

Regie op Registers Dure Geneesmiddelen

Case studie 1: [PROSPECTIEF LANDELIJK CRC COHORT](#)

Juni 2023

Regie op Registers Dure Geneesmiddelen

Register voor patiënten met darmkanker

Auteur

Anja van Gestel
Heidi van Doorne
Geraldine Vink
Patrick Lubbers

Opdrachtgever

Zorginstituut Nederland

Revisie- en goedkeuringshistorie

Versie	Datum	Beschrijving aanpassing / opmerkingen	Revisie / akkoord door
0.1	05-12-2022	Eerste bevindingen	PL
0.2	03-01-2023	Teksten toevoegen	PL, AvG
0.3	17-01-2023	Eerste uitkomsten analyse	PL, AvG, GV
0.4	31-01-2023	Feedback werkgroep verwerkt	PL, GV, HvD
0.5	27-02-2023	Uitkomsten werksessie ZINL verwerkt	PL, FM, GV
0.6	03-03-2023	Uitkomsten analyses toevoegen	PL, FM
0.7	04-04-2023	Feedback verwerken	PL, GV, HvD
0.8	12-04-2023	Laatste feedback ronde	PL, GV, HvD
1.0	19-04-2023	Definitieve versie	Projectteam
2.0	05-06-2023	Terugkoppeling ZINL verwerkt	Projectteam

Versie 2.0

Inhoudsopgave

1	Introductie	6
1.1	Aanleiding	6
1.2	Structuur en doelstelling	6
1.3	Leeswijzer	6
2	Achtergrond	7
2.1	Het Prospectief Landelijk CRC cohort (PLCRC)	7
2.2	Patient Reported Outcome Measures (PROMS)	7
2.3	De Nederlandse Kankerregistratie (NKR)	7
2.4	IKNL en de afdeling Research & Development	8
3	Onderzoeksvraag	9
3.1	Vaststellen onderzoeksvraag en vertaling naar Proof of Concept	9
3.1.1	Populatie en selectiecriteria bepalen	9
3.1.2	Gegevensset bepalen	11
3.1.3	Kwaliteit van leven	11
3.1.4	Behandellijnen	12
3.2	Groepering van middelen	13
4	Gegevensset en data-items	16
4.1	Bepalen kwaliteit, volledigheid en beschikbaarheid	16
4.2	Onderscheid user en systeem gegenereerde data	17
4.3	Noodzaak en uitdagingen bij koppelen databronnen	18
5	IT infrastructuur	20
5.1	Wat is federated learning?	20
5.2	Onderscheid verticale en horizontale data	21
5.3	Proof of Concept opstelling	21
6	Uitkomsten en conclusies	23
6.1	Aanpassing initiële PROMS onderzoeksvraagstelling	23
6.2	Analyse totale overleving populatie	25
6.3	Conclusie en vervolgstappen	27
6.3.1	Conclusies	28
6.3.2	Vervolgstappen	29
	Bijlagen	31
B1	Leden werkgroepen	32
B2	Specificatie data-items	33
B3	Uitkomsten probabilistisch matchen	34
B4	Combinaties van middelen	35

1 Introductie

1.1 Aanleiding

Voordat nieuwe oncologische geneesmiddelen tot de Nederlandse markt worden toegelaten vindt er in (klinische) studies onderzoek plaats naar hun effectiviteit en veiligheid. Op basis van de uitkomsten van deze studies wordt door de European Medicines Agency (EMA) bepaald of een geneesmiddel op de Nederlandse markt toegelaten wordt. Het Zorginstituut Nederland (ZiNL) kan middels een duiding of een advies bepalen welke geneesmiddelen in Nederland worden opgenomen in het basispakket.

Uit onderzoek van eerdere registratiestudies komt naar voren dat de uitkomsten over effectiviteit en veiligheid in de dagelijkse praktijk vaak afwijken van de resultaten in studieverband. Dit wordt mede veroorzaakt doordat de samenstelling van de studiepopulatie minder divers is dan die van de algemene patiëntenpopulatie. ZiNL wil nagaan of registers meer inzicht kunnen geven in de effectiviteit en veiligheid van middelen in de dagelijkse praktijk, en of zij kunnen ondersteunen in cyclisch pakketbeheer.

Om na te gaan of registraties ZiNL kunnen ondersteunen bij pakketbeheer is het programma Regie op Registers Dure Geneesmiddelen (RORDGM) gestart. Dat programma wil met een aantal casestudies ervaring opdoen, en vanuit daar de mogelijkheden en beperkingen van registers voor het evalueren van geneesmiddelen in de dagelijkse praktijk onderzoeken. Voor casestudie 1 (bestaande patiëntenregistratie in de oncologie) is het voorstel vanuit het Prospectief Landelijk ColoRectaal Carcinoom cohort (PLCRC) van de Dutch Colorectal Cancer Group (DCCG) in samenwerking met Integraal Kankercentrum Nederland (IKNL) geselecteerd.

1.2 Structuur en doelstelling

Het doel van de casestudie is om in het kader van pakketbeheer in de dagelijkse praktijk ervaring op te doen met het gebruik van een Real World Data (RWD)¹ dataregister voor het beoordelen van een middel op diens effectiviteit en veiligheid. De casestudie kijkt naar;

- 1) het proces rondom de totstandkoming van een dataset die van voldoende kwaliteit is om analyses uit te voeren,
- 2) de eisen waaraan RWD moeten voldoen om de dataset conform FAIR² principes op te zetten.

Om beide doelstellingen te realiseren is de projectorganisatie opgesplitst in twee werkgroepen³;

- *Werkgroep Inhoud*; focust zich op de methodologie en klinische relevantie van de vraagstelling die centraal staat in deze casestudie. Deze werkgroep bepaalt hierbij ook de benodigde dataset om de onderzoeksvraag(en) te kunnen beantwoorden.
- *Werkgroep Techniek*; focust zich op het vertalen van de dataset naar de benodigde RWD bronnen en op de kwaliteit van deze RWD bronnen. Het ontwerpen en realiseren van een IT infrastructuur om het verbinden en analyseren van de data te faciliteren, zal met een Proof of Concept (PoC) worden getoetst.

1.3 Leeswijzer

Dit eindrapport beschrijft de uitkomsten en ervaringen van de werkzaamheden binnen deze casestudie. Het eindrapport is opgebouwd uit een aantal hoofdstukken, elk met een centraal thema. De *lessons learned* per thema worden in grijze kaders beschreven.

¹ RWD wordt in deze casestudie gedefinieerd als data welke is geregistreerd binnen het reguliere zorgproces

² FAIR staat voor Findability, Accessibility, Interoperability en Reusability

³ Zie bijlage B1 voor de betrokken leden van beide werkgroepen

2 Achtergrond

2.1 Het Prospectief Landelijk CRC cohort (PLCRC)

De DCCG is een landelijke multidisciplinaire onderzoeksgroep. Het primaire doel van de DCCG is het verbeteren van de kwaliteit van diagnose en behandeling van patiënten met darmkanker door het initiëren en faciliteren van klinisch wetenschappelijk onderzoek. Ter ondersteuning aan dit doel is PLCRC opgericht. PLCRC biedt een infrastructuur voor wetenschappelijk onderzoek naar darmkanker, waar patiënten via een opt-in toestemming geven voor het opvragen en gebruiken van hun medische gegevens voor wetenschappelijk onderzoek⁴. Optioneel kan er toestemming worden gevraagd voor 1) het opvragen en gebruiken van (tumor)weefsel, 2) het afnemen van maximaal 10 extra buizen bloed per jaar tijdens een reguliere bloedafname, 3) het toesturen van (kwaliteit van leven) PROMS vragenlijsten, 4) het mogen benaderen voor toekomstige studies, en 5) het informeren over toevallsbevindingen.

PLCRC is WMO-plichtig getoetst en goedgekeurd door de METC Utrecht, waarbij de DCCG de verrichter is van deze studie. De dagelijkse leiding is in handen van de hoofdonderzoeker prof. dr. M. Koopman en programmamanager dr. G.R. Vink, en zij leggen verantwoording af aan het bestuur van de DCCG⁵. De wetenschappelijke raad beoordeelt aanvragen voor gebruik van de infrastructuur of levering van gegevens. De raad bestaat uit geïnteresseerde artsen, onderzoekers en methodologen en zij geven advies aan het bestuur van de DCCG.

Sinds de start van PLCRC is het aantal deelnemende ziekenhuisorganisaties gegroeid naar 63 van de 71 Nederlandse ziekenhuisorganisaties en hebben ruim 16.500 patiënten toestemming gegeven voor deelname aan PLCRC. Deze toestemming maakt het mogelijk om gegevens uit de Nederlandse Kankerregistratie (NKR) op te vragen, en eventueel te koppelen aan andere bronnen. Daarnaast kunnen door datamanagers of onderzoekers voor specifieke onderzoeksvragen aanvullende gegevens uit het zorgproces worden verzameld in een (losstaande) eCRF⁶ of in de NKR.

2.2 Patient Reported Outcome Measures (PROMS)

Patiënten kunnen bij deelname aan PLCRC aangeven of ze PROMS-vragenlijsten willen invullen, welke op meerdere momenten (longitudoonaal) worden toegestuurd. Dit is o.a. op baseline (moment van inclusie), 3 maanden, 6 maanden, 12 maanden, 18 maanden, 24 maanden en daarna jaarlijks. De set aan vragenlijsten die een patiënt ontvangt bestaat o.a. uit de gestandaardiseerde EORTC-QLQ-C30, EORTC-QLQ-CR29, EQ-5D, EORTC-CIPN en de HADS⁷. Deze set aan vragenlijsten kan worden aangepast voor een specifieke groep patiënten binnen PLCRC indien gewenst.

Bij het digitaal invullen van de vragenlijst krijgt de desbetreffende patiënt de mogelijkheid om zijn eigen scores op het gebied van kwaliteit van leven in te zien. Hierbij kan worden gekozen om alleen de eigen scores terug te krijgen, of de eigen scores ten opzichte van een gezonde populatie of andere patiënten met darmkanker. De terugkoppeling die de patiënt ontvangt kan na toestemming van de patiënt ook worden doorgestuurd naar het desbetreffende ziekenhuis van de patiënt.

2.3 De Nederlandse Kankerregistratie (NKR)

In de NKR staan gegevens van alle patiënten met kanker in Nederland, en is hiermee een “population based registratie”. De signalering voor solide tumoren vindt plaats via pathologische verslaglegging vanuit het Pathologisch-Anatomisch Landelijk Geautomatiseerd Archief (PALGA), waarna de registratie middels andere bronnen wordt aangevuld.

⁴ Voor meer informatie (o.a. inclusie en exclusie criteria) zie <https://plcrc.nl/veelgestelde-vragen>

⁵ Zie ook <https://dccg.nl/board-working-groups>

⁶ Zie ook https://en.wikipedia.org/wiki/Case_report_form

⁷ Zie ook <https://qol.eortc.org/questionnaires/>

In de NKR wordt voor darmkanker patiënten sinds 2015 een uitgebreidere set aan klinische gegevens opgeslagen, waaronder de demografische –, tumor –, diagnostische – en behandelgegevens⁸. Getrainde datamanagers, die toegang hebben tot de verschillende (elektronische zorginformatie) systemen waar deze informatie in te vinden is, leggen de gegevens vast. Afhankelijk van deelname aan (sub)studies kan er ook follow-up worden vastgelegd. De vitale status van de patiënt wordt jaarlijks via een koppeling met de Basisregistratie Personen (BRP) bijgewerkt.

