



Discussion Paper

# Ontwikkeling van een duurmodel voor onverwacht lange opnameduren

Jan van der Laan  
Corine Witvliet-Penning  
Agnes de Bruin

**Juni 2020**

# Inhoudsopgave

<b>1</b>	<b>Inleiding en samenvatting</b>	<b>3</b>
<b>2</b>	<b>Methoden</b>	<b>4</b>
2.1	Brongegevens	4
2.2	Variabelen in het OLO-model	4
2.3	Duurmodellen	6
2.4	Berekenen OLO-fracties	8
<b>3</b>	<b>Evaluatie van de OLO-indicator</b>	<b>8</b>
<b>4</b>	<b>Conclusie</b>	<b>14</b>
<b>A</b>	<b>Zwaarteklasse voor onverwacht lange opnameduren</b>	<b>17</b>
<b>B</b>	<b>Bepaling gebruikte duurmodel</b>	<b>20</b>

# 1 Inleiding en samenvatting

De huidige indicator Onverwacht Lange Opnameduur (OLO) geeft het percentage klinisch opgenomen patiënten met een opnameduur die minimaal 50% langer is dan verwacht. De Inspectie Gezondheidszorg en Jeugd (IGJ) gebruikt deze indicator voor hun toezicht op de kwaliteit en veiligheid van zorg in ziekenhuizen. Daarnaast gebruiken ziekenhuizen een afwijkende OLO score als signaal om verder onderzoek te doen naar verbeterpunten binnen de kwaliteit van zorg. De indicator wordt berekend op basis van de Landelijke Basisregistratie Ziekenhuiszorg (LBZ). De huidige berekeningswijze corrigeert voor hoofddiagnose, hoofdverrichting en leeftijd. Met de huidige indicator bestaan er verschillen tussen typen ziekenhuizen (UMC's, topklinische en algemene ziekenhuizen) in het percentage onverwacht lange opnamen, die mogelijk verklaard worden door onvoldoende case-mix correctie. Omdat ziekenhuizen kunnen verschillen in hun patiëntenpopulaties en daarmee in de zwaarte en duur van de opnamen, is het belangrijk om zoveel mogelijk voor deze verschillen te corrigeren. DHD heeft het CBS daarom gevraagd om een nieuw model voor deze indicator te ontwikkelen.

Het CBS gebruikt hiervoor een vergelijkbare methode als bij de berekening van de sterfte-indicator (HSMR) (van der Laan et al., 2019b) en de heropname-indicator (van der Laan et al., 2019a): het ontwikkelen van een regressiemodel dat gegeven de kenmerken van de opname de uitkomst (sterfte, heropname, of opnameduur) voorspelt. De voorspelde waarden kunnen vervolgens vergeleken worden met de waargenomen waarden. Een belangrijk verschil met de andere twee indicatoren is echter dat het hier om duren gaat terwijl het bij de andere indicatoren om binaire uitkomsten gaat (wel/niet overleden, wel/geen heropname). Daarom moet in plaats van logistische regressie gebruik gemaakt worden van duurmodellen.

Hoofdstuk 2 beschrijft de nieuwe berekeningsmethode waarmee de onverwacht lange opnamen (OLO) bepaald worden. Hierbij wordt voor een aantal details verwezen naar de bijlages bij dit rapport. In hoofdstuk 3 wordt de OLO-indicator geëvalueerd. Hierbij wordt de nieuwe indicator vergeleken met de uitkomsten op basis van een aantal andere eenvoudigere berekeningswijzen (waaronder de huidige methode) en wordt er gekeken in welke mate ziekenhuizen als significant afwijkend worden geclassificeerd.

Uit de evaluatie blijkt dat het merendeel van de ziekenhuizen een OLO-percentages heeft dat significant afwijkt van het landelijk gemiddelde. Zelfs bij een betrouwbaarheid van 99,8% is dat nog steeds ruim meer dan 50%. Daarnaast worden ongeacht de methode grotendeels dezelfde opnamen als OLO geclassificeerd. Met een relatief eenvoudige standaardisatie komen er nagenoeg dezelfde uitkomsten uit. Ten derde blijkt er een verband zichtbaar te zijn tussen de grootte van het ziekenhuis en het OLO-percentages en tussen het type ziekenhuis (UMC's, topklinische en algemene ziekenhuizen) en het OLO-percentages. Uit deze bevindingen kunnen twee conclusies getrokken worden:

1. Er zijn grote verschillen tussen de ziekenhuizen wat betreft hun OLO-percentages. Deze verschillen hangen ook samen met het type ziekenhuis. Een mogelijke oorzaak zou kunnen zijn dat de patiëntkenmerken uit de LBZ die meegenomen zijn in het model onvoldoende corrigeren voor verschillen in de patiëntenpopulaties. Andere factoren die wellicht meespelen zijn bijvoorbeeld de aard, inrichting of efficiëntie van de zorg. Voor een goede interpretatie van de OLO-uitkomsten zou nader onderzocht moeten worden waardoor de grote verschillen tussen de ziekenhuizen worden veroorzaakt.

2. Op dit moment voegt een complexe standaardisatie met een duurmodel weinig toe aan een eenvoudigere standaardisatie volgens de huidige methode of een naar diagnosegroep gestandaardiseerd model.

## 2 Methoden

### 2.1 Brongegevens

Tijdens het onderzoek is gewerkt met twee datasets. Eén voor de bepaling van de zwaarteklassen (zie bijlage A) en één voor de bepaling van de modellen voor de standaardisatie. De zwaarteklasse-indeling is gebaseerd op een dataset bestaande uit klinische opnamen uit de LBZ-jaren 2014 t/m 2016, voor de modelberekeningen is een dataset gebruikt die bestaat uit klinische opnamen uit de LBZ-jaren 2017 en 2018. Beide datasets zijn op dezelfde wijze samengesteld. De volgende opnamen zijn uit de datasets verwijderd:

- Dagopnamen en langdurige observaties zonder overnachting
- Klinische opnamen van niet-ingezetenen
- Klinische opnamen waarbij de patiënt is overleden (op basis van de variabele 'bestemming')
- Klinische opnamen die volledig bestonden uit een gastverblijf (op basis van hoofddiagnosen met de ICD-10 codes Z76.2, Z76.3 en Z76.4)
- Klinische eendaagse opnamen (zowel acuut als niet acuut) met bestemming eigen woonomgeving, instelling voor revalidatie of instelling voor verpleging/verzorging

Aanvullend is aan de hand van de nevenverrichtingen bepaald of er tijdens de opnamen sprake was van zogenaamde gastverblijfdagen. Deze dagen zijn op de opnameduur in mindering gebracht. Op basis van deze gecorrigeerde opnameduur zijn de volgende opnamen verwijderd:

- Klinische opnamen die volledig bestonden uit een gastverblijf
- Klinische opnamen met een gecorrigeerde opnameduur van 100 dagen of langer

Aan de dataset is tot slot een variabele toegevoegd die aangeeft of er sprake was van een operatieve, diagnostische of andersoortige hoofdverrichting. De overige covariaten zijn gelijk aan de covariaten die in het HSMR- en Heropnamenmodel zijn gebruikt, met uitzondering van de zwaarteklasse (zie 2.2).

### 2.2 Variabelen in het OLO-model

In het model worden alle covariaten meegenomen die ook in het model voor de heropnameratio (van der Laan et al., 2019a) en voor de HSMR (van der Laan et al., 2019b) gebruikt worden. Alleen wordt, zoals verderop besproken wordt, de zwaarteklasse anders bepaald. Verder is er een variabele 'verrichtingsgroep' toegevoegd die aangeeft wat voor type verrichting er tijdens de opname is uitgevoerd. De variabelen zijn in tabel 2.1 gegeven. Voor meer detail wordt verwezen naar de rapporten van de HSMR en de heropnameratio's.

**Tabel 2.1 Variabelen in het OLO-model**

Covariaat	Waarden
Geslacht	m, v
Herkomst van de patiënt	thuis, verpleging/verzorging, ander ziekenhuis
Jaar van ontslag	2017, 2018
Leeftijd	0-1, 1-5, 5-10, 10-15, .... , 90-95, 95+ jaar
Maand van opname	jan-feb, mrt-apr, mei-jun, jul-aug, sep-okt, nov-dec
Sociaaleconomische status	laagst, onder het gemiddelde, gemiddeld, boven het gemiddelde, hoogst
Urgentie	acuut, niet-acuut
Charlson comorbiditeiten 1 t/m 17:	
1. Acuut myocardinfarct	niet aanwezig, aanwezig
2. Hartfalen	niet aanwezig, aanwezig
3. Perifere vaatziekte	niet aanwezig, aanwezig
4. Cerebrovasculaire ziekten	niet aanwezig, aanwezig
5. Dementie	niet aanwezig, aanwezig
6. Chronische pulmonale aandoeningen	niet aanwezig, aanwezig
7. Bindweefselaandoening	niet aanwezig, aanwezig
8. Maagzweer	niet aanwezig, aanwezig
9. Leveraandoening	niet aanwezig, aanwezig
10. Diabetes	niet aanwezig, aanwezig
11. Diabetescomplicaties	niet aanwezig, aanwezig
12. Paraplegie en andere verlammingen	niet aanwezig, aanwezig
13. Nieraandoening	niet aanwezig, aanwezig
14. Kanker	niet aanwezig, aanwezig
15. HIV	niet aanwezig, aanwezig
16. Metastasen	niet aanwezig, aanwezig
17. Ernstige leveraandoening	niet aanwezig, aanwezig
Verrichtingsgroep*	operatief, diagnostisch, onbekend
Zwaarteklasse*	0-1, 2, 3, 4, 5-6, 7-9, 10-13, 14-16, 17-20, 21-27, 28-41, 42+ dagen, en 'other/onbekend'

\* Extra of afwijkende covariaten ten opzichte van het HSMR-model. In het HSMR-model is de verrichtingsgroep niet opgenomen. De hier beschreven covariaat 'zwaarteklasse' is gebaseerd op historische opnameduren per ICD10-diagnose, terwijl deze covariaat in het HSMR-model is gebaseerd op historische sterfte.

