



Audit achterstandsscores scholen en gemeenten: bevindingen per norm

projectnummer 206545/01

Kwaliteitszorg en Auditing

25 november 2019

Samenvatting

Deze nota bevat drie tabellen.

Tabel 1 bevat de bevindingen per onderzochte norm (exclusief norm 32) uit de Kwaliteitsrichtlijnen-2018, gebruikt bij het auditrapport *Audit achterstandsscores scholen en gemeenten* (CBS, 22 november 2019).

Tabel 2 bevat de bevindingen en aanbevelingen voor norm 32 uit de Kwaliteitsrichtlijnen-2018, gebruikt bij het auditrapport *Audit achterstandsscores scholen en gemeenten* (CBS, 22 november 2019).

Tabel 3 geeft de relatie weer tussen de drie geconstateerde fouten en de meest direct corresponderende normen uit de Kwaliteitsrichtlijnen-2018.

Tabel 1: Bevindingen per onderzochte norm (exclusief norm 32) uit de Kwaliteitsrichtlijnen-2018.

	Beschrijving norm		Bevinding	
1	Relatie en communicatie met de klant			
	1.1	Effectiviteit van de relatie en communicatie met de klant		
		1.1.1	Er is een adequate overlegstructuur	<ol style="list-style-type: none"> 1) Er is/was regelmatig extern overleg tussen de projectleiders en -medewerkers van SDI en SQS met hun evenknieën bij OCW. 2) Op managementniveau is er geen sprake van regulier extern overleg; er is vooral overlegd bij incidenten. 3) Het besef dat het project een politiek gevoelig onderwerp betreft was er bij het betrokkenen management (hoofddirectie, directeuren, projectleider, supervisor). Desondanks is er geen stuurgroep ingesteld. 4) Er is geen formele overdracht van ontwikkeling naar productie geweest.
		1.1.2	Wijzigingsverzoeken worden adequaat afgehandeld.	Ja. In de ontwikkelfase is op aanvullende vragen van m.n. OCW gereageerd, ook als dat tot tijdsdruk en uitstel van controles leidde. In de latere projectplannen/offertes is voor meerwerk steeds een procedure opgenomen.
1.1.3	Problemen worden adequaat afgehandeld.	<ol style="list-style-type: none"> 1) Bij de opgetreden fouten is er door de projectleiders snel, altijd binnen 1 dag, gereageerd naar OCW. 2) De opgetreden incidenten zijn door de projectmedewerkers snel opgelost. 		
2	Afspraken met klanten			
	2.3	Volledigheid van afspraken met klanten	<p>In de Werkafspraken komen een aantal zaken niet voor:</p> <ul style="list-style-type: none"> - Er is onvoldoende rekening gehouden met activiteiten die nodig zijn om processen en systemen te ontwikkelen voor een goed lopende en onderhoudbare productiestraat. - Er zijn geen afspraken gemaakt over de werkwijze bij de afhandeling van fouten of revisie van cijfers. - In de werkafspraken staat wel een bijlage 2 "Specificatie bestandslevering DUO" waarin de door het CBS "gebruikte" kenmerken worden genoemd die volgens afspraak minimaal in het ééncijferbestand moeten zitten. Dat is als specificatie echter nog niet compleet, want de codes en labels ontbreken. De codes en labels van die kenmerken moeten ook ergens gespecificeerd worden. 	

Beschrijving norm		Bevinding
		Bij navraag worden de codes en labels van deze (door het CBS gebruikte) kenmerken als “vanzelfsprekend” of “algemeen bekend” verondersteld. Deze aanvullende specificaties voor de codes en labels zijn echter niet formeel vastgelegd.
2.3.1	In alle afspraken met klanten is het product gespecificeerd (outputspecificatie).	Ja, de op te leveren producten zijn steeds gespecificeerd.
2.3.2	In alle afspraken met klanten zijn de kwaliteitsaspecten van het product gespecificeerd (outputspecificatie).	<ol style="list-style-type: none"> 1) In de Werkafspraken worden plausibiliteitsanalyses afgesproken; deze betreffen een vergelijking met cijfers van eerdere jaren. 2) Er is een directe relatie met richtlijn 13.2.3 en 16.1.4. In een projectvoorstel (23-11-2017, “proefdraaien schoolscores”) wordt op pagina 4 een meer expliciete beschrijving gegeven van de inhoud van diverse plausibiliteitsanalyses. Die beschrijving wordt nu met enige aanpassingen gebruikt. 3) Er zijn geen afspraken over de verwachte kwaliteit of over de foutmarges van de scores.
2.3.3	In alle afspraken met klanten zijn de gewenste kwaliteitsrapportages gespecificeerd (outputspecificatie).	In de Werkafspraken worden plausibiliteitsanalyses afgesproken. Deze worden mondeling besproken met de opdrachtgever; hier wordt geen verslag van gemaakt.
2.3.7	In alle afspraken met klanten is een wijzigingsprocedure beschreven.	In de Werkafspraken is geen wijzigingsprocedure opgenomen. Dit is wel onderdeel van de Algemene Voorwaarden, maar deze zijn geen onderdeel van de Werkafspraken.
2.3.8	In alle afspraken met klanten is een incidentenprocedure beschreven.	In de Werkafspraken is geen incidentenprocedure opgenomen. Dit is wel onderdeel van de Algemene Voorwaarden, maar deze zijn geen onderdeel van de Werkafspraken.
2.3.9	In alle afspraken met klanten is een evaluatieprocedure beschreven.	In de Werkafspraken is geen evaluatie afgesproken.
2.5	Haalbaarheid van afspraken met klanten	

Beschrijving norm			Bevinding
	2.5.1	Afspraken met klanten zijn haalbaar.	<p>1) De afgesproken termijnen zijn gehaald. Maar die termijnen lijken te krap om ook te voldoen aan de CBS-Kwaliteitsrichtlijnen. Ook hebben sommige medewerkers aangegeven dat ze tijdsdruk ervaren hebben.</p> <p>2) In de productiefase was er onvoldoende gelegenheid verbeteringen door te voeren. Gezien de complexiteit van het proces is dit onwenselijk.</p>
3	Statistisch concept		
	3.1	Relevantie van het statistisch concept	
	3.1.1	Vastgelegd is waarvoor bekende klanten de statistische informatie nodig hebben.	In de Werkafspraken met OCW ligt vast waarvoor de uitkomsten gebruikt gaan worden.
	3.1.2	Alle statistieken die zijn of worden gepubliceerd, zijn ook opgenomen in het Jaarplan van het CBS.	ASD opdrachten maken deel uit van het CBS werkprogramma maar zijn niet opgenomen in het CBS Jaarplan.
	3.1.6	Er is regelmatig overleg met klanten over de bruikbaarheid van de statistiek en over toekomstige behoeften en prioriteiten.	<p>Er is regelmatig overleg met OCW. Het is onvoldoende duidelijk geworden of er over een aantal aspecten die invloed hebben op de bruikbaarheid voldoende afstemming is geweest.</p> <p>1) De gebruikte aggregatieformule is gevoelig voor kleine inputwijzigingen. Kleine inputwijzigingen kunnen daardoor relatief grote invloed hebben op de resultaten.</p> <p>2) Het is onvoldoende duidelijk geworden of er voldoende gecommuniceerd is over alle aspecten van de betrouwbaarheid van de cijfers.</p> <p>3) Het is voor scholen en gemeenten niet mogelijk om na te rekenen of het aan hen toebedeelde budget klopt.</p> <p>Het CBS heeft andere verdeelmodellen ontwikkeld binnen een andere context. Bij deze verdeelmodellen is het mogelijk voor de gebruiker de cijfers na te rekenen. Dit maakt de toepassing van het model betrouwbaarder.</p>

