

Stappenplan Healthy AI (HAI)

Een stappenplan voor de kwaliteit van AI-software

Sade Faneyte, klinisch informaticus
versie 1.1



Artificial intelligence in de zorg

Introductie

Toekomstbestendige zorg vergt een transformatie van de zorg, ondersteund door technologische ontwikkelingen. Een van die technologische ontwikkelingen is *Artificial intelligence (AI)*, de toepassing van computersystemen die (humaan) intelligent gedrag vertonen. De inzet van AI is onvermijdelijk om de zorg efficiënt en betaalbaar te houden. Zo kan AI impact hebben op de gehele patiëntreis, van ziektepreventie, betere diagnostiek en behandelingen, tot de toegankelijkheid van zorg. Maar de toepassing van AI vindt nog niet op grote schaal plaats.

Binnen het **Maasstad Ziekenhuis (MSZ)** ontstaat steeds meer vraag naar de inzet van AI-gebaseerde *medische hulpmiddelen* in de vorm van software (hierna te noemen; *AI-software*). Patiënten vinden het belangrijk om goede afspraken te maken en voorwaarden te stellen aan het gebruik van AI. Maar het is onmogelijk om alle risico's weg te nemen met processen, wetten, en richtlijnen. **Verantwoorde implementatie van AI** vraagt dus vooral om een open gesprek, tussen patiënten en zorgverleners. Het MSZ zet daarom in op het creëren van optimale omstandigheden voor de inbedding van AI-technologie. Niet alleen het leervermogen van AI wordt hierdoor benut, maar ook die van de organisatie.

Datastrategie MSZ

De **datastrategie (2019-2024)** zet koers op het verhogen van de kwaliteit van leven van patiënten met de kracht van data. De strategie heeft als doel: zorgen voor doelmatige beschikbaarheid van data in bruikbare vorm en het ontwikkelen van middelen om deze data in te zetten voor waardegedreven zorg. De inzet van AI is een hulpmiddel hiervoor.

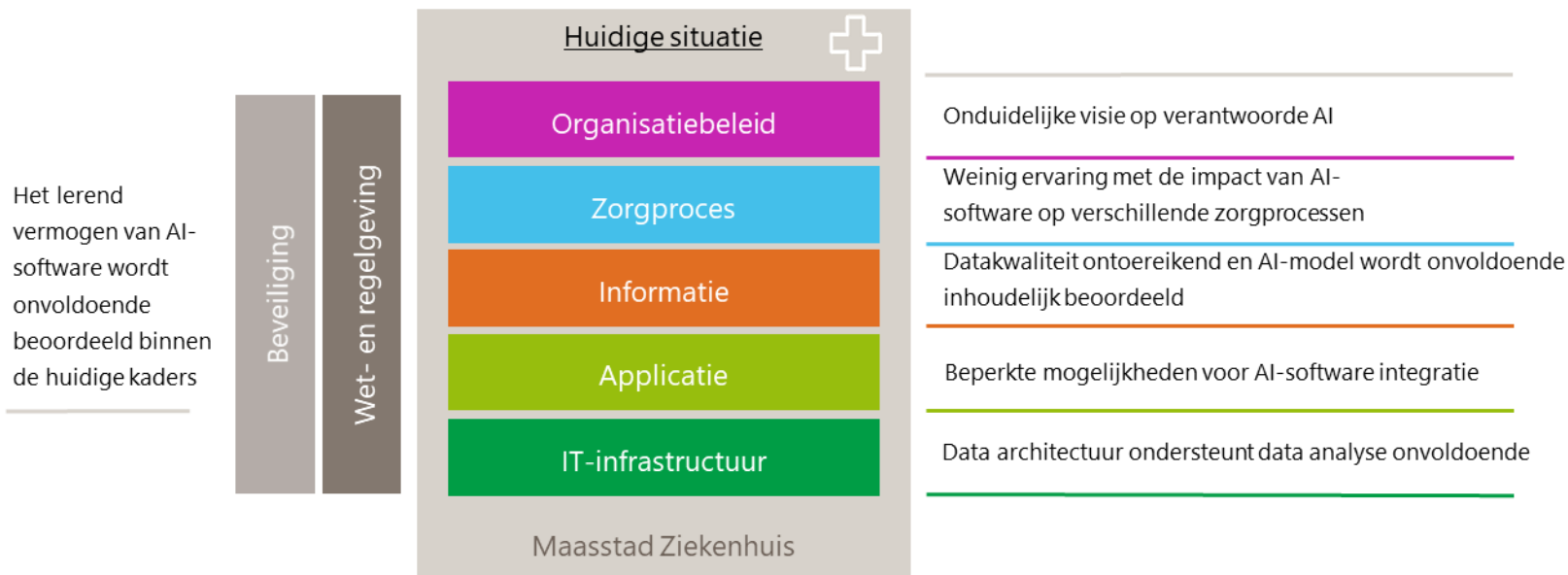


Uitdagingen bij de implementatie van AI-software

Stappenplan

De **grootste uitdagingen** liggen op het gebied van de kwaliteit en beschikbaarheid van data, transparantie en kwaliteitscontrole van AI, governance en organisatie, het juridisch kader, financiering, toegevoegde waarde van AI, en de educatie van zorgverleners. Maar ook sociale uitdagingen die vragen om een ethische reflectie, zoals vertrouwen en de rol van patiënten. Terwijl juist het **vertrouwen in AI** door patiënten toe neemt door voorbeelden en co-creatie van AI-toepassingen in de zorg.

Daarnaast is er nog weinig bekend over de **doelmatigheid** van AI-toepassingen in de zorg, waardoor de kosteneffectiviteit zich op een langere termijn nog moet uitwijzen. Desondanks worden er steeds meer AI-producten op de markt gebracht. Daarom is leren door te doen belangrijk, zodat er ook ruimte blijft om te leren van elkaar. Want juist de praktijkvoorbeelden helpen in het creëren van meer inzicht in hoe het werk van zorgverleners verandert.



Hoofdthema's van de uitdagingen bij de implementatie van AI-software (o.b.v. kwalitatief onderzoek)

Stappenplan HAI: Kwaliteit van AI-software



Stappenplan

Aanleiding

Het grootste deel van de AI-initiatieven in de zorg haalt de implementatiefase niet. Om adoptie te bevorderen, worden landelijk en internationaal kaders gevormd voor AI in de gezondheidszorg. Voor gebruik van AI als medisch hulpmiddel, zijn betrouwbaarheid en transparantie van groot belang voor de kwaliteit van zorg. Om de kwaliteit van AI-software beter te kunnen beoordelen is het **stappenplan Healthy AI (HAI)** ontwikkeld als eindontwerp voor de opleiding tot klinisch informaticus. Hiermee worden zorgprofessionals begeleid in het implementatieproces van AI-software en ontwikkelen zij kennis over AI in de zorg. Het stappenplan richt zich op voorspellende AI (*machine learning*) in *medische hulpmiddelen* en daarmee ook de klinische toepassing. De stappen sluiten aan bij **lokaal beleid** voor medische hulpmiddelen.

Doelgroep en gebruik

Het **stappenplan HAI** helpt **zorgverleners en ondersteuners** in de verantwoorde implementatie van AI-software. Met praktische handvaten voor iedere fase uit de levenscyclus van AI-software. Om het stappenplan te gebruiken stelt de zorgverlener een multidisciplinair **AI-team** samen. Met het team worden de stappen doorlopen in gezamenlijke sessies onder begeleiding van een klinisch informaticus. Aan de hand van **praktijkvoorbeelden** wordt de invulling van de stappen besproken.

In analogie met de gezondheid van patiënten die centraal staat in de zorg van de toekomst, is gekozen voor de term “**healthy**” AI. Want ook met de inzet van technologie moet de gezondheid van patiënten centraal staan, en AI kan daarin juist ondersteunend zijn.

Hoe gebruik je het stappenplan?

De [Home knop](#) navigeert naar het overzicht van het stappenplan. Elke stap is doorklikbaar naar een algemene beschrijving met een checklist, en naar de onderdelen die bij de stap horen. Daarin zijn praktijktips in blokken aangegeven en *definities* in schuingedrukte oranje tekst en **hyperlinks** naar andere onderdelen. Met de pijltoetsen kunnen de opvolgende onderdelen bereikt worden. Vóór de aanvraag van nieuwe AI-software moeten minstens stappen 1 t/m 6 worden doorlopen.

Stap 1: Bepaal de medische context

Stap 2: Bepaal het innovatieniveau

Stap 3: Bepaal de toegevoegde waarde

Stap 4: Bepaal de risico's en impact

Stap 5: Maak afspraken over taken, bevoegdheden en verantwoordelijkheden

Stap 6: Controleer toepasselijke wet- en regelgeving

Stap 7: Bereid voor op implementatie

Stap 8: Evalueer de uitkomsten

Definitielijst

Artificial Intelligence (AI): "Het soort systemen dat intelligent gedrag vertoont door hun omgeving te analyseren en – met enige graad van autonomie – actie te ondernemen om specifieke doelen te bereiken" – Europese Commissie

AI-software: ookwel Machine Learning-enabled Medical Devices (MLMD) genoemd (ref IMDRF) of Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) (ref FDA). Software als medisch hulpmiddel dat gebruik maakt van AI/machine learning

AI/ML-model: de relatie of functie als resultaat van een getraind machine learning algoritme met data

Algoritme: een set van regels en instructies die een computer toe kan passen op data (de wiskundige formules met het script)

Adaptieve algoritmes: in dit geval leert een algoritme van nieuwe input data waardoor de prestatie met de tijd verandert

Bias: systematisch verschil in de behandeling van objecten, mensen of groepen in vergelijking met anderen

Deep learning: subset van ML gebaseerd op de neurale structuur van het menselijk brein. Net als een neuraal netwerk gebruikt het verschillende abstractieniveaus om van de data te leren – vandaar de term 'diep' leren

Informed consent: geïnformeerde toestemming (informatieplicht en toestemmingsvereiste). Een verbijzondering van informed consent is 'shared decision making' (SDM), ofwel 'gezamenlijke besluitvorming'.

Intended use: het gebruik waarvoor een hulpmiddel is bestemd volgens de gegevens die door de fabrikant zijn verstrekt op het etiket, in de gebruiksaanwijzing, in reclame- of verkoopmateriaal of verklaringen en zoals dat door de fabrikant in de klinische evaluatie is gespecificeerd

Machine learning (ML): subset van AI waarbij er vooraf geen regels geprogrammeerd door de mens, maar sprake is van een lerend vermogen van de computer door patroonherkenning

Medisch hulpmiddel: een hulpmiddel dat door de fabrikant is bestemd om alleen of in combinatie te worden gebruikt bij de mens voor een of meer van de volgende specifieke medische doeleinden: diagnose, preventie, monitoring, voorspelling, prognose, behandeling of verlichting van ziekte (of compensatie van een letsel of een beperking). Dit is een verkorte versie van de volledige definitie uit de Medical Device Regulation

OTAP-omgeving: Een ontwikkel-, test-, acceptatie-, productieomgeving voor het testen en valideren van software

Pilot: een project opgezet om iets uit te proberen

Software validatie: controleren of de softwarespecificaties voldoen aan de behoeften van de gebruiker en het beoogde gebruik in een klinische omgeving. Meestal wordt met validatie van het AI-model verwezen naar de stap in het modelontwikkelingsproces waarbij een test dataset wordt gebruikt. Het is dus belangrijk om de term validatie altijd in context te benoemen. En validatie van het AI-model als onderdeel te zien van AI-software validatie

Statische algoritmes: een algoritme blijft niet leren van de data, maar wordt getraind op de grootst mogelijke dataset en vervolgens geschikt verondersteld voor implementatie in de klinische praktijk

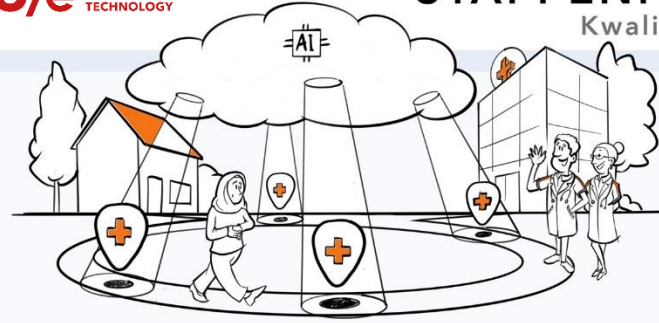
Temporale validatie: validatie methode waarbij prospectieve modelvoorspellingen op nieuwe data worden gedaan, zonder dat de eindgebruiker de resultaten gebruikt

Stappenplan Healthy AI

Kwaliteit AI software

STAPPENPLAN HEALTHY AI Kwaliteit AI software

 **Stap 1. Bepaal de medische context**



 **Stap 2. Bepaal het innovatieniveau**



 **Stap 3. Bepaal de toegevoegde waarde**



 **Stap 4. Bepaal de risico's & impact**



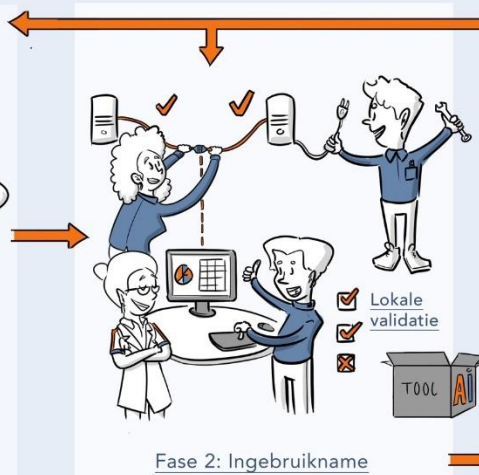
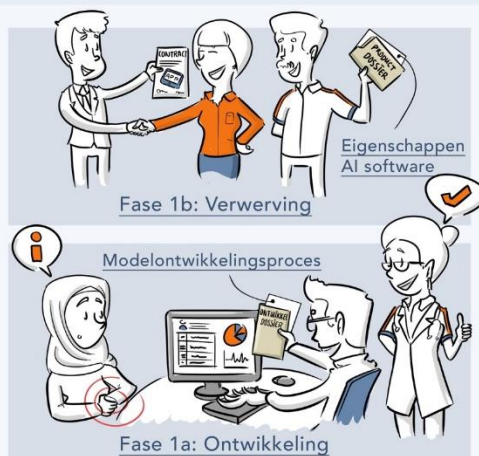
 **Stap 5. Maak afspraken over taken, bevoegdheden & verantwoordelijkheden**



