

Dataverwerking en functionaliteit monitor uitkomstindicatoren

Auteur(s)

Luc te Boekhorst DHD

Harm Nico Plomp DHD

Versie: 1.0

April 2018

Concept

Versiebeheer

Versie	Datum	Auteur(s)	Toelichting	Status
0.1	27-2-2018	Luc te Boekhorst	FO PUB en MUS zijn als basis gebruikt voor dit document	Concept intern DHD
0.2	2-3-2018	Luc te Boekhorst	Opmerkingen Harm Nico en Rick verwerkt	Concept intern DHD
0.3	7-3-2018	Luc te Boekhorst	Aanpassingen	Concept sub-werkgroep PUB / MUS
0.3	8-3-2018	Marielle Went	Commentaar	
0.4	13-3-2018	Luc te Boekhorst	Verwerking 8-3	Concept sub werkgroep PUB / MUS
0.5	20-03-2018	Harm Nico Plomp en LtB	Aanpassingen na overleg HM en LtB	Concept Versie voor projectgroep
1.0	4-4-2018	LtB	Aanpassingen n.a.v. projectgroep	Concept te goedkeuring



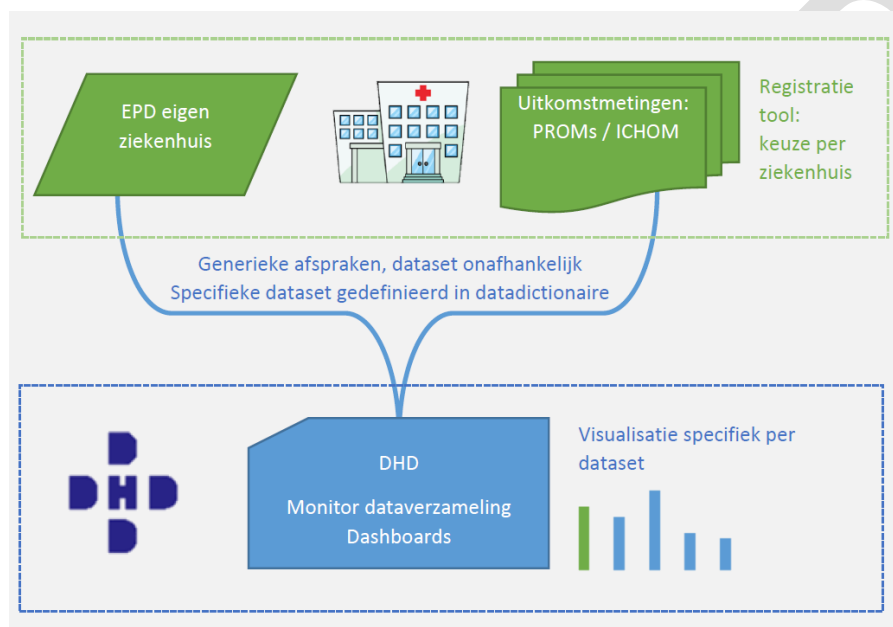
Inhoudsopgave

1	Inleiding.....	4
1.1	Achtergrond	4
1.2	Doel	4
1.3	Inhoud.....	5
2	Dataverwerking	6
2.1	Verzameling (ziekenhuis)	6
2.2	Aanlevering (ziekenhuis)	6
2.3	Controle (DHD).....	7
2.4	Goedkeuring (ziekenhuis).....	7
2.5	Opslag (DHD)	8
2.6	Bewerking (DHD).....	8
2.7	Verwerking (DHD)	9
3	Functionaliteit monitor uitkomstindicatoren	10
3.1	Lay-out en vormgeving	10
3.2	Toegankelijkheid.....	10
4	Bijlagen	12
4.1	Bijlage 1: Dataprotocol (generiek)	12
4.2	Bijlage 2: Datadictionaire (dataset specifiek)	12
4.3	Bijlage 3: Datastroom (generiek).....	13
4.4	Bijlage 4: Aanlevering (generiek).....	14
4.4.1	<i>Bijlage 4a: Aanleverspecificaties</i>	14
4.4.2	<i>Bijlage 4b: Voorbeeld aanleverbestand</i>	14
4.5	Bijlage 5: Missing values (generiek).....	15
4.5.1	<i>Bijlage 5a: Beslisboom</i>	15
4.5.2	<i>Bijlage 5b: statutabel controle dataset</i>	16
4.6	Bijlage 6: Visualisatie (dataset specifiek)	17
4.7	Bijlage 7: Toelichting missings (algemeen).....	17