Beschrijving norm			Bevinding
			Toekomstige behoeften: Om de onderwijsachterstandsindicator actueel te houden is OCW voornemens deze van tijd tot tijd bij te laten stellen. Omdat het verzamelen van nieuwe gegevens over intelligentie zeer kostbaar is wil OCW weten in hoeverre de huidige indicator zonder intelligentie op schoolniveau tot andere resultaten zou leiden. Hier is onlangs door het CBS onderzoek naar gedaan.
	3.1.7	Het is duidelijk op basis van welke afspraak een statistiek is samengesteld.	Er zijn Werkafspraken.
	3.2	Samenhang van het statistisch concept met concepten van andere statistieken	Het concept 'achterstandsniveau' is in het model geoperationaliseerd. Zowel het concept als de operationalisatie zijn uitgebreid beschreven in rapporten.
4	Statistische data (cijfers)		
	4.1	Nauwkeurigheid van cijfers	Zie ook 3.1.6
	4.1.1	Er is aandacht voor de variantie van het cijfer.	<ol style="list-style-type: none"> 1) Er was aandacht voor variantie. In de begeleidingscommissie (waarvan ook OCW lid was) is hier over gesproken. 2) In de rapportages over de methodologie wordt aandacht besteed aan betrouwbaarheid, met name aan de verklaarde variantie van het model. Daarnaast wordt aandacht besteed aan de betrouwbaarheid in relatie tot variaties door de imputatiemethode. Door het willekeurige karakter van de hot-deck imputatie van de leerlingsscores ontstaat variatie in de school- en gemeentescores. Deze variatie wordt onterecht aangeduid als de "marge" van de school- en gemeentescores. Andere aspecten van betrouwbaarheid zijn onderbelicht. 3) De aggregatieformule (van leerlingniveau naar school- en gemeenteniveaus) is gevoelig voor kleine wijzigingen in de data. Het kan zijn dat de relatief grote gevolgen van met name de eerste fout mede het gevolg zijn van de gevoeligheid van de gekozen formule. De gevoeligheid van de formule is besproken met OCW en in de begeleidingscommissie; ook alternatieven zijn aan de orde geweest. De keuze voor de aggregatieformule is gemaakt door OCW, in de CBS-rapportages is dit ook vermeld. De formule is vastgelegd in AMVB's. 4) Zie ook 19.1.1

Beschrijving norm			Bevinding
	4.1.3	De nauwkeurigheid is stabiel in de tijd.	<ol style="list-style-type: none"> 1) Er zijn pas voor 1 jaar cijfers geleverd; dus niet direct van toepassing. 2) Er zijn wel voorlopige of analoge berekeningen voor eerdere jaren gemaakt. De plausibiliteitsanalyses hebben zich met name gericht op een vergelijking met die eerdere jaren. Deze zijn besproken met de opdrachtgever en door hem mondeling akkoord bevonden. 3) Het model werkt met vastgestelde coëfficiënten voor de verschillende variabelen. Het is de bedoeling deze van tijd tot tijd opnieuw vast te stellen.
5	Statistische output		
	5.4	Reproduceerbaarheid van de output	
	5.4.1	Handmatige wijzigingen van de data zijn tot een minimum beperkt.	<p>De R-scripts binnen het rekenproces zijn ontworpen om zonder handmatige interventie uitgevoerd te worden, hoewel ze wél stuk voor stuk handmatig moeten worden gestart.</p> <p>De SPSS-scripts ondersteunen een handmatig verwerkingsproces, waarbij op basis van inspectie van de data en tellingen de nodige statements in SPSS worden uitgevoerd. De daadwerkelijke wijzigingen van de data vinden vanuit de scripts, dus geautomatiseerd, plaats.</p>
	5.4.2	Alle handmatige wijzigingen van de data zijn vastgelegd.	Zie 5.4.1, alle wijzigingen in de data wordt met behulp van scripts uitgevoerd en daarin als statements vastgelegd. Iedere wijziging is daardoor traceerbaar naar een script.
	5.4.3	Voor elke versie van de output is vastgelegd welke versie van de bestanden is verwerkt en welke versie van de programmatuur is gebruikt.	Een dergelijk register is niet aangetroffen.
	5.4.4	Output is reproduceerbaar vanaf de microbase.	Er is geen expliciete microbase onderscheiden, maar de “achterstandsscores per leerling” zouden als microbase kunnen worden gezien. Het proces om vervolgens van die microbase naar hogere aggregatieniveaus te komen is vanuit de methodebeschrijving voldoende reproduceerbaar.

Beschrijving norm			Bevinding
8	Rapportages		
	8.1	Beschikbaarheid van rapportages	
	8.1.1	Alle afgesproken kwaliteitsrapportages worden geleverd.	Er zijn plausibiliteitsanalyses. Deze zijn geleverd en besproken met OCW.
10	Publicatie van correcties, bijstellingen en herzieningen		
	10.1	Conformiteit correcties, bijstellingen en herzieningen met de richtlijnen	Er is conform de CBS richtlijnen gehandeld bij de correcties. Correcties zijn direct gecommuniceerd met OCW.
	10.1.1	Publicatie van correcties, bijstellingen en herzieningen vinden conform de richtlijnen plaats.	De fouten zijn niet in alle publicaties gecorrigeerd, zoals bijvoorbeeld in de methodologische rapporten waarin geen aanpassingen zijn gedaan naar aanleiding van fout 1. Het alternatief is dat in het rapport aangegeven wordt dat er naderhand nog aanpassingen gemaakt zijn.
12	Methodologie		
	12.6	Deugdelijkheid van het imputeren	
	12.6.2	De imputatiemethode is juist geïmplementeerd.	1) Er zijn hier geen fouten opgetreden. 2) De code voor de gekozen imputatiemethodiek was eerder ontwikkeld voor een voorafgaand onderzoek. De programmacode was daarin al eerder getest.
	12.6.3	De imputatiemethode wordt juist uitgevoerd.	Geen afwijkingen geconstateerd.
	12.8	Deugdelijkheid van het transformatieproces	
	12.8.1	De transformatie naar doeleenheden is juist ontworpen, juist geïmplementeerd en wordt juist uitgevoerd.	Geen afwijkingen geconstateerd.