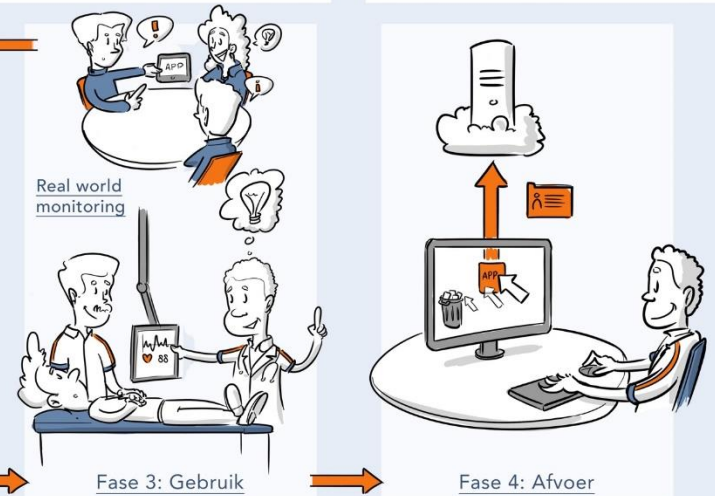
 **Stap 6. Controleer toepasselijke wet- & regelgeving**



 **Stap 7. Bereid voor op implementatie**



 **Stap 8. Evalueer de uitkomsten**

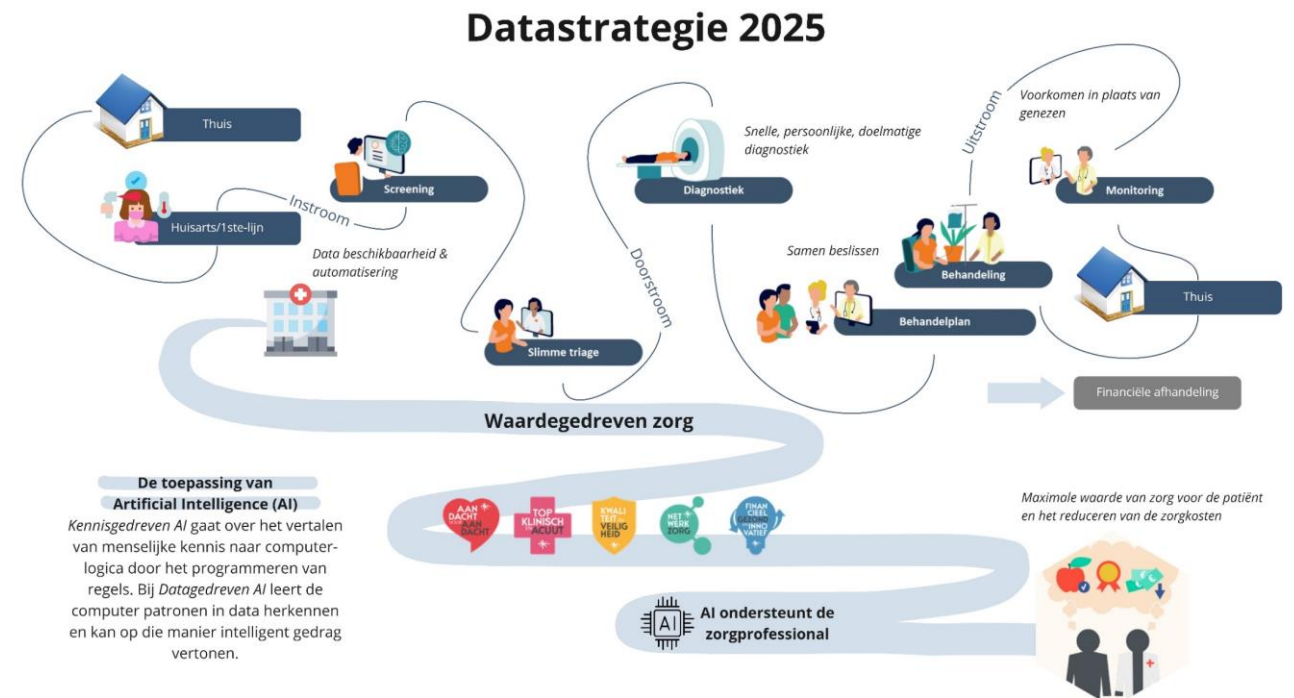


Stap 1: Bepaal de medische context



Stappenplan

In deze stap wordt inzicht verkregen in wat voor type AI-oplossing nodig is voor het **klinisch probleem**. Het doel en de **hoofdtak** moeten leiden tot een oplossing van het probleem. Daarbij moet ook rekening worden gehouden met de strategische koers van het ziekenhuis in relatie tot **waardegedreven zorg**. De medische context kan daarom het beste vanuit de patiëntreis bepaald worden. Het uitgangspunt is dat AI-software op verschillende momenten in de patiëntreis waarde kan toevoegen voor de patiënt.

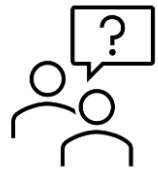


CHECKLIST STAP 1

- Er is een duidelijk beeld van het huidige zorgproces, de knelpunten en de ervaringen van eindgebruikers
- Het is duidelijk hoe het doel en de hoofdtak van AI-software leiden tot een oplossing voor het klinisch probleem

Klinisch probleem

Het **klinisch probleem** moet duidelijk zijn voordat de introductie van AI-software wordt overwogen. Het innovatieve karakter van AI-technologie geeft veel ruimte om vanuit een oplossing te denken, waardoor het klinisch probleem in de praktijk niet altijd duidelijk is. De vraag is dan of de beste oplossing het ontwikkelen of inkopen van AI-software is, of dat een simpelere oplossing ook volstaat. Gebaseerd op **best-practices** is een aantal overwegingen belangrijk:

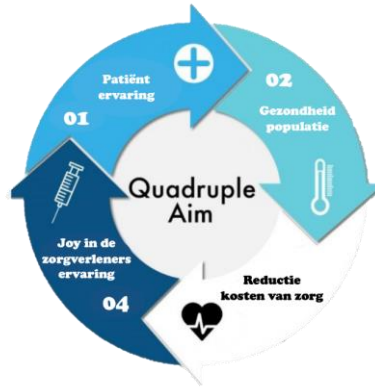


- Het klinisch probleem is geassocieerd met een **grote hoeveelheid data** waar een AI-model van kan leren
- Een analyse van die data door een mens is **niet efficiënt of haalbaar**
- De data zijn **beschikbaar** en mogen gebruikt worden
- De accuraatheid van een AI-model kan worden getest door de uitkomsten te vergelijken met een **gouden standaard** of menselijke beoordeling
- De uitkomsten van een AI-model leiden tot **acties** die het probleem oplossen in de klinische setting

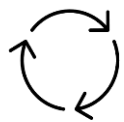
Hoofdtak

Doel

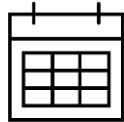
Het uiteindelijke doel waar AI-software aan bij moet dragen is het bereiken van waardegedreven zorg. Het streven naar betere gezondheidsuitkomsten tegen lagere kosten. Het **quadruple aim model** voegt daar nog een element aan toe, een verbetering in de ervaring van patiënten en zorgverleners. Zo kan de inzet van AI-software zorgverleners ontlasten, waardoor er meer tijd vrij komt voor contact met patiënten.



Hoofdtaken AI-software (aangepast uit literatuur)



Efficiëntere workflow



Snellere beoordeling



Eerdere detectie



Meer accurate diagnose/prognose



Gepersonaliseerde zorg

Efficiëntieverbetering

Verbeterde gezondheid

1 Efficiëntieverbetering

- Verbetering in de efficiëntie van de workflow of het zorgproces
- Vaak gaat het over een deel van de workflow, waar AI-software ondersteuning biedt. Dit kan een indirect effect hebben op patiëntuitkomsten
- Het is voor deze hoofdtak gebruikelijk om naar procesuitkomsten te kijken

2 Verbeterde gezondheid

- Bijvoorbeeld een verbetering in de diagnose of behandeling van een patiënt, of eerdere detectie van ziekte of herstel
- De informatie die de AI-software geeft heeft meestal een direct effect op een klinische actie, en indirect op patiëntuitkomsten
- Het is voor deze hoofdtak gebruikelijk om ook naar gezondheidsuitkomsten te kijken

Stap 2: Bepaal het innovatieniveau



Stappenplan

De volwassenheid van AI-software bepaalt de impact op de klinische praktijk. Hoe volwassener, hoe groter de klinische impact. De volgende indeling geeft een grove inschatting van het innovatieniveau ([Leidraad NIKP](#)):

- Experimenteel: AI-software die nieuw is voor Nederland en idealiter in studieverband worden toegepast
- Nieuw voor Nederland: AI-software waarvoor klinisch bewijs is verzameld (bijvoorbeeld in het buitenland), maar waarover nog geen vergoedingsafspraken bestaan
- Lokaal nieuw: AI-software die al in andere ziekenhuizen in Nederland wordt toegepast en waar vergoedingsafspraken over gemaakt worden

Met de **innovatiefunnel** en het **type product** wordt de wijze van introductie van AI-software bepaald. Deze stap bepaalt of er gestart wordt in de ontwikkelings-, pilot- of productiefase van een innovatie.

CHECKLIST STAP 2

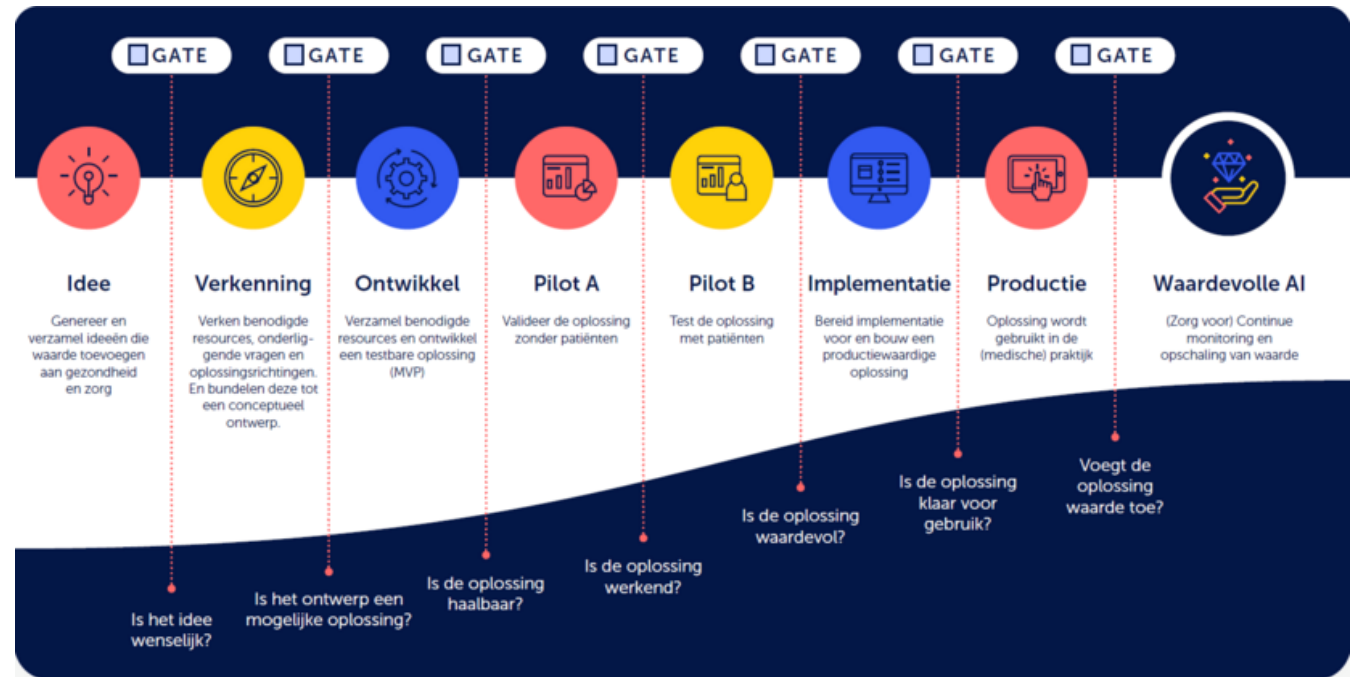
- Het is duidelijk in welke fase van innovatie de AI-software zich bevindt
- De overweging voor de ontwikkeling of inkoop van AI-software sluit aan bij de behoefte van de eindgebruikers
- Er zijn potentiële leveranciers geselecteerd vanuit een marktverkenning of de gewenste AI-software valt onder zelfbouw medische hulpmiddelen
- De marktverkenning is onderbouwd en toegevoegd aan het productdossier

Innovatiefunnel

Met de innovatiefunnel wordt het startpunt van de introductie van AI-software bepaald. Bij **zelfbouw** van AI-software wordt de innovatiefunnel geheel doorlopen. In voorbereiding op implementatie wordt dan gestart in **Fase 1A: Ontwikkeling**.

In geval van **inkoop** van AI-software wordt pilot A of B van de innovatiefunnel meestal nog een keer doorlopen als voorbereiding op het gebruik van AI-software in de reguliere patiëntenzorg. Daarom wordt gestart in **Fase 1B: Verwerving** en vormen pilot A, B onderdeel van **Fase 2: Ingebruikname**.

Innovatiefunnel Waardevolle AI



Type product

[Stappenplan](#)

De overweging voor de ontwikkeling of inkoop van AI-software heeft alles te maken met wat er nodig is om te voldoen aan de behoefte van de eindgebruiker:

- Is de **volwassenheid** van commercieel beschikbare AI-software zodanig dat deze aansluit bij de behoeftes van de eindgebruiker?
- Zijn de behoeftes **uniek** voor het eigen ziekenhuis of voldoet een generieke oplossing ook?
- Is het mogelijk om de AI-software te laten **integreren** met de bestaande IT-infrastructuur?

In tegenstelling tot zelfbouw van AI-software kan het AI-model bij inkoop meestal niet worden aangepast aan de lokale situatie. **Klik verder** om te zien waar je nog meer rekening mee moet houden.

[Inkoop](#)

[Zelfbouw](#)

[Markt
verkenning](#)

Inkoop

Bij inkoop van AI-software is de **productveiligheid** gewaarborgd door CE-certificering voor de kwaliteitseisen van de **MDR**. Met de **CE-klasse** wordt het risico voor de patiëntveiligheid aangegeven. De **klinische meerwaarde** is door CE-certificering nog niet aangetoond. De meerwaarde is ook deels context afhankelijk, zoals verschillen in de patiëntdoelgroep, systemen en werkwijze.

De *intended use* geeft aan hoe je het product mag gebruiken. Bij AI-software kunnen hierin voorwaarden worden gesteld voor de **interactie met de menselijke besluitvorming**. Bijvoorbeeld door de AI-software alleen in te zetten bij een groep patiënten waar de zorgverlener vooraf al een inschatting of keuze heeft gemaakt over het vervolgtraject.