### 2.2.1 Zwaarteklasse

De modellen worden geschat per hoofddiagnosegroep. Binnen een diagnosegroep kan er echter nog variatie zitten in het type opnamen. De zwaarteklasse heeft als doel onderscheid te maken tussen de verschillen in zwaarte binnen een diagnosegroep. Voor de HSMR en de heropnameratio wordt gebruik gemaakt van een zwaarteklassenindeling gebaseerd op sterfte. Het is echter niet waarschijnlijk dat de kans op sterfte heel veel zegt over de verwachte opnameduur. Daarom is er voor dit onderzoek voor gekozen om een aparte zwaarteklasse af te leiden voor opnameduur. In bijlage A worden verschillende opties hiervoor onderzocht. Uiteindelijk is gekozen om per ICD10-code het 85<sup>ste</sup> percentiel te bepalen van de opnameduren van alle opnamen met die ICD10-code als hoofddiagnose. Dit percentiel is vervolgens ingedeeld in de klassen zoals in tabel 2.1 weergegeven. ICD10-codes die minder dan 20 keer voorkomen in de dataset of die door minder dan 5 ziekenhuizen gebruikt worden krijgen de code 'other/onbekend'.

## 2.3 Duurmodellen

Een veelgebruikte manier om een duurverdeling weer te geven is de overlevingsfunctie (survivalfunctie),  $S(t)$ . Deze geeft de fractie personen weer waarvan de opname nog niet is afgelopen als functie van duur  $t$ . De overlevingsfunctie begint bij 1: iedereen heeft een duur van minimaal 0 dagen; vervolgens neemt de overlevingsfunctie af naar 0. In ons geval zijn de duren langer dan 100 dagen verwijderd, dus bij 100 is de overlevingsfunctie 0: niemand heeft een duur langer dan 100 dagen. Uit de overlevingsfunctie kunnen heel makkelijk allerlei statistieken over de duren bepaald worden. Zo is de gemiddelde duur de oppervlakte onder  $S(t)$ ; is de mediaan de duur  $t$  waarop  $S(t)$  gelijk is aan 0.5 en het 80<sup>ste</sup> percentiel het punt  $t$  waarop  $S(t)$  gelijk is aan 0.2.

De modellen die we gaan gebruiken zijn parametrische modellen van het type *accelerated failure time models*. In deze modellen wordt uitgegaan van een parametrische vorm  $S_0(t)$  van de overlevingsfunctie (meestal beschreven door enkele parameters die ook uit de data geschat worden). Afhankelijk van de kenmerken van de patiënt en opname loopt de klok langzamer of sneller dan de referentiegroep.

Dus van een patiënt  $i$  met kenmerken  $\mathbf{X}_i$  wordt de overlevingsfunctie gegeven door:

$$S(t|\mathbf{X}_i) = S_0(\theta(\mathbf{X}_i)t), \quad (1)$$

waarbij  $\theta(\mathbf{X}_i)$  een functie van  $\mathbf{X}_i$  is. Als  $\theta(\mathbf{X}_i)$  groter dan 1 is loopt de klok sneller voor deze persoon (de overlevingsfunctie wordt ingeperst); is  $\theta(\mathbf{X}_i)$  kleiner dan 1 (maar groter dan 0) dan loopt de klok langzamer (de overlevingsfunctie wordt uitgerekt). Voor  $\theta(\mathbf{X}_i)$  wordt de volgende functie gebruikt:

$$\theta(\mathbf{X}_i) = \exp(-(\beta_1 X_{i1} + \beta_2 X_{i2} + \dots)). \quad (2)$$

Dus voor de referentiepersoon (alle  $X_{ij}$  gelijk aan 0) is  $\theta X_i$  gelijk aan 1.

Aangezien het model afhankelijk van de kenmerken van de patiënt de klok sneller of langzamer laat lopen, is het ook eenvoudig om verwachte kwantielen uit te rekenen voor patiënten. Stel het kwantiel  $p$  waarin we geïnteresseerd zijn is voor de referentiepersoon  $q_0^{(p)}$  (dus  $S_0(q_0^{(p)}) = p$ ). Dan is te verwachten dat van patiënten met kenmerk  $\mathbf{X}$  een fractie  $p$  een duur korter heeft dan:

$$q^{(p)}(\mathbf{X}) = \frac{q_0^{(p)}}{\theta(\mathbf{X})} \quad (3)$$

**Tabel 2.2 Belang van de variabelen in de duurmodellen.**

Variabele	Deviance	DF	#Sig*
Zwaarteklasse	243647	611	135
Verrichtingsgroep	203524	298	157
Leeftijd	179102	2398	150
Urgentie	169558	156	139
Herkomst	50837	259	137
Geslacht	24772	150	94
Charlson 2	17270	114	113
Charlson 13	8300	142	118
Charlson 14	7247	137	109
Charlson 3	5573	135	106
Charlson 10	5220	153	109
Charlson 5	4577	101	70
Charlson 6	4270	150	96
Charlson 4	4192	106	85
Charlson 16	4023	128	75
Charlson 12	3988	71	53
Charlson 9	3865	112	92
Charlson 1	2991	142	79
Opnamemaand	2321	785	53
SES	2306	628	63
Charlson 11	2184	98	64
Ontslagjaar	1627	157	59
Charlson 7	1076	104	56
Charlson 17	540	28	19
Charlson 8	163	8	6
Charlson 15	29	3	3

\* Aantal van de 157 modellen waarin de variabele significant (95% betrouwbaarheid) bijdraagt.

In bijlage B wordt gekeken naar een aantal mogelijke parametrische vormen voor  $S_0(t)$ . Ter referentie wordt daar ook gekeken naar het Cox-model voor duren wat een flexibeler model is, maar minder geschikt voor de huidige toepassing omdat het model moeilijker te delen is. Uiteindelijk bleek het log-normale model het beste te passen bij de gegevens voor de meeste diagnosegroepen. Daarom wordt er verder gewerkt met het log-normale model. Dit model heeft de volgende vorm voor de overlevingsfunctie:

$$S_0(t) = \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\frac{\ln t - \mu}{\sqrt{2}\sigma}\right), \quad (4)$$

waarin  $\operatorname{erf}(x)$  de zogenaamde errorfunctie is.

Tabel 2.2 laat het belang van de verschillende variabelen in de geschatte duurmodellen zien. Voor ieder van de 157 modellen is voor iedere variabele uitgerekend wat de bijdrage aan de deviance (maat voor verklaarde variantie) is, het aantal vrijheidsgraden dat de variabele bijdraagt en of de variabele significant (95% betrouwbaarheid) is in het model. Dit is gesommeerd over alle 157 modellen. De totale deviance geeft een indicatie van het belang van een variabele. Een variabele met meer vrijheidsgraden zal ook meer bijdragen, dus de deviance gedeeld door het aantal vrijheidsgraden is een andere indicator voor het belang (niet opgenomen in de tabel). Hoe vaak een variabele significant is in het model zegt ook iets over het belang van de variabele. Naast leeftijd (waarvan het belang gedeeltelijk veroorzaakt wordt door het grote aantal categorieën) zijn vooral de zwaarteklasse, verrichtingsgroep en urgentie de variabelen die het meest bijdragen aan de modellen.

## 2.4 Berekenen OLO-fracties

Zoals hiervoor besproken wordt er voor iedere diagnosegroep een log-normaal duurmodel geschat. Hiermee is het mogelijk om voor iedere opname een verwachte waarde uit te rekenen. Aangezien het gaat om *onverwacht* lange duren ligt het voor de hand om hierbij voor een hoog kwantiel te gaan, bijvoorbeeld het 90<sup>ste</sup> percentiel. Vervolgens kan worden gekeken of de waargenomen duur bij een opname langer is dan dit voorspelde kwantiel. Als dat zo is dan is dit een OLO.

Dus voor iedere waargenomen duur,  $y_{hi}$ , van opname  $i$  uit ziekenhuis  $h$  voorspellen we met behulp van het model en de kenmerken van de opname een kwantiel  $q_{hi}^{(p)}$ . Een opname is onverwacht lang als de waargenomen duur langer is dan het voorspelde kwantiel:

$o_{hi}^{(p)} = [y_{hi} > q_{hi}^{(p)}]$ . Het aantal OLO voor ziekenhuis  $h$  is dan

$$o_h^{(p)} = \sum_i o_{hi}^{(p)} \quad (5)$$

en de fractie OLO lange duren

$$f_h^{(p)} = \frac{o_h^{(p)}}{n_h} \quad (6)$$

met  $n_h$  het aantal opnamen in ziekenhuis  $h$ . Voor de hele populatie is de fractie OLO

$$f^{(p)} = \frac{1}{n} \sum_{h,i} o_{hi}^{(p)}. \quad (7)$$

Vervolgens zouden we willen toetsen of de ziekenhuisfractie  $f_h^{(p)}$  significant hoger is dan de landelijke fractie,  $f^{(p)}$ . Aangezien de laatste gebaseerd is op een zeer groot aantal waarnemingen kunnen we de onzekerheid in de laatste verwaarlozen. Voor het waargenomen aantal onverwacht lange opnamen van ziekenhuis  $h$  kunnen we aannemen dat deze binomiaal verdeeld is met parameters  $n_h$  en  $f^{(p)}$ . Hieruit volgen dan het betrouwbaarheidsinterval en p-waarde.