Beschrijving norm		Bevinding	
	12.8.2	De transformatie naar de doelpopulatie is juist ontworpen, juist geïmplementeerd en wordt juist uitgevoerd.	Bij de implementatie is eerder de 3 ^e fout opgetreden. Deze fout is gecorrigeerd met aanpassing van de software.
	12.8.3	De transformatie naar doelvariabelen is juist ontworpen, juist geïmplementeerd en wordt juist uitgevoerd.	1) Er zijn hier geen fouten opgetreden. 2) Er is geen syntaxplan of stappenplan aangetroffen.
	12.8.4	De transformatie naar doelclassificatie is juist ontworpen, juist geïmplementeerd en wordt juist uitgevoerd.	Bij de implementatie is eerder de 1 ^e fout opgetreden. Deze fout is gecorrigeerd met aanvullende software. Deze fout had voorkomen kunnen worden door labels automatisch te laten genereren door de code. Ook had de fout gesignaleerd kunnen worden door voor en na kruistabellen te maken en te vergelijken. Daarnaast hebben andere factoren een rol gespeeld: 1. Code voldoet niet aan alle richtlijnen 2. Controles zijn niet volgens de richtlijnen uitgevoerd. 3. Er is een centraal overzicht van de controles aangelegd. De bevindingen uit de controles zijn niet gedocumenteerd.
	12.8.5	De transformatie naar doelperiode of -tijdstip is juist ontworpen, juist geïmplementeerd en wordt juist uitgevoerd.	Geen afwijkingen geconstateerd
	12.8.6	Berekeningen zijn juist ontworpen, juist geïmplementeerd en worden juist uitgevoerd.	Bij de implementatie zijn eerder de 2 ^e en 3 ^e fout opgetreden. Deze fouten zijn gecorrigeerd met aanvullende software.
12.9	Deugdelijkheid van het koppelen		
	12.9.1	De koppelmethode is juist ontworpen.	Geen afwijkingen geconstateerd.

Beschrijving norm			Bevinding
	12.9.2	De koppelmethode is juist geïmplementeerd.	De koppelmethode is geïmplementeerd met behulp van SSB-tools en SPSS-syntaxen. Het proces rond het koppelen is deels handmatig. Het proces is zeer bewerkelijk. Rapportages van acceptatietests van SPSS syntaxen zijn niet gevonden.
	12.9.3	De koppelmethode wordt juist uitgevoerd.	Er zijn hier geen fouten opgetreden.
13	Statistische processen		
	13.2	Bestaan van processen	
	13.2.1	Er zijn processen voor het beoordelen en accepteren input.	Het ééncijferbestand van DUO wordt beoordeeld. Voorafgaand aan de levering is er een overleg waarin conceptuele metadata van de input en eventuele wijzigingen daarin worden besproken. Een klein deel van de kenmerken in het ééncijferbestand van DUO wordt gebruikt. Er zijn ook in dat deel kenmerken, zoals STATUS_BEK, GEWICHT en TYPE_PO, waarvan de codes en labels niet expliciet zijn vastgelegd. Zij worden vanzelfsprekend of algemeen bekend verondersteld. De SSB-bestanden die worden gebruikt bij de bouw worden verwerkt tot een analysebestand. Als hierin onverwachte frequentieverdelingen worden gevonden dan wordt naar een verklaring gezocht. Als de verklaring is terug te voeren op een SSB-inputbestand dan wordt overlegd met de leverancier van het SSB-bestand.
	13.2.2	Er zijn processen voor het beoordelen en valideren van tussenresultaten.	1) Van het analysebestand worden diverse frequentietabellen gemaakt en gecontroleerd voordat ze aan de analyse worden geleverd. Soortgelijke frequentietabellen worden in veel tussenstappen gemaakt en gecheckt. In

Beschrijving norm			Bevinding
			<p>diverse tussenresultaten wordt dus gecontroleerd of frequentieverdelingen overeenkomen met de verwachtingen.</p> <p>2) Specificaties, bijvoorbeeld de specificaties tussen bouw en analyse, zijn onvoldoende vastgelegd. Het gaat hierbij met name om het analysebestand. Het gaat dan niet alleen om de frequentieverdelingen, maar vooral ook om gebruik van labels en coderingen of afrondingen van variabelen en hun waarden. Op dit soort specificaties wordt nog weinig gecontroleerd.</p> <p>3) De pre- en postcondities voor de uitvoering van de R-scripts zijn onvoldoende duidelijk.</p>
	13.2.3	Er zijn processen voor het beoordelen en valideren van statistische output.	<p>1) Als de output wordt gemaakt dan wordt daarnaast ook een rapportage opgesteld voor de plausibiliteit van de output. In de scripts die de output produceren zijn ook controles ingebouwd om de consistentie te controleren.</p> <p>2) De fout met wel of niet bekostigde leerlingen is niet ontdekt toen deze plausibiliteitsrapportages met opdrachtgever OCW werden besproken, maar pas later. De plausibiliteitsrapportages waren dus niet toereikend om de fout te ontdekken.</p>
	13.2.4	Er zijn processen voor het behandelen van kwaliteitskwesties.	<p>Er worden in het proces, vooral bij de bouw, veelvuldig frequentietabellen gemaakt. Daarin wordt gecontroleerd of de verdelingen plausibel zijn. Hiermee wordt de 'statistische' kwaliteit van de tussenresultaten stap voor stap in de gaten gehouden. Als de kwaliteit onvoldoende is worden de directe collega's eerst geraadpleegd. Op die manier worden al veel fouten gecorrigeerd. Als het probleem dan nog blijft wordt gecommuniceerd met de projectleider en/of met de leveranciers van de gegevens.</p> <p>In de periode waarin de fouten werden gevonden was er nog geen periodiek productieoverleg waarin kwaliteitskwesties werden besproken.</p>
	13.2.5	Er zijn processen voor het beoordelen of de werkinstructies gevolgd worden.	Nee, werkinstructies ontbreken