1 Inleiding

1.1 Achtergrond

Steeds meer ziekenhuizen vragen met behulp van uitkomstenregistratie op een gestandaardiseerde manier behandeluitkomsten uit bij patiënten. DHD zorgt ervoor dat de gegevens die binnen de verschillende ziekenhuizen worden verzameld, op één plek samengebracht worden. Aan de registratie is een interactief dashboard gekoppeld waarmee de deelnemende ziekenhuizen analyses kunnen verrichten op de door hen aangeleverde informatie en de eigen resultaten kunnen vergelijken met die van andere ziekenhuizen (benchmarking). Het onderling vergelijken vindt plaats op het ziekenhuis-niveau.



Een dataset bevat uitkomstindicatoren aangevuld met een beperkt aantal casemix variabelen en/of procesindicatoren rond een specifieke diagnose of aandoening. De casemix variabelen en procesindicatoren kunnen aan de dataset worden toegevoegd vanuit het EPD. De uitkomstindicatoren hebben betrekking op PROMS (bijvoorbeeld de vragenlijsten uit de ICHOM-dataset) en/of behandeluitkomsten (bijv. afkomstige uit kwaliteitsregistraties).

1.2 Doel

Het doel van dit document is om het proces van dataverwerking en de functionaliteiten van de monitor te definiëren, waarbij de rollen en verantwoordelijkheden tussen de ziekenhuizen en DHD helder zijn. Het document is van toepassing op de nu lopende projecten PUB (borstkanker) en MUS (Schisis) en moet toepasbaar zijn op nieuwe projecten rond andere ziektebeelden.

1.3 Inhoud

Dit document bestaat uit twee delen. Het eerste deel (H2) beschrijft hoe de **dataverwerking** plaats vindt. Wie doet wat, welke eisen worden er gesteld aan opslag en verwerking en hoe verloopt het proces?

In de datadictionary staat een gedetailleerde beschrijving en definitie van de betreffende dataset. Het bevat per item:

- de codering
- in welke fase van de behandeling (timing) de vraag aan de patiënt gesteld moeten worden
- de toegestane antwoordmogelijkheden
- de onderling afhankelijkheden tussen de vragen
- de eventuele afgeleide waarde bij een missing value
- of de vraag verplicht is
- of de vraag geschikt is voor imputatie
- welke scoringsformule gebruikt moet worden

Tevens wordt in de datadictionary vastgelegd welke patiëntengroep het betreft, vanaf welke datum er gegevens worden aangeleverd en in welke frequentie.

Het tweede deel (H3) beschrijft de **functionaliteit van de monitor** met de specificaties van het interactieve dashboard waarin de uitkomsten getoond worden.

Dit document bevat een aantal bijlagen met generieke afspraken en bijlagen met dataset-specifieke afspraken. Hieronder een overzicht van de bijlagen waarbij elke afzonderlijke dataset een eigen specifieke datadictionary en visualisatie heeft. Voor PUB en MUS is dit al ingevuld, voor een nieuw project / dataset zal dat aangevuld moeten worden.