## 3 Evaluatie van de OLO-indicator

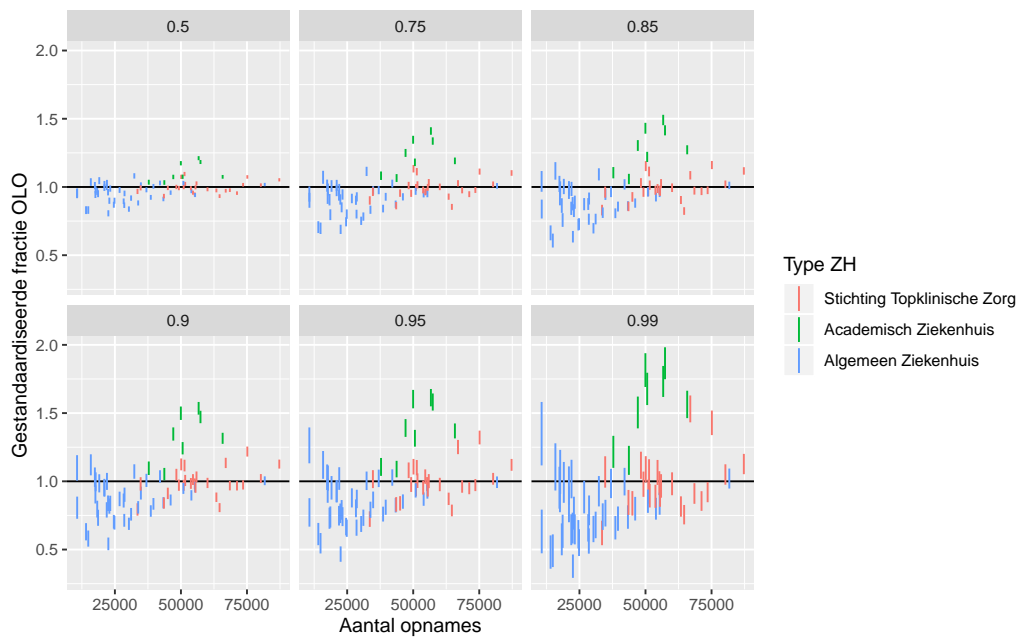
Om de OLO-indicator te evalueren wordt naar een aantal kenmerken van de indicator gekeken. Ten eerste wordt gekeken wat het effect is van de keuze van het percentiel dat gebruikt wordt voor het bepalen van OLO's. Voor iedere diagnosegroep wordt een log-normaal duurmodel geschat. Hiermee wordt voor iedere opname een verwacht percentiel voorspeld. Bijvoorbeeld, als we uitgaan van het 90<sup>ste</sup> percentiel, zegt de voorspelling dat gegeven de kenmerken van de patiënt en de diagnosegroep 90% van de opnamen een duur korter dan  $q$  heeft. Als de opname dan een duur heeft langer dan het voorspelde percentiel wordt het aangemerkt als OLO. De keuze voor het percentiel heeft dus effect op de fractie opnamen die als OLO wordt gekenmerkt. Bij een percentiel van 90% zal ongeveer 10% van de opnamen een OLO zijn. Ten tweede, wordt de indicator vergeleken met een aantal andere indicatoren waaronder de huidige OLO-indicator die DHD nu berekent voor de ziekenhuizen.

Figuur 3.1 toont de ziekenhuisfracties ( $f_h^{(p)}$ ) ten opzichte van het landelijk gemiddelde ( $f^{(p)}$ ) voor de verschillende onderzochte percentielen. Het landelijk gemiddelde is op 100 gesteld (de figuur toont dus  $100 \times f_h^{(p)} / f^{(p)}$ ). Deze figuur is vergelijkbaar met een funnelplot van de HSMR



**Figuur 3.1 Fracties OLO per ziekenhuis met 99,8% betrouwbaarheidsinterval vergeleken met het landelijk gemiddelde voor verschillende kwantielen.**

*Om de cijfers voor de verschillende kwantielen (0.5, 0.75, ..., 0.99) vergelijkbaar te maken is het landelijk gemiddelde op 100 gesteld.*



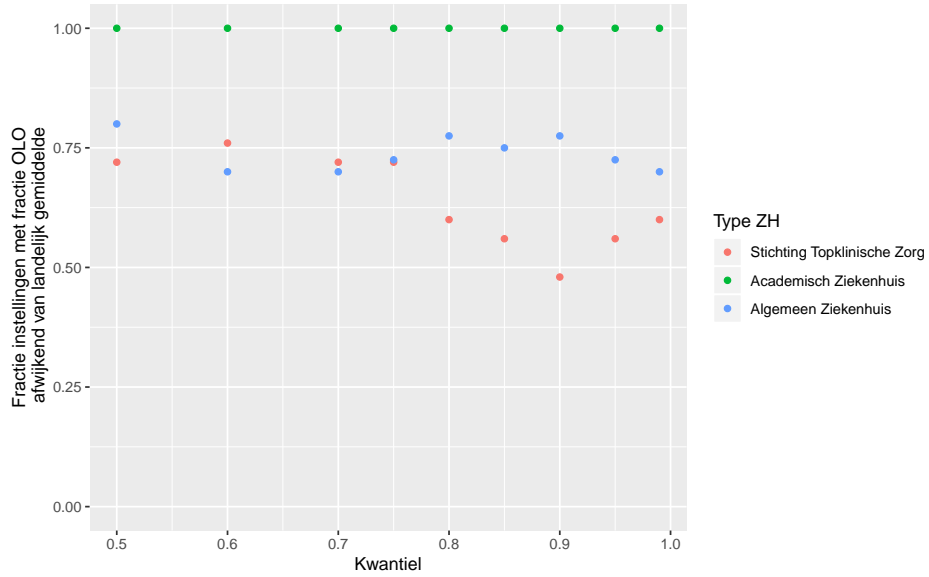
behalve dat op de x-as het aantal opnames staat in plaats van het aantal verwachte opnames (dit is echter ongeveer  $(1 - p) * n_h$ ). De lijntjes geven het 99,8% betrouwbaarheidsinterval van de ziekenhuisfracties aan. Een aantal dingen vallen op. Ten eerste, wijkt het merendeel van de ziekenhuisfracties significant af van het landelijk gemiddelde, ook na standaardisatie. Ten tweede, de betrouwbaarheidsintervallen worden groter als het percentiel groter wordt gekozen. Dit is logisch aangezien het aantal onverwacht lange duren dan afneemt waardoor de relatieve onzekerheid in dat aantal toeneemt. Ten derde, ongeacht de waarde van het percentiel zijn het in het algemeen telkens dezelfde ziekenhuizen die of significant hoger of significant lager scoren ten opzichte van het landelijk gemiddelde. En ten vierde, alle figuren laten hetzelfde patroon zien: de kleinere ziekenhuizen hebben relatief minder OLO's; bij de grotere ziekenhuizen zijn er twee groepen: de niet-academische (voornamelijk STZ-ziekenhuizen) die rond het gemiddelde liggen en de academische die daar boven liggen.

Figuur 3.2 toont de fractie van de ziekenhuisfracties die significant (99,8%) afwijken van het landelijk gemiddelde voor verschillende waarden van het gebruikte kwantiel. Het merendeel van de ziekenhuisfracties wijkt dus significant af van het landelijk gemiddelde. Voor de academische ziekenhuizen is dat zelfs 100%. Welke waarde van het kwantiel gebruikt wordt heeft hier ook weinig invloed op. De fractie neemt iets toe bij lagere waarden.

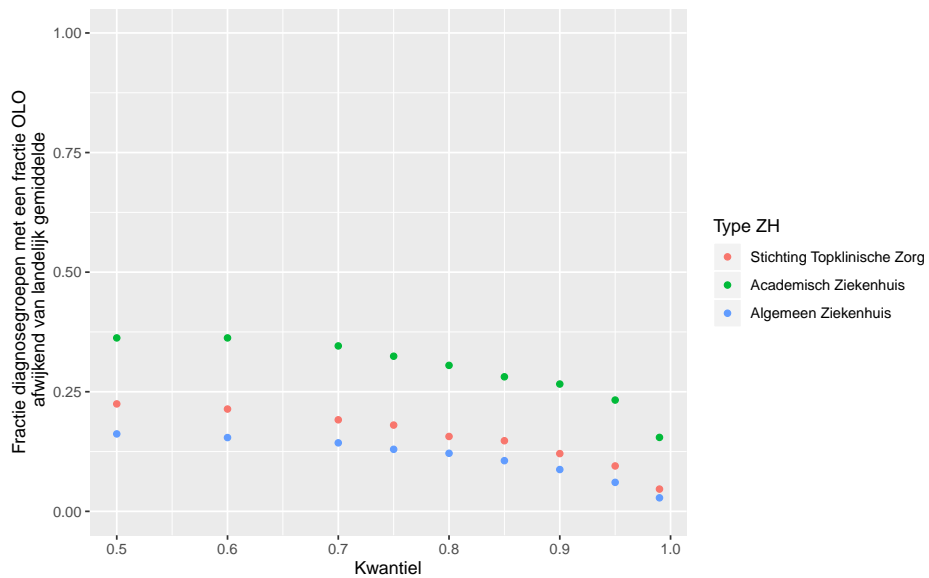
We kunnen bovenstaande ook uitvoeren per diagnosegroep. De resultaten hiervan staan in figuur 3.3. Door de kleinere aantallen per diagnosegroep en de bijbehorende grotere betrouwbaarheidsintervallen is de fractie significant hoge of lage ziekenhuisfracties wat kleiner. Maar verder zijn dezelfde patronen te zien en zijn de fracties significant afwijkende OLO-ratio's erg hoog gezien de betrouwbaarheid van 99,8%.