Beschrijving norm		Bevinding
13.6	Conformiteit van het proces met de businessarchitectuur	
	13.6.1 Het statistisch proces kent rustpunten en koppelvlakken.	<p>Het proces kent wel rustpunten of koppelvlakken.</p> <p>1) De rustpunten of koppelvlakken in het proces zijn niet duidelijk gespecificeerd.</p> <p>2) Als het om praktische redenen niet uitkomt wordt een "bypass" gemaakt en om een bestaand rustpunt of koppelvlak heen gewerkt.</p>
15	Overige processen	
	15.1 Bestaan van processen	
	15.1.1 Er is een besturingsproces.	Er is een besturingsproces. In de overgang van onderzoek naar productie werd besloten het proces en software ongewijzigd over te nemen. Men was voornemens om daar later verbeteringen in aan te brengen Dit besluit lijkt niet te passen bij de hoge eisen die aan de productie werden gesteld. Er was weinig aandacht voor de risico's van dat besluit.
	15.1.2 Er zijn verbeterprocessen.	<p>Er zijn in de afgelopen periode fouten geconstateerd. Er zijn daarna verbeteringen doorgevoerd om die fouten te corrigeren.</p> <p>Er bestond daarnaast een aparte lijst met gewenste verbeteringen van de medewerkers, een aantal verbeteringen zijn besproken met de teammanager/projectleider. Door de omstandigheden had men besloten de verbeteringen later door te voeren.</p> <p>Er is geen plan van aanpak gevonden om verbeteringen in het productieproces door te voeren. Er was geen periodiek productieoverleg om verbeteringen te bespreken en de voortgang van verbeteringen te monitoren.</p>
	15.1.3 Er zijn projectmanagementprocessen.	Ja, er was in iedere fase een projectmanagement proces. Dit geldt zowel voor de onderzoeksfases als de productiefase.
15.1.4 Er is een proces voor het afhandelen van incidenten.	Er hebben zich incidenten voorgedaan (zoals de 3 geconstateerde fouten). Deze incidenten zijn geregistreerd in de incidentendatabase van SQS. De fouten zijn gecommuniceerd en besproken in het DB. De incidenten zijn ook afgehandeld en	

Beschrijving norm			Bevinding
			<p>opgelost. Er was dus een proces om ze af te handelen. Er zijn vervolgens ook maatregelen genomen. Het gevolg daarvan is het uitvoeren van een audit en het uitvoeren van een verbeterproject.</p> <p>Het bovenstaande proces om incidenten af te handelen is uitgevoerd, maar een beschrijving daarvan is niet gevonden.</p> <p>In de Werkafspraken tussen het CBS en OCW is niet opgenomen hoe met incidenten moet worden omgegaan.</p>
15.3	Deugdelijkheid van het logistieke proces		
	15.3.1	Fouten met data-entry en/of inlezen van data, worden voorkomen of gecorrigeerd.	Er worden data ingelezen op veel plekken in het proces. Wanneer het inlezen volledig mislukt of frequentieverdelingen later in het proces niet plausibel zijn zullen de inleesfouten waarschijnlijk worden gevonden en worden gecorrigeerd. Wanneer een inleesfout niet leidt tot het mislukken van het inlezen en ook niet leidt tot implausibele frequentieverdelingen worden inleesfouten nu waarschijnlijk niet gedetecteerd. Zo'n situatie heeft zich eerder voorgedaan bij het inlezen van de status bekostiging (fout 3), wat heeft geleid tot een fout in de output. De fout met status bekostiging is gecorrigeerd met aanpassing van de software.
	15.3.2	Fouten met het selecteren en verplaatsen van bestanden, worden voorkomen of gecorrigeerd.	Geen afwijkingen geconstateerd
	15.3.3	Fouten in het publicatieproces worden voorkomen.	
16	Documentatie		
	16.1	Volledigheid van de documentatie	
	16.1.1	Het proces is beschreven.	In het kader van de VIR (Voorschrift Informatiebeveiliging Rijksdienst) zijn algemene procesbeschrijvingen gevonden, zowel voor onderwijsstatistieken

Beschrijving norm		Bevinding
		<p>(P.RegOnd) als voor beleidsstatistieken (P.BS). Beide beschrijven het proces van de onderwijsachterstandenindicator niet.</p> <p>Daarnaast is een PowerPoint presentatie gevonden waarin het proces van de onderwijsachterstandenindicator op hoofdlijnen wel wordt beschreven. Deze procesbeschrijving is niet bedoeld als procesdocumentatie.</p>
16.1.2	De informatiesystemen zijn benoemd.	De informatiesystemen voor de onderwijsachterstandenindicator worden niet benoemd.
16.1.3	De methodologie is beschreven.	De methodologie is beschreven tijdens het onderzoekstraject Er zijn 4 onderzoeksrapporten geschreven. Die rapporten beschrijven welke methoden worden gebruikt.
16.1.4	De kwaliteitsmetadata zijn beschreven.	De plausibiliteitsanalyses worden niet gespecificeerd in de Werkafspraken. Een eerdere beschrijving van de inhoud is gevonden. Er zijn R-scripts die plausibiliteitsanalyses voor OCW uitvoeren. Er is geen beschrijving gevonden van de methoden en het proces om de plausibiliteitsanalyses te produceren.
16.1.5	Er zijn handleidingen en werkinstructies.	Er zijn geen handleidingen en werkinstructies gevonden.
16.1.6	Er is systeemdokumentatie.	Er is geen aparte documentatie gevonden dat de informatiesystemen beschrijft om onderhoud van het systeem te faciliteren. De enige documentatie van de informatiesystemen is de programmacode zelf. Daarin staat beschrijvend commentaar. Er is geen apart overzicht gevonden van alle SPSS-syntaxen en R-scripts, met een beschrijving van hun functie en hun in- of output.
16.1.7	Er is een planning.	Er is een planning van het statistisch proces. Er is een tijdpad met duidelijke mijlpalen voor de productie van de cijfers.

Beschrijving norm			Bevinding
	16.1.8	Het self-assessment is toegepast.	Er is geen self-assessment toegepast op het proces voor de onderwijssachterstandenindicator.
16.2	Actualiteit van de documentatie		
	16.2.1	De documentatie heeft betrekking op het proces dat in productie is.	<ol style="list-style-type: none"> 1) Het PowerPoint document met de procesbeschrijving beschrijft op hoofdlijnen de huidige situatie maar is niet bedoeld als procesdocumentatie 2) De huidige methodologie is op hoofdlijnen zoals beschreven in de methodologische rapporten uit de onderzoeksfasen. <p>In de methodologische rapporten staat nog steeds een foutieve grens voor de verblijfsduur van de moeder. Deze grens is nog niet gecorrigeerd in die beschrijvingen. De methodologische beschrijvingen komen in de details dus niet overeen met het huidige proces.</p>
16.3	Duidelijkheid en eenduidigheid van de documentatie		
	16.3.1	Alle documenten zijn duidelijk en eenduidig.	<ol style="list-style-type: none"> 1) Procesdocumentatie ontbreekt. 2) De methodologische rapporten uit de onderzoeksfases zijn duidelijk en eenduidig.
16.4	Versie en status van de documentatie		
	16.4.1	Van alle documenten zijn de versie en de status duidelijk.	De versie en status van alle documenten zijn duidelijk.
	16.4.2	Alle documenten zijn goedgekeurd door de proceseigenaar.	Het proces van de onderwijssachterstandenindicator bestaat bijna een jaar. Van documenten over dit proces is geen formele goedkeuring gevonden van de proceseigenaar.
16.5	Overzichtelijkheid van de documentatie		
	16.5.1	Documentatie is op een geordende manier opgeslagen.	De documentatie die is gevonden is op geordende wijze opgeslagen.
16.6	Toegankelijkheid van de documentatie		
	16.6.1	Documentatie kan worden geraadpleegd door alle betrokken medewerkers.	Alle betrokken medewerkers hebben toegang tot de gevonden documentatie.