Documentnaam	Generiek	PUB specifiek	MUS specifiek	Nieuw project
Bijlage 1 Dataprotocol uitkomstenregistraties - Aanvullende Dienstverlenings-overeenkomst - Dataverwerking BCZHZ	x	x x	x	x
Bijlage 2 Datadictionary		x	x	x
Bijlage 3 Datastroom	x			
Bijlage 4a Aanleverspecificaties	x			
Bijlage 4b Voorbeeld aanlevering	x			
Bijlage 5a Missing value beslisboom	x			
Bijlage 5b Missing value statustabel	x			
Bijlage 6 Visualisatie		x	x	x
Bijlage 7 Toelichting missings	x			



2 Dataverwerking

Het hele proces van het verzamelen van gegevens in het ziekenhuis tot aan het tonen van de gegevens in de webapplicatie van DHD, staat beschreven in het stroomschema in [bijlage 3, datastroom](#).

Hieronder staat per processtap beschreven wat er moet gebeuren en wat er is afgesproken.

2.1 Verzameling (ziekenhuis)

1. Elk ziekenhuis is zelf verantwoordelijk voor het vastleggen en uitvragen van de items in de dataset volgens de specificaties vastgesteld in de datadictionaire.
2. Minimaal 1x per jaar vindt overleg plaats tussen de deelnemende ziekenhuizen of herziening van de datadictionaire nodig is, welke items dit betreft en de ingangsdatum (minimaal drie maanden na het vaststellen van de wijziging).
3. Elk ziekenhuis is zelf verantwoordelijk voor de inhoud en compleetheid van de dataset en controleert de set volgens de beslisboom missing values en de statustabel. ([Bijlage 5a en 5b](#))
4. Er vindt door het ziekenhuis in principe geen bewerking plaats op de data en geen selectie op de vragenlijsten zoals opgeslagen in de data-capturetool.
5. Alle opgeslagen vragenlijsten worden doorgestuurd naar DHD. Ook vragenlijsten die slechts gedeeltelijk zijn ingevuld. Op deze wijze kan de compliance betrouwbaar worden berekend.

2.2 Aanlevering (ziekenhuis)

1. Elk ziekenhuis heeft ten minste één persoon die de rol van datamanager heeft. Deze persoon heeft toegang tot alle onderdelen van de applicatie zoals deze bij DHD staat.
2. De verzamelde dataset zal volgens de gedefinieerde aanleverspecificaties en volgens de beschreven aanleverprocedure, door de datamanager worden aangeleverd aan DHD. ([Bijlage 4a, 4b en 4c](#))
3. Na uploaden van het bestand verschijnt er in de monitor een bevestiging dat het bestand ontvangen is.
4. Heraanleveringen zijn mogelijk. Dit kan op elk moment bij constatering van een fout.
5. Indien de nieuwe set niet direct beschikbaar is, is het mogelijk om via DHD een vorige versie terug te zetten.
6. Gegevens worden vanaf een afgesproken datum, periodiek aangeleverd volgens een in de stuurgroep (of ander besluitnemend orgaan) bepaalde frequentie. Deze afspraken zijn vastgelegd in de datadictionaire.

Bij elke data-aanlevering wordt de gehele dataset (de data van alle patiënten) opnieuw aangeleverd (dus geen incrementele aanlevering). Er kan voor gekozen worden om historische data over een bepaalde periode 'vast te zetten' en niet opnieuw aan te leveren. (Bijvoorbeeld als een ziekenhuis voor het eerst gebruik gaat maken van de monitor. Historische data kunnen dan in één keer worden aangeleverd.) Nieuwe data-aanleveringen bevatten alle data vanaf de startdatum van het gebruik van de monitor. Dit kan ook worden toegepast bij de overgang naar een nieuwe versie van een ICHOM-vragenlijst. Deze afspraken zijn vastgelegd in de datadictionaire.