TIPS

- Let op of de context van het gewenste gebruik past binnen de intended use van de CE-certificering
- Vraag hoe het kwaliteitssysteem is gecertificeerd, en hoe wijzigingen in het algoritme daarmee worden doorgevoerd
- Vraag een demo aan bij de leverancier om een indruk te krijgen van de functionaliteiten
- Spreek met de leverancier af hoe de klinische meerwaarde van AI-software wordt onderzocht samen met het ziekenhuis

Bij zelfbouw van AI-software wordt onderscheid gemaakt tussen in-house toepassing en toepassing buiten de eigen organisatie. Bij in-house toepassing hoeft niet aan de volledige **MDR** worden voldaan en wordt intern een ontwikkeldossier opgebouwd. Wanneer zelfbouw AI-software wordt gedeeld met andere organisaties, wordt het ziekenhuis verantwoordelijk gesteld als fabrikant. Een belangrijke voorwaarde voor zelfbouw is dat er geen andere AI-software met de vereiste prestatie op de markt verkrijgbaar is. Doe daarom altijd een **marktverkenning**.

Er kan een grijs gebied ontstaan in geval dat het eigen ziekenhuis een *algoritme* of *AI-model* van een ander ziekenhuis wilt overnemen. Dit hoeft dit nog niet gezien te worden als het verwerven van een *medisch hulpmiddel*. Zolang er geen data in en uit kunnen, en er geen toepassing omheen gebouwd is. Dit geldt ook voor het overnemen of aanpassen van open source algoritmes of AI-modellen.

TIPS

- Gebruik de CE-tool om te beoordelen of de gewenste AI-software een medisch hulpmiddel is
- Zorg dat er een duidelijke klinische vraagstelling is; algoritmes die worden gemaakt voor wetenschappelijke doeleinden sluiten niet altijd aan op de praktijk
- Naast de technische kennis en capaciteit voor ontwikkeling, moeten eindgebruikers ook beschikbaar zijn
- Overweeg opties voor een gezamenlijke ontwikkeling, met meerdere ziekenhuizen of een commerciële partij
- Er is informed consent nodig om patiëntdata te gebruiken voor het modelontwikkelingsproces

Marktverkenning

Een marktverkenning biedt mogelijke oplossingen voor het **klinisch probleem**. Hieruit volgt de keuze om AI-software leveranciers aan te schrijven of zelf AI-software te ontwikkelen. Naast de openbare bronnen die geraadpleegd kunnen worden, zijn referentieorganisaties ook een bron van informatie. Bijvoorbeeld het gebruik of de ontwikkeling van AI-software binnen samenwerkingsverbanden van de zorginstelling. Een overzicht van relevante bronnen is weergegeven in de tabel.

Overzicht bronnen voor marktonderzoek	
Bron	Link
Database	
European Database on Medical Devices (EUDAMED)*	https://ec.europa.eu/tools/eudamed/#/screen/home
FDA-gecertificeerde AI-software	https://medicalfuturist.com/fda-approved-ai-based-algorithms/
NVKI medische software*	Overview software MDR & IVDR (OSMI) (nvki.nl)
AIForRadiology - CE-gecertificeerde AI-software voor de radiologie	https://grand-challenge.org/aiforradiology/
Kenniscentra	
Expertisecentrum Zorgalgoritmen	https://www.saz-ziekenhuizen.nl/nieuws/start-expertisecentrum-zorgalgoritmen/
Erasmus MC Datahub	https://datahub.health/
LUMC CAIRELab - Clinical AI Implementation and Research Lab	https://campagne.lumc.nl/ai-in-de-zorg
NL AI Coalitie Use cases Gezondheid & Zorg	https://nlaic.com/use-cases/?fwp_filter_toepassingsgebieden=gezondheid-en-zorg
Thematic Technology Transfer - Artificial Intelligence (TTT.AI) Consortium	https://icai.ai/about-ttt-ai/
*in ontwikkeling	

Stap 3: Bepaal de toegevoegde waarde

[Stappenplan](#)

In deze stap wordt de (benodigde) **toegevoegde waarde van AI-software** aan de hand van beschikbare wetenschappelijke studies bepaald. Met het **effectiviteitsmodel** wordt het niveau van het bewijs voor de effectiviteit van AI-software bepaald. Hiermee wordt afgewogen of de effectiviteit van AI-software voldoende is aangetoond, of dat er meer bewijs nodig is. Deze afweging vindt plaats voor introductie van AI-software en wordt herhaald in de verschillende implementatiefases na evaluatie van de uitkomsten. Naast bewijs voor de effectiviteit, wordt de toegevoegde waarde ook ondersteund door bewijs over:

- De betrokkenheid (bij de ontwikkeling) en tevredenheid van gebruikers van de AI-software
- Succesvolle implementatie in de Nederlandse gezondheidszorg
- De bijdrage aan het oplossen van een maatschappelijk probleem

CHECKLIST STAP 3

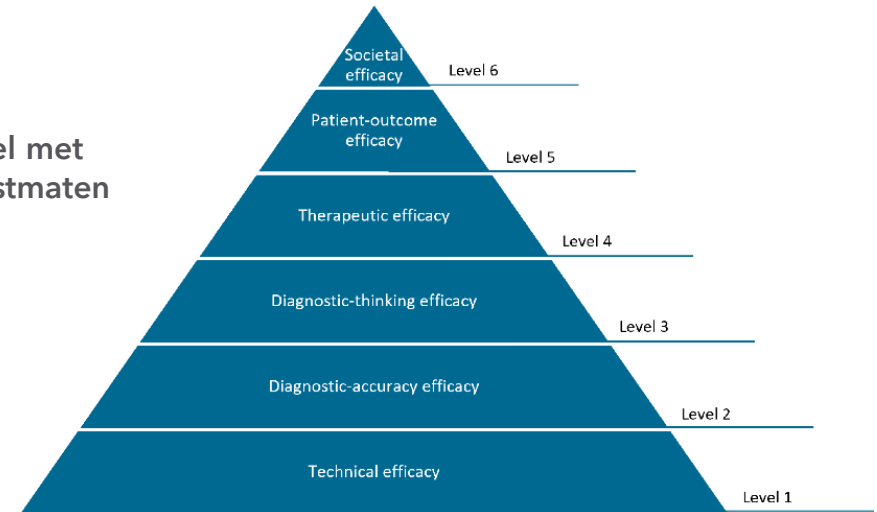
- Het (benodigde) bewijs over de toegevoegde van AI-software is beoordeeld
- De uitkomstmaten voor de effectiviteit zijn klinisch relevant

Effectiviteitsmodel

De meeste AI-software heeft op dit moment level 2/3 bewijs. Dit betekent dat bewijs over **daadwerkelijk gebruik in een klinische setting** in vele gevallen nog verzameld moet worden. De hoogste vorm van bewijs gaat over de **kosteneffectiviteit**; het effect van AI-software op de kosten en kwaliteit van leven van patiënten. Steeds vaker wordt met een **early-Health Technology Assessment** het beschikbare klinisch bewijs modelmatig onderzocht, om in een vroeg stadium een inschatting te maken of AI-software potentieel kosteneffectief is.

Met het effectiviteitsmodel wordt ook inzicht gecreëerd in relevante uitkomstmaten en of het mogelijk is om de benodigde data lokaal vast te leggen indien de zorginstelling zelf bewijs gaat verzamelen. Voor een vergelijkende studie opzet wordt verwezen naar **H5 van de Leidraad kwaliteit AI in de zorg**.

Hierarchisch effectiviteitsmodel met relevante uitkomstmaten (uit literatuur)



Level	Explanation	Typical measures
Level 1t	Technical efficacy Article demonstrates the technical feasibility of the software	Reproducibility, inter-software agreement, error rate
Level 1c	Potential clinical efficacy Article demonstrates the feasibility of the software to be clinically applied	Correlation to alternative methods, potential predictive value, biomarker studies
Level 2	Diagnostic accuracy efficacy Article demonstrates the stand-alone performance of the software	Standalone sensitivity, specificity, area under the ROC curve, or Dice score
Level 3	Diagnostic thinking efficacy Article demonstrates the added value to the diagnosis	Radiologist performance with/without AI, change in radiological judgement
Level 4	Therapeutic efficacy Article demonstrates the impact of the software on the patient management decisions	Effect on treatment or follow-up examinations
Level 5	Patient outcome efficacy Article demonstrates the impact of the software on patient outcomes	Effect on quality of life, morbidity, or survival
Level 6	Societal efficacy Article demonstrates the impact of the software on society by performing an economic analysis	Effect on costs and quality-adjusted life years, incremental costs per quality-adjusted life year

Level 1t level 1, technical; Level 1c level 1, clinical

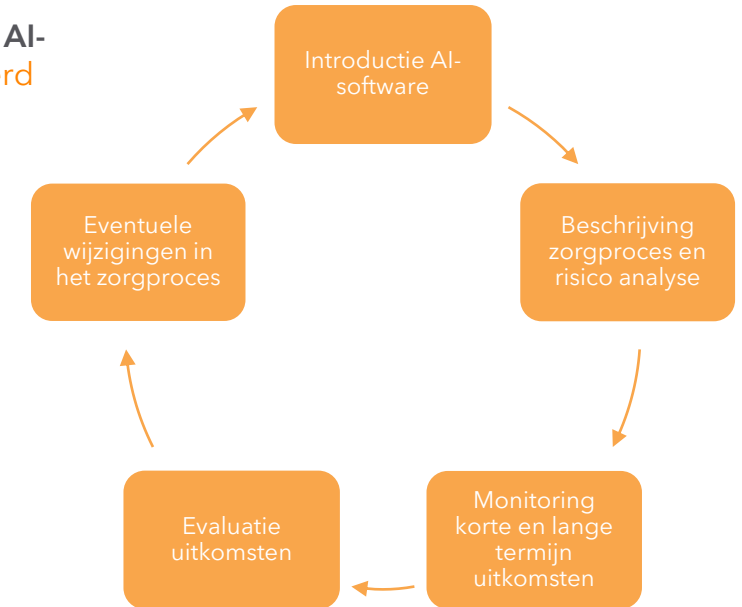
Stap 4: Bepaal de risico's en impact



In deze stap worden de veiligheid en impact van AI-software gedurende de levenscyclus in kaart gebracht. De introductie van AI-software heeft impact op het huidige zorgproces. Hierbij moeten de voor- en nadelen voor de patiënt worden afgewogen. In een **risico analyse**, worden de risico's voor de patiëntveiligheid beoordeeld. Maar de impact van AI-software wordt niet alleen bepaald door de risico's. De **verwachte effecten** en de **mens-machine interactie** worden ook in kaart gebracht, zodat AI-software zorgvuldig toegepast kan worden.

Stap 4 van het stappenplan gebeurt niet eenmalig, maar wordt herhaald in de verschillende implementatiefases na evaluatie van de uitkomsten. Afhankelijk van de uitkomsten en eventuele wijzigingen in het gebruik, wordt de impact op het zorgproces opnieuw beoordeeld.

Cyclus introductie AI-software (gebaseerd op Leidraad NIKP)



CHECKLIST STAP 4

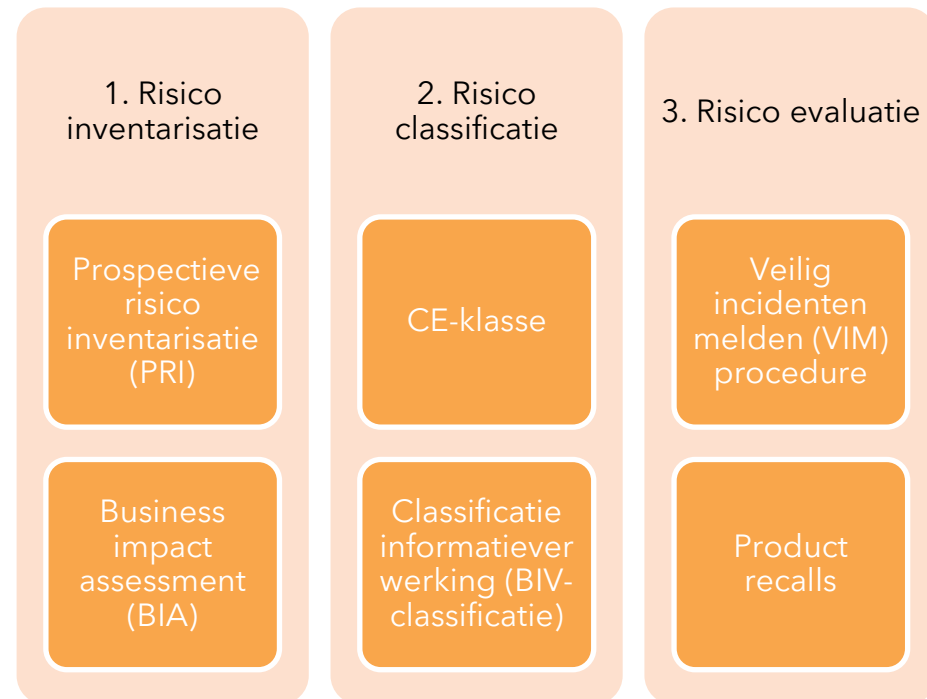
- De veiligheid van AI-software is beoordeeld met een risico analyse en toegevoegd aan het productdossier. Er is ook rekening gehouden met cybersecurity risico's.
- De impact van AI-software is ingeschat door de verwachte effecten en mens-machine interactie
- Data over de uitkomsten van AI-software kunnen lokaal worden vastgelegd

Risico analyse

De **risico analyse** bestaat uit de inventarisatie, de classificatie, en de evaluatie van risico's. Voordat de risico's geïnterpreteerd worden, is een algemene risico inschatting nodig. Onder meer op basis van het innovatieniveau, betrokken specialismen, patiënten doelgroep, de complexiteit in gebruik en veranderingen in het zorgproces. De risico's kunnen functioneel of technisch van aard zijn. Vanwege de toename in **cybersecurity risico's**, wordt hier extra aandacht aan besteed.