Beschrijving norm			Bevinding	
18	Conceptuele metadata			
	18.1	Beschikbaarheid van conceptuele metadata		
		18.1.1	Voor elk rustpuntbestand zijn conceptuele metadata beschikbaar.	Rustpunten zijn niet expliciet benoemd. Het analysebestand kan als zodanig gezien worden. Het GAP beschrijft het analysebestand.
	18.2	Juistheid en volledigheid van conceptuele metadata		
		18.2.1	De conceptuele metadata zijn juist.	Labels bij de categorieën van variabelen waren eerder foutief (de eerste fout). Deze fout is gecorrigeerd met aanvullende software.
		18.2.2	De conceptuele metadata zijn volledig.	Bij de inputdata van DUO zijn de metadata onvolkomen. Zie 13.2.1
	18.3	Duidelijkheid en eenduidigheid van conceptuele metadata		Bij fout 2 was het eerder bij de afronding van de verblijfsduur van de moeder niet duidelijk over en weer of er wel of niet afgerond was. Deze fout is gecorrigeerd met aanvullende software.
		18.3.1	De conceptuele metadata zijn duidelijk.	
		18.3.2	De conceptuele metadata zijn eenduidig.	
	18.4	Standaardisatie van conceptuele metadata		De labels behorende bij de codes worden niet overal zonder handmatige bewerking overgenomen uit het SSB.
18.4.1		Conceptuele metadata zijn gedocumenteerd volgens de richtlijnen van het Data Service Centrum (DSC).		
19	Kwaliteitsindicatoren			
	19.1	Relevantie en volledigheid van kwaliteitsindicatoren		
		19.1.1	Er is een adequate set van kwaliteitsindicatoren geïmplementeerd.	De belangrijkste en lastigste statistische processtappen zijn de verschillende imputatiestappen. Zij bepalen voor een belangrijk deel de betrouwbaarheid van de scores. 1) Er wordt in het productieproces gecontroleerd of al deze imputaties succesvol hebben plaatsgevonden.

Beschrijving norm			Bevinding
			<p>Ook worden de scores van de gemeenten en scholen van dit jaar vergeleken met die van voorafgaande jaren. Deze vergelijkingen zeggen ook iets over de kwaliteit van de imputaties. Over deze laatste kwaliteitsindicatoren wordt uitvoerig gerapporteerd in de plausibiliteitsanalyses voor OCW. Deze plausibiliteitsanalyses geven veel inzicht. Er is bovendien nog een heel eenvoudige kwaliteitsindicator afgeleid die met een "a" aangeeft dat de score van een school onbetrouwbaar is omdat de minder dan 41 leerlingen heeft.</p> <p>2) In het methodologisch rapport "<i>Herziening gewichtenregeling primair onderwijs, fase 3: bijschatting voor niet-ingeschreven leerlingen</i>" (augustus 2017) vinden we op pagina 15 de volgende formule:</p> $sd(\bar{S}) = \sqrt{\frac{0,27 + 2,27Q}{2n}}$ <p>Deze formule is niet perfect, maar kan gebruikt worden om per school of per gemeente apart een betere indicatie te geven van de onzekerheid in de score als gevolg van de verschillende imputaties. Een soortgelijke numerieke indicator voor de onzekerheid per school of per gemeente is dus mogelijk, maar wordt nu niet berekend.</p>
	19.2	Vergelijkbaarheid in de tijd van kwaliteitsindicatoren	
	19.2.1	Kwaliteitsindicatoren worden gevolgd in de tijd.	De vergelijkingen van scores van scholen en gemeenten wordt uitgevoerd van jaar op jaar. Opvallende grote jaarlijkse verschillen in scores worden verder uitgezocht in het productieproces.
24	Interne input		
	24.1	Conformiteit van de interne input met de afspraken	
	24.1.1	De interne input en de kwaliteit ervan, voldoet aan de afspraken.	Ja, geen afwijkingen geconstateerd.

Beschrijving norm			Bevinding
	24.1.2	De kwaliteit van de interne input kan worden beoordeeld op basis van een kwaliteitsrapportage.	<p>De verantwoordelijkheid voor de SSB componenten ligt bij de teams die de bestanden maken en in het SSB/DSC plaatsen. De gebruiker van de componenten beoordeelt de plausibiliteit van de gekoppelde gegevens. Als er twijfels zijn over de correctheid van de gebruikte component(en) wordt contact opgenomen met de verantwoordelijke team(s). In dit proces zijn tijdens de productiefase geen fouten in de gebruikte componenten geconstateerd.</p> <p>Nieuwe instanties van datasets behorende bij één databron worden niet voorzien van nieuwe (kwaliteits-)metadata. Per databron is wel de mogelijkheid om kwaliteitsmetadata op te nemen. De kwaliteit van invulling van deze kwaliteitsmetadata is variabel.</p>
29	Medewerkers		
	29.1	Beschikbaarheid van medewerkers (kwantitatief)	
	29.1.1	Er zijn voldoende medewerkers beschikbaar.	Voor het neerzetten van een goed geborgd productieproces en systeem was onvoldoende tijd ingeruimd in de Werkafspraken met OCW.
29.2	Beschikbaarheid van medewerkers (kwalitatief)		
	29.2.1	De groep van medewerkers voldoet aan de gestelde eisen.	<ol style="list-style-type: none"> 1) De projectgroep die tijdens de onderzoeksfase verantwoordelijk was voldeed aan de gestelde eisen. Wel moesten medewerkers soms taken, met name het controleren van code, vervullen zonder dat dit besproken is bij de start van het project, zonder dat dit tot hun normale takenpakket gerekend zou kunnen worden en zonder dat hun duidelijk is aan welke eisen dit moet voldoen. 2) Voor het in productie nemen van de onderzoeksresultaten waren de daarbij betrokken medewerkers onvoldoende voorbereid op wat er van hen verwacht werd. Zij geven ook aan niet op de hoogte te zijn gebracht van geldende richtlijnen.