Standaard worden binnen de NKR alleen de behandelingen t/m de 1^{ste} lijn geregistreerd. Dit betekent voor patiënten met lokale ziekte (stadium 1-3 darmkanker) dat gegevens over neoadjuvante therapie, operatie en adjuvante behandeling worden vastgelegd. Voor patiënten met gemetastaseerde ziekte zijn dit behandelgegevens tot aan 1e progressie. Gegevens over terugkeer van ziekte of over 2^{de} en volgende lijnen van therapie worden in basis alleen op projectbasis geregistreerd. Voor patiënten die toestemming hebben gegeven voor PLCRC worden zoveel mogelijk aanvullend geregistreerd, wat betekent dat gegevens tot 3 tot 5 jaar follow-up in een curatieve setting worden vastgelegd en behandelgegevens tot aan overlijden in een gemetastaseerde setting.

Patiënten worden gevolgd vanaf diagnose, waarbij essentiële informatie over ziekte, zorg en uitkomsten inzichtelijk worden gemaakt. Deze opzet zorgt ervoor dat de NKR een population based en ziekenhuis overstijgende kankerregistratie (patiëntvolgsysteem) is. Dit maakt de NKR uniek in de wereld, met betrouwbare gegevens door onafhankelijke vastlegging en volledig omdat de gehele (volwassen) populatie van patiënten met kanker in beeld komt.

2.4 IKNL en de afdeling Research & Development

IKNL is een onafhankelijk kennisinstituut voor de oncologische en palliatieve zorg. De missie van IKNL is het reduceren van de impact van kanker. Dat doet ze op basis van inzichten van real world data in de NKR. Aan het reduceren van de impact van kanker levert IKNL een belangrijke bijdrage door het verzamelen en ontsluiten van essentiële en betrouwbare gegevens en deze te transformeren naar informatie en inzichten voor zorgprofessionals, beleidsmakers en onderzoekers. Middels de NKR volgt IKNL al dertig jaar op landelijk niveau de zorg voor patiënten met kanker en worden gegevens in kaart gebracht van patiënten met kanker in Nederland, direct uit het zorgproces.

Binnen de afdeling Research & Development (R&D) van IKNL is een team Klinische Informatici aangesteld, dat zich o.a. bezighoudt met het ontwerpen en implementeren van informatiestandaarden binnen de primaire zorg. Met deze informatiestandaarden wordt meervoudig gebruik van geregistreerde items mogelijk. Bij de ontwikkeling van informatiestandaarden staan de FAIR principes centraal. Het team was nauw betrokken bij het inmiddels afgesloten landelijk Citrien programma Registratie aan de Bron⁹. Dit programma had als doel om registratielast in het primaire zorgproces te verminderen en het eenvoudiger te maken om zorginformatie te delen met collega-zorgverleners, zowel binnen de eigen zorginstelling als het zorgnetwerk. Het fundament hiervoor is het komen tot eenduidige en gestandaardiseerde registratie afspraken. Het gebruik van deze eenduidige en gestandaardiseerde afspraken biedt ook mogelijkheden voor secundair hergebruik voor wetenschappelijk onderzoek en registraties.

Naast het team Klinisch Informatici is er binnen R&D het team Clinical Data Science. Dit team houdt zich bezig met het (door)ontwikkelen van federated learning technieken. Deze techniek maakt het mogelijk om databronnen met elkaar te verbinden en vanuit een centraal netwerk inzichten te creëren, zonder dat de data de bron verlaat. Dit concept wordt elders in het document verder uitgelegd.

⁸ Zie ook <https://iknl.nl/nkr/uitleg-nkr-cijfers/itemsets>

⁹ Zie ook <https://www.nfu.nl/themas/zorg-op-de-juiste-plek/citrienfonds/registratie-aan-de-bron>

3 Onderzoeksvraag

3.1 Vaststellen onderzoeksvraag en vertaling naar Proof of Concept

De werkgroep *inhoud* heeft in een aantal consultatierondes de definitieve onderzoeksvraag en dataset vastgesteld die centraal staat in de Proof of Concept (PoC). Bij het vaststellen van de onderzoeksvraag en dataset is er continu afstemming geweest tussen de werkgroep *inhoud* en *techniek* om te bepalen of de gewenste dataset kan worden gevuld met data die in het primaire zorgproces worden vastgelegd, om extra registratielast voor de zorgverlener te voorkomen en te onderzoeken wat het potentieel is van de beschikbare RWD.

Voor de PoC is de volgende onderzoeksvraag vastgesteld;

“Wat is de overleving en de kwaliteit van leven van gemetastaseerde darmkanker patiënten met een BRAF mutatie die behandeld zijn met encorafenib-cetuximab in vergelijking met dezelfde groep patiënten die behandeld zijn met andere systemische therapie?”

Tijdens de casestudie wordt het hele proces van data verbinden, verwerken en analyseren doorlopen voor het beantwoorden van de onderzoeksvraag. Voor de beantwoording moet er data uit meerdere bronnen, die bij verschillende organisaties staan, worden samengevoegd. Dit maakt deze onderzoeksvraag interessant voor deze casestudie. De PROMS data, die nodig zijn voor bepaling van de kwaliteit van leven staan in de PROFIEL database, de signalering van encorafenib-cetuximab patiënten loopt via Dutch Hospital Data (DHD) en de klinische data staan geregistreerd in de Nederlandse Kankerregistratie (NKR).



Afbeelding 1: Workflow onderzoeksvraag via dataverzameling naar analyse en resultaten

Bij het opstellen van een onderzoeksvraag die gebruik maakt van RWD is het essentieel dat er continue afstemming is tussen de inhoudsdeskundigen en informatiedeskundigen. Door vroegtijdige afstemming a) kan worden voorkomen dat er onderzoeksvragen worden gedefinieerd die niet kunnen worden beantwoord met de beschikbare RWD op dat moment en b) verloopt de vertaalslag van onderzoeksvraag naar benodigde query's goed en hebben de uitkomsten een klinische relevantie.

Door deze continue afstemming voorkomen we teleurstellingen achteraf op het gebied van beschikbaarheid, kwaliteit en betrouwbaarheid van RWD en verminderen de kansen op een mismatch tussen onderzoeksvraag en gebruikte query's tijdens de analyse (en potentiële uitkomsten).

Om de vertaalslag te maken van onderzoeksvraag naar PoC is deze opgesplitst in een aantal deelgebieden;

1. Populatie: criteria voor selectie van de juiste patiënten
2. Gegevensset: bepalen van de juiste gegevensset om de onderzoeksvraag te beantwoorden
3. Kwaliteit van leven: definitie van kwaliteit van leven bij patiënten en de wijze van meten
4. Behandellijn: definitie van een behandellijn en het bepalen van het startmoment hiervan

3.1.1 Populatie en selectiecriteria bepalen

Wanneer er wordt gekozen om data uit de NKR te gebruiken voor onderzoek naar (nieuwe) geneesmiddelen zal er eerst nagegaan moeten worden of de benodigde data hiervoor beschikbaar is. In de NKR wordt voornamelijk data rondom diagnose en 1^e behandeling geregistreerd. Follow-up data wordt niet altijd toegevoegd, dat gebeurt voornamelijk op projectbasis. Er is dus een reële kans dat de gewenste

data niet direct beschikbaar is op het moment van aanvraag, indien het een latere lijn van behandeling betreft.

Het is belangrijk om de juiste databronnen te selecteren voor het verzamelen van data. In de NKR wordt niet vastgelegd welke patiënten er zijn behandeld met encorafenib-cetuximab, aangezien dit een tweede- of latere lijn behandeling is. De behandeling met encorafenib is geïndiceerd in combinatie met cetuximab voor de behandeling van volwassen patiënten met gemetastaseerd colorectaal carcinoom met een BRAF V600E mutatie die eerder al systemische therapie hebben ontvangen. Dit betekent dat encorafenib-cetuximab niet in de eerste lijn aan patiënten mag worden voorgeschreven. De NKR lijkt hierom niet de beste bron voor de selectie van patiënten. Voor accurate signalering is een andere bron noodzakelijk.

Dutch Hospital Data (DHD) is een landelijke bron die maandelijks (deels) wordt gevoed door de Landelijke Basisadministratie Ziekenhuizen (LBZ) aanleveringen, welke verplicht zijn voor ziekenhuizen om aan te leveren aan DHD¹⁰. De LBZ is een gegevensset die naast de DBC zorgproducten bestaat uit diagnoses, verrichtingen, complicaties, voorgeschreven geneesmiddelen, betrokken specialismen van patiënten die een Nederlands ziekenhuis bezochten of een digitaal contactmoment hadden.

Vanwege juridische beperkingen kon in de casestudie alleen gebruik worden gemaakt van DHD-data voor de signalering van patiënten. Het gebruik van gegevens omtrent zorgproducten anders dan voor signalering van patiënten viel buiten de bestaande verwerkersovereenkomsten die van kracht zijn tussen DHD, IKNL en de ziekenhuizen. Mede door de juridische complexiteit, doorlooptijd en beperkte tijd binnen de casestudie is ervoor gekozen om DHD-data alleen voor de signalering toe te passen. Hiervoor is er een gegevensaanvraag ingediend bij DHD om te achterhalen welke patiënten met encorafenib-cetuximab zijn behandeld. Aan de hand van deze selectie is het traject gestart om PROMS vragenlijsten op te vragen bij PLCRC. Ook startte het traject om NKR-data te koppelen, waarmee in een aanvullende registratieronde data-items zijn verzameld. De matching op patiëntniveau over verschillende databronnen is bewerkelijk en niet triviaal, aangezien het op dit moment nog niet is toegestaan om binnen cohorten en (kwaliteit)registraties het BSN te gebruiken. Er moet worden gekeken naar andere manieren om te koppelen. In paragraaf 4.3 wordt verder ingegaan op de wijze van koppelen die is toegepast in deze casestudie en bijbehorende uitdagingen en beperkingen.

De volgende criteria zijn gehanteerd bij de uiteindelijke gegevensaanvraag bij DHD om te komen tot een populatie waarvoor NKR en PROMS data is verzameld. Dit heeft geresulteerd in een initiële populatie van 167 patiënten;

¹⁰ DHD bewerkt de LBZ-data tot informatie die gebruikt kan worden bij de bedrijfsvoering van ziekenhuizen en het verbeteren van de kwaliteit van zorg. Zie ook <https://www.dhd.nl/>

Criteria	Waarde
Tumor: lokalisatie(s)	Colon en rectum (C18-C20)
Tumor: morfologie(en)	Carcinomen
Tumor: stadium/stadia	Alle
Tumor: gedrag	Invasief
Patiënt: geslacht	Mannen en vrouwen
Patiënt: leeftijd(indeling)	18+
Diagnosejaren/-periode	-
Niveau van selectie	Landelijk
Behandeling(en)	Encorafenib

Tabel 1: selectiecriteria DHD

3.1.2 Gegevensset bepalen

Na het vaststellen van de onderzoeksvraag en populatie is het belangrijk om de initiële gegevensset vast te stellen. Deze gegevensset is samengesteld met data-items die verzameld moeten worden om de onderzoeksvraag (op termijn) te kunnen beantwoorden. In hoofdstuk 4 wordt verder ingegaan op de activiteiten die zijn uitgevoerd m.b.t. het bepalen en verzamelen van de benodigde data-items.

