensiblen medizinischen Daten enthält, ist zudem XML-verschlüsselt.

### **Verschlüsselungsalgorithmen**

Für die Verschlüsselung der XML-Knoten wird die hybride Verschlüsselung nach dem W3C-Standard „XML Encryption Syntax and Processing“ verwendet.

---

<sup>6</sup> <https://tools.ietf.org/html/rfc5336>

Als Verschlüsselungsalgorithmen werden „AES128“ für die symmetrische Verschlüsselung der XML-Elemente und „RSA mit 2048-Bit“ für die asymmetrische Verschlüsselung des generierten symmetrischen Schlüssels verwendet.

## 6.3 Die Rückprotokollierung

Es existieren drei Arten von Rückmeldungen: Die Empfangsbestätigung für den Datenabsender durch den Datenempfänger als Bestätigung des korrekten Empfangs einer Lieferung, das Datenflussprotokoll als Ergebnis der Gesamtverarbeitung bzw. immer dann, wenn eine weitere Verarbeitung als Folge von Fehlern nicht möglich ist sowie das Miniprotokoll, wenn der gesamte Datensatz nicht bearbeitbar ist.

### 6.3.1 Empfangsbestätigung

Die Empfangsbestätigung wird vom Datenempfänger zurück an den Datenabsender geschickt, sofern das angenommene Datenpaket in Bezug auf die Prüfungen des Datenempfängers verarbeitet werden konnte. Im Fall einer segmentierten Transaktion wird eine Empfangsbestätigung je Segment erstellt.

Wie Abbildung 22 zeigt, enthält die Empfangsbestätigung lediglich eine Benachrichtigung, dass die vom Datenabsender übermittelte XML-Datei bearbeitbar war und an die nachfolgende Stelle weitergeleitet wurde. Die enthaltene Dokument-ID bezieht sich auf die ID des Segmentes und ermöglicht die Zuordnung der Empfangsbestätigung zur jeweiligen Datenlieferung:

```

1 <?xml version="1.0" encoding="iso-8859-1"?>
2 <root xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" content_version="1.0"
3   container_version="2.0"
4   xsi:noNamespaceSchemaLocation="../../../interface_KKR_VST/response_receipt_VST_KKR.xsd">
5   <header>
6     <document>
7       <id v="{55664537-5642-9056-8676-456468327446}" />
8     </document>
9     <provider email="vertrauensstelle@test-vertrauensstelle.de" name="Test-Vertrauensstelle"
10      registration="bu00000" function="Vertrauensstelle" address="Test Adresse" />
11   </header>
12   <body>
13     <nachricht>
14       Ihre Dateneinsendung konnte erfolgreich eingelesen und an die Vertrauensstelle weitergeleitet werden.
15       Ein umfassendes Datenflussprotokoll über die Prüfergebnisse der Vertrauensstelle
16       und der Bundesauswertungsstelle erhalten Sie von uns in den nächsten 24 Stunden.
17     </nachricht>
18   </body>
19 </root>
20

```

Abbildung 22: Beispiel einer Empfangsbestätigung der VST an ein Krebsregister

Die Vertrauensstelle stellt für die Übermittlung von Datensätzen eine REST-API zur Verfügung. Beim Eingang eines Datensatzes werden verschiedene Prüfungen ausgeführt. Der Ausgang dieser Prüfungen entscheidet welcher Protokolltyp erzeugt wird (siehe Tabelle 33). Die Protokolle werden via E-Mail bereitgestellt.

Tabelle 33: Rückmeldungen der Vertrauensstelle

Ausgang der Prüfung	Rückmeldung
Datensatz nicht bearbeitbar	Miniprotokoll
Eingangsverarbeitung mit hartem Fehler	Datenflussprotokoll
Erfolgreiche Eingangsverarbeitung	Empfangsbestätigung

### 6.3.2 Datenflussprotokoll

Die vorgenommenen Prüfungen werden in den dafür vorgesehenen Bereichen im XML-Code des übermittelten Dokuments auf Header- und auf Datensatzebene protokolliert. Das Protokoll des Dokuments wird damit von jeder entgegennehmenden Stelle um deren Prüfergebnisse ergänzt.

