

## ONDERZOEK

# Gestandaardiseerd ziekenhuissterftecijfer (HSMR): correctie voor ernst hoofddiagnose kan beter

Wim F. van den Bosch, Peter Spreeuwenberg en Cordula Wagner

- DOEL** De invloed van casemixverschillen binnen diagnosegroepen op het gestandaardiseerde ziekenhuissterftecijfers (HSMR) onderzoeken en het huidige HSMR-model verbeteren.
- OPZET** Beschrijvend, analyse van modellen.
- METHODE** Wij berekenden de HSMR's en de gestandaardiseerde sterftecijfers op diagnosegroepniveau (SMR's) van 84 Nederlandse ziekenhuizen over een periode van 5 jaren met behulp van 2 HSMR-modellen. Beide modellen corrigeerden we voor de ernst van de diagnose bij opname (hoofddiagnose volgens de 'International classification of diseases' (ICD)-9) aan de hand van zwaarteklassen. De indeling in zwaarteklassen was in het gebruikelijke model (model 1) gebaseerd op mortaliteitscijfers volgens de WHO, terwijl ons aangepaste 'Nederlandse' model (model 2) was gebaseerd op de Nederlandse ziekenhuismortaliteit. De HSMR- en SMR-uitkomsten en bijbehorende c-statistics van beide modellen werden vergeleken om te toetsen in hoeverre de modellen adequaat corrigeerden voor ICD-9-codes met hoog of laag sterfterisico.
- RESULTATEN** Model 1 corrigeerde bij 40 van de 48 geanalyseerde diagnosegroepen niet adequaat voor casemixverschillen op ICD-9-hoofddiagnoseniveau; model 2 corrigeerde niet adequaat voor 25 diagnosegroepen. De c-statistics voor model 2 waren beter dan voor model 1. De verschillen in SMR-uitkomsten tussen model 2 en model 1 varieerden van -63% tot 202%, afgemeten aan de uitkomst berekend met model 1. Bij de HSMR varieerden deze verschillen van -6,7% tot 4,3%.
- CONCLUSIES** Het huidige HSMR-model voor Nederland corrigeert onvoldoende voor casemixverschillen op het niveau van de ICD-9-hoofddiagnose. Een model met zwaarteklassenindeling van ICD-9-codes op basis van de geobserveerde ziekenhuismortaliteit in Nederland maakt een betere, zij het niet perfecte, correctie mogelijk. Ter verbetering van het HSMR-model wordt aanbevolen deze 'Nederlandse' zwaarteklassenindeling toe te passen.

De belangstelling in Nederland voor het gebruik van het gestandaardiseerde ziekenhuissterftecijfer, de 'hospital standardised mortality ratio' (HSMR), neemt de laatste jaren sterk toe. Dit hangt samen met de toenemende focus op het meten, verbeteren en transparant maken van de kwaliteit van zorg in ziekenhuizen. Het HSMR-model (zie uitlegkader) is voor Nederland ontwikkeld door Kiwa Prismant in samenwerking met Dr Foster Intelligence, het Engelse HSMR-kenniscentrum. Het model corrigeert voor patiëntkenmerken ('casemix') zoals leeftijd, geslacht en aandoening, ontleend aan de Landelijke Medische Registratie (LMR). In oktober 2010 verspreidde de stichting 'Dutch Hospital Data' (DHD) de (H)SMR-rapportage 2007-2009 met verdieping naar diagnosegroepen en patiëntencategorieën onder Nederlandse ziekenhuizen, met voor ieder ziekenhuis specifieke informatie. Ieder ziekenhuis kreeg zodoende inzicht in de eigen gestandaardiseerde sterfte ('standardised mortality ratios'; SMR's) behorende bij 50 diagnosegroepen (zie uitlegkader). Dit biedt de mogelijkheid gericht te werken aan kwaliteitsverbeteringen.