3.1.3 Kwaliteit van leven

Naast het verzamelen van klinische gegevens wordt voor het objectief meten van de kwaliteit van leven binnen het PLCRC cohort patiënten gevraagd om periodiek PROMS vragenlijsten in te vullen volgens het schema T0, T3, T6, T12, T18, T24, T36, T48, T60 en T72. Hier staat T0 voor tijdstip 0, dit is het moment dat de patiënt wordt geïncludeerd in PLCRC en de eerste keer een set aan vragenlijsten ontvangt. T3 staat voor 3 maanden na T0, etc.

De start van het aanbieden van vragenlijsten hangt samen met het moment van inclusie in PLCRC, en hangt dus niet samen met het moment van eerste incidentie. Dit zorgt ervoor dat het T0 moment per patiënt kan variëren, van het moment van incidentie tot ergens midden in een behandellijn, wat de analyse ingewikkelder maakt. Echter, omdat de invuldatum van de vragenlijst bekend is, kan deze aan het moment binnen de behandeling gekoppeld worden.

Per moment worden o.a. de EQ-5D, EORTC QLQ-C30 en EORTC QLQ-CR29 uitgevraagd bij de patiënt. De uitkomsten hiervan kunnen worden gebruikt voor de beantwoording van de onderzoeksvraag binnen deze casestudie.

Het merendeel van de patiënten wordt in geval van gemetastaseerde ziekte slechts enkele maanden met een specifiek geneesmiddel behandeld. Indien de patiënt al langer bekend is binnen PLCRC krijgt deze elke 6 maanden of jaarlijks een set van vragenlijsten toegestuurd. Dit heeft tot gevolg dat in geval van gemetastaseerde ziekte per lijn van therapie maar 0, 1 of 2 PROMS metingen gedaan zijn. Per 1 november 2021 wordt er daarom in de vragenlijsten aan patiënten gevraagd of zij gemetastaseerde ziekte hebben. Indien dit het geval is, gaan zij over op een ander schema en wordt hen elke 3 maanden een (bundel van) vragenlijst(en) toegestuurd. Dit schema wordt aangegeven met T3M, T6M enzovoort.

De EORTC QLQ-C30 bestaat uit een 30-tal vragen waaruit een overall kwaliteit van leven score kan worden berekend. Deze score toont het gemiddelde van 13 van de 15 QLQ-C30 schalen, de Global Quality of Life en Financial Impact schaal wordt buiten beschouwing gelaten¹¹. Voor een correcte

¹¹ Gebaseerd op Fayers et. al., De EORTC QLQ-C30 Scorehandleiding (3e editie) 2001

berekening moeten a) alle 13 schalen beschikbaar zijn, waarbij minimaal 50% van de items binnen de desbetreffende schaal zijn voltooid en b) de symptoomschalen worden getransformeerd om een uniforme inrichting van alle schalen te verkrijgen

3.1.4 Behandellijnen

Bij aanvang van de analyse is het belangrijk om een eenduidige definitie te hebben van de term behandellijn. In de hedendaagse literatuur is er nog geen consensus over een eenduidige definitie van deze term. Dit heeft o.a. te maken met het feit dat er een transitie gaande is van denken in *lijnen van therapie* naar *continuum of care*, wat het lastig maakt om specifieke en losstaande behandellijnen te definiëren.

“New agents for the treatment of metastatic colorectal cancer have extended median overall survival to more than 20 months, an increase that has changed the view of advanced colorectal cancer from an acute to a chronic condition. This article proposes a shift in treatment strategy from the concept of successive “lines” of therapy, in which chemotherapy is continued until disease progression, to that of a continuum of care, in which the use of chemotherapy is tailored to the clinical setting and includes switching chemotherapy prior to disease progression, maintenance therapy, drug “holidays,” and surgical resection of metastases in selected patients. In this approach, the distinction between lines of therapy is no longer absolute. This represents a paradigm shift in the management of metastatic colorectal cancer to that of a continuum of care approach that includes individualized planning, in which patients are given the opportunity to benefit from exposure to all active agents and modalities while minimizing unnecessary treatment and toxicity, with the ultimate goal of improving survival as well as quality of life.” (Goldberg et al.¹²)

Bovenstaande scherp geformuleerde conclusie uit 2007 geeft het dilemma weer rondom het definiëren van behandellijnen in de setting van gemetastaseerd colorectaal carcinoom. Voor de beantwoording van de onderzoeksvraag is er in de case studie voor gekozen om de volgende definitie van behandellijn te hanteren;

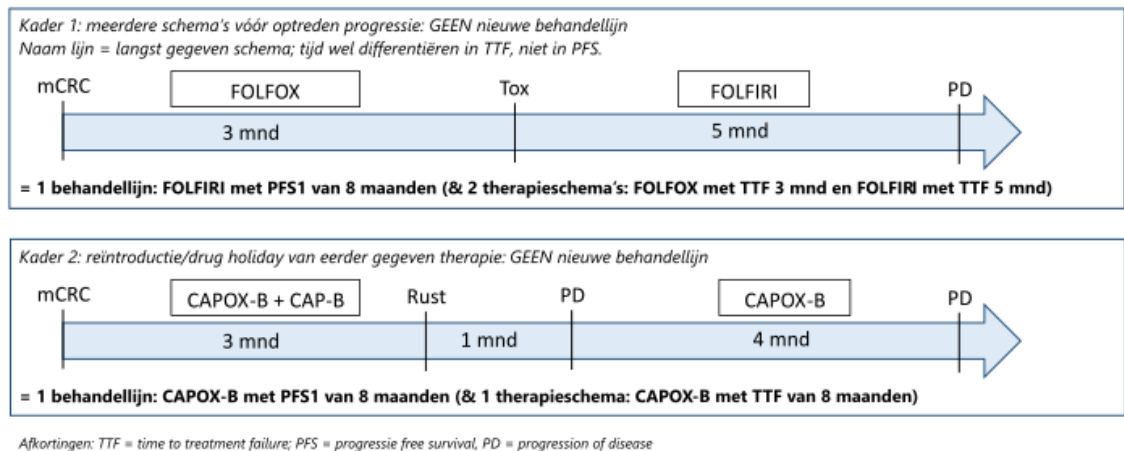
Progressie van ziekte en daarbij start nieuw chemotherapeutikum¹³ en/of nieuwe targeted therapie wordt gezien als een nieuwe behandellijn.

Houd er bij het formuleren van de onderzoeksvraag rekening mee dat voor alle aspecten in de onderzoeksvraag een duidelijke definitie wordt gehanteerd, om ambiguïteit te voorkomen en een succesvolle vertaalslag te maken naar de benodigde query's. Sommige termen die op het eerste gezicht eenduidig lijken, zoals behandellijn, blijken erg ingewikkeld om te bepalen op basis van de beschikbare data-items.

Concreet betekent dit dat bij een schemaswitch zonder progressie (kader 1) er dus geen sprake is van een nieuwe behandellijn, aangezien er geen progressie heeft plaatsgevonden. De naamgeving van de chemotherapie is o.b.v. de langst gegeven therapie(combinatie). Bij herintroductie of drug holiday (kader 2) is er ook geen sprake van een nieuwe behandellijn. In deze situatie is er namelijk geen nieuw chemotherapeutikum en/of targeted therapie gestart bij de patiënt.

¹² <https://ascopubs.org/doi/10.1200/JCO.2005.17.749>

¹³ Flouropyrimidines zijn uitwisselbaar



Afbeelding 2: weergave behandellijnen

Dit betekent voor de dataverzameling en analyse binnen deze casestudie dat;

1. Bij de niet-gemetastaseerde patiënt met adjuvante systeemtherapie die binnen 6 maanden na diagnose metastasen krijgt, de adjuvante systeemtherapie benoemd wordt als eerste lijn;
2. Er is progressie wanneer er een nieuwe metastase is, of wanneer de primaire tumor gegroeid is met een significante toename (>20%). In de NKR wordt er op dit moment geen expliciet item vastgelegd in relatie tot tumorgroei. Binnen de NKR kan door de toevoeging van een nieuwe episode of registratie van een nieuwe metastase worden vastgesteld of er sprake is van progressie bij de patiënt.

Uitzondering hierop geldt voor patiënten die de onderhoudsbehandeling CAP-B krijgen na eerdere CAPOX-B en er gaandeweg toch progressie van de tumor heeft plaatsgevonden, waarna er wordt herstart met CAPOX-B. Hoewel er progressie heeft plaatsgevonden en een nieuw chemotherapeutikum is gestart wordt dit toch niet gezien als een nieuwe behandellijn. Deze situatie is een uitzondering op de in de casestudie gehanteerde definitie.

De geformuleerde definitie geldt voor gemetastaseerde colorectaal carcinoom patiënten.

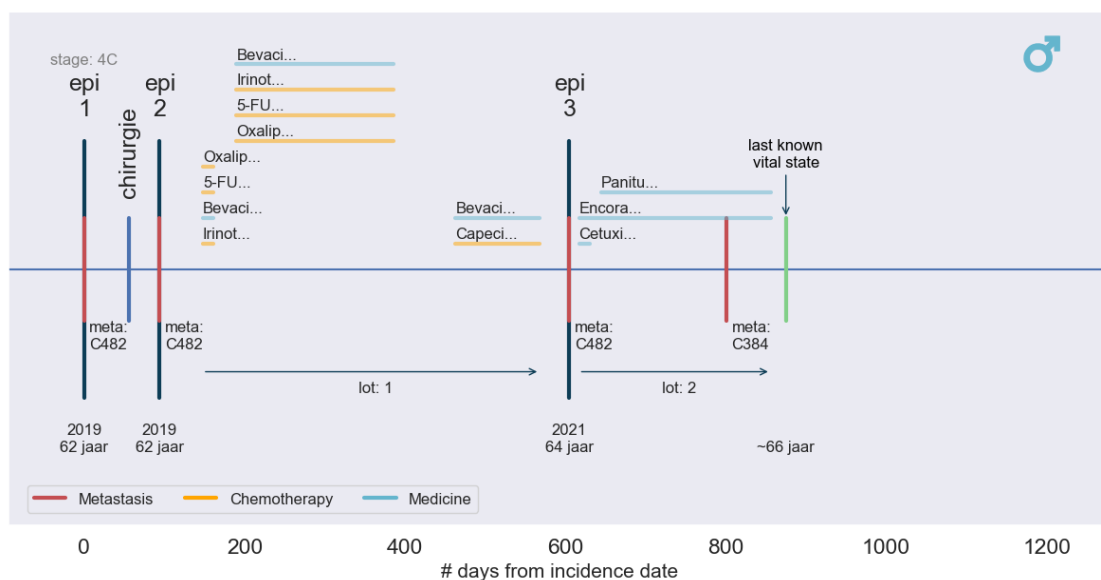
3.2 Groepering van middelen

Voor de behandeling van gemetastaseerd colorectaal carcinoom wordt een groot aantal verschillende oncolytics en combinaties hiervan ingezet (kuren). Afhankelijk van het unieke zorgtraject van de patiënt kunnen er meerdere behandelingen worden uitgevoerd, en daarbij verschillende (combinaties van) middelen worden gebruikt in één of meerdere behandellijnen. Zie paragraaf 3.1.4 voor de definitie van behandellijn zoals deze binnen de casestudie wordt gehanteerd.

Kijkend naar de data die beschikbaar zijn binnen de casestudie kunnen er 68 verschillende (of overlappende) combinaties van 21 unieke geneesmiddelen worden geïdentificeerd. Kijkend naar de behandellijnen waarin enkel encorafenib-cetuximab wordt toegediend blijven er nog 13 combinaties over. Deze combinaties, analyses en mogelijkheden zijn weergegeven in bijlage B4.