Verdieping cybersecurity risico's

Voor de risico analyse wordt een aantal instrumenten ingezet, voorafgaande en tijdens gebruik van AI-software:

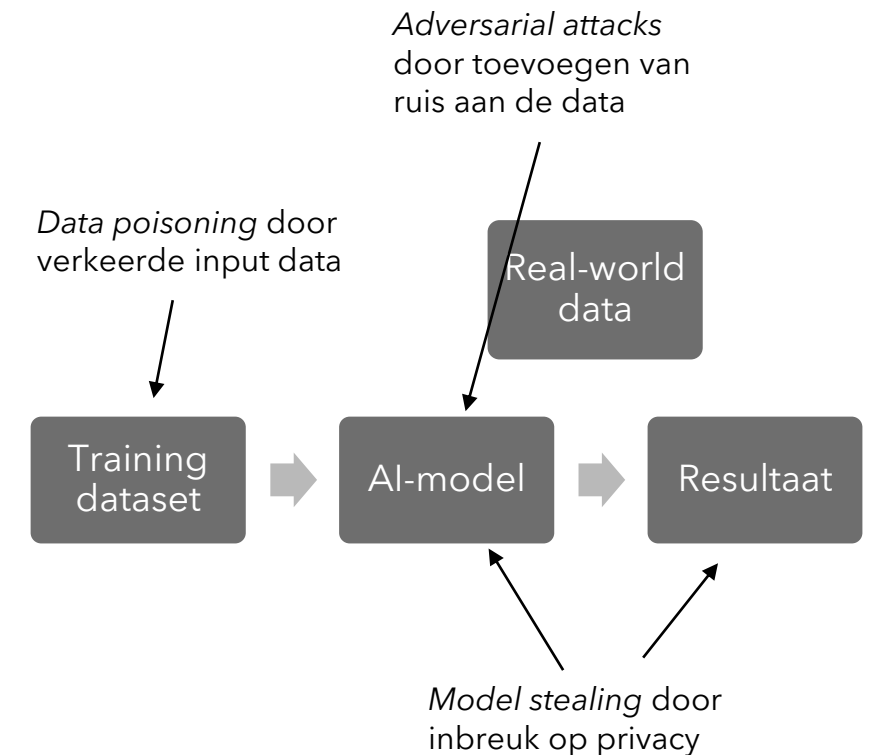


Cybersecurity risico's

Alle medische hulpmiddelen die verbonden zijn met het netwerk zijn kwetsbaar voor **cybersecurity aanvallen**. Dit kan gevolgen hebben voor de patiëntenzorg. Zo kan uitval van kritische systemen (EPD, PACS, labsysteem en mogelijk ook AI-software) uiteindelijk leiden tot sluiting van het ziekenhuis of een gevaar voor de patiëntveiligheid. Met een cybersecurity aanval wordt onbedoelde toegang verkregen tot patiëntdata. **Beschermende maatregelen** vormen de basis voor deze risico's, maar voor medische hulpmiddelen en AI-software specifiek, zijn er uitdagingen:

- **Software patches en updates** kunnen niet altijd op afstand doorgevoerd worden, waardoor dit op locatie moet plaatsvinden
- Software patches en updates vallen mogelijk buiten de omvang van de **CE-certificering**
- De documentatie over de **beveiligingseigenschappen van AI-software** is mogelijk onvoldoende om vast te stellen of een onderdeel kwetsbaar is
- **Extra cybersecurity risico's** voor AI-modellen in de training- en testfase zoals weergegeven in de figuur. Een overzicht van alle mogelijke cybersecurity risico's is te raadplegen in de [EU AI Threat Landscape](#)

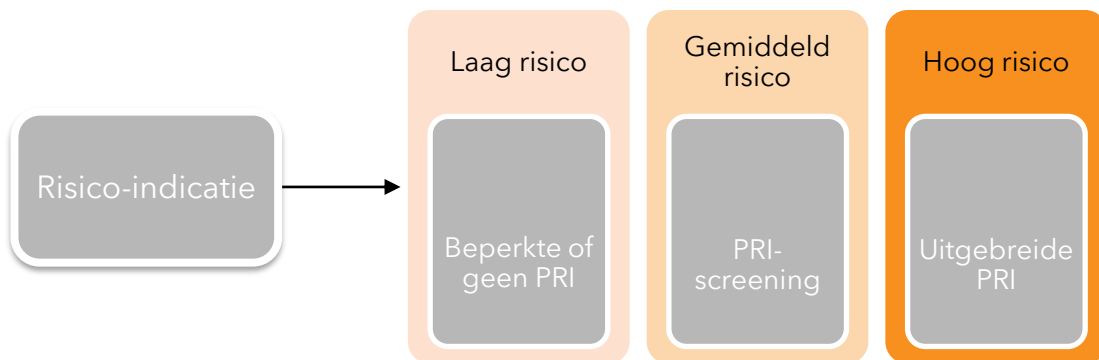
Extra cybersecurity risico's voor AI-software



Prospectieve risico inventarisatie (PRI)

Om AI-software veilig te introduceren worden de potentiële risico's van het gebruik in de dagelijkse klinische praktijk in kaart gebracht met een passende PRI. Hiertoe behoren onbedoelde effecten en verkeerd gebruik in het gewenste zorgproces. Met een inschatting van de impact en bijbehorende maatregelen. Zie de PRI als een dynamisch instrument, naarmate het gebruik van AI-software kunnen risico's en maatregelen worden bijgesteld.

Een uitgebreide PRI wordt volgens de **Healthcare Failure Mode and Effect Analysis (HFMEA)** methodiek uitgevoerd, wanneer er bijvoorbeeld sprake is van een grote verandering in het zorgproces en/of wanneer de AI-software een hoge **CE-klasse** heeft. In andere gevallen wordt volstaan met een lichtere variant van een PRI.



TIPS

- Bij commerciële AI-software heeft de fabrikant een algemene risico inventarisatie uitgevoerd, de belangrijkste risico's worden vermeld in de handleiding
- Technische risico's komen meestal voort uit de IT-omgeving van de AI-software, zoals het wegvallen van de verbinding met systeemcomponenten
- Denk ook aan mentale of emotionele risico's door de mens-machine interactie, zoals het stimuleren van bepaalde gedachten of ervaren stress bij patiënten door fout-positieve bevindingen
- Zolang de mens tussen de AI-software en de beslissing zit, kan het risico van "overtrust" op de AI-software een rol spelen

PRAKTIJKVOORBEELD

- PRI Siemens AI

CE-klasse

Volgens de **MDR** wordt AI-software in 4 risicoklassen ingedeeld: klasse I (laag risico), klasse IIa en IIb (gemiddeld risico) en klasse III (hoog risico). Hoe hoger het risico voor de patiënt als AI-software faalt, hoe hoger de klasse en hoe meer eisen er gesteld worden. De meeste AI-software wordt nu als klasse IIa of hoger beschouwd, wanneer informatie verstrekt wordt voor het nemen van beslissingen voor **diagnostische of therapeutische doeleinden**. De betekenis van de informatie en het effect op de patiëntveiligheid bepalen dan de risicoklasse. Er zit namelijk verschil in wanneer de informatie alleen getoond wordt ter ondersteuning van de gebruiker, of wanneer er bijvoorbeeld direct een diagnose wordt aanbevolen.

Bij **zelfbouw** van AI-software is een conformiteitsbeoordeling door een notified body niet nodig, maar moet het product wel voldoen aan de algemene vereisten voor veiligheid en prestatie in Annex I van de MDR.

Risico classificering voor stand-alone software (MDR Guide)

		Significance of Information provided by the MDSW to a healthcare situation related to diagnosis/therapy		
		High Treat or diagnose ~ <i>IMDRF 5.1.1</i>	Medium Drives clinical management ~ <i>IMDRF 5.1.2</i>	Low Informs clinical management (<i>everything else</i>)
State of Healthcare situation or patient condition	Critical situation or patient condition ~ <i>IMDRF 5.2.1</i>	Class III <i>Category IV.i</i>	Class IIb <i>Category III.i</i>	Class IIa <i>Category II.i</i>
	Serious situation or patient condition ~ <i>IMDRF 5.2.2</i>	Class IIb <i>Category III.ii</i>	Class IIa <i>Category II.ii</i>	Class IIa <i>Category I.ii</i>
	Non-serious situation or patient condition (<i>everything else</i>)	Class IIa <i>Category II.iii</i>	Class IIa <i>Category I.iii</i>	Class IIa <i>Category I.i</i>

BIA en BIV-classificatie

Vanuit de **AVG** moet er bij het ontwerp van AI-software al rekening worden gehouden met gegevensbescherming (privacy by design), het is dus van belang om de eisen rondom privacy en informatiebeveiliging in samenhang te zien.

- De **Business Impact Assessment (BIA)** beoordeelt de risico's en impact van een verwerking van persoonsgegevens. Daarbij wordt ook gekeken naar aanvullende kenmerken van de gegevensverwerking, zoals de systeemcomponenten die gebruikt worden (vb. cloud, mobiele hardware, webportaal).
- De classificatie van de **Beschikbaarheid, Integriteit en Vertrouwelijkheid (BIV-classificatie)** van de gegevensverwerking wordt tegelijk uitgevoerd, om de maatregelen voor informatiebeveiliging te bepalen. De maatregelen komen voort uit wet- en regelgeving, normenkaders en praktijkgidsen. Dit vormt input voor de beheerafspraken die in de verwerkersovereenkomst met leveranciers worden gemaakt.

TIPS

- Uit de BIV-classificatie kan blijken dat er een aanvullend toetsingsmiddel nodig is om de privacy van patiënten te waarborgen. Een Data Protection Impact Assessment (DPIA) wordt alleen uitgevoerd onder bepaalde voorwaarden uit de AVG.

PRAKTIJKVOORBEELD

- ❑ Classificatie informatieverwerking Teampay AI Companion

Veilig incidenten melden (VIM) en product recalls

1 Evaluatie van interne risico's

- Ongewenste gebeurtenissen rondom de zorgverlening aan de patiënt worden gemeld, onderzocht en in het vervolg voorkomen door VIM
- Het gaat om het melden van calamiteiten, complicaties, incidenten of onveilige situaties door het gebruik van AI-software
- Na een melding wordt onderzocht of er sprake is van een verband tussen de schade voor de patiënt en de kwaliteit van de geleverde zorg

2 Evaluatie van externe risico's

- Een recall is een veiligheidsmelding of terugroepactie van de leverancier, meestal met urgentie in verband met mogelijk patiëntrisico
- Voor fabrikanten is een recall procedure onderdeel van het post market surveillance systeem
- Recalls kunnen bijvoorbeeld gaan over kwetsbaarheden in cybersecurity

TIPS

- De traceerbaarheid van AI-software naar de individuele patiënt valt onder de verantwoordelijkheid van de zorginstelling. Denk dus ook aan vastlegging van het wel of niet meewegen van een voorspelling in een beslissing (en waarom) in het patiëntendossier
- Cybersecurity kwetsbaarheden in medische hulpmiddelen worden niet altijd als eerste door de leverancier opgemerkt, het [Nationaal Cyber Security Centrum \(ncsc.nl\)](https://www.ncsc.nl) meldt kwetsbaarheden en houdt lijsten bij met kwetsbare applicaties

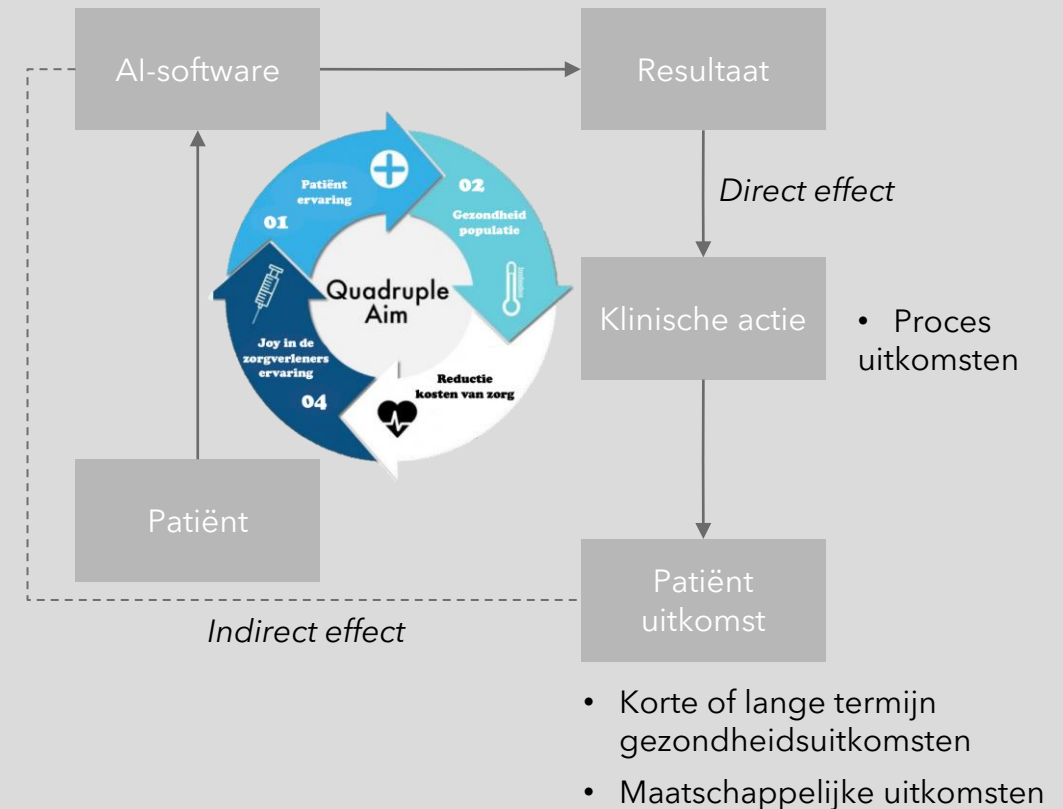
Verwachte effecten

In aanvulling op de *intended use* worden de verwachte effecten van de AI-software op het zorgproces in kaart gebracht. Denk hierbij ook aan (negatieve) effecten in specifieke patiëntengroepen, zoals kritische of acute patiënten. En effecten in de zorgketen. De informatie die de AI-software genereert, wordt gebruikt door zorgverleners en/of patiënten in het maken van keuzes en heeft op verschillende niveaus een effect:

- **directe effecten** zoals verbeterde werkprocessen, een snellere of betere diagnose/prognose of het eerder toepassen van een interventie.
- **indirecte effecten** zoals verbeterde individuele gezondheidsuitkomsten of preventie van ziekte.

Om de (patiënt)uitkomsten systematisch te evalueren moet **lokale dataregistratie** plaatsvinden. Het is dus belangrijk om vooraf te bepalen welke data nodig zijn en op welke manier de registratie van data plaatsvindt, bij voorkeur gekoppeld aan het patiëntendossier.

De relatie tussen AI-software en patiënt uitkomsten:



Mens-machine interactie

De impact van AI-software op het zorgproces is ook afhankelijk van de **interactie met het klinische besluitvormingsproces**. Hoe autonomer de AI-software acteert, hoe groter de uitdaging op het gebied van veilige toepassing, competenties, vertrouwelijkheid en aansprakelijkheid. En daarmee een groter risico voor de patiëntveiligheid. Stel daarom vast op welke manier de AI-software gebruikt wordt. Denk aan het verschil tussen:

- AI-software geeft alleen **informatieve tekst** als ondersteuning
- Er worden **risico-scores** gegenereerd die worden geïnterpreteerd door een zorgverlener
- AI-software geeft een **aanbeveling** die door de zorgverlener wordt beoordeeld
- Volledige **autonomie** bij het nemen van vervolgbeslissingen

De mens-machine interactie heeft dus ook invloed op de manier waarmee de informatie van AI-software gecommuniceerd wordt. Zoals de communicatie van resultaten naar **betrokken specialisten of patiënten**.