Als er geen verschillen zouden zijn tussen de ziekenhuizen zou ongeveer 0,2% van de OLO-ratio's significant moeten afwijken van het landelijk gemiddelde met de gebruikte betrouwbaarheid van

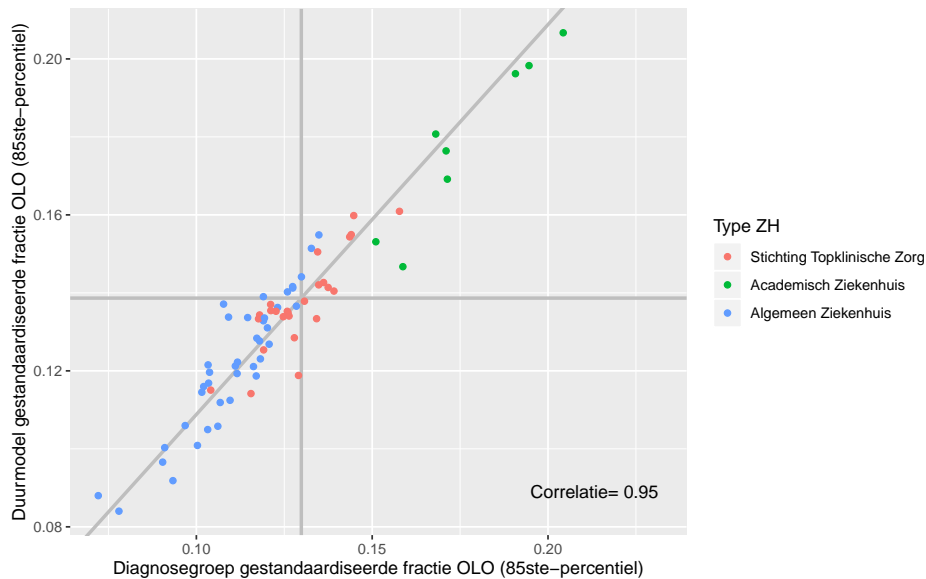
**Figuur 3.2 Fractie instellingen met OLO-ratio significant (99,8% betrouwbaarheid) afwijkend van het landelijk gemiddelde als functie van het kwantiel.**



**Figuur 3.3 Fractie van diagnosegroepen met een significante OLO-ratio (99,8% betrouwbaarheid) als functie van het kwantiel, per type instelling.**



**Figuur 3.4 Met duurmodel gestandaardiseerde fractie onverwacht lange opnameduren tegen de alleen naar diagnosegroep gestandaardiseerde fractie onverwacht lange opnameduren.**

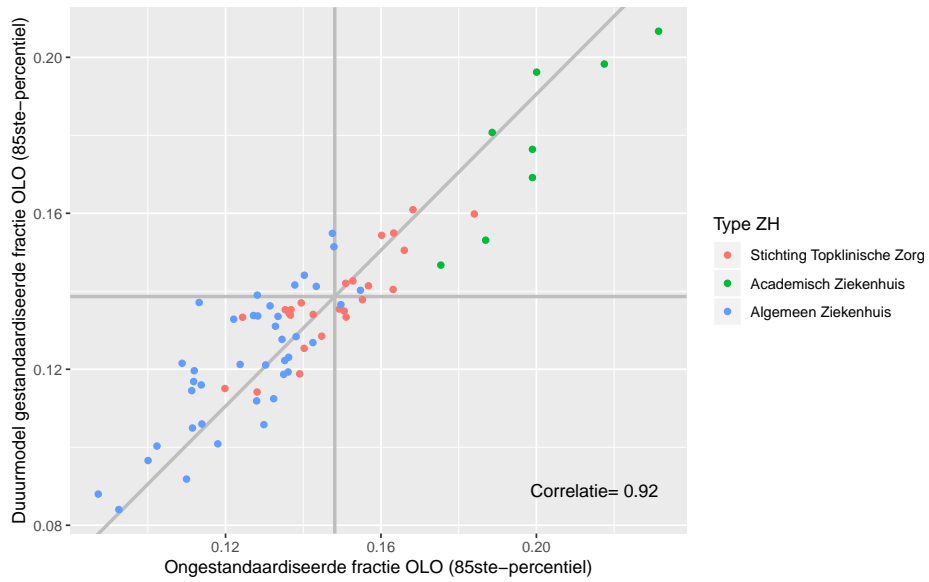


99,8%. Dat meer dan 50% van de instellingen afwijkt van het landelijk gemiddelde kan wijzen op onvoldoende correctie voor populatieverschillen tussen de ziekenhuizen of andere verschillen tussen ziekenhuizen die de ligduur beïnvloeden. In figuur 3.4 is de duurmodelgestandaardiseerde OLO-ratio uitgezet tegen de OLO-ratio die alleen gestandaardiseerd is naar diagnosegroep. Hieruit blijkt dat het duurmodel weinig toevoegt ten opzichte van een eenvoudige standaardisatie naar diagnosegroep. Voor de standaardisatie naar diagnosegroep is per diagnosegroep het 85<sup>ste</sup> percentiel uitgerekend; alle opnamen met een duur langer dan dat kwantiel zijn een OLO. De reden om van het 85<sup>ste</sup> percentiel uit te gaan is dat bij de huidige OLO-indicator circa 15% van de opnamen als OLO worden aangemerkt. Hierdoor is vergelijking met de huidige methode straks makkelijker. De horizontale lijn in de figuur geeft aan wat het landelijk gemiddelde is van de met het duurmodel berekende OLO-ratio en op dezelfde manier geeft de verticale lijn aan wat het landelijk gemiddelde is van de eenvoudig gestandaardiseerde OLO-ratio. De diagonale lijn is een lijn met helling één die door het snijpunt van de twee andere lijnen gaat. Als beide methodes exact dezelfde standaardisatie opleveren dan zouden alle ziekenhuizen op deze lijn moeten liggen. Dit is in de figuur ook bijna het geval. De standaardisatie met het duurmodel heeft dus weinig effect ten opzichte van een eenvoudige standaardisatie naar diagnosegroep en levert nagenoeg dezelfde OLO-ratio's op. Dit is ook te zien in de hoge correlatie coëfficiënt (0,95).

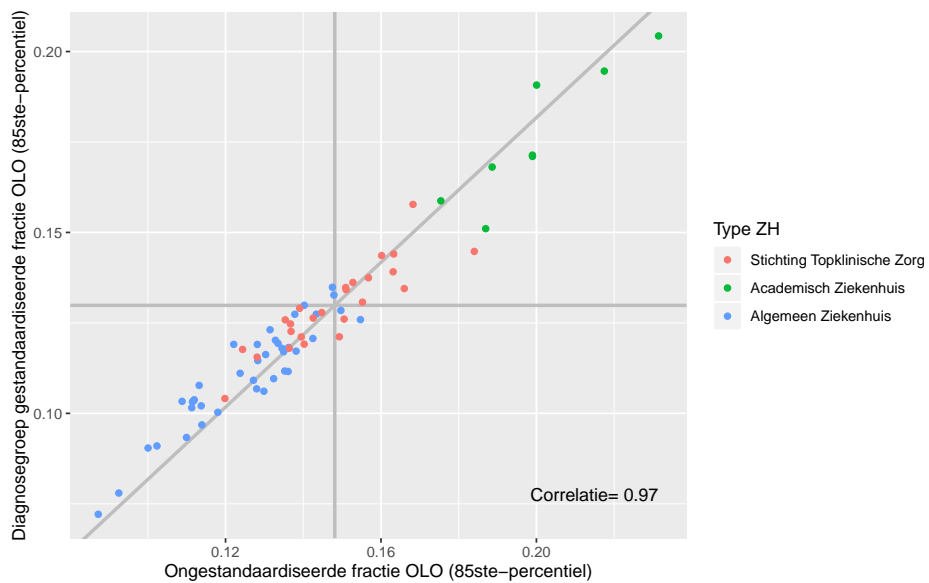
Figuur 3.5 toont het verband tussen de met het duurmodel gestandaardiseerde OLO-ratio en de volledig ongestandaardiseerde OLO-ratio. Figuur 3.6 toont hetzelfde voor de naar diagnosegroep gestandaardiseerde OLO-ratio met de ongestandaardiseerde OLO-ratio. In beide figuren zijn de verbanden weer bijna lineair wat ook te zien is in de hoge correlatiecoëfficiënten. Bij de met het duurmodel gestandaardiseerde met de ongestandaardiseerde OLO-ratio's is wel wat meer variatie te zien en is de correlatie ook iets lager dan in figuur 3.4. De standaardisatie heeft dus wel iets effect ook al is dit effect klein.