2.3 Controle (DHD)

1. DHD verwerkt na elke aanlevering de aangeleverde gegevens.
2. DHD controleert of het aanleverbestand voldoet aan de aanleverspecificaties en de datadictionaire.
3. De aangeleverde gegevens worden gecontroleerd zoals benoemd in de beslissboom missing values en zo nodig voorzien van een status.
([Bijlage 5a en 5b](#))
4. DHD stelt een aanleVERRAPPORTAGE op met daarin alle geconstateerde afwijkingen t.o.v. de aanleverspecificaties en datadictionaire
 - a. Aantal regels (huidige aanlevering en vorige aanlevering)
 - b. Aantal patiënten (huidige aanlevering en vorige aanlevering)
 - c. Aantal missings
 - d. Regels met status 'fout' en 'waarschuwing'
 - e. Verplichte velden met een melding fout
5. Zodra de aanleVERRAPPORTAGE gereed is (binnen 5 werkdagen nadat de data door DHD ontvangen is), wordt de rapportage via de upload-pagina van de monitor aan de datamanager beschikbaar gesteld. Alle datamanagers van het betreffende ziekenhuis, ontvangen hierover een mail.

2.4 Goedkeuring (ziekenhuis)

1. Nadat de aanleVERRAPPORTAGE beschikbaar is, is de datamanager verantwoordelijk om de data-aanlevering goed- of af te keuren. De betreffende datamanager van het ziekenhuis kan een toelichting invoeren waarom data-aanlevering is goed- of afgekeurd. Er staat geen limiet op de termijn waarop de datamanager de data-aanlevering beoordeelt.
2. Als een datamanager de data-aanlevering goedkeurt, wordt de data verwerkt in de monitor en kan de status van de aanlevering niet meer door een datamanager aangepast worden. Alle gebruikers (datamanagers en standaard gebruikers) van het betreffende ziekenhuis ontvangen een mail dat er nieuwe data van het ziekenhuis in de monitor is verwerkt.
3. Als een data-manager de data-aanlevering afkeurt, wordt de data niet verwerkt in de monitor. De data van de laatst goedgekeurde data-aanlevering blijft zichtbaar in de monitor. Er moet een nieuwe data-aanlevering plaatsvinden.
4. Door een data-aanlevering te selecteren worden de volgende gegevens per aanlevering getoond:
 - a. Datum aanlevering
 - b. Status van de aanlevering (ter beoordeling / goedgekeurd / afgekeurd)
 - c. Naam van de data-manager die de aanlevering heeft gedaan
 - d. Toelichting bij de aanlevering
 - e. Datum geaccordeerd of geweigerd
 - f. Naam van de data-manager die de goedkeuring/afkeuring heeft gedaan
 - g. Toelichting bij de goedkeuring/afkeuring
 - h. Beschikbare kengetallenDeze gegevens zijn voor andere gebruikers binnen hetzelfde ziekenhuis in te zien.

2.5 Opslag (DHD)

1. Alleen de data van goedgekeurde aanleveringen worden meegenomen in de verdere verwerking.
2. Zodra een data-aanlevering door het ziekenhuis is goedgekeurd, wordt de voorlaatste goedgekeurde data-aanlevering uit de monitor verwijderd. Gebruikers (zowel de data manager als de standaard gebruikers) kunnen geen eerdere goedgekeurde data-aanleveringen selecteren om te tonen in de monitor.
3. DHD kan, op verzoek van en in nauw overleg met de datamanager(s) van het ziekenhuis, de meest recent goedgekeurde data-aanlevering in de monitor vervangen door een eerdere goedgekeurde aanlevering terug te plaatsen in de monitor (rollback).
4. Gegevens van de ziekenhuizen die behoren tot een samenwerkingsverband kunnen als één anoniem ziekenhuis getoond worden.