TIPS

- De handelingen die een gebruiker moet uitvoeren met de AI-software moet aansluiten op de huidige workflow
- Bedenk of er grote veranderingen in het huidige zorgproces moeten plaatsvinden en beoordeel deze bij voorkeur in de vorm van een pilot
- Betrek andere specialisten en patiënten tijdig bij de veranderingen in het zorgproces

Stap 5: Maak afspraken over taken, bevoegdheden & verantwoordelijkheden



Stappenplan

Om AI-software veilig toe te passen gedurende de levenscyclus is het noodzakelijk om de **Taken (T) Bevoegdheden (B) en Verantwoordelijkheden (V)** af te spreken binnen de organisatie en met externe partijen. Dit wordt niet alleen gedaan op het niveau van de activiteiten in de levenscyclus, zoals wie uit het **AI-team** een rol heeft in het opstellen van een Pakket van Eisen of het borgen van **scholing**. Maar ook voor het beheer van de componenten van de AI-software, zoals afspraken over mobiele devices, een extra pc, het netwerk, en de cloud-omgeving (zie **lokaal beleid**). Externe afspraken worden vastgelegd in contracten over de diensten die een leverancier levert en de verwerking van persoonsgegevens. Stap 5 van het stappenplan gebeurt niet eenmalig, maar wordt herhaald in de verschillende implementatiefases na evaluatie van de uitkomsten. Afhankelijk van de uitkomsten en eventuele wijzigingen in het gebruik, worden de **beheerafspraken** opnieuw beoordeeld.

CHECKLIST STAP 5

- Er is een AI-team samengesteld met een aangewezen projectleider
- Interne en externe beheerafspraken zijn vastgelegd in het productdossier
- Er is een scholingsplan opgesteld en vastgelegd in het productdossier

AI-team

Het is belangrijk om een team samen te stellen die bestaat uit een **multidisciplinaire samenwerking** tussen:

- Eindgebruikers:



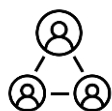
- Patiënten
- Medisch specialisten
- Verpleegkundigen

- Kennisexperts:



- Klinisch expert
- AI-expert (data scientist, klinisch fysicus, klinisch informaticus, medisch statisticus, epidemioloog, of klinisch technoloog);
- IT-specialist (functioneel beheerder, applicatiebeheerder, technisch beheerder, softwareontwikkelaar)
- Ethisch / juridisch expert

- Besluitvormers:



- Medisch managers
- Het implementatieteam wordt geleid door een projectleider.

Daarnaast zijn er voor de ontwikkeling en implementatie van AI-software specifieke rollen aan te wijzen:

Rol	Omschrijving
Data-steward	Verantwoordelijk voor het borgen en bewaken van de datakwaliteit
Data-architect	Verantwoordelijk voor het beheer van de data-architectuur
Data-engineer	Verantwoordelijk voor het opnemen van data uit de bronsystemen om via standaard processen de data naar de target database te verplaatsen. Kan ook verantwoordelijk zijn voor datatransformatie op ruwe data.
Data-scientist	Verantwoordelijk voor het uitvoeren van taken op data in de target database, als voorbereiding op het modelontwikkelingsproces. De taken omvatten het opschonen, selecteren van gewenste data en standaardisatie van data. Verantwoordelijk voor het bouwen van een AI-model gebaseerd op de voorbereide dataset, zodat het model klaar is voor implementatie
Product owner	Verantwoordelijk voor AI-modellen die gereed zijn voor productie

Beheerafspraken

In het ziekenhuis wordt met het **RACI model** de verschillende taken, bevoegdheden en verantwoordelijkheden beschreven. Hierin wordt een onderscheid gemaakt in de niveaus: verantwoordelijke (R), eindverantwoordelijke (A), geraadpleegd (C), en geïnformeerd (I). Externe afspraken over de dienstverlening worden vastgelegd in een **Service Level Agreement (SLA)** of een **SSO (Standaard Service Overeenkomst)** wanneer het specifiek gaat over het onderhoud van systemen (software updates/upgrades). Daarnaast is het sluiten van een verwerkersovereenkomst wettelijk verplicht als een zorgaanbieder een derde partij inschakelt die persoonsgegevens verwerkt namens de zorgaanbieder.

Een aanpassing van het RACI model op basis van nieuwe rollen en activiteiten wordt na inbedding van het stappenplan verder uitgewerkt.

Scholing

Eindgebruikers van AI-software zijn **verplicht** om scholing te volgen vooraf gebruik van een specifiek softwarepakket in de klinische praktijk. In geval van commerciële AI-software verzorgt de leverancier meestal scholing. Samen met de **Maasstad Academie** dient het scholingsplan (van de leverancier) te worden beoordeeld. Indien dit onvoldoende wordt geacht, wordt samen met de interne organisatie gekeken naar aanvullende scholing. Via de Maasstad Academie wordt ook algemene scholing over AI aangeboden. Het volgende aanbod is beschikbaar:

Overzicht algemene scholing AI

Bron	Link
De Nationale AI-zorg cursus	https://zorg.ai-cursus.nl/home
Online cursus Leidraad kwaliteit AI in de zorg	https://www.leidraad-ai.nl/
Digitale vaardigheidstraining - Artificial Intelligence	https://www.goodhabitz.com/nl-nl/online-trainingen/categorieen/digitale-vaardigheidstrainingen/artificial-intelligence/

TIPS

- Zorg dat de gebruiker op de hoogte is van de sterktes en beperkingen van de AI-software
- Scholing wordt vaak als mitigerende maatregel genoemd bij een risico analyse

Stap 6: Controleer toepasselijke wet- en regelgeving



Stappenplan

Er is veel **wet- en regelgeving** van toepassing voor AI-software. De belangrijkste is de wetgeving van de **MDR**. Deze beschrijft waar een medisch hulpmiddel minimaal aan moet voldoen. De privacywetgeving **AVG** is van toepassing omdat AI-software van de ontwikkelings- tot afvoerfase persoonsgegevens verwerkt. Voor de zorgverleners en patiënten is de wetgeving rondom de bescherming van de patiënt het belangrijkste. In termen van **aansprakelijkheid** heeft de zorginstelling de verantwoordelijkheid om kwalitatief goede zorg te verlenen. Daarnaast wordt gewerkt volgens (inter)nationale normen om de **kwaliteit en veiligheid** van een methode of product te garanderen. Bijvoorbeeld voor de informatiebeveiliging of het ontwikkelen van software. De vertaling van wet- en regelgeving naar concrete handvatten gebeurt in **praktijkguiden en veldnormen**. Voor AI-software wordt op internationaal en nationaal gewerkt aan nieuwe wet- en regelgeving. Het is dus belangrijk dat stap 6 van het stappenplan wordt herhaald indien zich nieuwe ontwikkelingen voordoen.

CHECKLIST STAP 6

- Het AI-team heeft kennis genomen van de relevante wet- en regelgeving
- Nieuwe ontwikkelingen in de wet- en regelgeving voor AI-software worden gevolgd

Medical Device Regulation (MDR)

1 Introductie Medische hulpmiddelen

De belangrijkste wetgeving voor productveiligheid van AI-software is de Medical Device Regulation (MDR). Voor een uitgebreide samenvatting wordt verwezen naar de literatuur. In de MDR valt meer AI-software onder de nieuwe definitie van een medisch hulpmiddel. Ook valt AI-software in een hogere CE-klasse dan voorheen, doordat de informatie gebruikt wordt bij het nemen van beslissingen voor diagnostische of therapeutische doeleinden. Hierdoor gelden meer eisen voor de klinische evaluatie en post market surveillance.

2 Klinische evaluatie De prestatie van AI-software

De klinische evaluatie van de prestatie van AI-software wordt nog belangrijker. Applicaties met een vergelijkbaar beoogd doel laten mogelijk verschillen zien in prestatie, omdat de trainingsdata niet hetzelfde is. Daarnaast is de certificering van adaptieve AI-software (continu lerend systeem) op dit moment moeilijk, omdat veranderingen in de prestatie van de software beoordeeld moet worden door een notified body.

3 In house ontwikkeling Zelfbouw van AI-software

Eigen ontwikkeling van AI-software is alleen toegestaan als er geen alternatief op de markt verkrijgbaar is. Er gelden minder eisen aan zorginstellingen die AI-software ontwikkelen voor eigen gebruik. Zoals de algemene veiligheids- en prestatie-eisen van Bijlage I en een passend kwaliteitssysteem. De invulling van een passend kwaliteitssysteem is organisatie- en toepassingsafhankelijk.

4 Wet op de Artificiële Intelligentie Een Europese verordening voor veilig gebruik van AI

De MDR richt zich niet specifiek op AI. De Wet op AI maakt onderscheid in risico's van AI-systemen ; medische hulpmiddelen vallen in de categorie hoog risico. Daarmee worden o.a. eisen gesteld aan risicobeheer, kwaliteit van data, technische documentatie, logging, transparantie, menselijk toezicht, nauwkeurigheid en robuustheid, en registratieplicht. Wanneer zorginstellingen zelf AI-software ontwikkelen, moet er volgens de nieuwe wetgeving dus aan meer eisen worden voldaan

De belangrijkste norm voor informatiebeveiliging is de NEN7510, en internationaal de ISO27001. Een overzicht van andere relevante normen voor softwareontwikkeling is te raadplegen in de [Hand-out wettelijke en normatieve kaders rondom AI in de zorg](#)

Inspectie Gezondheidszorg en Jeugd (IGJ)

Toetsingskader Inzet van e-health door zorgaanbieders Randvoorwaarden digitale zorg

De IGJ toetst of de zorgaanbieder bij de inzet van e-health zorgt voor de nodige voorwaarden voor goede en veilige zorg. Dit is ook van toepassing op AI-software. De fabrikant moet zorgen voor een veilig product, en de zorgaanbieder voor veilig gebruik.

Aandachtspunten:

- Het borgen van risicoanalyses over patiëntveiligheid tijdens de invoering van AI-software
- Het actief betrekken van patiënten vraagt ook de nodige kennis over AI
- Het versterken van de informatiebeveiliging in de zorginstelling

Zorgverzekeraars Nederland (ZN)

Leidraad applicaties en algoritmes in de zorg Digitale zorg innovaties

Zorgverzekeraars gaan gezamenlijk digitale zorg innovaties beoordelen op de kans rijkheid voor grootschalige toepassing in de zorg. Idealiter laten ontwikkelaars van AI-software dit al in een vroeg stadium toetsen. Zodat de waarde van de geleverde digitale zorg in termen van toegankelijkheid, kwaliteit en kosten onderzocht kan worden samen met zorgaanbieders.

Aandachtspunten:

- Zorgverzekeraars kunnen transformatiegelden vrij maken om AI-software in de klinische praktijk te implementeren. Dit is geld beschikbaar voor de transformatie naar de juiste zorg op de juiste plek
- De belangrijkste criteria om digitale zorg vergoed te krijgen is dat de zorg veilig en bewezen effectief moet zijn

Ministerie van Volksgezondheid en Welzijn (VWS)

Leidraad kwaliteit AI in de zorg

Leidraad voor kwalitatieve diagnostische en prognostische toepassingen van AI in de zorg

De leidraad helpt zorgverleners in alle sectoren om de kwaliteit van aangeboden voorspellende AI algoritmes te beoordelen en helpt ontwikkelaars om kwalitatieve, betrouwbare AI te ontwerpen en realiseren. De leidraad draagt ook bij aan uniformiteit en het proces van markttoelating en vergoeding door verzekeraars. De leidraad geeft een overzicht van de belangrijkste eisen en aanbevelingen per fase, van ontwikkeling tot implementatie. In de podcast vertellen drie specialisten uit het veld (waaronder Sade Faneyte) over de toepassing van de leidraad.



Beroepsverenigingen

Leidraad NIKP

(Nieuwe Interventies in de Klinische Praktijk)

De leidraad is een stappenplan voor de zorgvuldige introductie van nieuwe interventies in de klinische praktijk. Zodat medische innovaties vóór en tijdens de introductiefase goed worden onderzocht op veiligheid, effectiviteit en de voorwaarden voor gepast gebruik in de praktijk. Voor AI-software ligt de nadruk op het onderdeel evaluatie na introductie, iets wat bij medische innovaties nog te weinig gebeurt (**implementatie leidraad**).

Convenant Medische Technologie

Levenscyclus van medische hulpmiddelen

Het convenant geeft invulling aan de wettelijke verplichting vanuit de Wkkgz voor veilige toepassing van medische technologie door zorgaanbieders. Het convenant stelt eisen aan de risicobeheersing en de veilige toepassing van medische technologie binnen de ziekenhuiszorg. Door het ziekenhuis vertaalt naar de "Levensloop medische hulpmiddelen" met de fase invoer, gebruik en afvoer van het medische hulpmiddel. Voor software als medisch hulpmiddel geeft de **Praktijkgids Medische Informatietechnologie** praktische handvaten. De nadruk voor AI-software ligt op de invoer fase, en betreft de ontwikkelings- of verwervingsfase en ingebruikname fase zoals beschreven in de Praktijkgids.

Algemene Verordening Gegevensbescherming (AVG)

1 **Introductie** Verwerking van patiëntgegevens

De AVG stelt regels voor de verwerking van persoonsgegevens. Deze wetgeving is van toepassing in de ontwikkelingsfase van AI-software, omdat persoonsgegevens voor trainings- en testdata wordt verwerkt. En tijdens het gebruik van AI-software, omdat de verwerking van persoonsgegevens ook leidt tot nieuwe data. De AVG is technologie onafhankelijk, maar zorgt er ook voor dat de potentie van AI niet volledig benut kan worden. Er geldt in principe een verbod op uitsluitend geautomatiseerde besluitvorming zonder menselijke tussenkomst.

2 **Verantwoording door de verwerkingsverantwoordelijke** Grondslag en informatie over het AI-algoritme

De zorginstelling moet onderbouwen waarom patiëntgegevens gebruikt worden in een AI-algoritme, wat het doel van het gebruik van een algoritme is, en waarom het nodig is te werken met het gekozen algoritme. Het doel van de verwerking van patiëntgegevens kan onvoorspelbaar zijn, omdat een AI-model nieuwe correlaties kan voorspellen en daardoor het toekomstig gebruik vooraf niet altijd duidelijk is.

3 **Transparantie** Uitlegbaarheid van het AI-algoritme

Er is recht op uitleg over de gegevensverwerking, bijvoorbeeld in het geval dat een zorgverlener aan de patiënt moet kunnen uitleggen hoe hij/zij tot een beslissing is gekomen. Maar optimale transparantie van het AI-algoritme is vanwege intellectuele property rechten niet altijd mogelijk.

4 **Privacy-risico's** Maatregelen om privacy risico's te verkleinen

Belangrijke maatregelen om privacy risico's te verkleinen zijn dataminimalisatie en pseudonimiseren. Maar er is een bepaalde mate van pseudonimisatie nodig om de data bruikbaar te houden voor AI-software. Daarnaast wordt bij continu lerende AI-systemen het onderscheid tussen wetenschap en toepassing moeilijker, waardoor de grondslag van de verwerking kan veranderen. Afspraken over de verwerking van persoonsgegevens namens de zorgaanbieder door derde partijen moeten vastgelegd worden in een verwerkersovereenkomst.