Beschrijving norm			Bevinding	
	29.3	Verantwoordelijkheden en bevoegdheden van medewerkers		
	29.3.1	Rollen, verantwoordelijkheden en bevoegdheden zijn vastgelegd.	Is niet aangetroffen.	
30	Kennis			
	30.1	Beschikbaarheid van kennis bij medewerkers		
		30.1.1	Er wordt structureel aandacht besteed aan het onderhouden van relevante kennis.	De medewerkers zijn voorbereid doordat zij al tijdens de ontwikkeling deel uitmaakten van het team. Daar is veel aandacht voor kennisoverdracht geweest. Ook wordt er structureel aandacht besteed aan het opbouwen en onderhouden van R-expertise.
		30.1.2	Er is voldoende kennis aanwezig over de hardheid van de input.	1) SSB data: ja 2) DUO-data: onbekend. Een deel van de specificaties wordt niet meegeleverd; er kunnen dus onverwachte of onbekende codes voorkomen. Zie ook 13.2.1.
	30.1.4	Er is voldoende deskundigheid beschikbaar om fouten in het statistisch proces te verkleinen.	De SDI-richtlijnen en OTAP werkwijze waren niet bij alle leden van het projectteam bekend, ook niet bij codecontroles.	
	30.2	Spreiding van kennis over medewerkers		
30.2.1		Kennis van de materie, het proces en de systemen is voldoende geborgd.	Er wordt gewerkt aan spreiding van de kennis over het productiesysteem door een andere medewerker vertrouwd te maken met het systeem.	
31	Samenwerking met kenniscentra			
	31.1	Bestaan en effectiviteit van samenwerking met kenniscentra		
		31.1.1	Er wordt voldoende gebruik gemaakt van kennis bij andere kenniscentra.	Tijdens de onderzoeksfase is effectief samengewerkt met kenniscentra (CBS afdeling methodologie en begeleidingscommissie). Bij de ontwikkeling van het productieproces is onvoldoende gebruik gemaakt van interne expertise.

Tabel 2: Bevindingen en aanbevelingen voor norm 32 uit de Kwaliteitsrichtlijnen-2018.

	Beschrijving norm		Bevinding	Aanbeveling	
32	Informatiesystemen				
	32.2	Toegankelijkheid van informatiesystemen			
		32.2.1	De informatiesystemen zijn alleen voor geautoriseerde medewerkers toegankelijk.	De software wordt interactief uitgevoerd onder het gebruikersaccount van de medewerkers die betrokken zijn in het statistisch proces. De programmatuur is niet beschikbaar voor andere gebruikers. De scripts staan in een productiemap onder de standaard Varonis mappenstructuur waardoor alleen geautoriseerde medewerkers van het team hier toegang toe hebben. De algemene maatregelen m.b.t. logische toegangsbeveiliging van het CBS (conform BIR) zijn hier van toepassing.	
	32.3	Integriteit van informatiesystemen			
		32.3.1	De programmatuur is correct.	Er zijn geen onopgeloste fouten bekend. Enkele eerder ontdekte fouten zijn aangepast middels een apart script, met aanvullende software waarmee een correctie wordt uitgevoerd. Zie 002-190103 Operationalisering.R, regel 290.	Los de fout structureel op zodat de aparte correctiestap kan komen te vervallen.
32.3.2		De broncode is afgeschermd voor gebruikers.	Er is geen scheiding tussen ontwikkelaars en gebruikers, de in het productieproces werkzame statistici passen zelf de scripts aan.		

Beschrijving norm			Bevinding	Aanbeveling
			De software staat op beveiligde netwerkmappen (toegang beperkt middels de Varonis systematiek) waardoor de toegang beperkt is.	
	32.3.3	Data worden alleen via de programmatuur gewijzigd.	Ja, geen overtreding aangetroffen.	
	32.3.4	Er wordt versiebeheer toegepast op de programmatuur.	Beperkt. Er worden versienummers en datums in bestandsnamen gehanteerd, maar geen geautomatiseerd versiebeheersysteem waarin wijzigingen traceerbaar zijn.	Zie bevinding 32.5.3-R.V.1.
	32.3.5	Er wordt versiebeheer toegepast op de databestanden.	Alle bestanden binnen het SSB hebben een versie-aanduiding.	
	32.3.6	Er is sprake van een overzichtelijke mappenstructuur voor de gegevensbestanden.	Ja, de scripts maken gebruik van de gestandaardiseerde CBS mappenstructuur, waarbij data per deelproces en daarbinnen naar Input, Werk (throughput) en Output is georganiseerd. Er zijn geen afwijkingen van deze structuur aangetroffen.	
32.5	Onderhoudbaarheid van informatiesystemen			
	32.5.1	Er wordt gebruik gemaakt van standaard tools.	Ja, R, SPSS en SSB-tooling.	
	32.5.2	Informatiesystemen zijn modulair opgebouwd.	Ja, voor stappen in het proces worden aparte scripts gebruikt.	

Beschrijving norm		Bevinding	Aanbeveling
32.5.3	Informatiesystemen zijn volgens de voorschriften gebouwd.	Er zijn enkele afwijkingen aangetroffen, zie de gedetailleerde bevindingen bij normen S.* (voor SPSS) en R.* (voor R).	<p>Communiceer duidelijk waar de richtlijnen staan en dat ze verplicht zijn;</p> <p>Pas richtlijnen zo veel mogelijk geautomatiseerd toe en biedt hiervoor de middelen aan, de tidyverse style guide wordt door RStudio ondersteund en code kan automatisch aan deze stijl worden aangepast met het lintr package;</p> <p>Controleer regelmatig, zowel door self assessment als onafhankelijk, dat de richtlijnen worden toegepast;</p> <p>Verwerk de richtlijnen in de bestaande software (zie onderstaande detailaanbevelingen).</p>
32.5.3-S.R.1	Gebruik de syntax voor de reproduceerbaarheid van je resultaten.	Ja, de syntax is bruikbaar om te reproduceren.	
32.5.3-S.R.3	Breng nooit veranderingen direct in de Data Editor aan, altijd via de syntax.	Ja, geen afwijkingen geconstateerd.	
32.5.3-S.R.4	Elke nieuwe versie van een SAV-file moet je later kunnen reproduceren vanuit de	Ja, geen afwijkingen geconstateerd.	

Beschrijving norm		Bevinding	Aanbeveling
	originele SAV-file m.b.v. een syntaxjob.		
32.5.3-S.R.5	Zet alle datatransformaties en labeling in de syntax-file. Staan alle transformaties in de syntax? Dus: is de output een exacte afspiegeling van de syntax?	Ja, geen afwijkingen geconstateerd.	
32.5.3-S.R.6	Wordt een bestaand sav-bestand niet overschreven?	Alle aangetroffen "output save outfile" commando's schrijven hun output naar nieuwe bestanden.	
32.5.3-S.L.1	Heldere opbouw en naamgeving van data-, syntax- en outputbestanden.	<p>Over het algemeen is de programmatuur goed leesbaar. De code is opgedeeld naar deelstappen in de totale bewerking en is van boven naar beneden te lezen. In de code worden substappen met commentaar duidelijk aangegeven.</p> <p>Echter, het feit dat letterlijke teksten en codelijsten in de code zijn opgenomen in plaats van constanten of apart gedefinieerde tabellen is van negatieve invloed op de leesbaarheid; de logica van de programmatuur wordt daardoor deels verhuld door inhoudelijke data die door de programmatuur wordt verwerkt.</p> <p>Zie bijvoorbeeld 006-190111 Imputeren onderwijsscores.R, regel 128.</p>	Gebruik zo veel mogelijk (standaard of maatwerk) samengestelde functies in plaats van losse statements die inhoudelijk aan elkaar verbonden zijn. Indien het niet mogelijk is om stukken code die altijd tegelijk aangepast moeten worden in één statement neer te zetten, zet ze dan lay-out technisch zo dicht mogelijk bij elkaar. Bijvoorbeeld het volgende codefragment...