2.6 Bewerking (DHD)

1. DHD zal de gegevens van patiënten die verplichte vragen hebben met een status 'fout' verwijderen uit de database.
2. Er wordt imputatie uitgevoerd als een geldige waarde wordt toegekend aan een ontbrekende waarde in een dataset. Ten aanzien van het imputeren van variabelen gelden de volgende uitgangspunten:
 - a. Imputatie wordt alleen toegepast op case-mix variabelen (zie datadictionaire)¹.
 - b. Imputatie kan alleen plaatsvinden op variabele met de status 'Onbekend' of 'Ontbrekend veld'.
3. Missende case-mix variabelen worden geïmputeerd met behulp van multiple imputation. (PM)
4. DHD zal de antwoorden die missen volgens de 'beslisboom missing value' en die volgens de in de datadictionaire vastgestelde definitie een afgeleide waarde hebben, vullen met de afgeleide waarde.
5. DHD zal een aantal vragen mogelijk combineren zodat de afgesproken filters kunnen worden toegepast.
6. DHD zal, indien nodig, de antwoordwaarden bewerken met berekeningen volgens aangeleverde scoringstabellen die in de datadictionaire ([bijlage 2](#)) zijn gedefinieerd.
7. DHD zal statische bewerkingen uitvoeren om waarden zoals kwartielen, standaard deviatie, gemiddeldes, etc. te bepalen.
8. Bij voldoende aangeleverde data wordt een analyse gemaakt van de 'missing values' vanuit statistisch oogpunt of deze relevante informatiewaarde hebben.

Regressie (PM)

1. De ziekenhuizen worden met elkaar vergeleken door middel van random effect modellen.
2. Ordinale uitkomsten worden geanalyseerd met proportional odds logistische regressie.

¹ Het imputeren van niet cases-mix variabelen is theoretisch mogelijk. Of dat wenselijk en noodzakelijk is kan pas bepaald worden als er van een groot aantal verschillende instellingen data beschikbaar is. Dat zal op een later moment bepaald moeten worden en zal verwerkt worden in dit document.



2.7 Verwerking (DHD)

Nadat de data is bewerkt, zal deze worden verwerkt in de monitor. Alle gebruikers (datamanagers en standaard gebruikers zie 3.2 voor afspraken toegankelijkheid) van het betreffende ziekenhuis ontvangen een mail dat er nieuwe data van het ziekenhuis in de monitor is verwerkt.

CONCEPT



3 Functionaliteit monitor uitkomstindicatoren


3.1 Lay-out en vormgeving

De monitor bij DHD is opgebouwd uit een aantal mogelijke onderdelen, afhankelijk van de dataset.

1. De monitor is toegankelijk via de website van DHD (www.dhd.nl) en wordt ondersteund door de servicedesk van DHD.
2. DHD ondersteunt weergave via de gangbare internetbrowsers. DHD behoudt zich het recht voor verouderde versies van browsers niet meer te ondersteunen. Gebruikers worden daarover tijdig geïnformeerd.
3. De monitor is desktop en tablet vriendelijk.
4. Opmaak, kleurstelling, pagina-hiërarchie, navigatie-mogelijkheden van de monitor zijn conform DHD stijl.
5. Standaard is er een homepage waarin ruimte is voor nieuwsberichten en een algemene beschrijving en toelichting op het gebruik van de monitor.
6. Er is een pagina met gegevens over de aangeleverde data zoals beschreven in 3.2.
7. Er is een aantal tabbladen die qua inhoud (filters en visualisatie) specifiek per dataset worden beschreven in één of meerdere documenten die als bijlagen worden toegevoegd aan dit document.
 - a. Er is een analysetool in een Excel-omgeving die bestaat uit interactieve draaitabellen en grafieken waarmee o.a. eigen analyses gemaakt kunnen worden.
 - b. Dashboardpagina's met filters, tabellen en grafieken die informatie tonen uit de beschikbare dataset.
 - c. Downloadpagina: Hier wordt de mogelijkheid geboden om (eigen) gegevens te downloaden voor verdere analyses.