Da das Datenflussprotokoll vom ursprünglichen XML-Dokument abgeleitet wird, müssen der `qs_data`-Block mit den medizinischen Daten und der `patient`-Block mit den patientenidentifizierenden Daten entfernt werden. Das übrigbleibende XML enthält innerhalb der ursprünglichen Struktur des Dokuments die bis dahin protokollierten Prüfungen und die sich daraus ergebenden Statusmeldungen der Datensätze und des Dokuments. Der Aufbau des Protokolls ist im Abschnitt 6.3.4 detailliert beschrieben.

Zur Visualisierung des Datenflussprotokolls kann dieses mithilfe einer Template-Definition z. B. nach HTML transformiert werden.

#### Miniprotokoll

Die Erstellung eines vollständigen Datenflussprotokolls ist nur möglich, wenn die ursprüngliche XML-Datei lesbar ist und nach Entfernung der PID und der QS-Datenblöcke schemakonform bleibt. Andernfalls ist ein reduziertes Protokoll (Miniprotokoll) zu erstellen, das die ID (GUID) des Dokuments (sofern diese zur Verfügung stand und lesbar war, sonst wird diese nicht angegeben) und die konkrete Fehlermeldung enthält. Die Struktur dazu ist Abschnitt 6.1.5 zu entnehmen.

### 6.3.3 Fehlermeldungen

Eine Fehlermeldung ist das Ergebnis einer negativ verlaufenen Prüfung. Die vorhandenen Prüfungen können der Spezifikationsdatenbank (Tabelle `Pruefung`) entnommen werden.

Sollte keine (administrative) Prüfung definiert sein, kann in begründeten Ausnahmefällen statt einer standardisierten Fehlermeldung eine vorläufige Fehlermeldung benutzt werden.

Fehlermeldungen sind Ausdruck des `<error_message>`-Elements innerhalb des `<error>`-Elements. Im Folgenden werden Aufbau und Attribute (Tabelle 34) sowie Kind-Elemente (Tabelle 35) des `<error>`-Elements beschrieben:

Tabelle 34: Aufbau und Attribute des Elements `<error>`

Name	Typ	Use	Beschreibung
------	-----	-----	--------------



originator	enum_organisation	optional	Mögliche Werte: Bundesauswertestelle, Vertrauensstelle, Krebsregister, undefined
------------	-------------------	----------	--

Tabelle 35: Kind-Elemente des <error>-Elements

Kind-Element	Beschreibung
<rule_id>	ID der Prüfung
<rule_type>	Werte H (=hart), D (=weich, eigentlich hart) oder W (= weich).
<list>	Pfad zu den an der Prüfung beteiligten Attributen (als Trennzeichen wird das „ “ Zeichen verwendet).
<error_message>	Fehlermeldung

Die Abbildung 23 stellt ein <error>-Objekt exemplarisch dar:

```
<protocol>
  <status_case V="WARNING"/>
  <validation_item V="Spezifikation" id="1">
    <status V="WARNING">
      <error originator="Krebsregister">
        <rule_id V="26"/>
        <rule_type V="W"/>
        <list V="/root/header/provider/@function"/>
        <error_message
          V="Die angegebene Funktion des Absenders (&lt;providerFunktion&gt;) stimmt
            nicht mit dem bei der VST hinterlegten Wert überein."
        />
      </error>
    </status>
  </validation_item>
</protocol>
```

Abbildung 23: Beispiel eines <error>-Elements

Dem Wert im Attribut `originator` ist zu entnehmen, dass der Fehler mit weicher Strenge (`<rule_type V="W" />`) mit der beispielhaften ID 26 (`<rule_id V="26" />`) durch die Krebsregister verursacht wurde. Die in der Fehlermeldung (`<error_message>`) erwähnten Datenfelder werden im Element `<list>` als Pfad referenziert.

Referenzierung der Datenfelder

Die vom Fehlerfall betroffenen Datenfelder werden als Pfadangabe mittels XPath beschrieben. Der Pfadausdruck geht dabei vom Enkelelement des zugehörigen `<case>`-Elements als aktuellem Kontextknoten aus und verwendet nur einfache Element- und Attributknoten, die jeweils mit „/“ getrennt sind. Die immer vorhandenen Containerelemente `<qs_data>` und Modulkennungs-Elemente (z. B. `<DARMK>`) werden also nicht explizit angegeben.