*St. Antonius Ziekenhuis,  
Bestuursstaf, Nieuwegein.*

*Ir. W.F. van den Bosch, kwaliteitsmanager.*

*Onderzoeksinstituut NIVEL, Utrecht.*

*Drs. P. Spreeuwenberg, statisticus.*

*VU Medisch Centrum,*

*afd. Sociale Geneeskunde, Amsterdam.*

*Prof.dr. C. Wagner, fysiotherapeut en socioloog*

*(tevens: NIVEL, Utrecht).*

*Contactpersoon: ir. W.F. van den Bosch*

*(w.bosch@antoniuziekenhuis.nl).*

## UITLEG

**De HSMR in het kort**

De 'hospital standardised mortality ratio' (HSMR) is de verhouding tussen de in een ziekenhuis waargenomen sterfte en de verwachte sterfte, die bepaald is door een vergelijking van patiëntkenmerken met het landelijke gemiddelde. De uitkomst wordt genormeerd op 100. Een HSMR boven de 100 duidt op een hogere sterfte dan gemiddeld, onder de 100 op een lagere sterfte. Als de berekening wordt toegepast op één van de 50 diagnosegroepen dan spreken we van de 'standardised mortality ratio' (SMR).

Het HSMR-model dat in 2010 in Nederland gebruikt werd, corrigeert voor de volgende **patiëntkenmerken**: jaartal van ontslag, leeftijd bij opname, geslacht, opname-urgentie, comorbiditeit (charlsonindex), sociaaleconomische status, maand van opname, verblijfplaats patiënt voorafgaand aan de opname, diagnosegroep bij opname en gedeeltelijk voor zwaarte van de aandoening binnen de diagnosegroep.

Iedere **diagnosegroep** is samengesteld uit een aantal onderliggende aandoeningen, aangeduid met 'International classification of diseases' (ICD)-9-codes zoals vastgelegd door de 'Clinical classification software' (CCS). Dit hulpmiddel clustert verschillende ICD-9-diagnosen in een hanteerbaar aantal betekenisvolle categorieën, aangeduid als 'CCS-diagnosegroepen'. Deze hebben een relatief hoge mortaliteit en omvatten tezamen meer dan 80% van de totale ziekenhuissterfte.

De Inspectie voor de Gezondheidszorg stimuleert dit HSMR-gebruik zoals beschreven in de basisset prestatie-indicatoren.<sup>1</sup> Verder is er een publieke behoefte aan het vergelijken van ziekenhuizen middels openbaar gemaakte ziekenhuissterftecijfers ([www.nfu.nl/index.php?id=166](http://www.nfu.nl/index.php?id=166)).<sup>2-4</sup> Er zijn echter ook twijfels over de betrouwbaarheid,<sup>5,6</sup> validiteit,<sup>7,8</sup> en bruikbaarheid van het model.<sup>9,10</sup> In hoeverre zijn de huidige HSMR- en SMR-uitkomsten nog vertekend door meetfouten?

In deze publicatie beschrijven wij voor Nederland het effect van meetfouten die kunnen optreden als niet adequaat wordt gecorrigeerd voor casemixverschillen binnen diagnosegroepen. In Medisch Contact is een voorbeeld van de invloed van casemixverschillen op de SMR beschreven voor de diagnosegroep 'cerebrovasculaire aandoening' (CVA).<sup>11</sup> Hierbij werd onvoldoende gecorrigeerd voor casemixverschil tussen bloedige CVA (hersensbloeding of subarachnoïdale bloeding; mortaliteit circa 35%) en niet-bloedige CVA (herseneninfarct; mortaliteit circa 10%). Bij inadequate correctie hiervoor bleken ziekenhuizen met verhoudingsgewijs veel patiënten met een hersensbloeding of subarachnoïdale bloeding in het nadeel ten opzichte van ziekenhuizen die veel patiënten met een herseneninfarct opnamen.

De oorzaak van dit verschijnsel hangt samen met de huidige correctiemethode, die is gebaseerd op een zwaarte-classificatie van de WHO. Deze vormt vermoedelijk geen adequate weergave van de feitelijke casemix in Neder-

land. In ons onderzoek vergeleken wij daarom de uitkomsten van het traditionele model met die van een model waarin de correctie was gebaseerd op de Nederlandse casemix. De centrale vraag was in hoeverre het aangepaste, Nederlandse model een verbetering zou inhouden. Dat onderzochten wij niet alleen voor de CVA-diagnosegroep, maar ook voor de overige diagnosegroepen.