3.2 Toegankelijkheid

1. De monitor is alleen toegankelijk voor personen met een gebruikersaccount. Per account wordt aangegeven welke rol de gebruiker heeft (datamanager of standaard gebruiker).
2. Toegang tot monitor is beveiligd met SMS-verificatie. Gebruikers moeten dus beschikken over een eigen mobiele telefoon.
3. Elke ziekenhuis heeft één of meer lokale beheerders van de monitor. De RvB moet hiervoor schriftelijk toestemming geven, waarna DHD deze accounts aanmaakt.
4. De lokaal beheerder moet de rol van datamanager hebben.
5. De lokaal beheerder kan zelf accounts aanmaken voor personen met de rol van standaard gebruiker.
6. Accounts voor datamanagers kunnen alleen worden aangemaakt door DHD, op verzoek van de lokaal beheerder.

- 
7. Een standaard gebruiker heeft toegang tot de monitor en kan de data van de eigen instelling bekijken en gebruiken.
 8. Aantal accounts per ziekenhuis met toegang tot de monitor is niet gelimiteerd.

CONCEPT



4 Bijlagen

4.1 Bijlage 1: Dataprotocol (generiek)

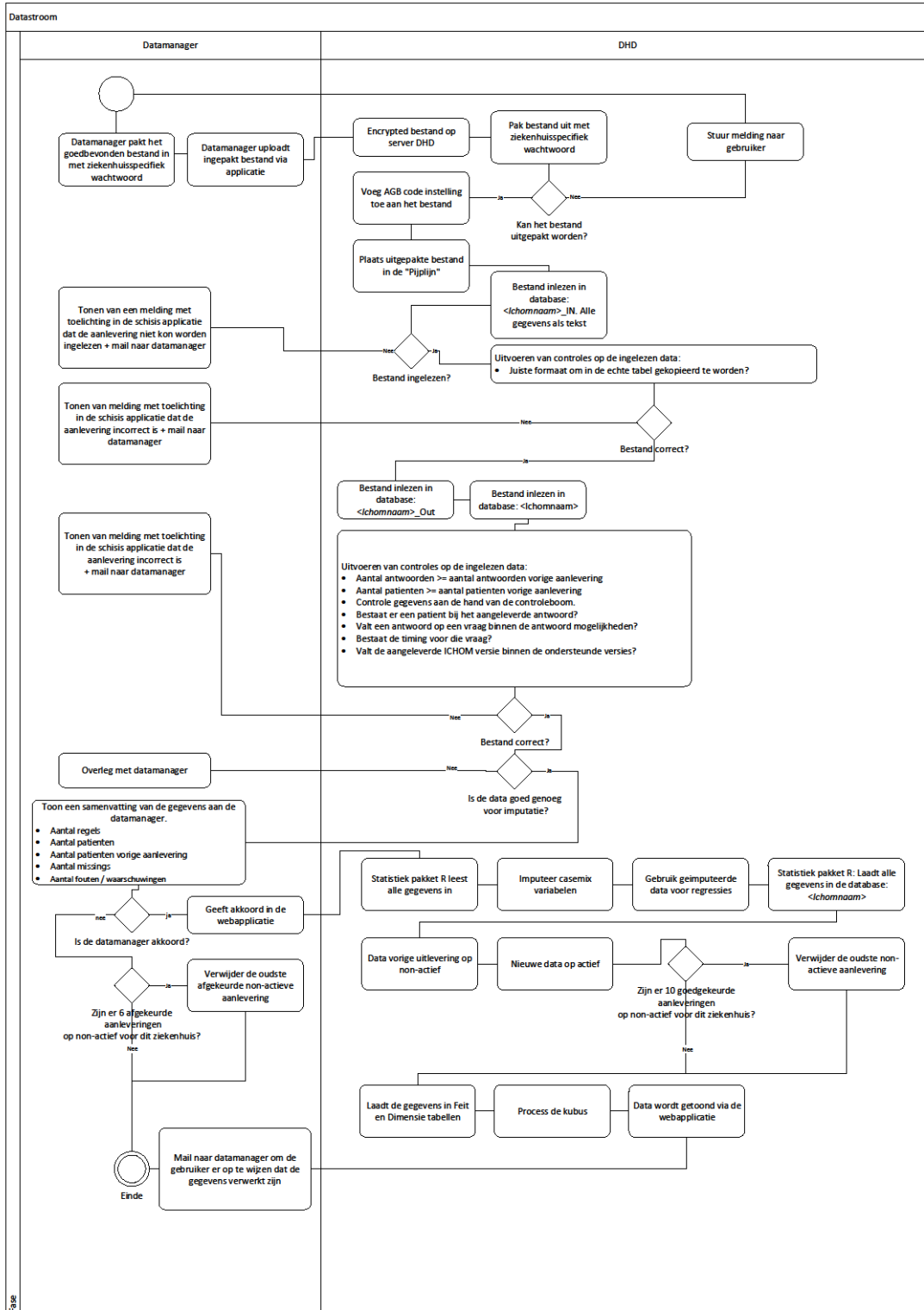
Generiek bestand, geldend voor elke dataset, ondertekend door de RvB, met daar waar nodig een specifieke bijlage.