De huidige methode, die DHD momenteel gebruikt, standaardiseert naar een combinatie van hoofddiagnose, leeftijd en hoofdverrichting. Figuren 3.7 en 3.8 tonen de met het duurmodel gestandaardiseerde en de alleen naar diagnosegroep gestandaardiseerde OLO-ratio tegen de

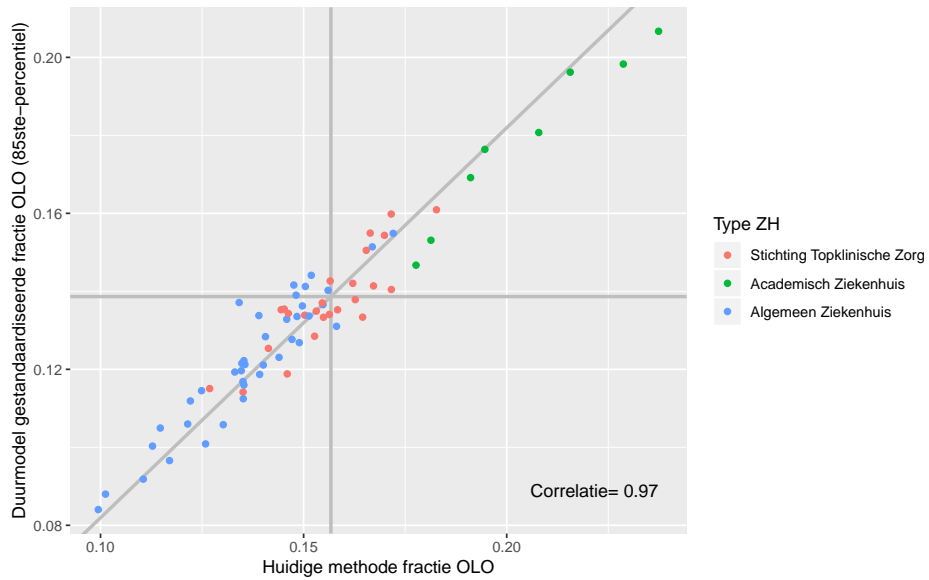
**Figuur 3.5 Met duurmodel gestandaardiseerde fractie onverwacht lange opnames tegen de ongestandaardiseerde fractie onverwacht lange opnames.**



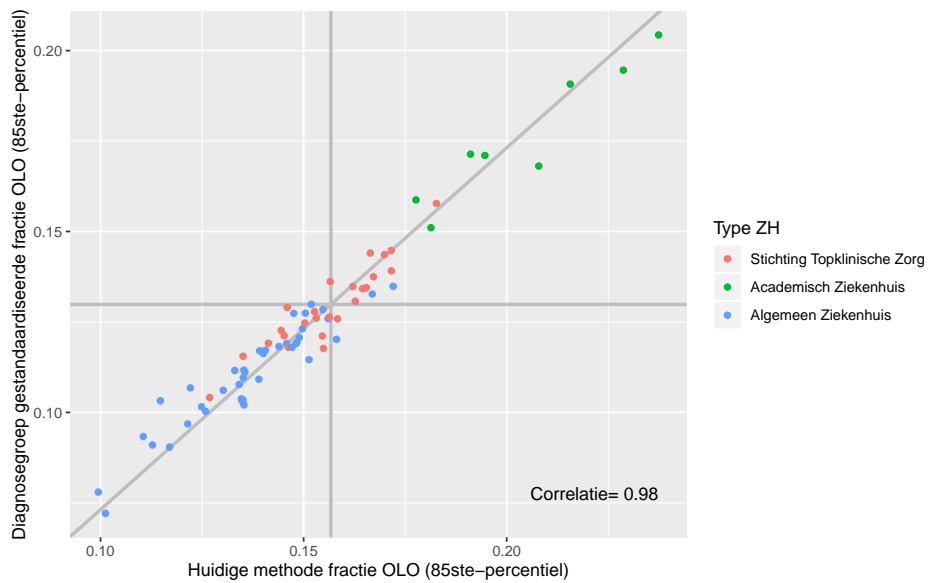
**Figuur 3.6 De alleen naar diagnosegroep gestandaardiseerde fractie onverwacht lange opnames tegen de ongestandaardiseerde fractie onverwacht lange opnames.**



**Figuur 3.7 Met duurmodel gestandaardiseerde fractie onverwacht lange opnames tegen de fractie OLO bepaald door de huidige methode.**



**Figuur 3.8 De alleen naar diagnosegroep gestandaardiseerde fractie onverwacht lange opnames tegen de fractie OLO bepaald door de huidige methode.**



**Tabel 3.1 Percentage opnamen geclassificeerd als OLO (85<sup>ste</sup> percentiel) of geen OLO met de verschillende methodes en de samenhang daartussen.**

		Huidige methode		Naar diagnose groep gest.		Ongestand.	
		Geen OLO	OLO	Geen OLO	OLO	Geen OLO	OLO
Duurmodel	Geen OLO	81,8%	4,3%	83,0%	3,2%	80,1%	6,0%
	OLO	2,5%	11,4%	4,0%	9,8%	5,1%	8,9%
Huidige methode	Geen OLO	-	-	81,7%	2,6%	79,3%	5,0%
	OLO	-	-	5,3%	10,4%	5,9%	9,8%

**Tabel 3.2 C-index**

Standaardisatie methode	C-index
Duurmodel	0,77
Huidige methode	0,81
Standaardisatie naar diagnosegroep	0,66
Geen standaardisatie	0,50

OLO-ratio bepaald met de huidige methode. Deze laten hetzelfde beeld zien als eerder: de verschillen tussen de verschillende manieren van standaardiseren zijn minimaal. De methode van standaardiseren heeft dus weinig effect op de gevonden fracties OLO. Dit wordt bevestigd door de resultaten in tabel 3.1 die voor combinaties van methodes laat zien welke gedeeltes van de opnamen als OLO worden geclassificeerd. Er is te zien dat iedere methode dezelfde ca. 80% van de opnamen als geen OLO classificeert en ook dezelfde ca. 10% van de opnamen als OLO classificeert. Dus circa 90% van de opnamen worden hetzelfde geclassificeerd ongeacht de methode.

Tabel 3.2 toont de overall C-indexen van alle methodes. Deze zijn bepaald door van combinaties van opnamen te kijken of de opnamen met een langer voorspelde duur ook daadwerkelijk een langere duur heeft (dit is gedaan voor een willekeurige steekproef van 500 000 opnamen). De fractie opnameparen waarbij dit het geval is, is de C-index. Als het model niets verklaart is deze gelijk aan 0,5 (dit is in de tabel inderdaad te zien als er geen standaardisatie is). De maximale waarde is 1. Hoe hoger de waarde hoe beter het model in staat is verschillen tussen de opnamen te verklaren. De tabel laat zien dat er wat betreft verklaring wel degelijk verschillen zijn tussen de methodes. Daarnaast zijn C-indexen van het duurmodel en de huidige methode (resp. 0,77 en 0,81) niet laag. De modellen verklaren dus best wel wat van de variatie in de opnameduren. De iets hogere C-index van de huidige methode wordt waarschijnlijk veroorzaakt doordat bij die methode de opnames in heel veel en daardoor soms kleine strata (groepen) worden verdeeld. Ondanks dat de modellen dus een redelijke verklarende kracht hebben, classificeren ze grotendeels dezelfde opnamen als OLO. Het lijkt er dus op dat er een groep opnamen is, die ongeacht de standaardisatiemethode als OLO worden geclassificeerd en blijkbaar zijn er substantiële verschillen tussen de ziekenhuizen in hoe groot deze groep is.

## 4 Conclusie

In het voorgaande hoofdstuk zijn een aantal methoden om onverwacht lange opnamen te bepalen vergeleken. Hieruit zijn twee hoofdconclusies te trekken. Ten eerste, heeft de methode

waarmee gestandaardiseerd wordt weinig effect op de berekende fractie OLO. Een eenvoudige standaardisatie naar alleen hoofddiagnosegroep levert bijna dezelfde fractie OLO op als een duurmodel met standaardisatie naar een groot aantal kenmerken. Ongeacht de methode worden voor een groot gedeelte dezelfde opnamen als OLO geïnclassificeerd. Ten tweede, blijken er grote verschillen te zijn tussen de ziekenhuizen wat betreft hun fracties OLO. Het merendeel van de ziekenhuizen wijkt significant af van het landelijk gemiddelde. Dus hoewel er ongeacht de methode min of meer dezelfde opnamen als OLO geïnclassificeerd worden, blijken deze opnamen heel verschillend verdeeld te zijn over de ziekenhuizen. Deze verschillen hangen o.a. samen met de grootte van het ziekenhuis en het type ziekenhuis. Het is onduidelijk waardoor deze verschillen veroorzaakt worden. Zijn er patiëntkenmerken die de verschillen tussen de ziekenhuizen verklaren die in dit onderzoek nog niet in het model zijn meegenomen? Is het daadwerkelijk zo dat dezelfde patiënt in een ander ziekenhuis gemiddeld langer ligt? In dat laatste geval is het interessant om na te gaan waardoor dit komt. Mogelijk dat ook andere factoren hierbij een rol spelen zoals de aard, inrichting of efficiëntie van de zorg.

Voor een goede interpretatie van de OLO-uitkomsten zou nader onderzocht moeten worden waardoor de grote verschillen tussen de ziekenhuizen worden veroorzaakt. Gezien dit punt en de kleine verschillen tussen de methoden, lijkt het op dit moment niet nuttig om een ingewikkelde standaardisatiemethode met een duurmodel te kiezen. De huidige methode of een eenvoudige standaardisatie naar diagnosegroep lijkt vooralsnog te volstaan.

## Referenties

van der Laan, J., Penning, C., and de Bruin, A. (2019a). *Hospital Readmission Ratio: Methodological report of the 2017 model*. CBS, Den Haag.

<https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/overzicht/methodologische-rapportage-heropname-ratio-ziekenhuizen>.

van der Laan, J., Witvliet, C., Peters, M., and de Bruin, A. (2019b). *HSMR 2018: Methodological report*. CBS, Den Haag. <https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/overzicht/>

[methodologische-rapportages-bij-hospital-standardised-mortality-ratios--hsmr--](https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/overzicht/methodologische-rapportages-bij-hospital-standardised-mortality-ratios--hsmr--).