**METHODE EN DATA**

Voor de modelanalyse werd, met toestemming van DHD, gebruik gemaakt van een geanonimiseerde LMR-dataset over de jaren 2005-2009 van 84 Nederlandse ziekenhuizen; ziekenhuizen werden geïncludeerd als de gegevens voldoende volledig in de LMR waren ingevoerd. Voor de HSMR-berekeningen gebruikten wij het 'Dr Foster Intelligence'-model zoals Kiwa Prismant dat vanaf 2009 heeft toegepast.<sup>12</sup> Vanaf 2010 tellen hierbij alleen nog maar klinische opnamen en dagopnamen met sterfte mee.

**CASEMIXCORRECTIE BINNEN DIAGNOSEGROEPEN**

Een goed standaardisatiemodel corrigeert onder andere voor de verschillen in ernst van aandoening die tot uiting komen in verschillen in de hoofddiagnosen volgens de 'International classification of diseases' (ICD)-9. Er kunnen meer dan 1000 verschillende ICD-9-diagnosecodes voorkomen bij een HSMR-berekening. Deze worden geclusterd in 50 diagnosegroepen volgens de 'Clinical classification software' (CCS; [www.hcup-us.ahrq.gov/toolsoftware/ccs/ccsfactsheet.jsp#what](http://www.hcup-us.ahrq.gov/toolsoftware/ccs/ccsfactsheet.jsp#what)). De casemixcorrectie vond plaats op het niveau van deze diagnosegroepen, waardoor binnen een CCS-diagnosegroep geen onderscheid gemaakt werd tussen ernstige en minder ernstige aandoeningen op ICD-9-niveau.

Om vertekening door de clustering tegen te gaan, heeft Dr Foster Intelligence een extra correctievariabele ingevoerd die aan iedere diagnose bij opname een zwaarteklasse toekent, afhankelijk van de ICD-9-hoofddiagnose. De zwaarte varieert van 1 (licht) tot 7 (zwaar) en is gebaseerd op een indeling naar sterfterisico van de WHO. Voor een aantal ICD-9-codes heeft Dr Foster Intelligence om onbekende redenen geen extra correctie ingevoerd; opnamen met deze ICD-9-codes krijgen in de classificatie een waarde 0. Bij deze waarde vindt geen extra correctie plaats.

**DE MODELLEN VERGELEKEN**

In ons onderzoek draaide het om de indeling van ICD-9-codes in zwaarteklassen. Wij onderzochten in hoeverre de zwaarteclassificatie in het 'Dr Foster Intelligence'-model spoort met de geobserveerde ziekenhuismortaliteit van ICD-9-codes in Nederland en in hoeverre dit

invloed heeft op de HSMR en SMR. Daartoe berekenden wij met 2 modellen de HSMR- en SMR-waarden van de 84 ziekenhuizen over 2005-2009. De modellen waren identiek op het volgende verschil na:

- (a) model 1 maakt gebruik van de zwaarteclassificatie volgens Dr Foster Intelligence;
- (b) model 2 maakt gebruik van een zwaarteclassificatie die gebaseerd was op de ziekenhuismortaliteit in Nederland in de jaren 2005-2009 per voorkomende ICD-9-code.

Tabel 1 toont de mortaliteitsgrenzen die wij kozen voor de indeling in 7 zwaarteklassen. Voor een beperkt aantal ICD-9-codes stelden wij de zwaarteklasse in verband met representativiteit op 0, omdat die codes door te weinig ziekenhuizen (minder dan 5) waren gebruikt.

**BEOORDELING CASEMIXVERSCHILLEN**

Om de invloed van casemixverschillen binnen een CCS-diagnosegroep aan te kunnen tonen, verdeelden wij per CCS-diagnosegroep alle in die groep meetellende opnamediagnosen over:

- (a) een ‘hoogrisico’-subgroep van ICD-9-hoofddiagnosen met hogere ruwe ziekenhuismortaliteit;
- (b) een ‘laagrisico’-subgroep van ICD-9-hoofddiagnosen met lagere ruwe ziekenhuismortaliteit.

De ICD-9-codes werden zodanig ingedeeld in ‘hoog risico’ of ‘laag risico’ dat het aantal sterfgevallen binnen een diagnosegroep bij benadering evenredig verdeeld was over deze 2 subgroepen. De vraag was in hoeverre een onevenredig groot aantal patiënten in een van de 2 subgroepen invloed had op de SMR van een individueel ziekenhuis.