Aansprakelijkheid

Wetgeving rondom aansprakelijkheid

1 **Introductie** Bescherming van de patiënt

De eindverantwoordelijke in het gebruik van AI-software is altijd de behandelend arts, wanneer er sprake is van een behandelrelatie met de patiënt. Dit vraagt dus om een bepaalde verantwoordelijkheid van de zorgverlener in het toezien op het juiste gebruik van AI-software. Is er sprake van schade aan de patiënt door fouten, dan is de zorgverlener in de eerste instantie het aanspreekpunt. Maar er is een gedeelde verantwoordelijkheid met de leverancier/fabrikant in de productketen.

2 **Via de behandelrelatie (Wgbo)** Aansprakelijkheid van de zorgverlener

De Wet op de geneeskundige behandelovereenkomst (Wgbo) biedt bescherming aan de patiënt:

- Nadruk op informed consent en beroepsgeheim
- Informatieverplichting van de zorgverlener aan de patiënt over de toegepaste AI-software

3 **Via de zorginstelling (Wkkgz)** Aansprakelijkheid van de zorginstelling

De Wet kwaliteit, klachten en geschillen zorg (Wkkgz) stelt eisen aan de kwaliteit van zorg, en de verplichting om goede zorg te verlenen:

- Goede zorg is veilig en doeltreffend
- Als leveranciers andere diensten leveren dan uitsluitend het afleveren van AI-software (zoals advies), kan ook sprake zijn van het verlenen van zorg

4 **Via de fabrikant (Wet productaansprakelijkheid)** Aansprakelijkheid van de fabrikant

De Wet productaansprakelijkheid biedt bescherming tegen onveilige en gebrekkige producten. Afspraken met leveranciers over schade vergoedingen worden opgenomen in contracten. Maar de huidige wetgeving is verouderd, software wordt in juridische termen niet altijd als product gezien. Recent is een standpunt over aansprakelijkheid en de verplichting tot verzekeren voor AI-systemen uitgebracht door het Europees Parlement. De toekomstige wetgeving zal patiënten meer bescherming moeten bieden.

Stap 7: Bereid voor op implementatie

[Stappenplan](#)

In deze stap wordt de implementatie van AI-software voorbereid op klinisch gebruik, in **Fase 1A: ontwikkelingsfase** of **Fase 1B: verwervingsfase** en **Fase 2: ingebruikname**. Per fase worden de onderdelen beschreven waar je aan moet denken, met de documenten die vastgelegd worden in het ontwikkel- of productdossier. Afhankelijk van de keuze in AI-software of uitkomsten van de fases, wordt stap 3-6 herhaald.

CHECKLIST STAP 7

- Binnen het AI-team is inzicht verkregen in de eigenschappen van de AI-software
- Het PvE, ontwerp, implementatieplan en validatieprotocol zijn toegevoegd aan het ontwikkel- of productdossier

Fase 1A: Ontwikkeling

In geval dat de AI-toepassing nog ontwikkeld moet worden, wordt aanbevolen dit volgens vaste stappen te doen. Het **modelontwikkelingsproces** vormt de basis. De documenten die opgesteld worden tijdens de ontwikkelingsfase worden vastgelegd in het **ontwikkeldossier**. Het gaat om dezelfde documenten die ook voor commerciële AI-software gelden. Zoals het **PvE, ontwerp, risico-analyse**, en het **implementatieplan**. Aanvullende producten, zoals een openbare verklaring volgen uit de **Checklist Zelfbouw medisch hulpmiddel**. Het ontwikkelproces uit de **Leidraad kwaliteit AI in de zorg** is hieronder weergegeven. In fase 1-3 wordt het AI-algoritme gebouwd en getest. In fase 4 wordt er van het algoritme een hulpmiddel gemaakt; de AI-software.



Fase 1:
Verzameling en
beheer van de data



Fase 2:
Ontwikkeling van
het voorspellend
AI-algoritme



Fase 3:
Validatie van het
voorspellende
algoritme

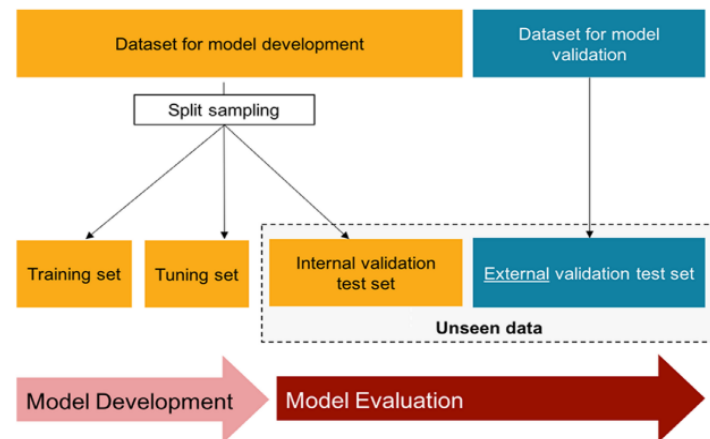


Fase 4:
Ontwikkeling
van de software-
toepassing

Modelontwikkelingsproces

Voor de ontwikkeling van het AI-model wordt de dataset gesplitst in een training, tuning en interne validatie test set. De prestatie van een AI-model is sterk afhankelijk van de data waarmee het getraind is.

- **Training dataset:** een subset van de data die worden gebruikt om een AI-model te trainen
- **Tuning dataset:** een subset van de data die worden gebruikt om de hyper-parameters van het AI-model af te stemmen
- **Test dataset:** een subset van de data die tijdens de training nooit aan het AI-model wordt getoond, maar wordt gebruikt om te testen wat het model heeft geleerd. Een externe validatie test set wordt gebruikt om de generaliseerbaarheid van het AI-model te testen.



TIPS

- Let op:
 - Is de ziekte prevalentie in de test dataset representatief voor de patiënten doelgroep in de echte wereld?
 - Is er sprake van onder- of over representatie van subgroepen in de training dataset?
 - Welke inclusie of exclusie criteria zijn gebruikt?
 - Wat is de gouden standaard waarmee de uitkomsten van het AI-model is vergeleken?
- De Area under the curve (AUC) is de meest bekende uitkomstmaat om de prestatie van AI-modellen te evalueren voor classificaties. Een lage AUC (<0.8) kan alsnog gunstiger zijn dan de huidige situatie
- Voor het ontwikkelen van een AI-model met patiëntgegevens moet informed consent gevraagd worden
- Voor eventuele publicatie van het AI-model dient het protocol door een wetenschapscommissie beoordeeld te worden

Verdieping prestatie uitkomsten

Pakket van Eisen (PvE)

Het **PvE** beschrijft de eisen waaraan de AI-software moet voldoen. De **functionele eisen** worden door de ontwikkelaar samen met de eindgebruikers vastgesteld. **Niet-functionele eisen** worden met ondersteunende specialismen vastgesteld. De eisen worden in de volgende categorieën beschreven: functioneel, technisch, kwaliteit & veiligheid, implementatie & opleiding, en data & ICT. Op basis van de gestelde eisen wordt een (software)ontwerp gemaakt. Zie het onderdeel **eigenschappen AI-software** voor meer aandachtspunten.

TIPS

- Bouw functionaliteiten in die de transparantie en monitoring van de AI-software ondersteunen
- Denk hierbij aan de mate van model-uitleg en informatie die gewenst is, het overrulen van resultaten, de registratie van feedback, fouten, en melding van onbetrouwbare input en resultaten
- Maak grenswaardes en bijbehorende referentiewaardes inzichtelijk bij categorische uitkomsten
- Toon patiënt-individuele trends van de uitkomsten
- Ondersteun lokale integratie van AI-software door het model in een klinisch dashboard te integreren

PRAKTIJKVOORBEELD

- ☐ Siemens AI-Rad Companion PvE

Ontwerp

1. Functioneel ontwerp	Gewenste zorgproces	<ul style="list-style-type: none">• Op functioneel niveau wordt de integratie met het zorgproces beschreven• De momenten in de workflow waar AI-software geïntegreerd kan worden:<ul style="list-style-type: none">➢ Voor een handeling, bijvoorbeeld triage van patiënten of het prioriteren van belangrijke bevindingen➢ Tijdens een handeling als waarnemingshulpmiddel (tweede lezer) of als interactieve beslisondersteuning
	Data management	<ul style="list-style-type: none">• De specificaties van de dataset, meta-data, en IT-standaarden worden vastgesteld in een datamanagementplan• Het modelontwikkelingsproces wordt hierin ook beschreven, met versiebeheer
2. Technisch ontwerp	Applicatie en integratie	<ul style="list-style-type: none">• Voor het ontwerp van de AI-software applicatie zelf wordt het onderdeel softwareontwikkelproces uit de Checklist zelfbouw medisch hulpmiddel gevolgd• Op applicatie niveau wordt beschreven hoe de systemcomponenten met elkaar communiceren
	IT-infrastructuur integratie	<ul style="list-style-type: none">• Op database en server niveau wordt beschreven hoe de systeemcomponenten met elkaar communiceren• De data-aanlevering gebeurt vanuit een lokaal datawarehouse of datalake
3. Security ontwerp	Security maatregelen	<ul style="list-style-type: none">• De beveiligingsmaatregelen worden beschreven, belangrijke onderwerpen zijn de authenticatie en autorisatie van gebruikers, het privacy niveau van de data en de weerbaarheid van systeemcomponenten

TIPS

- Zorg dat de data-aanlevering aansluit bij de tijdigheid die nodig is; een vertraging op de data maakt real-time voorspellingen onmogelijk
- De HL7-FHIR standaard kan gebruikt worden voor transport van informatie, maar ook om de inhoud van de data te structureren
- Breng in kaart hoe een applicatie of beslissingsondersteuning vanuit het bronsysteem aangeroepen kan worden en resultaten terug geeft

PRAKTIJKVOORBEELD

- Zorgproces workshop Radiologie
- Functioneel en technisch ontwerp AI-Rad Companion Chest CT
- Checklist zelfbouw medisch hulpmiddel

Implementatieplan

In het implementatieplan wordt de **training van eindgebruikers** beschreven en hoe de **evaluatie van AI-software** wordt uitgevoerd. Er wordt afgesproken wie waar verantwoordelijk voor is (**beheerafspraken**) gedurende de opvolgende fases. Een plan voor de **lokale validatie** en **monitoring van AI-software** kan ook onderdeel zijn van het implementatieplan. Om de AI-software op gecontroleerde wijze te introduceren in het zorgproces kunnen 2 vormen gekozen worden:

- een *pilot* waarin de AI-software voor een bepaalde periode wordt gebruikt
- **schaduwdraaien** naast het traditionele zorgproces, de AI-software wordt in “stille modus” gebruikt en nog niet in de dagelijkse klinische praktijk

TIPS

- Bedenk een communicatieplan om bijvoorbeeld verwijzende specialisten of patiënten te informeren over het gebruik van AI-software en de interpretatie van resultaten
- Voer dialoog met stakeholders over ethische dilemma's rondom bias, privacy en vertrouwen gedurende opvolgende fases met de [Handleiding aanpak begeleidingsethiek voor artificiële intelligentie in gezondheid en zorg](#) | [Publicatie](#) | [Data voor gezondheid](#)

PRAKTIJKVOORBEELD

- ❑ Implementatieplan pilot AI-Rad Companion Chest CT

Fase 1 B: Verwerving

[Stappenplan](#)

Verwerving van AI-software is in verschillende vormen mogelijk, waarvan aanschaf, proefplaatsing, huur/lease en een periode gratis gebruik het meest voorkomen. De producten die opgesteld worden tijdens de verwervingsfase worden vastgelegd in het **productdossier**. Zoals het **PvE**, **ontwerp, risico analyse** en het **implementatieplan**. De informatie die nodig is volgt uit het **modelontwikkelingsproces** en de **eigenschappen van AI-software**.

Eigenschappen AI-software

De **eigenschappen van de AI-software** worden (bijvoorbeeld in een **PvE**) uitgevraagd bij de leverancier of ontwikkelaar met hulp van 10 vragen. Hiermee wordt het doelgebruik duidelijk, wordt inzicht gecreëerd in het modelontwikkelingsproces, de potentiële meerwaarde, implementatie uitdagingen en de mogelijkheden voor de integratie met het zorgproces. Andere relevante hulpmiddelen zijn de koophulpen van **ECLAIR** en de **NHS**.

10 vragen aan de ontwikkelaar van de AI-software (gebaseerd op **literatuur**):

Intended use

- Wat zijn het doel en de context van het algoritme?

Data

- Hoe goed zijn de data waarmee het algoritme is ontwikkeld?
- Is het algoritme met voldoende data ontwikkeld?

Effectiviteit

- Hoe goed presteert het algoritme?
- Zijn de uitkomstmaten klinisch relevant?
- Is het algoritme toepasbaar in een nieuwe klinische setting?
- Welk bewijs is er dat de effectiviteit van het algoritme aantoon?

Integratie

- Hoe integreert het algoritme in het huidige zorgproces?
- Welke risico's zijn er bij het gebruik van het algoritme?

Adoptie

- Welke ethische, juridische of sociale overwegingen spelen een rol bij de adoptie van het algoritme?

Pakket van Eisen (PvE)

Het **PvE** beschrijft de eisen waaraan de AI-software moet voldoen. De **functionele eisen** worden door de projectleider samen met de eindgebruikers vastgesteld. **Niet-functionele eisen** worden met ondersteunende specialismen vastgesteld. Zie het onderdeel **eigenschappen AI-software** voor meer aandachtspunten. De eisen, wensen en de behoefte aan informatie worden in de volgende categorieën beschreven: functioneel, technisch, kwaliteit & veiligheid, data & ICT, documentatie, implementatie & opleiding, service, inkoopvoorwaarden, en referenties. Op basis van de beantwoording wordt een keuze gemaakt in AI-software pakket.