# Bijlagen



# A Zwaarteklasse voor onverwacht lange opnameduren

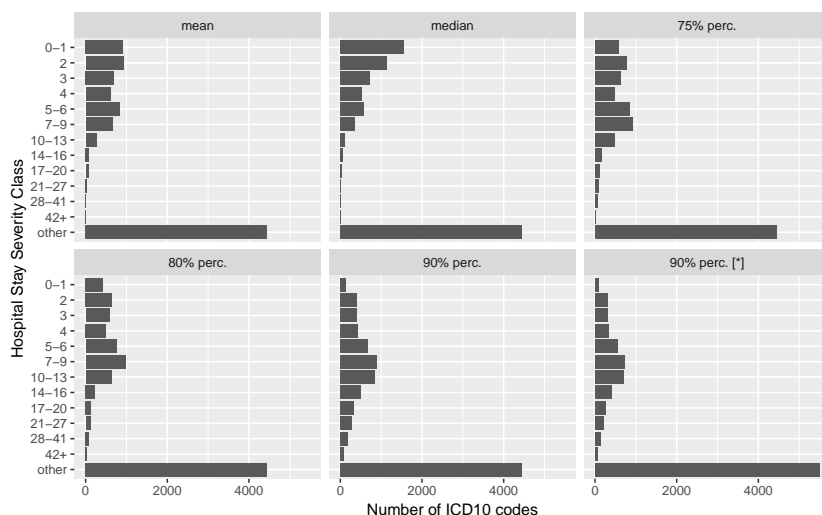
Alle opnamen zijn ingedeeld in 157 diagnosegroepen. Net als bij het HSMR- en het heropnamenmodel wordt voor iedere diagnosegroep een apart model geschat. Echter, binnen diagnosegroepen kan er nog steeds sprake zijn van een variatie in de zwaarte van de opnamen. Daarom wordt voor iedere opname aan de hand van de in ICD-10 gecodeerde hoofddiagnose een zwaarteklasse bepaald. Dit wordt gedaan aan de hand van historische data die niet overlappen met de data die in het model zelf worden gebruikt, zoals is beschreven in de sectie 2.1. Voor de HSMR en de heropnamenratio's van de ziekenhuizen wordt hiervoor gebruik gemaakt van de historische sterfte. Echter, voor opnameduren is het onwaarschijnlijk dat veel sterfte bij een diagnose ook een indicator is voor lange opnameduren. Daarom is besloten om voor het modelleren van opnameduren een aparte zwaarteklasse-indeling af te leiden voor de hoofddiagnose.

Om te bepalen wat de beste indicator is zijn voor iedere hoofddiagnose de volgende indicatoren afgeleid:

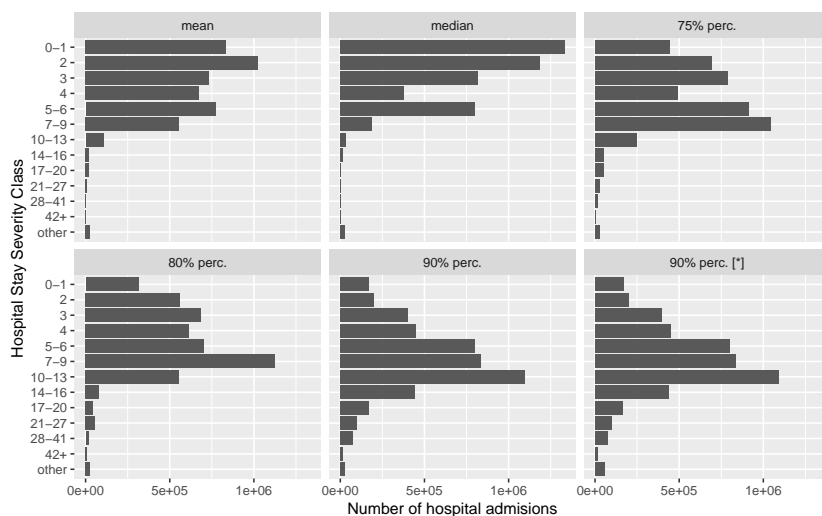
1. Gemiddelde opnameduur
2. Mediane opnameduur
3. Zeventigste percentiel van de opnameduren. Dit betekent dat 70% van de opnamen een duur heeft korter dan dit percentiel.
4. Tachtigste percentiel van de opnameduren.
5. Negentigste percentiel van de opnameduren.
6. Negentigste percentiel van de opnameduren met een hogere drempel (40) voor het minimale aantal opnamen. Deze variant is in de resultaten met een '['\*]’ gemarkeerd.

Voor alle indicatoren (behalve 6) geldt dat er per ICD-10 code minimaal 20 opnamen moeten zijn geweest. Verder moet een ICD-10 code door minimaal 5 ziekenhuizen gebruikt zijn. Als niet aan deze eisen wordt voldaan, krijgt de ICD-10 code de zwaarteklasse 'other' toegewezen.

**Figuur A.1** Verdeling van de zwaarteklassen over de ICD-10 codes voor elk van de zwaarteklasse-indelingen.



**Figuur A.2 Verdeling van de zwaarteklassen over het aantal opnamen voor elk van de zwaarteklasse-indelingen.**



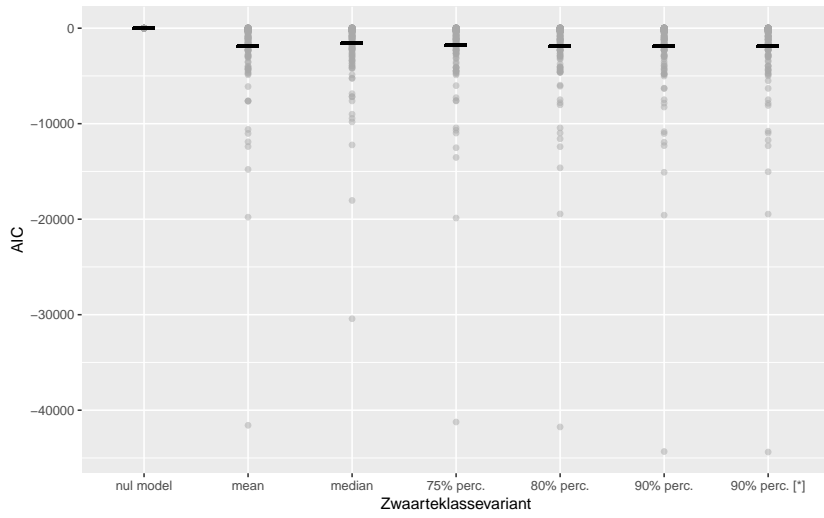
Figuur A.1 toont de verdeling van de zwaarteklassen voor ieder van de indicatoren. Er zijn veel ICD-10 codes die de zwaarteklasse ‘other’ krijgen. Echter, dit zijn codes die weinig gebruikt worden als hoofddiagnose. Figuur A.2 toont de verdeling van de zwaarteklassen over het aantal opnamen.

Om te evalueren welke variant van de zwaarteklasse-indelingen het beste is, zijn voor ieder van de 157 diagnosegroepen Weibull-modellen<sup>1)</sup> geschat met daarin alle covariaten zoals deze ook in de HSMR worden meegenomen en de eerdergenoemde verrichtingcode. De modellen zijn geschat zonder zwaarteklasse in het model en met ieder van de varianten van de zwaarteklasse-indeling. De modellen zijn aan de hand van de volgende criteria beoordeeld:

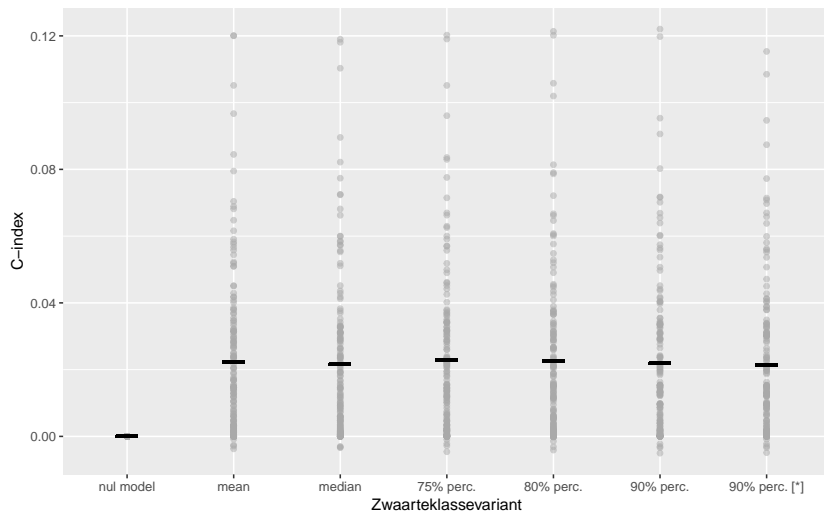
1. AIC (Akaike Informatie Criterium). Het AIC meet hoeveel het model verklaart gegeven het aantal parameters. In het algemeen verklaart het model met de laagste AIC het meest (gegeven het aantal parameters).
2. C-index (ook wel Concordance Index of C-statistic genoemd). De C-index meet hoe goed het model dueren voorspelt. Het is de kans dat voor twee willekeurige opnamen voor de opname met de langste waargenomen duur ook de langste duur wordt voorspeld door het model. Aangezien de modellen gebruikt gaan worden om te voorspellen is dit een belangrijk criterium.
3. Calibratie van het model. Aangezien we voornamelijk geïnteresseerd zijn in onverwacht lange dueren, is vooral de staart van de duurverdeling interessant. Om te controleren of het model de staart van de verdeling goed voorspelt, voorspellen we voor iedere opname het 95<sup>ste</sup> percentiel gebruikmakend van de kenmerken van de opname. Als het model de duurverdeling goed beschrijft zou 5% van de opnamen een duur langer dan dit percentiel moeten hebben. Om dit te controleren vergelijken we de fractie opnamen met een duur langer dan dit percentiel met de verwachte 5%. We berekenen de gemiddelde afwijking en de wortel van de gemiddelde kwadratische afwijking (RMSQE). Beiden zouden dicht bij nul moeten zijn.