Bij de analyse keken we naar de verhouding (ratio) van de mortaliteit van de hoogrisico-subgroep en de laagrisico-subgroep per CCS-diagnosegroep voor alle ziekenhuizen tezamen. Zonder standaardisatie is deze ratio groter dan 1, maar na perfecte standaardisatie zou de ratio 1 moeten bedragen. Een ziekenhuis ondervindt dan geen voor- of nadeel bij relatief veel of weinig opnamen in een van de 2 subgroepen.

Wij bepaalden voor beide modellen per CCS-diagnosegroep in hoeverre deze ratio van 1 afweek. Als een model statistisch significant afweek van 1 (dat wil zeggen: 1 lag buiten het 95%-BI van de ratio), dan concludeerden wij dat dat model niet adequaat corrigeerde voor het casemixverschil binnen die CCS-diagnosegroep.

**VERGELIJKING HSMR'S EN SMR'S**

Wij vergeleken de HSMR- en SMR-uitkomsten van beide modellen op 3 punten:

- (a) procentueel in hoeverre verschillen de SMR's berekend volgens model 1 en model 2 per CCS-diagnosegroep, per ziekenhuis?

**C-statistic**

Deze statistische maat geeft aan hoe goed een regressiemodel voorspelt. Daarbij wordt de sterftevoorspelling van iedere opname vergeleken met de uitkomst: wel of niet overleden. Bij een waarde van 0,5 is er geen voorspellende waarde: bij 50% van de opnamen was de voorspelling goed, bij 50% niet goed. Bij 1 is de voorspellende waarde perfect. De c-statistic van het HSMR-model dat in Nederland wordt toegepast ligt boven de 0,85.

- (b) significantiescore, dat wil zeggen: wijst de SMR op een sterfte die statistisch significant hoger, lager of niet afwijkend van het gemiddelde is?

- (c) voorspellende waarde, uitgedrukt in ‘c-statistic’ per CCS-diagnosegroep (zie uitlegkader).

De procentuele verschillen in uitkomst tussen de modellen 1 en 2 werden berekend met de formule  $SMR_{\text{delta}} = (SMR_{\text{model 2}} / SMR_{\text{model 1}} - 1) \times 100\%$ . Deze verschillen werden weergegeven in een frequentieverdeling. Hetzelfde gebeurde voor de HSMR's.

**RESULTATEN**

Wij analyseerden 48 CCS-diagnosegroepen die een zinvolle splitsing in hoogrisico- en laagrisico-ICD-9-codes toelieten. Het ging hierbij om 2.851.973 opnamen met 149.279 sterfgevallen. Van 2 CCS-diagnosegroepen, ‘prostaatcancer’ en ‘aspiratiepneumonie door voedsel of braken’, was splitsing niet zinvol en deze vielen dus af.

**TABEL 1** Indeling in zwaarteklassen\* voor ICD-9-codes, gebaseerd op de mortaliteit van aandoeningen in Nederlandse ziekenhuizen.

zwaarteklasse	mortaliteitsgrenzen in %‡	
	ondergrens	bovengrens
0§		
1	0,00	0,41
2	0,41	1,02
3	1,02	2,56
4	2,56	6,40
5	6,40	16,00
6	16,00	40,00
7	40,00	100,00

\*Bij de berekening van de ‘hospital standardized mortality ratio’ wordt gecorrigeerd voor de zwaarteklasse van de diagnose bij opname.  
 ‡De ondergrens is steeds 40% van de bovengrens, beginnend bij een bovengrens van 100%; de indeling van onder- en bovengrenzen is arbitrair, andere indelingen zijn ook mogelijk.  
 §Bij opnamediagnosen in deze zwaarteklasse vindt geen extra correctie plaats.

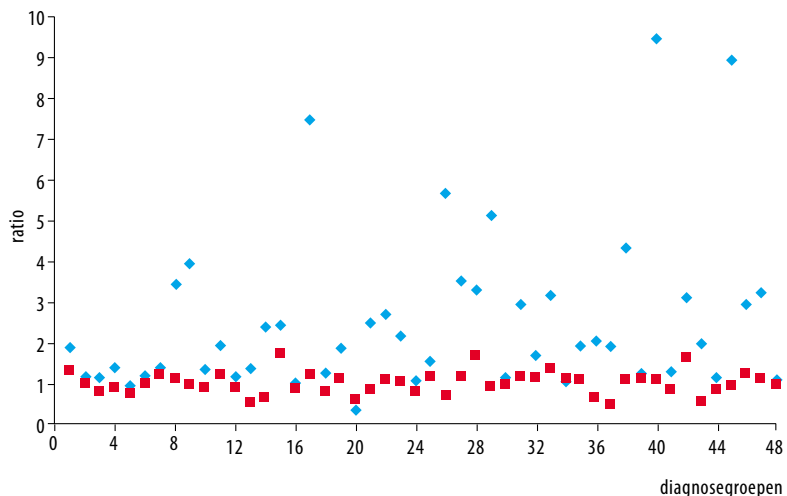
**RATIOS VAN GESTANDAARDISEERDE STERFTECIJFERS**