4.2 Bijlage 2: Datadictionaire (dataset specifiek)

Specifiek bestand per vastgestelde dataset. Zie bestand <dataset>_Datadictionaire.xlsx

CONCEPT

4.3 Bijlage 3: Datastroom (generiek)





4.4 Bijlage 4: Aanlevering (generiek)

4.4.1 *Bijlage 4a: Aanleverspecificaties*

Aanleverspecificaties.pdf

Generiek bestand met de definities over hoe de gegevens technisch aangeleverd moeten worden. Inhoudelijk zullen de gegevens per dataset verschillen.

4.4.2 *Bijlage 4b: Voorbeeld aanleverbestand*

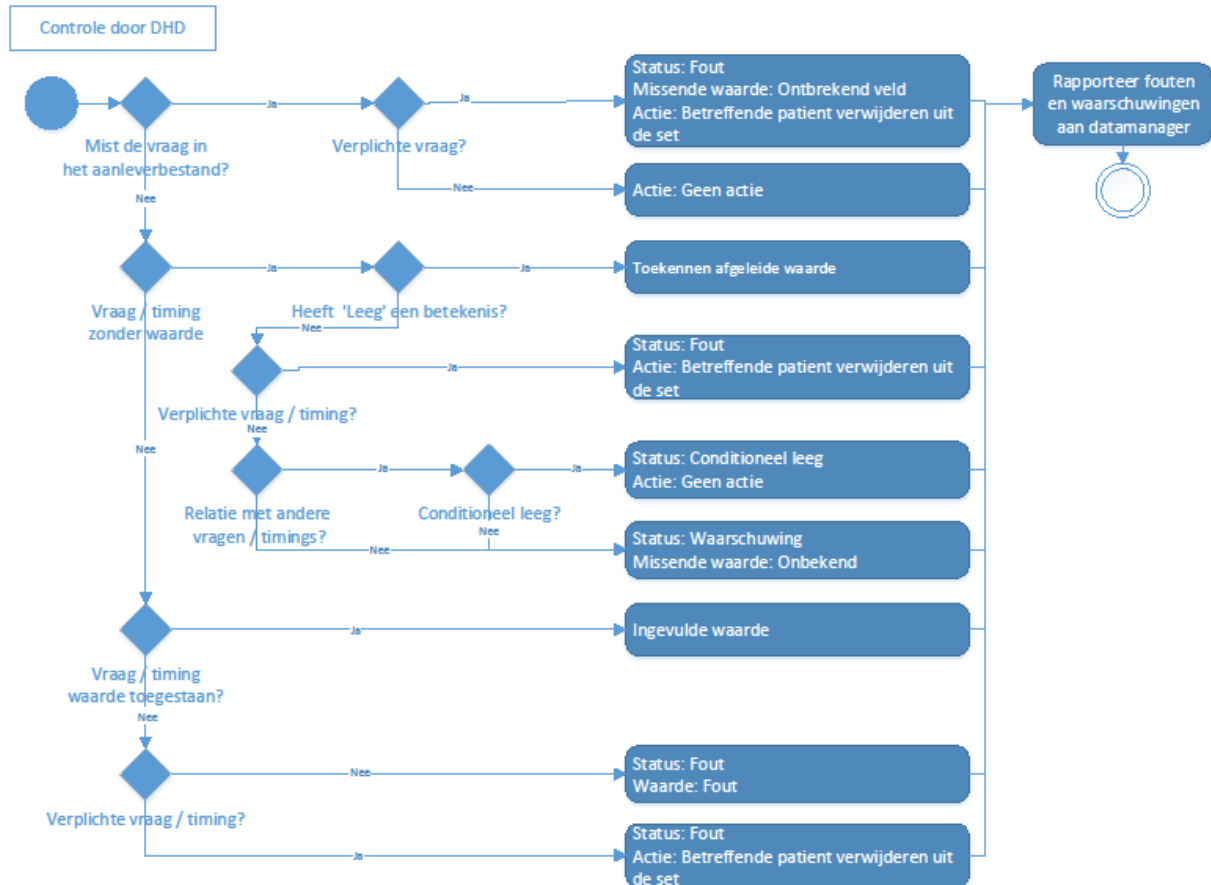
AnswerExample.csv

Generiek voorbeeld van hoe het aanleverbestand eruit moet zien.

CONCEPT

4.5 Bijlage 5: Missing values (generiek)

4.5.1 Bijlage 5a: Beslisboom



4.5.2 Bijlage 5b: statustabel controle dataset

Status uitkomsten controle	Toelichting	Toegekende waarde (wel/niet imputeren)	Vermelding foutenrapportage
Correct ingevuld	Ingevulde waarde is juiste en volledig (voldoet aan kenmerken datadictionaire)	Ingevulde waarde	Geen
Afgeleide waarde	Het veld ontbreekt of het veld bevat bij aanlevering geen waarde. Aan missing value kan een betekenis worden toegekend, al dan niet in combinatie met de waarden van andere velden (zie beschrijving datadictionaire)	Toekennen afgeleide waarde	Geen