Beschrijving norm			Bevinding	Aanbeveling
			<p>Daarnaast is de code binnen scripts niet verder opgedeeld naar functies, waardoor statements die visueel ver van elkaar af staan elkaar kunnen beïnvloeden.</p> <p>Zie bijvoorbeeld variabele werk in 006-190111 Imputeren onderwijsscores.R, regels 75 en 749.</p>	<pre>VariabeleX <- ifelse(VariabeleY >= -1,1, ifelse(VariabeleY >= -5,2, ifelse(VariabeleY >= -10,3, ifelse(VariabeleY >= 15,4,5))))</pre> <pre>VariabeleX <- factor(VariabeleX, levels = c(1,2,3,4,5), labels = c("<1jr", "1-5jr", "5-10jr", "10-15jr", ">15"))</pre> <p>...kan omgeschreven worden met behulp van een functie die de aanroep van de factor-functie uit het eerste codefragment en het toevoegen van labels aan de verschillende levels uit hetzelfde codefragment samen neemt. Dat resulteert in het volgende codefragment:</p> <pre>dataset[, variabeleY := cut(variabeleY, breaks = c(0, 5, 15, Inf), labels = c("0-5jr", "5-15jr", ">15"), include.lowest = TRUE)]</pre> <p>Het voordeel van de laatste schrijfwijze is dat de veranderlijke aspecten van de programmacode die inhoudelijk onlosmakelijk aan elkaar verbonden zijn en altijd tegelijk aangepast</p>

Beschrijving norm		Bevinding	Aanbeveling
			moeten worden, te weten de labels en de bijbehorende variabelewaarden, visueel dicht bij elkaar staan. Dit ondersteunt zowel een review van de code als het aanpassen ervan
32.5.3-S.L.2	Begin de syntax met het vermelden van je naam, datum (op maandniveau is ok), naam van setup (is gemakkelijk als je deze uitprint) en project.	Alle scripts binnen het rekenproces voldoen aan deze norm. In de scripts voor het bouwproces ontbreekt de bestandsnaam, zie bijvoorbeeld 03.TOTALEPOPULATIE.sps.	Pas de broncode volgens de genoemde richtlijn aan.
32.5.3-S.L.3	Geef zoveel mogelijk labels aan variables en values. Gebruik duidelijke labels. Lange labels van variabelen en hun waarden maken output onoverzichtelijk.	Ja, geen afwijkingen geconstateerd.	
32.5.3-S.L.5	Commentaar: is er voldoende commentaar in de syntax toegevoegd en is deze duidelijk?	De scripts bevatten structureel commentaar. Echter, vanuit de workshop is gebleken dat er behoefte bestaat aan beter commentaar.	Voeg conform wat in de workshop besproken is commentaar toe waar dat nodig is.
32.5.3-S.L.7	We streven ernaar steeds als er een nieuw bestand wordt geopend eerst het commando 'DISPLAY labels' te geven. Dit geeft een text-output waarin alle	Gebruik van DISPLAY labels niet aangetroffen.	Pas de broncode volgens de genoemde richtlijn aan.

Beschrijving norm		Bevinding	Aanbeveling
	variabelen en hun labels van het geopende bestand worden opgesomd.		
32.5.3-S.A.1	Houd bij het schrijven van een syntax in je achterhoofd dat deze zo gemakkelijk mogelijk aan te passen is in het geval van bijvoorbeeld een herhalingsproject. Is waar mogelijk gebruik gemaakt parametriseren?	Nee, de scripts bevatten gecodeerde waarden voor bijvoorbeeld jaren.	Zie bevinding 32.5.4.
32.5.3-S.I.2	Maak bij het hercoderen van een variabele altijd eerst een nieuwe variabele aan (d.m.v. compute of recode ... into ...). Laat de originele variabelen intact.	Enkele schendingen aangetroffen. Voorbeeld: 01.DEELPOP1BO.sps, regel 48.	Pas de broncode volgens de genoemde richtlijn aan.
32.5.3-S.I.3	Maak keuzes voor populatieafbakeningen en hercoderingen stapsgewijs en expliciet. Als bijvoorbeeld de doelpopulatie is personen van 15-64 jaar in gemeenten met meer	Afwijkingen aangetroffen. Selecties worden soms samengenomen. Zie bijvoorbeeld 01.DEELPOP1BO.sps, regel 84.	Pas de broncode volgens de genoemde richtlijn aan.

Beschrijving norm		Bevinding	Aanbeveling
	dan 100 000 inwoners met bijstand op 1 januari 2004, maak dan drie indicatoren aan voor leeftijd, gemeentegrootte en bijstand, en selecteer vervolgens de populatie met behulp van deze drie indicatoren.		
32.5.3-S.I.5	Missings: als een variabele missings heeft, dan wil dit bij een datatransformatie of een aggregatie nog wel eens mis gaan. Daarom is de regel: altijd missings direct hercoderen naar een 8- of 9-score.	Afwijkingen aangetroffen. Op sommige plaatsen worden missings geconstateerd en geteld, maar niet omgecodeerd. Voorbeeld: 01.DEELPOP1BO.sps, regel 270.	Pas de broncode volgens de genoemde richtlijn aan.
32.5.3-S.I.6	Zoek in de output naar 'error' en 'warning'. Open ook de gedeelten van de output die niet geheel zichtbaar zijn in de spv-file. Let ook op het gebruik van punten: ontbreken er geen punten, zodat een deel	Dit is over het algemeen gebeurd en is ook gebruikelijk, soms is door tijdsdruk deze analyse overgeslagen.	Introduceer middelen om (geautomatiseerd) efficiënt de logging op errors en warnings te analyseren.

Beschrijving norm		Bevinding	Aanbeveling
	van de syntax onbedoeld niet wordt uitgevoerd?		
32.5.3-S.I.7	Zijn alle variabelen en hun waarden voorzien van labels? Zijn de labels duidelijk? Er zijn uitzonderingen mogelijk op deze regel (bv. 0-1 variabelen waar alleen de waarde 1 een label heeft).	In 04.ACHTERSTANDBO2018MMDDANAVV.SPS krijgen alle variabelen en waarden meteen een label.	
32.5.3-S.I.8	Worden alle datatransformaties zoals een recode of (DO) IF gecontroleerd met een frequency en/of crossing vooraf én achteraf?	De code bevat zeer systematische frequentietellingen bij recodes of do if statements.	
32.5.3-S.I.10	Controleer bij een 'DO IF' statement of alle voorkomende gevallen behandeld worden.	Deze controle vindt niet structureel plaats en deze richtlijn is niet bij de geïnterviewde ontwikkelaars bekend.	Voeg deze controle als standaardonderdeel toe aan het ontwikkel- en reviewproces.
32.5.3-S.I.13	Controleer naar aanleiding van het eerste blok commentaar in de setup (met doel en samenvatting van setup) of dit overeenkomt met het	Er is geen stappenplan / syntaxplan / stroomschema gemaakt waartegen deze controle plaats had kunnen vinden.	Maak en onderhoud een stappenplan / syntaxplan / stroomschema.