TIPS

- Vraag welke stakeholders betrokken zijn geweest bij de ontwikkeling van de AI-software
- Vraag de wetenschappelijke studies op die met de AI-software zijn uitgevoerd
- Vraag hoe de leverancier om gaat met updates/upgrades van de AI-software en in welk format patiëntdata gearchiveerd kan worden
- Vraag naar mogelijkheden om een voorspelling af te wijzen en de reden ervan te registreren
- Vraag een demo aan bij de leverancier om een indruk te krijgen van de resultaten die de AI-software toont
- Breng samen met de leverancier de mogelijkheden in kaart om de klinische meerwaarde aan te tonen

PRAKTIJKVOORBEELD

- ☐ Siemens AI-Rad Companion PvE

Ontwerp

1. Functioneel ontwerp	Gewenste zorgproces	<ul style="list-style-type: none">• Op functioneel niveau wordt de integratie met het zorgproces beschreven• De momenten in de workflow waar AI-software geïntegreerd kan worden:<ul style="list-style-type: none">➢ Voor een handeling, bijvoorbeeld triage van patiënten of het prioriteren van belangrijke bevindingen➢ Tijdens een handeling als waarnemingshulpmiddel (tweede lezer) of als interactieve beslisondersteuning
	Data management	<ul style="list-style-type: none">• De specificaties van de dataset, meta-data, en IT-standaarden worden vastgesteld in een datamanagementplan
2. Technisch ontwerp	Applicatie en integratie	<ul style="list-style-type: none">• Op applicatie niveau wordt beschreven hoe de systemcomponenten met elkaar communiceren
	IT-infrastructuur integratie	<ul style="list-style-type: none">• Op database en server niveau wordt beschreven hoe de systeemcomponenten met elkaar communiceren• De data-aanlevering gebeurt vanuit een lokaal datawarehouse of datalake
3. Security ontwerp	Security maatregelen	<ul style="list-style-type: none">• De beveiligingsmaatregelen worden beschreven, belangrijke onderwerpen zijn de authenticatie en autorisatie van gebruikers, het privacy niveau van de data en de weerbaarheid van systeemcomponenten

TIPS

- Zorg dat de data-aanlevering aansluit bij de tijdigheid die nodig is; een vertraging op de data maakt real-time voorspellingen onmogelijk
- De HL7-FHIR standaard kan gebruikt worden voor transport van informatie, maar ook om de inhoud van de data te structureren
- Breng in kaart hoe een applicatie of beslissingsondersteuning vanuit het bronsysteem aangeroepen kan worden en resultaten terug geeft

PRAKTIJKVOORBEELD

- Zorgproces workshop Radiologie
- Functioneel en technisch ontwerp AI-Rad Companion Chest CT

Implementatieplan

In het implementatieplan wordt de **training van eindgebruikers** beschreven en hoe de **evaluatie van AI-software** wordt uitgevoerd. Er wordt afgesproken wie waar verantwoordelijk voor is (**beheerafspraken**) gedurende de opvolgende fases. Een plan voor de **lokale validatie** en **monitoring van AI-software** kan ook onderdeel zijn van het implementatieplan. Om de AI-software op gecontroleerde wijze te introduceren in het zorgproces kunnen 2 vormen gekozen worden:

- een **pilot** waarin de AI-software voor een bepaalde periode wordt gebruikt
- **schaduwdraaien** naast het traditionele zorgproces, de AI-software wordt in “stille modus” gebruikt en nog niet in de dagelijkse klinische praktijk

TIPS

- Bedenk een communicatieplan om bijvoorbeeld verwijzende specialisten of patiënten te informeren over het gebruik van AI-software en de interpretatie van resultaten
- Voer dialoog met stakeholders over ethische dilemma's rondom bias, privacy en vertrouwen gedurende opvolgende fases met de [Handleiding aanpak begeleidingsethiek voor artificiële intelligentie in gezondheid en zorg](#) | [Publicatie](#) | [Data voor gezondheid](#)

PRAKTIJKVOORBEELD

- ❑ Implementatieplan pilot AI-Rad Companion Chest CT

Fase 2: Ingebruikname

[Stappenplan](#)

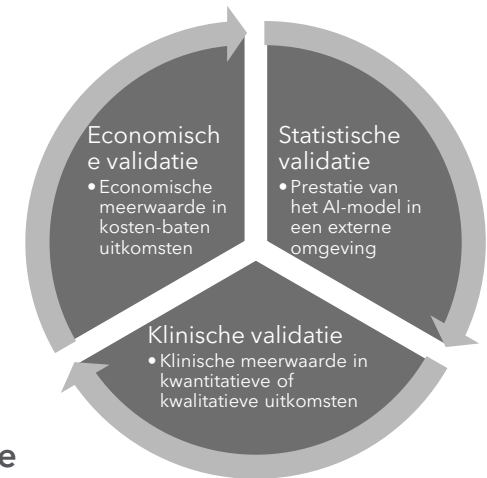
De **fase ingebruikname** start nadat de AI-software is verworven of ontwikkeld. In deze fase worden alle voorbereidingen voor het klinisch gebruik getroffen. De AI-software wordt geïnstalleerd, geconfigureerd, getest, gevalideerd en vrijgegeven na succesvolle afronding. De aanpak voor de **lokale validatie** wordt in een **validatieprotocol** beschreven. **Scholing** en **beheerafspraken** moeten geregeld zijn vooraf klinisch gebruik.

Lokale validatie

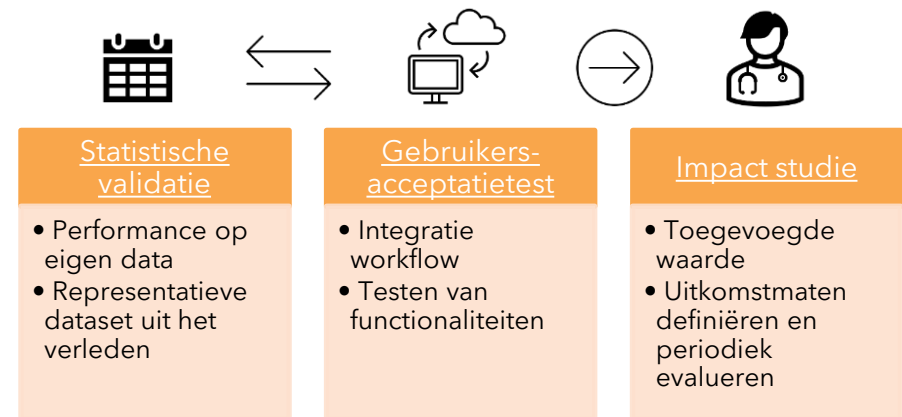
Voor de ingebruikname van AI-software is het noodzakelijk om het functioneren van de software te valideren. Om te controleren of de softwarespecificaties voldoen aan de behoeften van de gebruiker en het beoogde gebruik in een klinische omgeving. Dit wordt meestal gedaan door een **gebruikers-acceptatietest**, waarin de workflow integratie en de functionaliteiten van de AI-software getest worden.

Maar in de context van het AI-model gebeurt externe validatie in 3 dimensies: statistisch, klinisch en economisch. Afhankelijk van het **level of evidence** en de haalbaarheid wordt bepaald hoe de lokale validatie wordt ingevuld. Uiteindelijk is de validatie dat de AI-software waarde toevoegt voor zorgverleners en patiënten, het belangrijkste.

Dimensies van externe validatie van AI-software (uit literatuur)



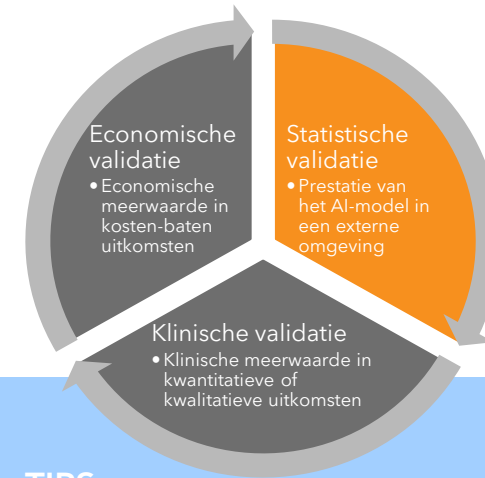
Volgorde lokale validatie van AI-software



Statistische validatie

Een statistische validatie wordt bij voorkeur altijd lokaal uitgevoerd bij inkoop van AI-software. Hierbij wordt de voorspellende waarde van het AI-model op een eigen dataset getest. Dit is nodig wanneer eerdere effectiviteitsstudies onvoldoende indicatie geven van de betrouwbaarheid en toepasbaarheid van het AI-model binnen de lokale context. Maar het kan ook handig zijn om de limitatie van het AI-model inzichtelijk te maken voordat het klinisch gebruikt wordt. Zoals verschillende vormen van algoritmische *bias* uit het **modelontwikkelingsproces**. Bij zelfbouw van AI-software wordt de voorspellende waarde van het AI-model op een externe dataset getest.

De dataset is meestal een representatieve set uit het verleden. Maar een retrospectieve validatie is niet altijd mogelijk: soms wordt de data nog niet vastgelegd of er is te weinig data, of de data is niet compatibel als input voor het AI-model. Dan is *temporale validatie* ook mogelijk, waarbij eerst nieuwe data wordt verzameld.



TIPS

- De gouden standaard waarmee de uitkomsten van het AI-model worden vergeleken is vaak de beoordeling van een groep experts; houd rekening met opleidingsniveau, beschikbare tijd, en discrepanties in de beoordeling
- In het algemeen geldt: hoe groter de dataset hoe beter (maar hangt wel af van de kwaliteit van data)

Gebruikers-acceptatietest

Met een **gebruikers-acceptatietest** wordt de opmaak van de AI-software, de benodigde input en eventuele handelingen die worden gevraagd van de eindgebruiker getoetst aan de functionele eisen die vooraf zijn gesteld. Daarvoor worden alle koppelingen van de AI-software met andere systemen getest in een test/acceptatieomgeving, gescheiden van de productieomgeving (*OTAP-omgeving*). Door vast te stellen hoe er met de AI-software gewerkt gaat worden kan de juiste configuratie getest worden. Wanneer deze test succesvol is, vindt technische vrijgave plaats door de (technisch-, applicatie- en functioneel)beheerders aan de gebruiker. Na functionele vrijgave door de eindverantwoordelijke gebruiker mag de AI-software klinisch ingezet worden. Om de test uit te voeren kunnen de volgende hulpmiddelen worden gebruikt:

- Data van test patiënten of vrijwilligers
- Synthetische data
- Referentie-datasets, bijvoorbeeld openbare datasets

TIPS

- Test de functionaliteiten van de AI-software ook met extreme of afwijkende cases
- Controleer of notificaties en foutmeldingen getoond worden
- Test de beveiliging van AI-software om cybersecurity kwetsbaarheden te signaleren, en laat dit bij zelfbouw AI-software onafhankelijk testen door een externe partij

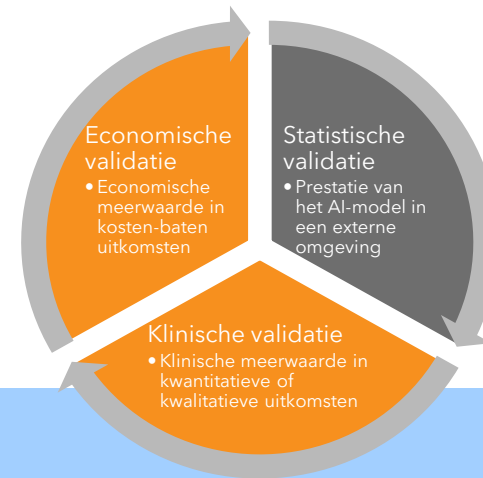
PRAKTIJKVOORBEELD

- ❑ Gebruikers-acceptatietest AI-Rad Companion Chest CT

Impact studie

In een impact studie wordt de klinische of economische meerwaarde van de AI-software gevalideerd. Binnen het ziekenhuis is een **haalbaarheidsstudie** een belangrijke stap om in een vroeg stadium de impact van AI-software te bepalen. Dit gebeurt in **Fase 3: gebruik**. Met een haalbaarheidsstudie wordt een **verbetering van de gezondheidszorg** onderzocht tijdens gebruik van de AI-software in de echte klinische omgeving. Met **uitkomstmaten** voor het bestuderen van processen en de effecten van de AI-software.

Naast kwantitatieve uitkomsten, vormt de **gebruikerservaring** ook een belangrijk bron van informatie. Bijvoorbeeld of de AI-software bijdraagt aan **beter inzicht** of **meer vertrouwen** in de beoordeling van een patiënt, en andere waargenomen effecten door de eindgebruiker. De impact van AI-software op het zorgproces is dan wel afhankelijk van de ervaring van eindgebruikers, en de keuzes in integratie met bronsystemen.



TIPS

- Denk aan een verbetering van de kwaliteit, veiligheid, effectiviteit, patiëntgerichtheid, tijdigheid, kosten, efficiëntie en gelijkheid van gezondheidszorg
- Zolang de zorgverlener geen meerwaarde ervaart, zal de AI-software niet gebruikt worden

PRAKTIJKVOORBEELD

- [Impact of Concurrent Use of Artificial Intelligence Tools on Radiologists Reading Time: A Prospective Feasibility Study](#)

Validatieprotocol

Het validatieprotocol beschrijft de **lokale validatie** van de AI-software. Wat minimaal beschreven moet worden is het doel van de uitkomstmaten en hoe deze berekend worden, de meetfrequentie, wie de data verzameld en de resultaten beoordeelt, en de bron van de data. Voor het uitvoeren en vastleggen van de externe validatie van AI-software kunnen de volgende richtlijnen geraadpleegd worden:

Externe validatie	Richtlijn (*in ontwikkeling)
Statistische validatie	TRIPOD , TRIPOD-AI*
Impact studie	NHS Handbook of quality and service improvement tools, H5 SQUIRE 2.0 guideline DECIDE-AI*
* In ontwikkeling	

TIPS

- Voor eventuele publicatie van de resultaten uit de lokale validatie dient het protocol door een wetenschapscommissie beoordeeld te worden

PRAKTIJKVOORBEELD

- Validatieprotocol AI-Rad Companion Chest CT

Stap 8: Evalueer de uitkomsten

Stappenplan

In deze stap wordt het gebruik van AI-software in de klinische praktijk geëvalueerd, in **Fase 3: gebruiksfase**, en afgevoerd in **Fase 4: afvoerfase**. Per fase worden de onderdelen beschreven waar je aan moet denken, met de documenten die vastgelegd worden in het ontwikkel- of productdossier. Afhankelijk van de uitkomsten en eventuele wijzigingen in het gebruik van AI-software, wordt stap 3-6 herhaald. Indien het een grote wijziging betreft wordt **Fase 1** opnieuw doorlopen. In andere gevallen wordt alleen **Fase 2** herhaald.

CHECKLIST STAP 8

- De uitkomsten van het gebruik van AI-software worden periodiek geëvalueerd
- Ervaringen over het gebruik van AI-software worden gedeeld binnen de samenwerkingsverbanden van het ziekenhuis
- Maatregelen worden getroffen om AI-software veilig af te voeren

In de **gebruiksfase** wordt AI-software toegepast in de klinische praktijk. Daar hoort het proces omtrent **wijzigingen** en periodieke evaluatie van **real-world monitoring** bij.

Real-world monitoring

Het gebruik van AI-software in de dagelijkse klinische praktijk wordt gemonitord aan de hand van positieve en negatieve effecten. In **Fase 2** wordt vastgelegd op welke uitkomstmaten gestuurd wordt. Na verloop van tijd (bijvoorbeeld elke 6 maanden) wordt op basis van de verzamelde gegevens de **veiligheid en effectiviteit** van AI-software geëvalueerd. Eventuele **VIM's of product recalls** spelen hierin ook een rol. Maar fouten, risico's of verkeerd gebruik kunnen niet altijd real-time technisch gemonitord worden in de software zelf. De monitoringsuitkomsten kunnen dan ook periodiek in een overlegstructuur worden besproken.