Figuur 1 toont een puntendiagram met de ratios van hoogrisico- en laagrisico-gestandaardiseerde sterftecijfers voor 48 CCS-groepen voor beide modellen. Bij model 1 weken 40 van de 48 ratios statistisch significant af van de waarde 1 en bij model 2 weken 25 ratios af. De afwij-

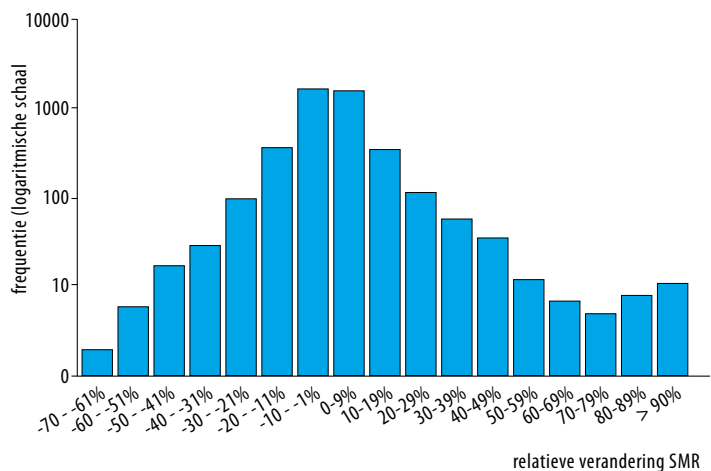
kingen van model 1 waren zichtbaar groter dan die van model 2.

**VERSCHILLEN IN HSMR EN SMR**

Wij berekenden de procentuele verschillen tussen SMR's ( $SMR_{\text{delta}}$ ) voor 48 diagnosegroepen en 84 ziekenhuizen;



**FIGUUR 1** Puntendiagram van de verhouding (ratio) in gestandaardiseerd sterftecijfer tussen patiënten met een hoogrisico-diagnose en die met een laagrisico-diagnose, berekend volgens een landelijk gebruikt model (blauwe ruit) of volgens een model aangepast aan de ziekenhuismortaliteit in Nederland (rood vierkant). De ratios werden berekend per diagnosegroep; binnen elk van de 48 diagnosegroepen waren meerdere diagnoses (ICD-9-codes) geclusterd. De ratio is 1 als het model optimaal corrigeert voor het sterfterisico van elke ICD-9-code. Zonder standaardisatie is de ratio groter dan 1.



**FIGUUR 2** Frequentieverdeling van het verschil in de 'standardized mortality ratio' ( $SMR_{\text{delta}}$ ) tussen de SMR berekend met het gebruikelijke model (gebaseerd op sterfterisico's volgens de WHO) en de SMR berekend met een 'Nederlands' model, dat wil zeggen: een model gebaseerd op ziekenhuismortaliteit in Nederland. Het verschil werd uitgedrukt als procentuele verandering ten opzichte van de SMR berekend met het gebruikelijke model. Verschillen werden berekend voor 48 diagnosegroepen met gegevens van 84 ziekenhuizen (totaal: 4032 verschillen).

dit leverde 4032 uitkomsten op. Figuur 2 geeft de frequentieverdeling van de verschillen. Op dezelfde wijze maakten wij een frequentieverdeling van de verschillen tussen de HSMR's van de 84 ziekenhuizen, berekend met model 1 en model 2 (figuur 3). De uiterste waarden van de  $SMR_{\text{delta}}$  waren -63% en 202% (SD: 14,6). De uiterste waarden van de HSMR's waren -6,7% en 4,3% (SD: 2,0).