Beschrijving norm		Bevinding	Aanbeveling
		stappenplan/stroomschema en met de productspecificatie in de offerte. Dit gedeelte van de setup wordt in de projectgroep besproken.	
	32.5.3-S.I.14	Aggregate: bij dit commando moet voor iedere variabele die je ook in het nieuwe bestand opgenomen wilt hebben, aangegeven worden welke waarde je wilt kiezen. Dit kan zijn de som, het minimum, maximum, gemiddelde, de waarde uit eerste of laatste record. Bedenk van tevoren goed welke scores een variabele aan kan nemen en of er inderdaad gebeurt wat je zou willen. Extra aandacht verdienen bijvoorbeeld datumvariabelen met scores voor 'niet van toepassing'.	Alle aangetroffen Aggregates zijn adequaat gedocumenteerd en gestructureerd.

Beschrijving norm		Bevinding	Aanbeveling
32.5.3-S.I.15	Controleer bij SELECT- en FILTER-commando's of de juiste deelpopulatie wordt geselecteerd.	SELECT-commando's worden vergezeld van frequentietellingen of crosstabs. FILTER-commando's zijn niet aangetroffen.	
32.5.3-R.C.1	We adviseren de tidyverse style guide te gebruiken	Nee, deze guide is niet toegepast.	Pas de style guide automatisch toe door het desbetreffende package te gebruiken.
32.5.3-R.L.1	Productiescripts dienen een log aan te maken, zodanig dat uit de logbestanden is te reconstrueren hoe en wanneer de eindresultaten tot stand zijn gekomen.	Er wordt logging weggeschreven vanuit alle scripts.	
32.5.3-R.L.2	Het gebruik van het package logging is verplicht.	Nee, dit package wordt niet toegepast.	Overweeg of gebruik van dit package meerwaarde heeft.
32.5.3-R.I.1	Bepaal wat voor jouw systeem de belangrijkste punten zijn waar (in interactie van een applicatie met andere systemen) fouten op kunnen treden, en documenteer dit.	Nee, een documentatie van de bedoelde risico-analyse dan wel een testplan is niet uitgevoerd en het proces bevat meerdere punten waarop data in verschillende formaten tussen twee software-componenten worden uitgewisseld waarbij inleesfouten op zouden kunnen treden.	<u>Gebruik bij voorkeur inputs met formele, afgedwongen technische en inhoudelijke meta.</u> Een database kan bijvoorbeeld unieke waarden, het niet-leeg zijn van velden en referentiële integriteit (met bijvoorbeeld een codelijst) afdwingen en bevat geen onzekerheden wat betreft bestandsformaten en –opmaak. Indien dit niet mogelijk is, voeg dan expliciete controles aan de code toe op de integriteit van de input.

Beschrijving norm		Bevinding	Aanbeveling	
	32.5.3-R.T.1	De automatische testen staan in de tests directory en worden gespecificeerd en uitgevoerd met behulp van het testthat package.	Nee, er zijn geen automatische testen ontwikkeld.	Gebruik regressietests, minimaal voor risicovolle onderdelen van de programmatuur.
	32.5.3-R.T.2	Test in elk geval de functionaliteit met het hoogste risico voor je applicatie. Maak daarvoor een inschatting eventueel in samenspraak met de afnemer/gebruiker. Houd hierbij rekening met de impact en bij een fout en de kans op een fout. Documenteer deze inschatting.	Nee, er is geen documentatie van deze inschatting gemaakt. Door tijdsdruk zijn minder controles en tests uitgevoerd.	Bespaar bij tijdsdruk op scope, niet op de ontwikkeling van noodzakelijke tests. Zie verder bevinding 32.5.3-R.T.1.
	32.5.3-R.T.3	Laat je code tussentijds reviewen door collega's of R-kenners.	Ja, code reviews hebben plaatsgevonden.	
	32.5.3-R.S.1	Gegevens worden van R naar andere software doorgegeven via een (tijdelijk) bestand in Nederlands csv format	Deels, de input voor de R-programmatuur is een SPSS-bestand.	Zie aanbeveling bij 32.5.3-R.I.1. Merk op dat in de tekst bij de richtlijn alleen gesproken wordt over uitwisseling vanuit R met andere systemen, maar de richtlijn zou, gezien

Beschrijving norm		Bevinding	Aanbeveling
	(dat wil zeggen puntkomma scheiding en een komma als decimaal scheidingsteken).		het aspect dat het afdekt (interactie met andere systemen), ook vice versa moeten werken. Aanbeveling is de richtlijn in die zin aan te scherpen.
32.5.3-R.V.1	Tenzij er redenen zijn om dit niet te doen wordt de broncode beheerd met GIT	Nee, er wordt geen versiebeheersysteem voor de broncode gebruikt.	Gebruik tooling voor versiebeheer.
32.5.4	Informatiesystemen zijn instelbaar door gebruikers.	Nee, veranderlijke onderdelen van de programmatuur zijn in de code zelf opgenomen. Bijvoorbeeld: 002-190103 Operationalisering.R, regels 74, 314, 350 of 001-181218 Imputeren in SPSS.sps, regels 54, 410, 497.	<u>Gebruik constanten in plaats van hard gecodeerde waarden</u> , zodat wat hetzelfde moet zijn gegarandeerd hetzelfde is, code beter onderhoudbaar en leesbaar is doordat constanten dicht bij elkaar en op een duidelijke plek hun waarde kunnen krijgen en scripts beter herbruikbaar zijn

Tabel 3: De relatie tussen de 3 fouten en de meest direct corresponderende normen in de Kwaliteitsrichtlijnen-2018.

Fout	Richtlijnen in de Kwaliteitsrichtlijnen-2018		
Fout 1	12.8.4 12.8.6 18.2.1 32.3.1	32.5.3 32.5.3.S.L.1 32.5.3-S.L.5 32.5.3.S.A.1	32.5.3-R.C.1 32.5.3-R.T.1 32.5.3-R.T.2 32.5.4
Fout 2	12.8.6 13.2.1	13.2.2 18.3	18.3.1
Fout 3	12.8.2 12.8.6 13.2.1 13.2.2	15.3.1 32.3.1 32.5.3 32.5.3-S.I.6	32.5.3-R.L.1 32.5.3-R.I.1 32.5.3.R.T.1 32.5.3.R.T.2