Conditioneel leeg	Het veld ontbreekt of het veld bevat bij aanlevering geen waarde. De lege waarde heeft geen afgeleide waarde. De missing value wordt afgedwongen of is in verband te brengen met de waarde in een of meerdere andere velden.	Geen	Geen
Onbekend	Het veld bevat bij aanlevering geen waarde. De lege waarde heeft geen afgeleide waarde. De lege waarde van het veld wordt niet afgedwongen of is niet in verband te brengen met de waarde in een of meerdere andere velden.	'Onbekend' (als case-mix var, dan mogelijk imputatie)	waarschuwing
Fout	Het veld is aangeleverd en bevat een waarde die niet voldoet aan de kenmerken uit de datadictionaire	'Fout'	Foutmelding
Ontbrekend veld	Het veld is niet meegeleverd in de aanlevering.	'Ontbrekend veld' (als case-mix var, dan mogelijk imputatie)	Foutmelding (bij verplichte velden) Geen (bij niet verplichte velden)



4.6 Bijlage 6: Visualisatie (dataset specifiek)

Een aantal dataset specifieke bestanden met daarin uitgewerkt hoe de visualisatie van de gegevens vorm gegeven wordt. Deze bestande bevatten informatie over Filters / Dashboard / Downloads / Mock-ups.

4.7 Bijlage 7: Toelichting missings (algemeen)

Achtergrond Missings.pdf

Algemeen document met een toelichting op missings.

CONCEPT