In afbeelding 3 is een voorbeeld van een representatieve patiënt die de complexiteit van de verschillende combinaties van oncolytics en behandellijnen inzichtelijk maakt. De verschillende start- en stopmomenten van de behandellijnen voor deze patiënt zijn weergegeven met de verticale epi lijnen; waar epi hier staat voor episode. De horizontale lijnen geven de combinaties en start- en stopmomenten van oncolytics weer die binnen één behandellijn zijn gegeven. De term meta geeft aan wanneer er een metastase is

vastgesteld met bijbehorende topologie. Dit alles weergegeven op de horizontale as van het moment van incidentie (0) tot laatst bekende vitale status.



Afbeelding 3: voorbeeld zorgtraject met afwisselende oncolytica (ook binnen één behandellijn)

Om de effectiviteit van een specifiek geneesmiddel (of combinatie hiervan) te kunnen bepalen is het belangrijk om rekening te houden met de combinaties van behandelingen en de complexiteit van verschillende behandellijnen. Idealiter zou men alleen patiënten selecteren die enkel het desbetreffende middel krijgen toegediend binnen één specifieke behandellijn. Dit zou echter niet de werkelijkheid in de dagelijkse praktijk weerspiegelen. Het is daarom belangrijk om bij aanvang van de analyses rekening te houden met de verschillende middelen en combinaties van middelen, en te kijken naar mogelijkheden om middelen te groeperen.

De introductie van moleculaire diagnostiek en het bijbehorende personalized medicine concept zorgt ervoor dat er steeds kleinere populaties komen van patiënten die een specifiek geneesmiddel toegediend krijgen. Dit maakt het moeilijker om voldoende patiënten te vinden om een onderzoeksvraag goed te kunnen beantwoorden. Wees hiervan bewust en denk na over mogelijkheden om geneesmiddelen te groeperen om te komen tot grotere patiënten populaties.

Er is besloten om in deze casestudie de geneesmiddelen op te delen in 11 groepen;

1. CAPOX/FOLFOX (met al dan niet Bevacizumab)
2. Capecitabine/5-FU (met al dan niet Bevacizumab)
3. Irinotecan (met al dan niet Bevacizumab)
4. FOLFIRI/CAPIRI (met al dan niet Bevacizumab)
5. FOLFOX/FOLFIRI met Panitumumab/cetuximab
6. FOLFOXIRI (met al dan niet Bevacizumab)
7. TAS-102 (Tifluridine/Tipiracil)
8. S1 (Tegafur, Gimeracil en Oteracil)
9. Centuximab en Panitumumab monotherapie
10. Encorafenib met Cetuximab
11. Alle behandelingen met Nivolumab of Pembrolizumab

Door deze groepering is het mogelijk om een grotere populatie te behouden om de onderzoeksvraag te beantwoorden.

4 Gegevensset en data-items

4.1 Bepalen kwaliteit, volledigheid en beschikbaarheid

Na het vaststellen van de onderzoeksvraag is er door de werkgroep *inhoud* een initiële gegevensset samengesteld van data-items die verzameld moeten worden om de onderzoeksvraag (op termijn) te beantwoorden. De definitieve gegevensset bestaat uit 24 data-items die opgesplitst kunnen worden in de categorieën patiënt/ziekte kenmerken, behandeling, PROMS, zorggebruik en adverse events. Zie bijlage B2 voor de data-item definities.

Bij het opstellen van de gegevensset is uitgegaan van de D4 procedure. Voor darmkanker bestaan er reeds meerdere ziekte specifieke gegevenssets, zoals de NKR, PLCRC, DCRA en ICHOM set. Op 13 september 2021 is de werkgroep *inhoud* een eerste maal bijeengeweest. Doel van deze bijeenkomst was om feedback te verzamelen op de initiële PICO en bijbehorende concept gegevens. Om zo te komen tot een breed gedragen gegevensset en onderzoeksvraag. In totaal namen er 18 personen (6 mensen vanuit RORDGM – HemoNED en 12 stakeholders) deel aan deze sessie. Zie bijlage B1 voor de namen van de werkgroepleden. Na het schetsen van de context werd er een constructieve en inhoudelijk goede discussie gevoerd. Tijdens deze discussie werd aan alle deelnemers gevraagd om te kijken naar de initiële gegevensset aan de hand van een drie vragen; 1) Kunt u zich hierin vinden?, 2) Mist u specifieke data-items? en 3) Hoe schat u de haalbaarheid in?

Na deze eerste bijeenkomst is om schriftelijke input gevraagd, welke is verwerkt voorafgaande aan de tweede bijeenkomst. Deze tweede bijeenkomst vond plaats op 2 november 2021. Tijdens deze bijeenkomst is het voorstel voor de definitieve gegevensset gepresenteerd en was er ruimte voor aanvullende adviezen hierover. Aan het einde is de gegevensset definitief vastgesteld door de werkgroepleden.

Deze gegevensset is het startpunt voor de werkgroep *techniek* om na te gaan in welke bronnen deze data-items zijn opgeslagen en welke inspanning het kost om deze data-items op landelijke schaal te verzamelen. Belangrijk uitgangspunt bij deze analyse is de kwaliteit, tijdigheid, volledigheid en beschikbaarheid van de bron.

Tijdens de analyse is er op het onderdeel 'beschikbaarheid' onderscheid gemaakt tussen lokale (binnen het ziekenhuis zelf), regionale en landelijke bronnen om de dataset te vullen. Idealiter zou men alle data-items vullen vanuit een lokale bron, aangezien dit de ruwe data betreft uit het zorgproces welke nog niet op enige wijze is verwerkt of getransformeerd. Een dergelijke verwerkings- of interpretatieslag vindt wel plaats als lokale data worden doorgestuurd naar een regionale of landelijke bron. Uit de analyse blijkt dat nog niet ieder data-item op een gestandaardiseerde of gestructureerde wijze beschikbaar is binnen het EPD. De geselecteerde data-items worden nu nog vaak als vrije tekst, niet gestandaardiseerd, willekeurig en afhankelijk van de inrichtingskeuzes van het desbetreffende ziekenhuis op een bepaalde manier vastgelegd. Dit alles maakt het op dit moment nog lastig om effectief de lokale bron te gebruiken om de data-items (geautomatiseerd) te verzamelen.

Om in de toekomst effectief informatie uit lokale bronnen te kunnen gebruiken voor het vullen van registers is het belangrijk om bij de totstandkoming van de gegevensset zo goed mogelijk aan te sluiten op de datadefinities van zorginformatiebouwstenen (zibs), bijbehorende informatiestandaarden. Ook is het aan te bevelen dat de gegevensset vanuit de originele bron wordt gevuld. Deze zibs en informatiestandaarden zijn het fundament voor het vastleggen en uitwisselen van informatie binnen de zorg. Naast het ondersteunen van de reguliere zorg kunnen deze informatiestandaarden ook het secundaire hergebruik bevorderen. Aansluiting hierop biedt kansen om eenvoudiger data-items te verzamelen. Dit betekent wel een switch in het denkproces bij het vaststellen van een onderzoeksvraag

voor onderzoekers, waarbij men moet uitgaan van wat er nu reeds wordt vastgelegd i.p.v. wat wenselijk zou zijn om vast te leggen.

De problematiek rondom lokale bronnen maakt het opschalen van de extractie van de benodigde data-items zeer lastig. Er is vanuit de werkgroep *techniek* daarom gekeken naar bronnen met een landelijke dekkinggraad om de dataset te vullen. Uit deze analyse blijkt dat PLCRC, NKR, PROFIEL en DHD potentieel interessante bronnen zijn. Door aansluiting op deze bronnen is het mogelijk om de informatiebehoefte te voeden vanuit vier bronnen, i.p.v. de 71 ziekenhuisgroepen in Nederland. Deze keuze heeft er wel in geresulteerd dat het data-item *adverse event > graad 3* is komen te vervallen in de dataset. Er is op dit moment geen landelijke bron die een volledig overzicht heeft van alle adverse events van een patiënt, wat mede veroorzaakt wordt door de niet-eenduidige vastlegging hiervan in het EPD. De (meeste van de) overige data-items zijn terug te vinden in één van deze vier bronnen.

4.2 Onderscheid user en systeem gegenereerde data

Hoewel het gebruik van landelijke bronnen het verzamelen van data-items vereenvoudigd, kan dit wel een negatief effect hebben op de tijdigheid en beschikbaarheid van informatie. Als voorbeeld; de DHD registratie loopt maximaal 3 maanden achter op de werkelijkheid, terwijl de NKR registratie tot wel 6 tot 12 maanden kan achterlopen.

De tijdigheid heeft te maken met het type data wat in deze bronnen wordt verwerkt. DHD verzamelt voornamelijk systeem gegenereerde data zoals verrichtingscodes en diagnosecodes. Dergelijke informatie wordt door het systeem zelf gegenereerd n.a.v. acties van de medisch specialist en op eenduidige wijze in het EPD vastgelegd. Dit maakt het eenvoudiger om deze gegevens te extraheren en te verwerken in een landelijke bron. De NKR verzamelt daarentegen met name essentiële user gegenereerde klinische informatie en observaties die in het EPD nu nog vaak op een ongestructureerde en niet gestandaardiseerde wijze zijn vastgelegd. Dit bemoeilijkt de (geautomatiseerde) extractie en verwerking van deze cruciale gegevens wat resulteert in een langere doorlooptijd. Zolang de klinische gegevens niet op een gestructureerde en gestandaardiseerde wijze worden vastgelegd blijft het nodig dat datamanagers deze gegevens verzamelen.

Systeem gegenereerde data heeft als voordeel dat het automatisch wordt gegenereerd o.b.v. handelingen door de eindgebruiker (b.v. het voorschrijven van oncolytica) en altijd op eenzelfde wijze wordt opgeslagen in het systeem. Deze informatie is zeer waardevol en kan relatief eenvoudig uit het EPD worden geëxtraheerd. Het is belangrijk om bewust te zijn dat dergelijke informatie niet altijd een weerspiegeling hoeft te zijn van de werkelijkheid. Deze informatie hangt vaak samen met de financiële afhandeling van verrichtingen binnen het EPD.

Als voorbeeld; de verrichtingscode dat bepaalde oncolytica is voorgeschreven betekent niet per definitie dat de patiënt deze kuur ook volledig heeft afgerond. De toediening en de registratie hiervan is een voorbeeld van user gegenereerde data door een verpleegkundige. Er is hier sprake van een paradox tussen de "papieren" werkelijkheid en de "real-world" werkelijkheid. Bij het opstellen van de onderzoeksvraag en het selectieproces van bronnen is het belangrijk om hier scherp op te zijn. Het zoeken naar een goede verhouding tussen systeem en user gegenereerde data is essentieel.

De tijdigheid en beschikbaarheid van informatie die nodig is om de onderzoeksvraag te beantwoorden bepaalt in hoge mate welke landelijke bronnen wenselijk zijn.

Binnen deze use case was het voor de aanvullende gegevensregistratie essentieel om vroegtijdig patiënten te signaleren waarbij het middel encorafenib is voorgeschreven. Gezien de verwachte kleine patiëntpopulatie was het doel om alle gemetastaseerde colorectaal carcinoom patiënten die behandeld zijn met encorafenib-cetuximab te identificeren. Vanwege de afhankelijkheid met het onderdeel tijdigheid en dekkingsgraad is ervoor gekozen om de DHD databron te gebruiken. Uit de analyse van de werkgroep *techniek* blijkt dat DHD bij het eerste aanlevermoment van één maand na het moment van registratie al ongeveer een dekkingsgraad heeft van 80% van alle potentiële patiënten.