Tabelle 36: Verwendete XPath-Syntax

Syntax	Semantik	Beispiel
<code>element[n]</code>	<b>Elementknoten</b> n-tes Kindelement (beginnend mit [1]) mit dem angegebenen Namen. Wenn nur ein Kindelement dieses Namens vorhanden ist, kann die Angabe „[n]“ entfallen.	<b>op[2]</b> Bezeichnet das zweite <code>&lt;op&gt;</code> -Element unterhalb des Modulkennungs-Elements
<code>/</code>	<b>Trennzeichen</b> Trennt zwei Knoten im Pfadausdruck. Der rechte Knoten ist ein Kindelement des linken Knotens.	<b>op[2]/ops</b> Bezeichnet das (einzige) <code>&lt;ops&gt;</code> -Element unterhalb des zweiten <code>&lt;op&gt;</code> -Elements unterhalb des Modulkennungs-Elements
<code>@attribut</code>	<b>Attributknoten</b> Bezeichnet ein Attribut des vorherstehenden Elementknotens	<b>op/ @datum</b> Bezeichnet das Attribut „datum“ des (einzigen) <code>&lt;op&gt;</code> -Elements unterhalb des Modulkennungs-Elements

6.3.4 Prüfergebnisse und Ergebnisprotokollierung

Für die Protokollierung der Prüfungen und deren Ergebnisse gibt es auf Dokumentenebene im Header und auf Fallebene im `<case_admin>` das Element `<protocol>`.

Auf Dokumentenebene sind alle Prüfungen zu dokumentieren, einschließlich der Prüfungen, die ausschließlich die Datensätze betreffen. Die prüfende Einrichtung trägt sich jeweils als `<validation_provider>` in die entsprechende Auflistung ein und dokumentiert dann ihre durchgeführten Prüfungen in der Auflistung `<validation_item>`.

Prüfungen, die – wie z. B. die Schemakonformität – das Dokument insgesamt betreffen, sind ausschließlich im Headerbereich einzutragen.

Prüfungen, die auf Patientenebene erfolgen, müssen folgendermaßen protokolliert werden:

Die Auswirkung des Prüfungsergebnisses in Bezug auf das gesamte Dokument muss im `<header>` eingetragen werden.

Das Ergebnis der Fallprüfung muss in `<case_admin>` eingetragen werden.

Alle Ergebnisse einer Prüfung, die auf Patientenebene erfolgt, müssen mit einer gemeinsamen, dokumentweit eindeutigen ID im Attribut ID des Elements `<validation_item>` eingetragen

werden. Dadurch ist es möglich, über die ID eines Prüfergebnisses, die man auf Patientenebene findet, auf Dokumentenebene den `<validation_provider>` eindeutig zu identifizieren (siehe Abbildung 24):