<sup>1)</sup> Er is hier met Weibull-modellen gewerkt hoewel later in het project gekozen is voor een log-normaal model. De verschillen tussen het log-normale model en het Weibull-model zijn echter klein en het is onwaarschijnlijk dat de uitkomsten zoals in deze bijlage gepresenteerd sterk zullen veranderen als er hier een log-normaalmodel gebruikt zou zijn.

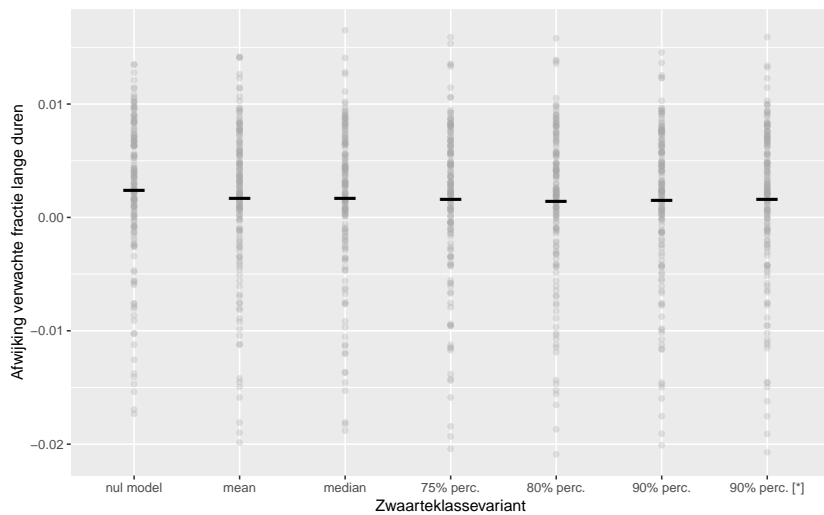
**Figuur A.3 AIC (lager is beter)**



**Figuur A.4 C-index (hoger is beter)**



**Figuur A.5 Afwijking van de voorspelde fracties (dichter bij nul is beter)**



**Tabel A.1 Vergelijking van de zwaarteklasse-indelingen**

Variant	Criterium			
	Mean AIC	Mean C-index	Mean Error	RMSQE
Nul model	89 772	0,675	0,0024	<b>0,0068</b>
Mean	87 894	0,696	0,0017	<b>0,0068</b>
Median	88 152	0,696	0,0017	<b>0,0068</b>
75% perc.	87 921	<b>0,697</b>	0,0016	0,0069
80% perc.	87 875	<b>0,697</b>	<b>0,0014</b>	<b>0,0068</b>
90% perc.	<b>87 824</b>	0,696	0,0015	0,0069
90% perc. [*]	87 849	0,696	0,0016	0,0069

Figuren A.3 en A.4 tonen de verdelingen van respectievelijk de AIC en de C-index van de verschillende varianten van de zwaarteklasse-indeling voor de 157 diagnosegroepen. Beiden zijn ten opzichte van het model zonder zwaarteklasse (nulmodel). Figuur A.5 toont de afwijking van de voorspelde fracties langer dan het 95<sup>ste</sup> percentiel ten opzichte van de verwachte 5%. Tabel A.1 vat de resultaten van de figuren verder samen. De meest optimale uitkomsten per criterium zijn vetgedrukt.

De verschillen tussen de verschillende varianten van de zwaarteklasse-indeling zijn erg klein. Alleen de mediaan scoort vooral op het criterium AIC slechter. Voor de zwaarteklasse-indeling gebaseerd op het 90<sup>ste</sup> percentiel is de grens van het minimaal 20 opnamen laag. Dat zou betekenen dat deze waarde gebaseerd is op de twee opnamen met de langste duur. Dit introduceert waarschijnlijk instabiliteiten bij het herberekenen van de zwaarteklassen in de toekomst. Vandaar dat ook een zwaarteklasse-indeling gebaseerd op het 90<sup>ste</sup> percentiel is genomen met een grens van 40 opnamen. De verschillen tussen de overgebleven drie varianten (75/80/90 [\*]) zijn erg klein. Aangezien de variant gebaseerd op het 80<sup>ste</sup> percentiel net iets vaker beter scoort, is besloten om deze variant te kiezen.

## B Bepaling gebruikte duurmodel

Hoewel er bij aanvang dit onderzoek besloten was om in principe uit te gaan van een Weibull-model, bleek al redelijk snel dat deze niet altijd de waargenomen duurverdeling goed beschrijft. Daarom is besloten om een aantal modellen met elkaar te vergelijken om tot een best passend model te komen. Deze vergelijking wordt in deze bijlage beschreven.

Drie van de onderzochte modellen zijn van het type *accelerated failure time models*: Weibull-model, log-logistische, log-normale model. Bij deze modellen wordt aangenomen dat de duurverdeling een bepaalde parametrische verdeling volgt (resp. Weibull, log-logistische en log-normale verdeling) waarbij voor sommige opnamen de klok langzamer of sneller loopt afhankelijk van de kenmerken van de patiënt. Voor meer informatie zie paragraaf 2.3. Daarnaast is er ook gekeken naar het Cox-proportionele-hazard-model (Cox-model). In dit model wordt er geen duurverdeling aangenomen. De duurverdeling kan worden geschat met behulp van een niet-parametrische methode. Dit maakt dit model veel flexibeler. Dit model is echter minder geschikt voor het einddoel waarbij andere partijen dan het CBS het model gaan gebruiken voor het maken van predicties. Dit model is namelijk moeilijker te delen omdat er geen parametrische

beschrijving van de duurverdeling is. Het Cox-model wordt daarom meegenomen als referentie waarmee de andere modellen vergeleken kunnen worden.

Voor ieder van de 157 diagnosegroepen zijn ieder van de modellen geschat. Allereerst zijn de modellen grafisch geëvalueerd door naar de residuen te kijken. Twee soorten residuen zijn bekeken: de reguliere residuen (verschil tussen voorspelde duur en waargenomen duur) en de devianceresiduen. Deze laatste zijn het genormaliseerde verschil tussen het aantal verwachte gebeurtenissen op het moment dat de patiënt het ziekenhuis verlaat en het waargenomen aantal gebeurtenissen. In ons geval is het laatste altijd gelijk aan één aangezien een patiënt altijd het ziekenhuis verlaat en het ziekenhuis ook maar één keer kan verlaten voor een gegeven opname.

In het geval van reguliere lineaire regressie wordt vaak aangenomen dat de residuen normaal verdeeld zijn. In het geval van een duurmodel geldt deze aanname niet; de residuen zijn dan scheef verdeeld. Ze volgen echter nog steeds wel een bekende verdeling waarmee de waargenomen verdeling vergeleken kan worden. In het subfiguur linksboven in figuur B.1 is de overlevingsfunctie van de residuen van het Weibull-model geplot met in rood de verwachte verdeling van één van de diagnosegroepen. Door de normalisatie zouden de devianceresiduen een normale verdeling moeten volgen. In het subfiguur linksonder zijn de devianceresiduen geplot voor het Weibull-model. De rode lijnen geven het gebied aan waarin ongeveer 99% van de residuen zouden moeten liggen als deze een standaard normale verdeling zouden volgen.

Dat het model perfect bij de data past is niet te verwachten. Een gecompliceerd proces zoals het ontslaan uit een ziekenhuis zal niet te beschrijven zijn met een verdeling met slechts een paar parameters. Verder zijn de waargenomen duren discreet (in dagen) terwijl het model een continue verdeling aanneemt. Dit verklaart ook de stapvorm die te zien is in de overlevingsfunctie van de residuen (bovenste rij in figuur B.1) en de banden die te zien zijn in de plots met de devianceresiduen (onderste rij). Omdat het model nooit perfect zal passen en omdat we flinke aantallen waarnemingen hebben zal een toets of het model goed past deze aanname altijd verwerpen. Vandaar dat gekozen is voor een visuele inspectie van de residuen.

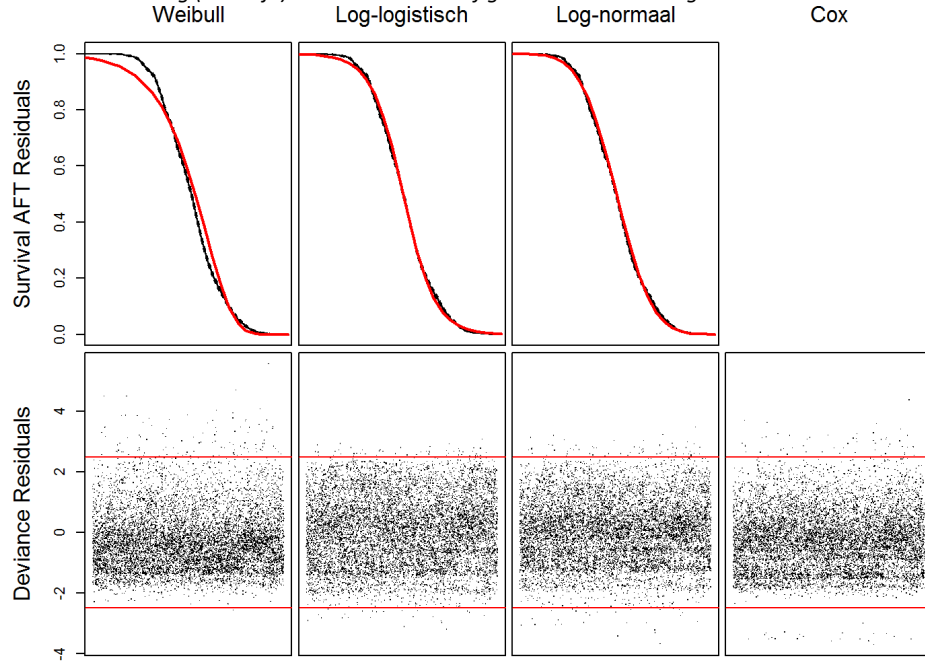
De linker kolom in figuur B.1 laat zien dat het Weibull-model de waarnemingen niet heel goed beschrijft voor deze diagnose groep. Dit geldt ook voor de meeste andere diagnosegroepen (soms past het model iets beter, maar vaak ook nog wel slechter dan in dit voorbeeld). De residuen van de log-logistische, log-normale, en Cox-modellen staan in respectievelijk de tweede, derde en vierde kolom van de figuur.