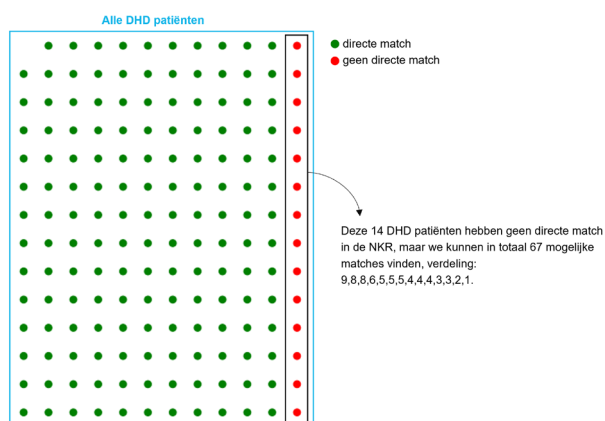
4.3 Noodzaak en uitdagingen bij koppelen databronnen

Voor het verzamelen van de data-items moest er informatie uit vier verschillende databronnen worden samengebracht. Op dit moment is het nog niet toegestaan om binnen cohorten en (kwaliteits)registraties het BSN van de patiënt te registreren en moet er naar een andere manier worden gekeken om te koppelen.

Voordeel van de geselecteerde landelijke bronnen is dat PLCRC, NKR en PROFIEL al door een interne IKNL koppelmatrix met elkaar zijn verbonden. Voor het koppelen van de DHD data is binnen deze use case gebruik gemaakt van probabilistische matching methodiek. Deze technologie maakt gebruik van de patiëntkenmerken geslacht, geboortedatum, (laatst bekende) viercijferige postcode en tumor typologie om een match te vinden met patiënten die in de NKR bekend zijn.

Afhankelijk van de complexiteit van een onderzoeksvraag zullen er bronnen met elkaar moeten worden gekoppeld om de bijbehorende dataset te vullen. Hoe meer bronnen aan elkaar moeten worden gekoppeld, hoe lastiger het wordt om probabilistisch matchen goed toe te passen in de praktijk. Er dient een juridische grondslag te komen om ook voor wetenschappelijk onderzoek het BSN te mogen toepassen, al dan niet middels een Trusted Third Party (TTP) constructie. Zo niet, dan zal het in de toekomst lastig worden om bepaalde relevante onderzoeksvragen goed te kunnen beantwoorden.

Hoewel probabilistisch matchen een bewezen tool is om patiënten over bronnen heen met elkaar te koppelen, is deze niet 100% waterdicht. De initiële dataset die DHD heeft aangeleverd voor het signaleren van patiënten waarbij encorafenib is voorgeschreven bestaat uit 167 patiënten. Met deze methodiek kon bij 153 patiënten een 1-op-1 match worden gevonden. De overige 14 patiënten hadden een mogelijke match met 67 NKR patiënten, zie bijlage B3 voor de uitkomsten. Handmatige check van een datamanager was noodzakelijk om ook deze patiënten op een correcte wijze te koppelen.



Afbeelding 4: uitkomsten probabilistisch matchen

Deze uitkomsten laten zowel de kracht van het toepassen van probabilistisch matchen bij koppelen zien, maar ook de zwakte bij verdere opschaling over meer bronnen heen. Als er in de toekomst complexere onderzoeksvragen worden geformuleerd waarbij informatie uit veel bronnen moet worden samengebracht, wordt het steeds moeilijker om door probabilistisch matchen de patiëntgegevens over de bronnen heen te koppelen. Deze complexiteit heeft o.a. te maken met de onzekerheid of dezelfde patiënt in alle bronnen die moeten worden gekoppeld, beschikbaar is. Aangezien er geen unieke identifier is die mag worden gebruikt (b.v. BSN) voor onderzoeksdoeleinden is dit niet vooraf met 100% vast te stellen. Wanneer er een breed scala aan bronnen moet worden gekoppeld aan één patiënt is de kans groter op mismatches en dat achteraf handmatige correcties noodzakelijk zijn. Er zal moeten worden gekeken naar een andere vorm van identificatie van patiënten over de bronnen heen.

In deze casestudie is gekeken naar een middel voor gemetastaseerde patiënten dat in de tweede of latere lijn wordt ingezet, hierbij is de kans groot dat de desbetreffende patiënt in de NKR aanwezig is. Bij andere oncolytica kan de NKR achterlopen op de werkelijkheid en is het lastiger om middels probabilistische matching een match te vinden omdat niet met zekerheid te zeggen is dat de patiënt al in de NKR is opgenomen.

5 IT infrastructuur

5.1 Wat is federated learning?

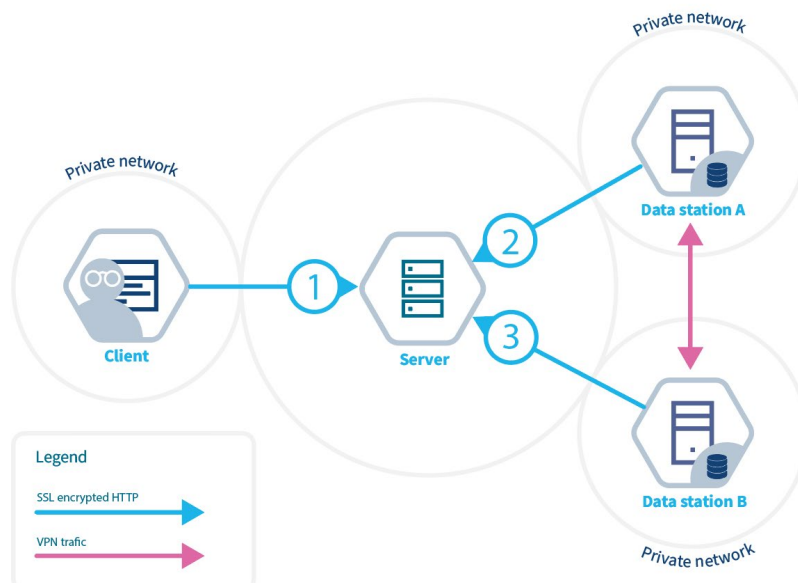
Traditioneel wordt bij onderzoek data uit verschillende bronnen samengebracht in één centrale databank en wordt er vanuit daar onderzoek gedaan. Door de introductie van strengere nationale en internationale privacywetgeving wordt dit steeds lastiger, en eigenlijk is het niet langer wenselijk om centrale databanken op te zetten. Federated learning is gebaseerd op het privacy-by-design principe wat de mogelijkheid biedt om op verantwoorde wijze data te delen en te analyseren.

Hierbij blijft de data decentraal bij de bron opgeslagen en worden enkel de (geaggregeerde) uitkomsten op de onderzoeksvraag gedeeld met de onderzoeker. Door federated learning is het technisch gezien niet meer noodzakelijk om grote hoeveelheden data te verzamelen binnen één centrale databank om waardevolle analyses en onderzoek te doen. De data worden door partijen enkel toegankelijk gemaakt via een beveiligd federated learning netwerk waarin een onderzoeker zijn query's kan uitvoeren.

Voor het succesvol opzetten van een federated learning netwerk is een aantal concepten van belang; server, client, node (als data station) en query.

Centraal in het netwerk staat de server. Deze server zorgt voor de communicatie tussen de client en node, prioriteert de afhandeling van de taken binnen het netwerk en controleert of de onderzoeker toegangsrechten heeft tot bepaalde nodes binnen het netwerk. Er wordt op de server zelf geen patiëntdata opgeslagen, enkel tijdelijk geaggregeerde uitkomsten. De node is het koppelstuk met een decentrale databron en voert de query's uit die vanuit de centrale server worden verstuurd. In de node zelf wordt de berekening uitgevoerd op de data, de data zelf blijft enkel opgeslagen in de decentrale databron. Alleen binnen de collaboratie goedgekeurde query's kunnen worden uitgevoerd op de node. Deze controle wordt vanuit beveiligingsoogpunt door de node zelf uitgevoerd. De client is de software applicatie waarmee een gebruiker of externe applicatie verbinding kan maken met de server om de analyses te starten.

Door de concepten server, client en node is het mogelijk om een federated learning netwerk op te zetten en de gewenste analyses op de data uit te voeren. In paragraaf 5.3 wordt verder ingegaan op de architectuur die is gehanteerd voor de PoC.



Afbeelding 5: Architectuur federated learning netwerk

5.2 Onderscheid verticale en horizontale data

Vanuit het data-oogpunt kan er onderscheid worden gemaakt tussen horizontale en verticale data query's. Door horizontale query's wordt per node in het netwerk dezelfde data-items bevroegd. Dit is met name bedoeld om populatieselecties te vergroten. Als voorbeeld; in node A staan gegevens van 25 unieke patiënten en in node B staan exact dezelfde data-items voor 35 unieke patiënten. Dit zorgt ervoor dat de query een uitkomst kan berekenen over een populatie van 60 patiënten. Omdat er in dit voorbeeld sprake is van allemaal unieke patiënten die slechts in één van beide nodes voorkomt, is er geen koppeling nodig tussen patiënten over bronnen heen.

Voor verticale query's is het wel noodzakelijk om patiënten en diens gegevens over bronnen heen te identificeren en te koppelen. De kracht van deze verticale query's is om van één patiënt over de nodes heen verschillende data-items te bevroegen en deze aan elkaar te koppelen voor analyses. Als voorbeeld; in node A staan gegevens over de oncologische aandoening van de patiënt en in node B de socio-economische gegevens. De query zorgt ervoor dat beide type data-items voor de desbetreffende patiënt met elkaar worden gekoppeld om daarna de gewenste analyse uit te voeren.

De kracht van verticale query's is de mogelijkheid om data-items van één patiënt die in verschillende nodes zijn opgeslagen te combineren en te analyseren. Dit maakt het mogelijk om complexe analyses uit te voeren zonder eerst zelf alle data centraal te verzamelen. Randvoorwaardelijk hieraan is dat de gegevens van de patiënt over de nodes heen worden geïdentificeerd. Hoewel probabilistisch matchen in de praktijk een goede basis blijkt, wordt het steeds lastiger om dit goed uit te voeren wanneer de hoeveelheid nodes groeit.

Om de query te kunnen uitvoeren is het essentieel dat de unieke patiënt en diens gegevens over de nodes heen wordt geïdentificeerd en gekoppeld. Zoals aangegeven in paragraaf 4.3 wordt het koppelen, vanwege het ontbreken van een unieke identifier over de nodes heen, op basis van probabilistisch matchen gefaciliteerd. Dit zorgt ervoor dat er nu nog beperkingen zijn in de hoeveelheid nodes die per query kan worden bevroegd. Deze complexiteit heeft o.a. te maken met de onzekerheid of dezelfde patiënt binnen alle nodes die moeten worden gekoppeld beschikbaar is. Binnen het initiatief LANCELOT¹⁴ is IKNL samen met TNO druk bezig om het probabilistisch matchen binnen een federated learning netwerk verder te optimaliseren om deze complexiteit te reduceren. Zo is het nu bijvoorbeeld mogelijk om een Kaplan Meier analyse ook verticaal uit te voeren.