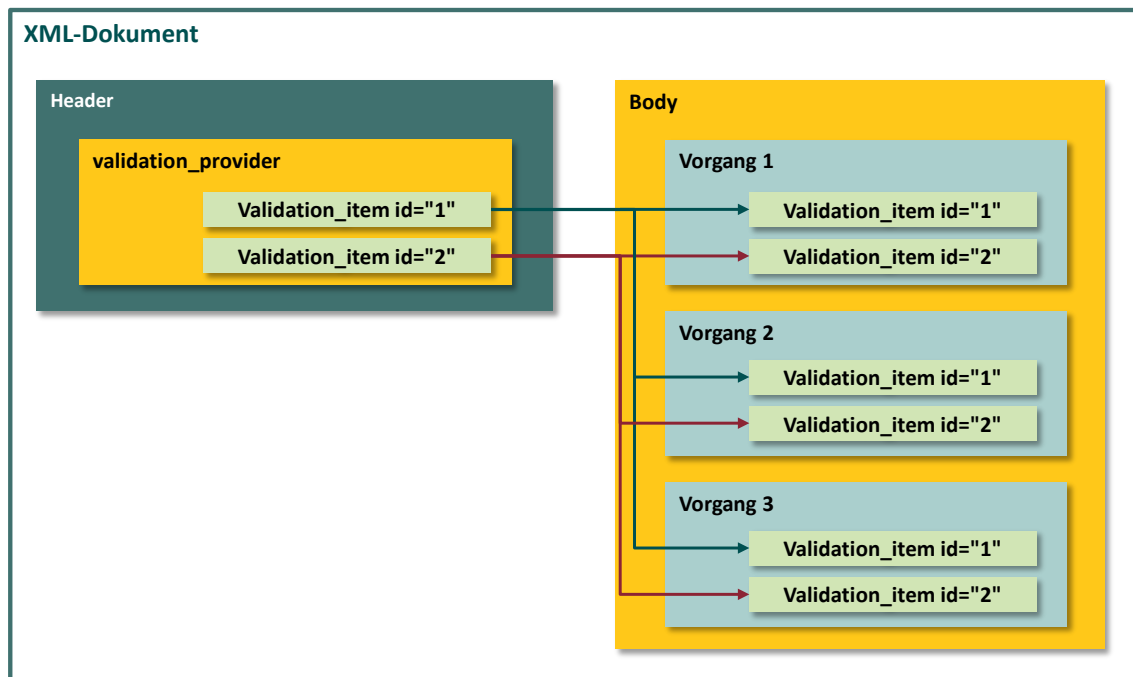


Abbildung 24: Beziehungen zwischen `<validation_item>` im header und `<validation_item>` im body über die id

Zur Veranschaulichung dieser Konstruktion wird im Folgenden eine Analogie aus dem relationalen Datenbank-Modell bemüht: die Dokumentenebene kann als Master-Tabelle, die Fallebene als Detail-Tabelle bezeichnet werden. Letztere enthält die zum Master gehörenden Detail-Datensätze, auf die über das Attribut „ID“ referenziert werden kann.

## Prüfergebnisse

Prinzipiell wird als Ergebnis jeder Prüfung eine der folgenden Aussagen über das geprüfte Objekt getroffen:

- Keine Auffälligkeiten
- Auffälligkeiten, die einer Weiterverarbeitung nicht im Weg stehen
- Auffälligkeiten bzw. Fehler, die eine Weiterverarbeitung des Objekts ausschließen.

In der datentechnischen Übersetzung wird dieses durch

- OK
- WARNING
- ERROR

ausgedrückt, die das Ergebnis der Prüfung im Attribut „V“ des Elements `<status>` im Element `<validation_item>` zusammenfassen.

Darüber hinaus gibt es die Möglichkeit, eine beliebige Anzahl von `<error>` Elementen mit einer `<error_message>` im `<status>` Element unterzubringen.

### Beziehung zwischen Vorgangsebene und Dokumentenebene

Es gibt zwei Kategorien geprüfter Objekte, die zueinander in einer hierarchischen Beziehung stehen:

- Erste Hierarchieebene: das gesamte Dokument
- Zweite Hierarchieebene: der `case` (jeweilige Patientendatensatz)

Jedes dieser Objekte hat einen Status in Bezug auf die Weiterverarbeitung, der sich aus den Ergebnissen der durchgeführten Prüfungen ergibt.

Auf Dokumentenebene ist dieser Status im Unterelement `<status_document>` von `<protocol>` im Attribut „V“ abgelegt.

Auf Patientenebene ist dieser Status ebenfalls in einem Attribut „V“ eines Unterelements von `<protocol>` abgelegt, welches hier aber `<status_case>` benannt wird.

In Bezug auf die Weiterverarbeitung gibt es folgende Regeln:

- Ein „ERROR“ als Ergebnis einer der Prüfungen verhindert die Weiterverarbeitung des geprüften Objekts.
- Eine oder mehrere Auffälligkeiten („WARNING“, „ERROR“) auf Fallebene bedeuten ein „WARNING“ in dem korrespondierenden Eintrag auf Dokumentenebene.
- Wenn bei einer fallbezogenen Validierung in allen Fällen auf Status „ERROR“ erkannt wird, muss auch für das Dokument abweichend von der Regel unter 2. der korrespondierende Eintrag auf Dokumentenebene auf Status „ERROR“ gesetzt werden.



#### Hinweis

Aufgrund der Übergangsregelung mit dem Status „D“ werden Prüfungen die eigentlich „hart“ geprüft werden nur „weich“ ausgewertet und der Status auf „WARNING“ statt „ERROR“ gesetzt.

---

- Der Status (`<status_case>` / `<status_document>`) eines Objekts kann nicht „besser“ sein als sein schlechtestes Prüfergebnis.

### Szenarien

Aus diesen Regeln abgeleitet, soll der Status jedes geprüften Objekts nach jeder Prüfung entsprechend dem Prüfergebnis aktualisiert werden. Daraus ergibt sich folgender Aktualisierungs- und Protokollierungsplan:

**Vor der Prüfung und Protokollierung:**

- ( 1 ) → Feststellen der höchsten ID in Bezug auf vorhandene `<validation_item>`-Elemente.
- ( 2 ) → Festgestellte ID um 1 inkrementieren und als ID der anstehenden Prüfung festlegen.