Monitoringsuitkomsten AI-software (gebaseerd op **Leidraad AI in de zorg**):

Foute voorspellingen	•De accuraatheid van het AI-algoritme, denk aan foutpositieve en foutnegatieve voorspellingen
Technische fouten	•Technische fouten in de AI-software, belangrijk bij updates/upgrades van software
Fairness	•Ongewenste verschillen in uitkomsten bij patienten groepen
Risico's	•Eerder vastgestelde risico's en onzekerheden
Deployment bias	•Of de AI-software op de juiste manier wordt gebruikt
Benefits	•Periodieke evaluatie van de meerwaarde van AI-software

Wijzigingen

1 Externe wijzigingen

- Externe wijzigingen komen voort uit **updates, upgrades of recalls** van AI-software
- De meeste AI-software wordt ontwikkeld met *statische algoritmes*. Maar AI-software ondergaat updates/upgrades als er nieuwe gegevens beschikbaar zijn en/of als er verbeteringen aan het algoritme moeten worden aangebracht
- Bij **grote wijzigingen** moet de prestatie van het AI-model en de datakwaliteit aantoonbaar gelijk blijven; zeker wanneer het gaat om *adaptieve algoritmes*

2 Interne wijzigingen

- Interne wijzigingen komen o.a. voort uit **VIM's** over calamiteiten, complicaties, incidenten of onveilige situaties door het gebruik van AI-software
- Ook kan de **evaluatie van de uitkomsten** van AI-software leiden tot wijzigingen in AI-software of het zorgproces

TIPS

- Leg alle wijzigingen vast in het softwarebeheersysteem
- Lokaal hertrainen van het AI-algoritme is een specifieke wijziging, waarbij het modelontwikkelingsproces opnieuw wordt doorlopen

Fase 4: Afvoer

[Stappenplan](#)

In de **afvoerfase** is besloten om het gebruik van AI-software te stoppen. Omdat de patiëntdata na afvoer bewaard blijven, worden **afvoercontroles** gedaan om de **data veilig te stellen**.

Afvoer controles

Wanneer AI-software wordt afgekeurd, buiten gebruik wordt gesteld, en afgevoerd vinden een aantal controles plaats:

- Impact en risico's bepalen voor het zorgproces bij de afvoer van AI-software (**stap 4**)
- Buiten gebruik stellen AI-software
- Veiligstellen patiëntdata door archivering (bijvoorbeeld in een datawarehouse)
- Archivering van AI-software in het softwarebeheersysteem

En in geval van commerciële AI-software:

- Verwijderen van data bij de leverancier
- Bewijs van vernietiging opvragen bij leverancier
- Beëindigen servicecontract en/of licentieovereenkomst

1 Veiligstellen van patiënt data

- Bij afvoer van AI-software moet de data veiliggesteld worden
- De bewaartermijn van patiëntdata is **20 jaar**
- Wanneer er data bij de leverancier staan en deze niet verwijderd kunnen worden door de zorginstelling, wordt een **vernietigingsverklaring** opgesteld en gezamenlijk ondertekend met de leverancier. Hiermee wordt de verantwoordelijkheid voor het verwijderen van patiëntgegevens overgedragen naar de leverancier.

2 Veiligstellen van gegevens over AI-software

- Voor fabrikanten geldt een wettelijke bewaartermijn van tenminste **10 jaar** voor gegevens over AI-software, zoals de software versies en de staat waarin een AI-algoritme zich bevond
- Voor zorginstellingen geldt **geen wettelijke bewaartermijn**, zodra er geen grondslag meer is om de gegevens te bewaren moet het verwijderd worden
- Voor **zelfbouw van AI-software** is het nog belangrijker om een bewaartermijn vast te stellen, omdat ook het ontwikkeldossier moet worden bewaard
- De archivering van de gegevens over AI-software in de zorginstelling is geborgd met het **softwarebeheersysteem**, waarin alle wijzigingen zijn vastgelegd.

Productdossier Siemens AI:

- PvE
- Workshop zorgproces
- Functioneel & technisch ontwerp
- Classificatie informatieverwerking Teamplay
- PRI
- Implementatieplan
- Validatieprotocol en gebruikers-acceptatietest

Overig:

- Checklist zelfbouw medisch hulpmiddel (MDR)

Op te vragen via: klinischeinformatica@maasstadziekenhuis.nl

Key publicaties:

- Toolkaart hulpmiddelen Waardevolle AI
- The clinical artificial intelligence department: A prerequisite for success (Cosgriff et al., 2020)
- Key challenges for delivering clinical impact with artificial intelligence (Kelly et al., 2019)
- The practical implementation of artificial intelligence technologies in medicine (He et al., 2019)
- Rapport health data infrastructuur en afsprakenstelsel (NL AIC, 2020)
- Evaluation framework to guide implementation of AI systems into healthcare settings (Reddy et al., 2021)



[Use case 1](#)



[Use case 2](#)

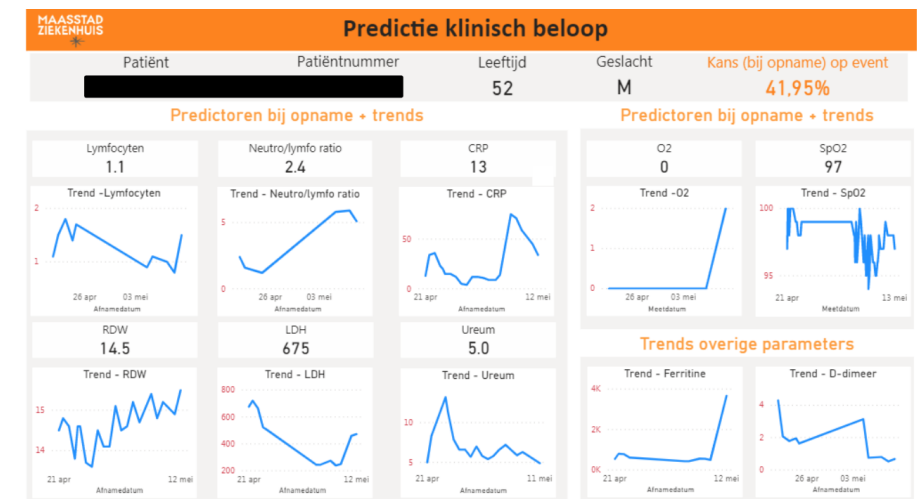
Use case 1: Covid-19 severity score

In het MSZ is een Covid-19 severity score ontwikkeld waarmee de ernst van het klinische beloop (respiratoir falen) van Covid-19 patiënten vroegtijdig kan worden voorspeld. Met *machine learning* op basis van klinische en laboratoriumparameters gemeten op het moment van opname in het ziekenhuis. Het algoritme is ontwikkeld met data van Covid-19 patiënten die tussen 5 maart en 7 mei 2020 op een klinische verpleegafdeling werden opgenomen. De data omvatten klinische data uit het EPD, waaronder vitale parameters, leeftijd en lab uitslagen. De uitkomstmaat van het algoritme is het optreden van ernstig respiratoir falen, door behandeling met high-flow nasal cannula support (HFNC), opname op de intensive care afdeling of overlijden tijdens of aansluitend aan de opname. Op basis van de voorspelde kans, worden patiënten verdeeld in laag, midden en hoog-risico categorieën. De Covid-19 severity score helpt artsen om de zorg tijdig en beter af te stemmen. Bij opname kan er een betere inschatting gemaakt worden van het klinische beloop, waardoor de zorg voor de patiënt op maat geboden wordt. De Covid-19 severity score is geïntegreerd in een dashboard en daarmee een praktische tool voor samen beslissen.

Het stappenplan HAI is gevalideerd met deze use case.

De resultaten zijn op te vragen via:

klinischeinformatica@maasstadziekenhuis.nl



Use case 2: Siemens AI-Rad Companion

Siemens AI-Rad Companion ondersteunt de radiologische workflow bij het beoordelen van medische beelden. Het doel van dit hulpmiddel is het reduceren van de last van repetitieve taken, waardoor de diagnostische precisie verbeterd kan worden bij het interpreteren van medische beelden. De AI-Rad Companion CT Chest module biedt kwantitatieve ondersteuning bij de beoordeling van CT scans van de thorax, zoals het hart, de aorta, de longen en de vertebrae. Met *deep learning* algoritmes geeft AI-Rad Companion automatisch afwijkingen, segmentaties, en metingen aan. Zoals het markeren van longnoduli en het berekenen van de longdensiteit. Daarnaast wordt er een gestructureerd verslag gemaakt van de resultaten. De AI-software wordt aangeboden via het digitale platform Teamply. Het implementeren van AI-Rad Companion in de radiologie workflow kan de werkdruk voor de radioloog te verlagen en aanvullende bevindingen in kaart brengen. Zo kunnen patiënten de best mogelijke zorg ontvangen en krijgen de radiologen meer tijd voor andere taken. Zie de **bevindingen** die tijdens het project zijn opgedaan.



Bevindingen use case 2: Siemens AI-Rad Companion

Stap 1: Bepaal de medische context

- Bij de **toekomstige ontwikkelingen voor de CT** is de verwachting dat AI erg belangrijk wordt voor een **efficiënte workflow** voor medische beeldvorming. Gezien de toename in productiecijfers van CT thorax onderzoeken is er behoefte aan betere ondersteuning bij de routine beeldanalyse en rapportage. Hierdoor wordt de kwaliteit van de geleverde patiëntenzorg gewaarborgd en kunnen radiologen zich beter focussen op meer complexe casuïstiek. De inzet van AI-software moet bijdragen aan het **verminderen van de werkdruk** en een efficiënte workflow van de radioloog. Daarnaast draagt het bij aan een verbetering van de strategische samenwerking met externe partners, en wordt er binnen de vakgroep Radiologie ervaring opgedaan met AI als beslissingsondersteuning.

Stap 2: Bepaal het innovatieniveau

- AI-Rad Companion** is een **commercieel product** en komt met het Teamplay digital health cloud platform van Siemens
- Het MSZ is het **eerste ziekenhuis** dat AI-Rad Companion implementeert

Stap 3: Bepaal de toegevoegde waarde

- De algoritmes zijn **getraind en getest** met een dataset tot in totaal 3 miljoen images. Daarnaast **multi-vendor (extern) gevalideerd** met Siemens, GE en Philips scanners. De reproduceerbaarheid van resultaten is gemeten middels cross-validation. De trainingsdataset bevatte diverse populaties (voor vrijgave getest op de betreffende populatie (Asian, Caucasian, etc..)) gebaseerd op landelijke voorschriften.
- De **sensitiviteit** voor de detectie van longnoduli ligt tussen de 90-100%, afhankelijk van type patiënt cohort
- Het huidige wetenschappelijke bewijs naar de **klinische effectiviteit** van de AI-software laat geen consequente tijdswinst zien, wel aanvullende klinisch relevante bevindingen

Stap 4: Bepaal de risico's en impact

- Een verlenging van de **proeflicentie** duurde langer dan verwacht, waardoor de AI-software een paar dagen buiten gebruik was
- Uit de **PRI** blijken de grootste risico's te zitten in de suboptimale presentatie van resultaten, overnemen van verkeerde resultaten in de verslaglegging, en de communicatie over het gebruik van AI-software naar aanvragers en patiënten
- Met een **CE-klasse IIa** (MDR) is het risico voor de patiëntveiligheid middelhoog. De AI-software wordt ingezet als 2de lezer waarbij de resultaten door de radioloog zelf worden geïnterpreteerd.

Stap 5: Maak afspraken over taken, bevoegdheden en verantwoordelijkheden

- De **AI-kartrekker** binnen de vakgroep Radiologie kreeg ondersteuning van het implementatie team, bestaande uit eindgebruikers, kennisexperts, en besluitvormers. Er werd een **laborant** als functioneel beheerder van de AI-software aangesteld, die de lokale validatie begeleidde. Twee studenten hebben verdiepend onderzoek gedaan naar de verwachtingen van eindgebruikers.
- Er bleek veel afstemming nodig te zijn over welke resultaten uit het **AI-rapport** wel en niet nuttig zijn
- Eindgebruikers hebben nog **onvoldoende scholing** gevolgd

Stap 6: Controleer toepasselijke wet- en regelgeving

- De afstemming over de verwerkersovereenkomst leverde veel vertraging op in het project. Een ISO27001 certificering bleek namelijk onvoldoende, het MSZ vraagt om een NEN7510 certificering specifiek voor de **informatiebeveiligingsnormen** voor patiënt data.
- AI-Rad Companion Chest CT bestaat uit **3 medische hulpmiddelen** met functionaliteiten voor de long-, cardiovasculaire-, en musculoskeletale analyses
- Gepseudonimiseerde beelden** worden 7 dagen in de cloud omgeving bewaard, langer is onnodig

Stap 7: Bereid voor op implementatie

- De **gewenste functionaliteit** voor follow-up longnoduli is in de huidige softwareversie nog niet beschikbaar
- Uit de **gebruikers-acceptatietest** bleek de koppeling tussen het PACS en de cloud-omgeving niet te werken, door problemen in de configuratie van de firewall
- Statistische validatie** op data uit het verleden was niet mogelijk, vanwege andere instellingen voor de reconstructie van CT-beelden. Er werd besloten om in de loop van de tijd testcases te verzamelen (d.m.v. schaduwdraaien) om te laten beoordelen door verschillende radiologen.
- De **impactstudie** naar aanvullende (kritische) bevindingen en tijdswinst wordt samen met Siemens vormgegeven.

Stap 8: Evalueer de uitkomsten

- De **AI-resultaten** in het PACS werden suboptimaal getoond:
 - Het was niet duidelijk hoe normaalwaarden zijn gevalideerd, en bij welke grenswaardes betrokken specialisten acteren. Ook voor automatische scores was niet te zien of corrigerende factoren, zoals leeftijd en geslacht werden meegenomen
 - Extra series werden onoverzichtelijk in het PACS opgeslagen
- Om **patiënten en betrokken specialisten** te informeren over de AI-software, is een disclaimer toegevoegd aan het verslag van de radioloog in het EPD (en patiëntenportaal)
- Er wordt een **goede werkwijze** ontwikkeld die toepasbaar is voor andere use cases, mede door inzicht in implementatie benodigdheden. De gebruikerstevredenheid wordt nog geëvalueerd

Het stappenplan HAI beweegt mee met de toekomst en staat open voor verandering. Niet alleen het leervermogen van AI-algoritmes wordt hierdoor benut, maar ook die van de organisatie. Suggesties en opmerkingen kunnen worden gestuurd naar klinischeinformatica@maasstadziekenhuis.nl