Een visuele inspectie van de plots met residuen van alle 157 diagnosegroepen (waarvan figuur B.1 één voorbeeld geeft) laat zien dat in het algemeen het log-logistische en log-normale model beter bij de waargenomen verdelingen passen dan het Weibull-model. De verschillen tussen het log-logistische en het log-normale model zijn klein. De devianceresiduen hebben bij het Cox-model vaker een iets kleinere spreiding dan bij het log-logistische en het log-normale model.

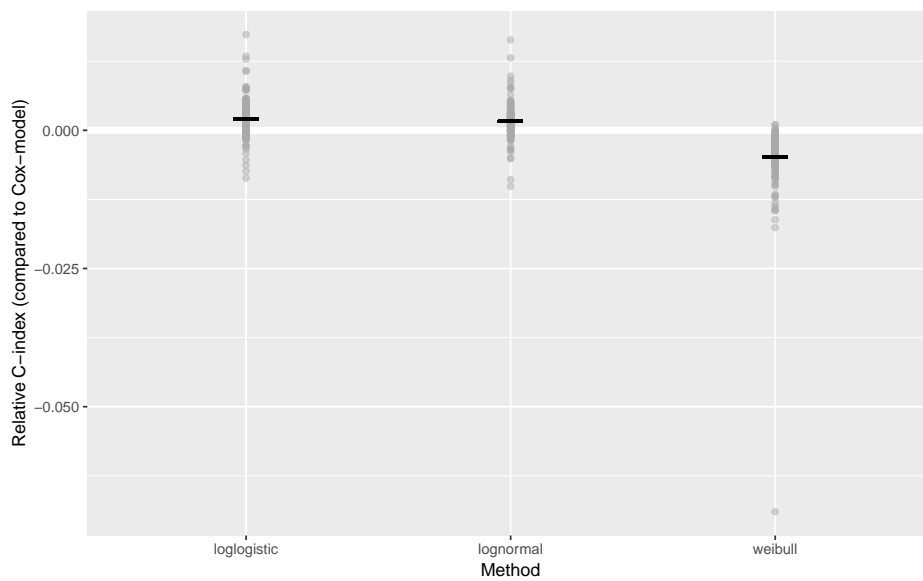
De reden dat het Weibull-model niet altijd goed bij de waargenomen duren past, is waarschijnlijk dat het Weibull-model het geval van een stijgende hazard gevolgd door een dalende hazard niet goed modelleert. In het geval van ziekenhuisopnamen is dit verloop van de hazard niet onwaarschijnlijk: er is eerst een korte periode waarin weinig patiënten het ziekenhuis verlaten (lage hazard) waarna veel van de patiënten uitstromen (hoge hazard) er is een kleine groep patiënten die langer blijft liggen. Hier is de hazard dan weer laag.

**Figuur B.1 Residuën voor alle onderzochte methoden: Weibull, log-logistisch, log-normaal en Cox-proportionele-hazard-model.**

*De residuën zijn voor een redelijk typische diagnosegroep (30). Aangezien het Cox-model geen aannamen doet over de verdeling, kan de verdeling van de residuën (zwarte lijn) niet vergeleken worden met de verwachte verdeling (rode lijn). Daarom is de sub-figuur recht-boven leeg.*



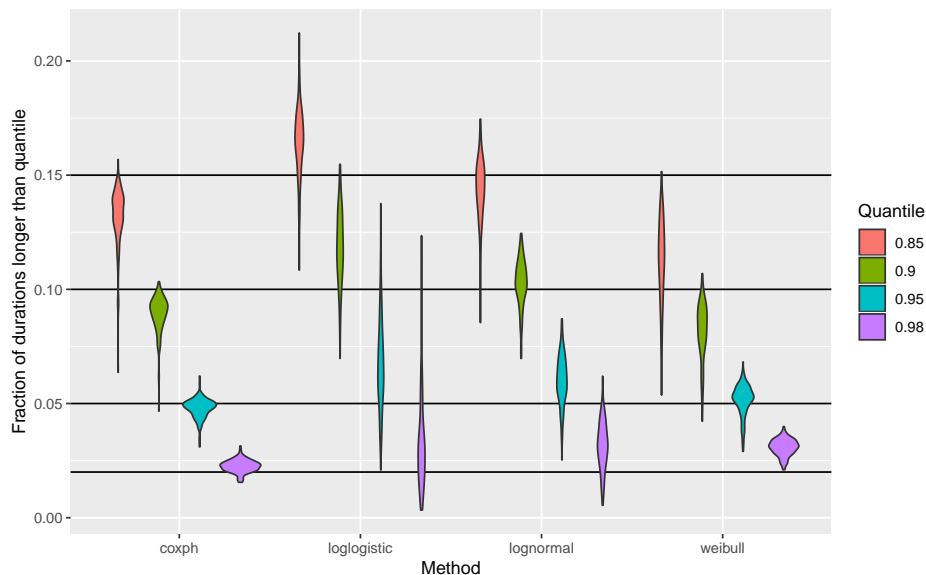
**Figuur B.2 C-indexen van de 157 modellen voor ieder van de diagnosegroepen vergeleken met de C-index van het Cox-model.**



**Tabel B.1 Gemiddelde C-index voor de vier methodes.**

Methode	Gemiddelde C-index
Cox	0,699
Log-logistisch	0,701
Log-normaal	0,700
Weibull	0,694

**Figuur B.3 Fractie opnamen met een duur langer dan het gegeven kwantiel.**



Figuur B.2 vergelijkt de C-index van ieder van de parametrische methodes met die van het Cox-model voor ieder van de 157 diagnosegroepen. De C-index beschrijft hoe goed het model in staat is onderscheid te maken tussen lange en korte duren. Het log-logistische en log-normale model presteren in het algemeen iets beter dan het Cox-model. Het Weibull-model presteert het minst. Tabel B.1 met daarin de gemiddelde C-indices bevestigt dit.

In principe zou het zo moeten zijn dat als we het  $q$ de percentiel voorspellen voor opnamen, er  $(100 - q)\%$  van de opnamen een duur langer dan het voorspeld zouden moeten hebben. Voor ieder van de modellen en voor ieder van de diagnosegroepen zijn deze fracties uitgerekend voor verschillende kwantielen. Figuur B.3 toont de verdeling van de fracties opnamen langer dan het voorspelde kwantiel. De horizontale lijnen geven de verwachte fracties aan voor de verschillende kwantielen. Het Cox-model ligt in het algemeen het dichtst bij de verwachte waarde. De spreiding van het log-logistische model is erg groot. Het log-normale model heeft een grotere spreiding dan het Cox-model, maar de fracties liggen wel rondom de verwachte waarden. Het Weibull-model heeft iets minder spreiding, maar de fracties liggen niet altijd rondom de verwachte waarde. Tabel B.2 bevestigt dit. Voor de 85<sup>ste</sup> en 90<sup>ste</sup> percentielen presteert het log-normale model eigenlijk altijd beter dan de andere methoden.

Het Weibull-model presteert op bijna alle maten minder goed dan het log-logistische of log-normale model. Deze twee modellen presteren vergelijkbaar met het Cox-model en de plots van de residuën laten zien dat deze twee modellen in het algemeen goed bij de data passen. Het log-normale model lijkt soms net iets beter te passen. Ook bij de voorspelde fracties duren presteert het log-normale model net iets beter dan het log-logistische: de spreiding in de fracties is kleiner.

**Tabel B.2 Gemiddelde afwijking en root mean square afwijking van de fractie opnamen met een duur langer dan een gegeven kwantiel ten opzichte van het verwachte aantal.**

Kwantiel	Methode	Gemiddelde fout	RMSQE
0.85	Cox	-0,018	0,018
	Log-logistisch	0,015	0,018
	Log-normaal	-0,007	0,011
	Weibull	-0,037	0,037
0.90	Cox	-0,011	0,011
	Log-logistisch	0,018	0,021
	Log-normaal	0,003	0,008
	Weibull	-0,018	0,018
0.95	Cox	-0,002	0,003
	Log-logistisch	0,017	0,021
	Log-normaal	0,010	0,013
	Weibull	0,002	0,006
0.98	Cox	0,003	0,003
	Log-logistisch	0,012	0,015
	Log-normaal	0,011	0,013
	Weibull	0,011	0,011

Het log-normale model lijkt dus goed te passen bij de data en heeft hiermee de voorkeur. Bij de evaluatie van de OLO-indicator (zie hoofdstuk 3) is daarom gewerkt met het log-normale duurmodel, op basis van de 85<sup>ste</sup> percentiel.



## **Colophon**

### *Publisher*

Statistics Netherlands  
Henri Faasdreef 312, 2492 JP The Hague  
[www.cbs.nl](http://www.cbs.nl)

### *Prepress*

Statistics Netherlands, Grafimedia

### *Design*

Edenspiekermann

### *Information*

Telephone +31 88 570 70 70, fax +31 70 337 59 94  
Via contact form: [www.cbs.nl/information](http://www.cbs.nl/information)

© Statistics Netherlands, The Hague/Heerlen/Bonaire 2020.  
Reproduction is permitted, provided Statistics Netherlands is quoted as the source