5.3 Proof of Concept opstelling

Voor de PoC is een omgeving opgezet met twee demo nodes die gekoppeld zijn aan het al bestaande federatieve netwerk van IKNL. De nodes simuleren de organisaties IKNL en ZiNL. Aan de node van IKNL worden de databronnen NKR, PROFIEL aangeboden die nodig zijn om de onderzoeksvraag te beantwoorden. De centrale server zorgt voor de authenticatie van de gebruiker en toegang tot de nodes en bijbehorende databronnen.

In deze POC is de clientsoftware geïnstalleerd in een virtuele machine die draait op een separate laptop. Via deze client worden de benodigde analyses gestart over het netwerk en de resultaten hiervan samengevoegd en weergegeven. In figuur 6 is een voorbeeld van de client code weergegeven.

De uitkomsten van de analyses die zijn uitgevoerd in kader van deze casestudie zijn weergegeven in paragraaf 6.1.

¹⁴ Zie ook: <https://iknl.nl/en/news/lancelot-new-collaboration-between-iknl-and-tno-to>

```

patient_selection_criteria = {
    'drugs': ['L01EC03', 'L01FE01'],
    'lot': [2, 3],
    'drugs_combination': 'at least'
}

my_task = client.task.create(
    collaboration=COLLABORATION_ID,
    organizations=[ORGANIZATION_ID],
    name='ROR-demo',
    image="harbor2.vantage6.ai/algorithms/ror:11",
    description='',
    input={"master":False,
          "kwargs": {'patient_selection_criteria': patient_selection_criteria},
          "method": 'master'}
)
task_id = my_task.get('id')
wait_for_complete_msg(task_id)

```

Afbeelding 6: voorbeeld client code voor QoL bepaling

6 Uitkomsten en conclusies

6.1 Aanpassing initiële PROMS onderzoeksvraagstelling

Zoals in paragraaf 4.1 is beschreven is het helaas niet mogelijk gebleken om alle initieel vastgestelde data-items te verzamelen. Dit werd mede veroorzaakt vanwege juridische beperkingen omtrent het automatisch verkrijgen via DHD en omdat bijvoorbeeld bepaalde data-items niet tot nauwelijks werden geregistreerd in het EPD.

Echter was conform de huidige bewerkersovereenkomst het wel mogelijk om via een DHD gegevensaanvraag de signalering te laten plaatsvinden van alle gemetastaseerde CRC patiënten die zijn behandeld met encorafenib i.c.m. cetuximab. Door probabilistische matching zijn deze patiënten gekoppeld aan de NKR om aanvullend te worden geregistreerd. Tijdens het aanvullend registeren zijn klinische gegevens omtrent behandellijnen, (beste) response en tumor kenmerken verzameld door de IKNL datamanagers. Vervolgens is binnen PLCRC nagegaan of er voor deze specifieke populatie patiënten ook PROMS data beschikbaar waren.

De initiële populatie van 167 patiënten is in de laatste datalevering verhoogd naar 179 patiënten. Voor de gehele populatie patiënten behandeld met encorafenib i.c.m. cetuximab zijn de volgende gegevens beschikbaar:

```
Van de in totaal 179 patiënten,  
zitten er 47 in PLCRC  
en hebben er 35 PROMS ingevuld,  
in totaal zijn dit 125 PROMS:
```

```
T0 30  
T3 24  
T6 25  
T12 15  
T18 9  
T24 6  
T36 3  
T48 1  
T60 <NA>  
T72 <NA>  
T3M 4  
T6M 4  
T9M 3  
T12M 1  
T12PM <NA>
```

Afbeelding 7: Dataset statistieken algemeen

C30
 Van de 30 vragen zijn er het volgende aantal vragen ingevuld:

# ingevulde vragen	# vragenlijsten
30	113
29	3
28	2
0	7

Er kan voor 113 vragenlijsten een summary score worden berekend.

EQ5D
 Van de 6 vragen zijn er het volgende aantal vragen ingevuld:

# ingevulde vragen	# vragenlijsten
6	119
5	1
0	5

CR29
 Van de 36 vragen zijn er het volgende aantal vragen ingevuld:

# ingevulde vragen	# vragenlijsten
33	1
28	8
27	5
26	35
25	62
24	2
22	2
7	1
1	3
0	6

Afbeelding 8: Dataset statistieken omtrent PROMS

PROMS & Behandellijnen
 Er zijn 13 patiënten die PROMS hebben die matchen met een (of meer) behandellijn(en).

	# vragenlijsten
T0	3
T3	2
T6	6
T12	2
T18	1
T24	1
T36	1
T3M	2

Afbeelding 9: Dataset statistieken omtrent PROMS i.c.m. behandellijnen

Uit de uitkomsten kan worden geconcludeerd dat er 13 patiënten zijn die PROMS rond de start van encorafenib-cetuximab hebben ingevuld (afbeelding 9). Een dergelijke kleine populatie maakt het lastig om een gedegen analyse te doen naar de (overall) kwaliteit van leven bij start van een behandellijn. 'Rond het begin van een behandellijn' is hier gedefinieerd als 1 maand voor tot 2 weken na start van de behandellijn.

Voor de POC is er ondanks de geringe populatie gekeken of het mogelijk was om kwaliteit van leven te berekenen binnen het federatieve netwerk – als praktijktoets voor de client code. De gemiddelde score van kwaliteit van leven, tijdens de behandeling met encorafenib-cetuximab was 75.34. Dit resultaat toont aan dat het (technisch) mogelijk is om middels een federatief model onderzoeksvragen te beantwoorden over meerdere bronnen heen. De uitkomsten zelf hebben geen (klinische) relevantie vanwege de beperkte populatie.

Onderzoeksvraag

Er zijn 179 patiënten die encorafenib + cetuximab hebben gehad.

Er zijn 156 patiënten die encorafenib + cetuximab hebben gehad in de 2e of 3e behandellijn, waarvan er in totaal 5 PROMS hebben die matchen met de 2e of 3e behandellijn.

C30 kwaliteit van leven score over deze 5 patienten: 75.34

Afbeelding 10: Uitkomsten QoL overall score

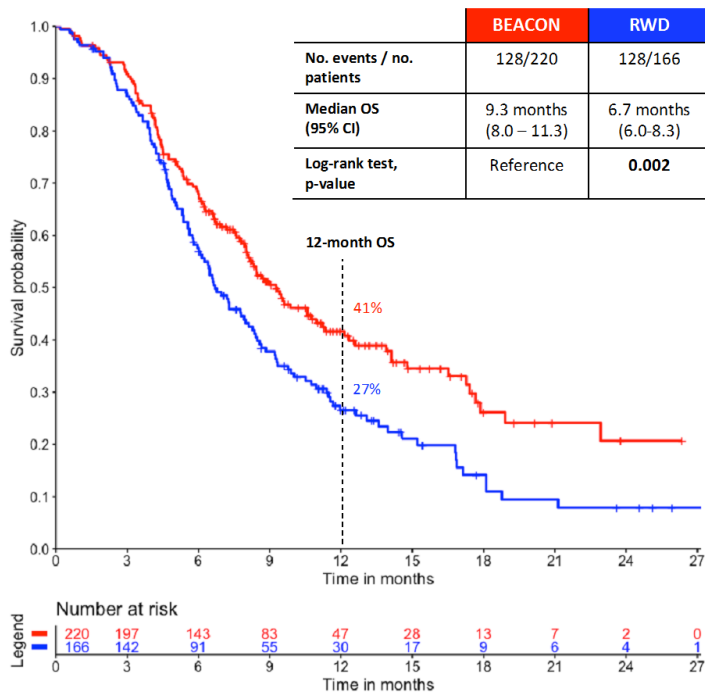
6.2 Analyse totale overleving populatie

Bij nieuwe middelen die op de markt komen en in de dagelijkse zorgpraktijk gebruikt gaan worden is er vaak sprake van een *efficacy-effectiveness gap* – een verschil in effectiviteit tussen de homogene studiepopulatie en de effectiviteit in de heterogene patiëntenpopulatie in de dagelijkse praktijk. De werkzaamheid van de combinatie encorafenib en cetuximab bij gemetastaseerde CRC patiënten is onderzocht in de internationale BEACON trial. De belangrijkste inclusie criteria voor de BEACON studie bestond uit gemetastaseerd colorectaal carcinoom met een BRAFV600E mutatie, behandeld met één of twee eerdere behandellijnen, waarbij er geen sprake was van hersenmetastasen of andere maligniteit in de 5 jaar voorafgaand aan de start van encorafenib-cetuximab, niet eerder behandeld met een RAS/RAF/MEK inhibitor, een neutrofielen count ≥ 1.5 en een WHO score 0 of 1.

Met de klinische gegevens die binnen deze casestudie zijn verzameld van de patiënten die zijn behandeld met encorafenib i.c.m. cetuximab is het mogelijk om de volgende onderzoeksvraag te beantwoorden;

“Wat is de totale overleving (OS) bij patiënten met een BRAF-gemuteerde gemetastaseerde colorectale kanker die in de dagelijkse praktijk worden behandeld met encorafenib-cetuximab?”

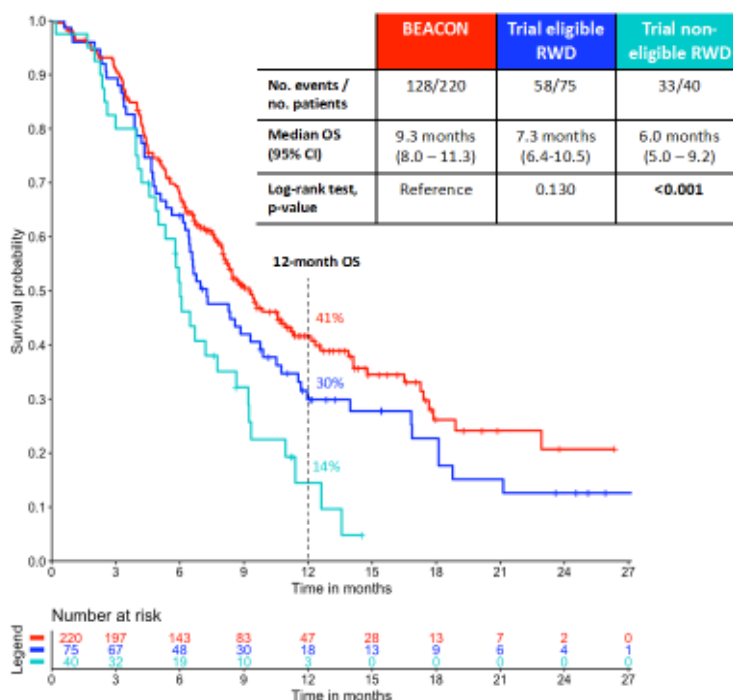
De overleving van patiënten in de dagelijkse praktijk in Nederland is vergeleken met de patiënten in de BEACON trial. In totaal konden 166 patiënten meegenomen worden in de analyse. Gemiddelde leeftijd van de patiënten was 64,0 jaar (versus 60,2 jaar in de BEACON). De overleving bleek in de dagelijkse praktijk statistisch significant korter: 6,7 maanden in de dagelijkse praktijk versus 9,3 maanden binnen de BEACON trial (afbeelding 11).



Afbeelding: 11: Overleving BEACON vs. Real World Data (cohort data)

Het totale cohort van 166 patiënten die behandeld zijn in de dagelijkse praktijk is vervolgens zo goed mogelijk opgedeeld naar patiënten (n=75) die voldeden aan de inclusie criteria van de BEACON trial en patiënten (n=41) die niet voldeden aan de inclusie criteria. Bij 51 patiënten waren er onvoldoende gegevens beschikbaar om deze goed in te delen.