Voor beide modellen werd de 'c-statistic' per CCS-diagnosegroep bepaald (tabel 2). Bij 38 diagnosegroepen vertoonde model 2 een hogere c-statistic dan model 1; de verschillen varieerden van 0,1% tot 17,5%. Bij 8 diagnosegroepen was de c-statistic van model 2 wat lager.

Tot slot stelden wij per diagnosegroep vast bij hoeveel ziekenhuizen de SMR volgens het ene model wees op een statistisch significant hogere of lagere sterfte en volgens het andere model niet (zie tabel 2). Model 1 gaf bij 391 van de 4032 SMR's aan dat de sterfte statistisch significant verhoogd was ten opzichte van het gemiddelde. Bij 87 hiervan (22%) was de sterfte volgens model 2 niet hoger dan gemiddeld. Model 2 gaf bij 352 SMR's significant hogere sterfte aan; bij 48 hiervan (14%) was de sterfte volgens model 1 niet verhoogd.

## BESCHOUWING

In dit onderzoek werden de SMR- en HSMR-uitkomsten van 2 modellen onderling vergeleken. Deze uitkomsten bleken substantieel te verschillen op het niveau van CCS-diagnosegroepen, zoals te zien aan de volgende waarnemingen:

(a) De relatieve verschillen tussen de SMR's van model

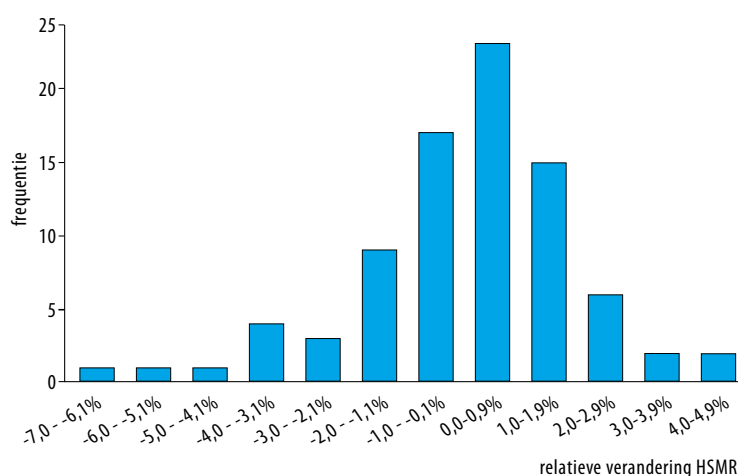
1 en die van model 2 varieerden aanzienlijk, van -63% tot 202%. Dit betekent dat het modeltype dat werd toegepast grote invloed had op de waarde van de SMR's van ziekenhuizen, onafhankelijk van de kwaliteit van zorg. Het volgende rekenvoorbeeld laat zien wat de grootte van de relatieve verschillen betekent: stel  $SMR_{\text{model 1 op 100 SMR-punt}}$ ; als  $SMR_{\text{model 2}} = 50$  SMR-punt, dan bedraagt het verschil -50%, en als  $SMR_{\text{model 2}} = 200$  SMR-punt, dan is het verschil +100%. De standaarddeviatie van de verdeling van relatieve verschillen bedroeg 14,6%; ter vergelijking: de standaarddeviatie van de verdeling van de 4032 SMR's bedroeg 45 SMR-punt.

(b) De modellen verschilden bij het aanwijzen van significant hoge sterfte. 22% van de SMR's die volgens model 1 op significant verhoogde sterfte wezen, deden dat niet volgens model 2. Omgekeerd was 14% van de SMR's die volgens model 2 op verhoogde sterfte wezen, volgens model 1 juist niet verhoogd. Dit betekent dat 'best practices' – net als rangordes van ziekenhuizen – die gebaseerd zijn op SMR-berekeningen, kunnen verschillen afhankelijk van de modelkeuze.

Op HSMR-niveau waren de verschillen minder sterk dan op SMR-niveau.

Welk model is nu 'beter' en waarom? Wij menen dat model 2 een betere, zij het niet perfecte correctie bewerkstelligt in de Nederlandse situatie, op grond van de volgende overwegingen:

(a) De verdeling van de ratios van hoogrisico- en laagrisico gestandaardiseerde sterfte had bij model 2 een geringere spreiding (0,47-1,7) dan bij model 1 (0,35-9,5)



**FIGUUR 3** Frequentieverdeling van het verschil in de 'hospital standardized mortality ratio' (HSMR) van 84 ziekenhuizen berekend volgens 2 modellen: het gebruikelijke model (gebaseerd op sterfterisico's volgens de WHO) en een 'Nederlands' model gebaseerd op ziekenhuismortaliteit in Nederland. Het verschil werd uitgedrukt als procentuele verandering ten opzichte van de HSMR berekend met het gebruikelijke model.