**Protokollierung der fallbezogenen Prüfung:**

Nachdem die fallbezogene Prüfung erfolgt ist, ist dies auf der Fallebene und der Dokumentenebene folgendermaßen zu protokollieren:

**Protokollierung auf Fallebene:**

- ( 3 ) → `<validation_item>` der Liste hinzufügen, dabei die unter (2) ermittelte ID verwenden.
- ( 4 ) → `<status_case>` des Falls auslesen.
- ( 5 ) → Ergebnis der Prüfung mit dem Status des Falls vergleichen. In den Fällen, in denen das Ergebnis der Prüfung schlechter ist als der aktuelle Status des Falls den Status mit dem Ergebnis der Prüfung aktualisieren.
- ( 6 ) → Falls ein Ergebnis der Prüfung schlechter ist als „OK“, muss dieses als dokumentbezogenes Ergebnis „WARNING“ vermerkt werden.

Protokollierung auf Dokumentenebene:

- ( 7 ) → `<validation_item>` mit dem unter (4) ermittelten Prüfungsergebnis der fallbasierten Prüfung unterhalb des Elements `<validation_provider>` eintragen. Falls `<validation_provider>` für die eigene Einrichtung noch nicht besteht, muss er angelegt werden.
- ( 8 ) → `<status_document>` auslesen.
- ( 9 ) → Das unter (6) ermittelte Gesamtergebnis der Prüfung muss mit dem Status des Dokuments verglichen werden. Falls das Ergebnis der Prüfung schlechter ist als der aktuelle Status des Dokuments, muss dessen Status mit dem Ergebnis der Prüfung aktualisiert werden.

**Protokollierung der dokumentenbezogenen Prüfung:**

Nachdem die dokumentenbezogene Prüfung erfolgt ist, ist dies auf der Dokumentenebene folgendermaßen zu protokollieren:

- ( 10 ) → `<validation_item>` mit dem Prüfungsergebnis unterhalb des Elements `<validation_provider>` eintragen. Falls `<validation_provider>` für die eigene Einrichtung noch nicht besteht, muss er angelegt werden.
- ( 11 ) → `<status_document>` auslesen.
- ( 12 ) → Das Ergebnis der Prüfung mit dem Status des Dokuments vergleichen und in dem Fall, in dem das Ergebnis der Prüfung schlechter ist als der aktuelle Status des Dokuments, dessen Status mit dem Ergebnis der Prüfung aktualisieren.

## 7 Releaseplanung

Die Spezifikation unterliegt einem definierten, mit dem G-BA, den klinischen Krebsregistern, der Gesundheitsforen Leipzig GmbH (GFL) und Vertrauensstelle abgestimmten Releasezyklus. Sollte es zu Abweichungen zwischen beschlossenen Spezifikationen oder Änderungen daran und den Regelungen der oKFE-RL kommen, sollten Richtlinie und Spezifikation zum nächstmöglichen Zeitpunkt wieder in Übereinstimmung gebracht werden. Bis dies erfolgt, folgt die technische Spezifikation im Zweifel den Vorgaben des Richtlinientextes.

### 7.1 Termine für die Releaseplanung

Um Planungssicherheit zu gewährleisten und angemessen auf Fehler reagieren zu können, werden die Termine zur Veröffentlichung von Spezifikationspaketen (Versionen) eines Spezifikationsjahres und zu Rückmeldefristen in der Technischen Dokumentation angekündigt. ~~Tabelle 37 stellt eine entsprechende Übersicht für das Paket der Spezifikation im ersten Jahr dar. Die Releaseplanung folgt dabei den Zielwerten aus Tabelle 37.~~

Die genannten Termine sind als Zielwerte zu betrachten und basieren auf Vorgaben des G-BA, Abstimmungen mit den Krebsregistern sowie Anforderungen aus der Umsetzung.