De totale overleving van de patiënten uit de dagelijkse praktijk die volgens de beschikbare gegevens aan de inclusie criteria van BEACON voldeden was niet significant verschillend aan de overleving van patiënten binnen de BEACON trial. De overleving van patiënten uit de dagelijkse praktijk die niet aan de inclusie criteria voldeden was wel significant korter (afbeelding 12).



Afbeelding 12: Totale overleving patiënten BEACON vs. trial eligible vs. trial non-eligible

Op basis van de uitkomsten van deze studie naar de overleving van gemetastaseerde CRC patiënten in de dagelijkse praktijk versus overleving van patiënten in de BEACON trial kan worden geconcludeerd dat;

- Er een significant verschil is in mediane totale overleving tussen patiënten behandeld in de BEACON trial ten opzichte van de dagelijkse praktijk
- Er subgroepen zijn die minimale overlevingswinst hebben van het gebruik van encorafenib-cetuximab, zoals de patiënten met hersenmetastasen en patiënten met een WHO performance status van >2

6.3 Conclusie en vervolgstappen

Het programma Regie op Registers Dure Geneesmiddelen (ROR DGM) is gestart met een aantal casestudies om ervaring op te doen met de mogelijkheden en beperkingen van (real world) registers voor het evalueren van geneesmiddelen in de dagelijkse praktijk. Met deze casestudie is ervaring opgedaan met alle noodzakelijk stappen; het vaststellen van een gegevensset, het signaleren van de benodigde patiënten middels probabilistisch koppeling, het aanvullend registeren van klinische gegevens, het aanpassen van de frequentie voor het uitsturen van vragenlijsten (o.a. PROMS) en het federatief analyseren. Daarnaast is er onderzoek gedaan naar de *efficacy-effectiveness gap*.

Doordat binnen deze casestudie is uitgegaan van de reeds bestaande PLCRC cohort konden er gedurende het project snel stappen worden gezet. Het merendeel van de ziekenhuizen deed al mee aan PLCRC en het vragen van informed was reeds geïmplementeerd. Echter, in de praktijk blijkt dat in de meeste ziekenhuizen nog niet iedere patiënt wordt gevraagd voor deelname, niet iedere patiënt wil deelnemen of wil niet altijd een PROMS vragenlijst invullen. Als gevolg hiervan zijn er nog onvoldoende PROMS gegevens beschikbaar om een relevante en betrouwbare analyse uit te voeren binnen deze casestudie naar de kwaliteit van leven van gemetastaseerde CRC behandeld met encorafenib-cetuximab.

6.3.1 Conclusies

Uit de uitkomsten van deze casestudie kan worden geconcludeerd dat uitkomsten van trials voor (nieuwe) geneesmiddelen niet (altijd) representatief zijn om ook uitkomsten in de dagelijkse praktijk te voorspellen. Deze discrepantie wordt mede veroorzaakt door het verschil in de homogene studiepogaties en de heterogene patiëntenpopulatie in de dagelijkse praktijk.

Binnen de case studie is er ervaring opgedaan met het verzamelen, verbinden en analyseren van RWD. Ondanks dat er vanaf het begin gebruik kon worden gemaakt van bestaande infrastructures en bronnen voor het verzamelen en verbinden van (klinische) gegevens van patiënten bleek het nog steeds een uitdaging te zijn om deze gegevens bij elkaar te brengen. Deze uitdagingen spelen op een groot aantal deelgebieden, van wetgeving tot informatiestandaardisatie en dossierhygiëne.

Vanwege het ontbreken van een juridische grondslag is het op dit moment niet mogelijk voor registers om het BSN te gebruiken en moeten alternatieve manieren worden gebruikt om gegevens aan elkaar te koppelen. In deze casestudie is gebruik gemaakt van probabilistisch matchen, met een success rate van 91,7%. De overige patiënten zijn achteraf handmatig, na dossieronderzoek, gekoppeld. Voor kleine populaties en beperkte bronnen is deze manier van koppelen nog behapbaar. Bij onderzoeken waar meerdere bronnen en grote populaties moeten worden gekoppeld is deze manier steeds lastiger toe te passen en is het noodzakelijk om een unieke identifier over de bronnen heen te hanteren zoals het BSN. Het is de hoop dat hier, vanuit deze initiatieven of daarbuiten, snel duidelijkheid komt over de juridische grondslag voor gebruik van het (gepseudonimiseerde) BSN of een ander type koppelsleutel. De koppeling van het BSN maakt effectiviteits- en veiligheidsonderzoek van geneesmiddelen in real world populaties immers eenvoudiger.

Idealiter zou men gebruik willen maken van gegevens uit een lokale bron (EPD), aangezien hier de ruwe gegevens staan welke nog niet op enige wijze is verwerkt of getransformeerd. Een dergelijke verwerkingsslag vindt wel plaats als lokale gegevens worden opgeslagen in een regionale of landelijke bron. In het EPD worden op dit moment nog veel gegevens vastgelegd middels vrije tekst, welke niet gestandaardiseerd of gestructureerd is en waarbij binnen het ziekenhuis eigen inrichtingskeuzes zijn gemaakt. Dit maakt het complex om data uit lokale bronnen te extraheren, zeker bij opschaling naar landelijke dekking over alle ziekenhuizen heen. Het gebruik van regionale of landelijke bronnen is nu nog de beste optie om op landelijke schaal informatie te verzamelen. Deze bronnen beperken zich tot een beperkte set aan gegevens, waardoor het niet mogelijk is om via dit kanaal specifieke (klinische) gegevens eenvoudig en op grote schaal te verzamelen. Om deze gegevens daadwerkelijk te mogen gebruiken dienen er wel additionele verwerkersovereenkomsten te worden opgesteld, wat mogelijk een lange doorlooptijd kan hebben. Dit is ook de reden dat de DHD-data enkel voor signalering is gebruikt binnen deze casestudie en de gegevensaanvraag geen additionele gegevens bevatte over zorgproductie.

Door de introductie van moleculaire diagnostiek en bijbehorende personalized medicine in de dagelijkse praktijk krijgen steeds kleinere patiënt populaties een specifieke behandeling bestaande uit één of meerdere middelen toegediend. Daarom is er in deze casestudie ook voor gekozen om middelen te groeperen om voldoende grote populaties te creëren om onderzoek te doen naar de effectiviteitsvraag.

De uitgevoerde PoC heeft laten zien dat het middels een federatief learning model mogelijk is om RWD te gebruiken om onderzoeksvragen te beantwoorden zonder hiervoor één centraal register op te zetten, maar juist te werken met decentrale bronnen binnen een beveiligd netwerk.

6.3.2 Vervolgstappen

Bij het verbinden van bronnen en het koppelen van gegevens van patiënten over deze bronnen heen is het essentieel dat er één landelijke methodiek komt voor patiënt identificatie. Idealiter zou dit het BSN zijn, maar vanwege het ontbreken van de juridische grondslag is dat op dit moment niet mogelijk. Vanuit meerdere initiatieven wordt nu gekeken naar mogelijkheden om deze juridische grondslag aan te passen voor registers en wetenschappelijk onderzoek zodat het BSN wel mag worden toegepast. Advies vanuit de casestudie richting ZiNL is om bij deze initiatieven aan te sluiten en te kijken of het mogelijk is om het BSN, of een andere koppelsleutel, te mogen hanteren¹⁵. Een ander traject zou het opzetten van een landelijke koppeldienst kunnen zijn die werkt volgens het trusted third party (TTP) principe.

Goede afstemming en een duidelijke governance structuur is wenselijk bij de uitgifte van gegevens. Idealiter is er binnen Nederland één informatieloket waarbij geïnteresseerde partijen gegevensaanvragen kunnen indienen. Binnen dit informatieloket worden deze aanvragen getoetst, de benodigde gegevens aan elkaar gekoppeld en beschikbaar gesteld – al dan niet binnen een federatief model. Dit sluit aan op de initiatieven vanuit de EHDS wetgeving, Health-RI¹⁶ en visiedocument ROR.

Op het gebied van gegevensvastlegging, en dan specifiek de user gegenereerde data, is het van belang dat er meer wordt ingezet op gestructureerde en gestandaardiseerde vastlegging voor zowel primair als secundair gebruik. Om dit te realiseren moeten er, naast het vaststellen van de benodigde gegevensset voor de dagelijkse praktijk, ook governance afspraken gemaakt worden met de desbetreffende wetenschappelijke verenigingen en landelijke onderzoeksgroepen¹⁷. In de periode dat deze casestudie heeft gelopen zijn er in Nederland meerdere initiatieven gestart om meer inzicht te krijgen in de behandeling van CRC patiënten en om het samen beslissen en het leren en verbeteren te ondersteunen. Bij al deze initiatieven staan gegevens van de patiënten centraal en worden steeds nieuwe gegevenssets vastgesteld. Deze initiatieven worden gebundeld tot één multidisciplinaire gegevensset binnen het harmonisatieproject CRC. Het advies is om in samenhang met het harmonisatieproject CRC stappen te zetten in het verbeteren van de bronregistratie. Dit in lijn met het gedachtegoed vanuit visiedocument ROR registratie aan de bron, FAIR en OpenEHR¹⁸.

In de afgelopen periode stappen steeds meer (commerciële) partijen in de markt met een breed scala aan producten om het extraheren van gegevens uit het EPD te vereenvoudigen. Waarbij Natural Language Processing (NLP), Machine Learning (ML) of Artificial Intelligence (AI) technieken worden gebruikt om zowel systeem als user gegenereerde data op een gestructureerde wijze uit het EPD te halen. De meeste producten focussen zich nu nog op het laaghangend fruit van systeem gegenereerde data zoals lab waarden of verrichtingscodes. Aangezien deze gegevens reeds gestandaardiseerd worden vastgelegd aan de bron en ongeacht de inrichtingskeuzes van het ziekenhuis altijd hetzelfde zijn. Dit komt de schaalbaarheid ten goede en vergt slechts beperkte inspanning om snel resultaten te zien.

De uitdaging ligt voornamelijk bij het extraheren van user gegenereerde data, waar vrije tekst notities worden omgezet naar gestructureerde gegevens en metadata. IKNL doet samen met partnerorganisaties hier onderzoek in. De eerste uitkomsten hiervan zijn positief, maar de onderliggende modellen moeten nog

¹⁵ In lijn met kamerbrief Visie en strategie secundair datagebruik gezondheidszorg d.d. 13-04-2023, zie ook: <https://www.rijksoverheid.nl/documenten/kamerstukken/2023/04/13/kamerbrief-over-visie-en-strategie-secundair-datagebruik>

¹⁶ Zie ook: <https://www.health-ri.nl/news/bnc-fiche-european-health-data-space>

¹⁷ Conform koers ingezet vanuit Ministerie Volksgezondheid, Welzijn en Sport middels de Wegiz (en bijbehorende meerjarenagenda) en het Integraal Zorg Akkoord, zie ook: <https://www.informatieberaadzorg.nl/actueel/nieuws/2023/04/18/eerste-kamer-stemt-voor-wetsvoorstel-elektronische-gegevensuitwisseling-in-de-zorg>

¹⁸ Zie ook: <https://www.openehr.org/>

verder worden getraind om vals positieve en vals negatieve uitkomsten zoveel mogelijk te ondervangen. Afhankelijk van de inrichtingskeuzes is de kans aanwezig dat deze modellen op individueel ziekenhuisniveau nog steeds moeten worden gefinetuned en aangepast om goed te werken. Dit belemmert de mogelijke schaalbaarheid van de modellen. Deze belemmering kan deels worden weggenomen door EPD data van ziekenhuizen te transformeren naar bijvoorbeeld een lokale OpenEHR of OHDSI OMOP¹⁹ datamodel node en hierop de modellen te laten draaien. Dit bevordert de schaalbaarheid en efficiëntie. De transitie van EPD data naar deze datamodellen, en het onderhoud hiervan, vergt daarentegen weer inspanningen van de ziekenhuizen zelf. De vraag is of ziekenhuizen hiervoor de tijd en middelen heeft om dit te kunnen uitvoeren.