*Tabelle 37: Meilensteine der Die Releaseplanung der Spezifikation 2024 für das erste Jahr*

Frist	Spezifikation	Bereitstellung	Inhalte
15.05.2023	2024-V01	Veröffentlichung auf der IQTIG-Webseite ( <a href="http://www.iqtig.org">http://www.iqtig.org</a> )	Finale Version Frist abhängig vom Plenumsbeschluss des G-BA
November 2023	2024-V02	Veröffentlichung auf der IQTIG-Webseite ( <a href="http://www.iqtig.org">http://www.iqtig.org</a> )	Aktualisierung der ICD-Kodes, ggf. Fehlerkorrekturen; Frist abhängig vom Veröffentlichungszeitpunkt der endgültigen ICD-Kataloge für das Jahr 2024 durch das BfArM <sup>7</sup>
Februar 2024	2024-V03	Bei Bedarf	Fehlerkorrekturen

~~Ab dem zweiten Spezifikationsjahr (ab der Spezifikation 2025-V01) der Krebsregisterspezifikation folgt die Releaseplanung den Angaben aus Tabelle 38 (am Beispiel Spezifikation 2025). Regulär wird einem regulären Zyklus, wobei~~ die erste Version der Spezifikation eines neuen Spezifikationsjahres am 30. Juni des Vorjahres, nach Beschluss durch das Plenum, veröffentlicht, ~~um wird.~~ Dies ermöglicht den klinischen Krebsregistern und ggf. deren Softwareanbietern sowie allen im

<sup>7</sup> BfArM: Bundesinstitut für Arzneimittel und Medizinprodukte

Datenfluss beteiligten Instanzen den notwendigen Zeitraum für Implementierung und Testung-~~zu~~ ermöglichen.

Tabelle 37: Meilensteine der Releaseplanung ab der Spezifikation 2025

Spezifikation	Veröffentlichung	Inhalte
2025 V01	nach Beschluss durch das Ple-num (Juni 2024)	Finale Version (Frist abhängig vom Ple-numsbeschluss des G-BA)
2025 V02	November 2024	Update der finalen Version (Aktualisierung der ICD-Kodes, ggf. Fehlerkorrekturen); Frist abhängig vom Veröffentlichungszeit-punkt der endgültigen ICD-Kataloge für das jeweilige Jahr durch das BfArM <sup>8</sup>
2025 V03	Bei Bedarf	Fehlerkorrekturen

Über die aufgeführten Meilensteine zu Veröffentlichungsterminen hinaus erfolgt eine regelmä-Bige Abstimmung mit Krebsregistern und weiteren Beteiligten (z. B. Vertrauensstelle). Zudem wurden Möglichkeiten zur Bereitstellung von Alphaversionen getroffen, die die Qualität der Spe-zifikation erhöhen und die Richtlinienkonformität sicherstellen. Wesentliche Änderungen werden nur im Rahmen finaler Versionen (Version V01 einer Spezifikation) berücksichtigt.

Typen von Veröffentlichungen

- **Release**  
Grundsätzlich neue Funktionalität. Abstimmung erfolgt mit Verfahrensteilnehmern und dem G-BA.
- **Service Release**  
Korrekturen/kleine Erweiterungen zu Releases. Abstimmung erfolgt mit Verfahrensteilneh-mern und dem G-BA.
- **Patch**  
Aufgrund eines Notfalls (Fehler) kurzfristig erzeugtes Reparatur-Paket. Abstimmung erfolgt mit Teilnehmern und mit dem G-BA.

7.2 Fehler- und Changemanagement

Das IQTIG empfiehlt für die Optimierung der Zusammenarbeit mit den beteiligten Stellen die fol-genden Aktivitäten:

- Meldung von festgestellten Fehlern
- Verbreitung von Änderungsvorschlägen
- Abstimmung von Terminen und Umsetzungen im Rahmen der Releaseplanung

<sup>8</sup> BfArM: Bundesinstitut für Arzneimittel und Medizinprodukte

- Erfahrungsaustausch, um eine möglichst einheitliche Vorgehensweise zu ermöglichen
- Abstimmung der Spezifikationsänderungen

Vorschläge, Fehlermeldungen und Diskussionspunkte können per E-Mail an den Verfahrenssupport oder über die Kommunikationsplattform mitgeteilt werden.

Sollten Sie keine Zugangsdaten zur **Kommunikationsplattform** haben, lassen Sie sich bitte bei uns registrieren.

**Ihr Ansprechpartner:**

Herr Klaus Rittmeier

Telefon: (+4930) 58 58 26-340

Telefax: (+4930) 58 58 26-341

[verfahrenssupport@iqtig.org](mailto:verfahrenssupport@iqtig.org)



# Impressum

## HERAUSGEBER

---

IQTIG – Institut für Qualitätssicherung  
und Transparenz im Gesundheitswesen  
Katharina-Heinroth-Ufer 1  
10787 Berlin

Telefon: (030) 58 58 26-0

[info@iqtig.org](mailto:info@iqtig.org)

[iqtig.org](https://www.iqtig.org)