Helaas bleek binnen de tijd van de case studie het niet mogelijk om alle vragen op te pakken en te beantwoorden die er zijn rondom de behandeling van CRC patiënten met een BRAF V600E mutatie. De komend tijd zullen er aanvullende analyses worden uitgevoerd waarbij de verschillende behandelingen naast elkaar worden gezet en vergeleken. Daarbij zal er ook worden gekeken naar de progressie vrije overleving en/of time to treatment failure.

¹⁹ Zie ook: <https://www.ohdsi.org/data-standardization/>

Bijlagen

B1 Leden werkgroepen

Werkgroep Inhoud

Naam	Organisatie
Wim Goettsch (Voorzitter ROR DGM)	ZINL
Alet Jansen-Mager	NVMO
Ly Tran	WAR-CG
Ann van Ganse	Pierre-Fabre
Gerhard ten Hoor	Pierre-Fabre
Martijn van Gils	CBG
Naomi Jessurun	Lareb
Merel Kimman	MUMC
Geraldine Vink	PLCRC/IKNL
Miriam Koopman	PLCRC/IKNL
Heidi van Doorne	PLCRC/IKNL
Dorien Lobbezoo	ZINL
Marjan van den Bosch	St. Darmkanker
Doranne Hilarius	NVZA
Floris Lochman	ZN
Ghislaine Steenberghe	ZN

Werkgroep Techniek

Naam	Organisatie
Anja van Gestel	PLCRC/IKNL
Frank Martin	PLCRC/IKNL
Patrick Lubbers	PLCRC/IKNL
Geraldine Vink	PLCRC/IKNL
Heidi van Doorne	PLCRC/IKNL

B2 Specificatie data-items

#	Item	Bron
#1	Geboortedatum	NKR (standaard)
#2	Geslacht	NKR (standaard)
#3	Comorbiditeit	NKR (standaard)
#4	Performance status	NKR (standaard)
#5	Datum diagnose primaire tumor	NKR (standaard)
#6	TNM stadium primaire tumor	NKR (standaard)
#7	Lokalisatie primaire tumor	NKR (standaard)
#8	Datum diagnose metastasen	NKR (standaard)
#9	Lokalisatie metastasen	NKR (standaard)
#10	Moleculaire status	NKR (standaard)
#11	Systemische behandeling (middelen)	NKR (standaard)
#12	Datum start systemische behandeling	NKR (standaard)
#13	Datum stop systemische behandeling	NKR (standaard)
#14	Reden stop systemische behandeling	NKR (standaard)
#15	Vitale status	NKR (standaard)
#16	EQ-5D(-5L)	PROFIEL (PLCRC)
#17	EORTC QLQ-C30	PROFIEL (PLCRC)
#18	EORTC QLQ-CR29	PROFIEL (PLCRC)
#19	Aantal opnames	DHD
#20	Aantal ligdagen	DHD
#21	Aantal SEH bezoeken	DHD
#22	Aantal operaties	DHD
#23	Aantal dagbehandelingen	DHD
#24	Dosering encorafenib-cetuximab	DHD

B3 Uitkomsten probabilistisch matchen

#	Regie-op-Registers ID	Aantal
#1	22	11
#2	144	10
#3	81	10
#4	61	9
#5	139	8
#6	122	7
#7	9	6
#8	40	6
#9	131	5
#10	133	5
#11	165	4
#12	106	4
#13	115	3
#14	166	2

Het probabilistisch matchen is uitgevoerd a.d.h.v. de kenmerken geslacht, geboortedatum, viercijferige postcode en topografie van de tumor. Dit maakt het mogelijk om 153 van de 167 patiënten direct te koppelen aan de NKR. De overige 14 patiënten hebben in totaal 67 mogelijke matches zoals weergegeven in bovenstaande tabel. Voor de correcte matching is een handmatige actie van een datamanager noodzakelijk.

Bij het onderdeel topografie van de tumor is enkel gekeken naar de *basis ICDO-10* codering van de hoofdlocatie van de tumor CXX en niet naar de *complete ICDO-10* codering met bijbehorende sublocatie CXX.X. Er is hiervoor gekozen omdat bij de *basis ICDO-10* codering een veel hogere succes rate werd gehaald met het matchen met de NKR, zoals weergegeven in onderstaande tabellen.

Basis ICDO-10 codering (o.b.v. 153 patiënten)

Match	136 (89%)
Geen match	17 (11%)

Complete ICDO-10 codering (o.b.v. 153 patiënten)

Match	44 (29%)
Geen match	109 (71%)

Mogelijke reden voor dit verschil is dat DHD zich baseert op verrichtingen data en de NKR op klinische data uit het EPD.

B4 Combinaties van middelen

In de data komen 21 verschillende geneesmiddelen voor;

Geneesmiddel

5-FU
Bevacizumab
Binimetinib
Capecitabine
Cemiplimab
Cetuximab
Dabrafenib
Encorafenib
Ipilimumab
Irinotecan
Nivolumab
Olaparib
Panitumumab
Pembrolizumab
Spartalizumab
Tegafur
Tegafur/Gimeracil/Oteracil
Trametinib
Trifluridine/Tipiracil
Vemurafenib

Per behandellijn worden middelen vaak in combinaties en/of kuren toegediend, in de data komen er 68 verschillende (of overlappende) combinaties voor; Een deel van deze combinaties komt door de definitie van een behandellijn: *Progressie van ziekte en daarbij start nieuw chemotherapeuticum²⁰ en/of nieuwe targeted therapie wordt gezien als een nieuwe behandellijn.* (zie paragraaf 3.1.4.)

Combinaties

5-FU
5-FU, Bevacizumab, Oxaliplatin
Capecitabine
Capecitabine, Bevacizumab, Oxaliplatin
Capecitabine, Irinotecan, Bevacizumab
Capecitabine, Irinotecan, Bevacizumab, Oxaliplatin
Capecitabine, Irinotecan, Oxaliplatin
Capecitabine, Oxaliplatin
Capecitabine, Tegafur/Gimeracil/Oteracil, Irinotecan, Bevacizumab, Oxaliplatin
Capecitabine, Tegafur/Gimeracil/Oteracil, Irinotecan, Oxaliplatin
5-FU, Irinotecan
5-FU, Irinotecan, Bevacizumab
5-FU, Irinotecan, Bevacizumab, Oxaliplatin
5-FU, Irinotecan, Cetuximab
5-FU, Irinotecan, Encorafenib, Cetuximab
5-FU, Irinotecan, Encorafenib, Cetuximab, Bevacizumab

²⁰ Flouropyrimidines zijn uitwisselbaar

5-FU, Irinotecan, Encorafenib, Cetuximab, Bevacizumab, Oxaliplatin
5-FU, Irinotecan, Oxaliplatin
5-FU, Irinotecan, Panitumumab,
5-FU, Oxaliplatin
5-FU, Panitumumab, Oxaliplatin
5-FU, Tegafur, Capecitabine, Irinotecan, Bevacizumab
5-FU, Vemurafenib, Cetuximab, Bevacizumab, Oxaliplatin
5-FU, Tegafur/Gimeracil/Oteracil, Irinotecan, Bevacizumab, Oxaliplatin
Capecitabine
Capecitabine, Bevacizumab
Capecitabine, Bevacizumab, Oxaliplatin
Capecitabine, Encorafenib, Cetuximab, Bevacizumab, Oxaliplatin
Capecitabine, Encorafenib, Cetuximab, Oxaliplatin
Capecitabine, Irinotecan
Capecitabine, Irinotecan, Bevacizumab, Oxaliplatin
Capecitabine, Irinotecan, Encorafenib, Cetuximab, Bevacizumab, Oxaliplatin
Capecitabine, Oxaliplatin
Capecitabine, Pembrolizumab, Bevacizumab
Capecitabine, Tegafur/Gimeracil/Oteracil, Bevacizumab
Capecitabine, Tegafur/Gimeracil/Oteracil, Bevacizumab, Oxaliplatin
Capecitabine, Tegafur/Gimeracil/Oteracil, Oxaliplatin
Capecitabine, Trifluridine/Tipiracil, Bevacizumab, Oxaliplatin
Cetuximab
Chemotherapie in buitenland, Capecitabine
Dabrafenib, Spartalizumab
Dabrafenib, Trametinib
Encorafenib, Binimetinib, Cetuximab
Encorafenib, Cetuximab
Encorafenib, Cetuximab, Bevacizumab
Encorafenib, Cetuximab, Panitumumab
Irinotecan
Irinotecan, Bevacizumab
Irinotecan, Encorafenib, Cetuximab
Irinotecan, Nivolumab
Irinotecan, Panitumumab
Nivolumab
Nivolumab, Ipilimumab
Olaparib
Oxaliplatin
Panitumumab
Pembrolizumab
Studiemedicatie, immunotherapie, Trifluridine/Tipiracil, Dabrafenib
Studiemedicatie, Targeted therapy, Cemiplimab
Targeted therapy, Encorafenib, Cetuximab
Tegafur, Capecitabine, Bevacizumab
Tegafur, Oxaliplatin
Tegafur/Gimeracil/Oteracil
Tegafur/Gimeracil/Oteracil, Bevacizumab

Tegafur/Gimeracil/Oteracil, Bevacizumab, Oxaliplatin
Trifluridine/Tipiracil
Trifluridine/Tipiracil, Encorafenib, Cetuximab
Trifluridine/Tipiracil, Irinotecan

Kijkend naar de patiënten waarbij encorafenib-cetuximab is voorgeschreven blijven er van de initiële 21 geneesmiddelen nog 17 over, waarbij 13 verschillende combinaties in de data zijn terug te vinden. Door de gebruikte definities worden er middelen in één behandeling samengevoegd die niet per definitie tegelijk gegeven zijn. Bij de middelen is een start en stopdatum bekend;

Combinaties

5-FU, Irinotecan, Encorafenib, Cetuximab
5-FU, Irinotecan, Encorafenib, Cetuximab, Bevacizumab
5-FU, Irinotecan, Encorafenib, Cetuximab, Bevacizumab, Oxaliplatin
Capecitabine, Encorafenib, Cetuximab, Bevacizumab, Oxaliplatin
Capecitabine, Encorafenib, Cetuximab, Oxaliplatin
Capecitabine, Irinotecan, Encorafenib, Cetuximab, Bevacizumab, Oxaliplatin
Encorafenib, Binimetinib, Cetuximab
Encorafenib, Cetuximab
Encorafenib, Cetuximab, Bevacizumab
Encorafenib, Cetuximab, Panitumumab
Irinotecan, Encorafenib, Cetuximab
Targeted therapy, Encorafenib, Cetuximab
Trifluridine/Tipiracil, Encorafenib, Cetuximab