**TABEL 2** Voorspellende waarde van 2 modellen voor het berekenen van 'standardized mortality ratios' (SMR), uitgedrukt als 'c-statistic'\* per diagnosegroep, en de significantiescores van deze 2 modellen

CCS-code	CCS-diagnosegroep	c-statistic* per diagnosegroep		verschillen in SMR-uitslagen per diagnosegroep; n‡			
		model 1†	model 2†	lage sterfte		hoge sterfte	
				1: niet 2: wel	1: wel 2: niet	1: niet 2: wel	1: wel 2: niet
2	sepsis (behalve tijdens bevalling)	0,720	0,760	0	2	2	2
12	slokdarmkanker	0,723	0,723	1	0	0	1
13	maagkanker	0,735	0,731	0	1	2	2
14	colonkanker	0,779	0,776	1	3	2	1
15	kanker van rectum of anus	0,775	0,772	0	0	0	2
17	alvleesklierkanker	0,688	0,688	0	0	0	0
19	longkanker/bronchus kanker	0,793	0,792	0	0	0	4
24	borstkanker	0,921	0,918	3	6	3	3
32	blaaskanker	0,859	0,873	1	4	2	1
38	non-hodgkinlymfoom	0,792	0,793	0	1	2	0
39	leukemie	0,772	0,775	0	0	0	1
42	secundaire maligniteiten	0,744	0,748	0	1	0	1
44	niet nader gespecificeerde tumoren of tumoren van onzekere aard	0,808	0,812	3	1	2	3
50	diabetes mellitus met complicatie	0,804	0,821	0	2	0	2
55	aandoeningen van vocht- en elektrolythuishouding	0,770	0,796	1	1	0	1
59	deficiënties en andere anemie	0,741	0,757	1	0	0	0
85	coma, stupor en hersenbeschadiging	0,615	0,790	0	2	0	1
96	hartklepaandoeningen	0,759	0,781	3	3	0	0
100	acuut myocardinfarct	0,716	0,743	2	4	3	4
101	coronaire atherosclerose en overige hartaandoeningen	0,751	0,786	2	4	0	1
103	pulmonale hypertensie	0,741	0,755	0	2	1	3
106	hartritmestoornis	0,792	0,828	1	5	4	3
107	hartstilstand en ventrikelfibrilleren	0,624	0,742	3	4	2	3
108	hartfalen, zonder hypertensie	0,636	0,638	5	2	1	4
109	acute cerebrovasculaire aandoening	0,730	0,744	0	2	1	1
114	perifere en viscerale atherosclerose	0,876	0,909	1	3	3	13
115	aneurysmata van aorta en perifere en viscerale arteriën	0,839	0,874	1	2	4	3
116	trombose van aorta en perifere en viscerale arteriën	0,869	0,881	0	1	0	3
117	overige circulatoire aandoeningen	0,778	0,841	2	3	0	1
122	pneumonie	0,750	0,752	0	1	0	1
127	COPD en bronchitis	0,666	0,698	1	4	0	2
130	pleuritis, pneumothorax, klaplong	0,805	0,815	0	1	0	0

en bevatte minder vaak significante afwijkingen van 1 (25 bij model 2 versus 40 bij model 1).

(b) De c-statistics van model 2 zijn aanzienlijk beter dan die van model 1.

(c) De indrukvaliditeit ('face validity') van model 1 is beperkt. Zo is in figuur 4 op het eerste gezicht al duidelijk dat de zwaarteklassen volgens model 1 niet correleren met de mortaliteit in Nederlandse ziekenhuizen voor de afzonderlijke diagnoses binnen de CCS-diagnosegroepen 'Overige circulatoire aandoeningen' en 'Overige leveraandoeningen'; model 2 vertoont wel een goede correlatie. Andere CCS-diagnosegroepen vertonen soortgelijke patronen. Mogelijk verschillen Nederlandse mortaliteitspatronen van die waarop de WHO-zwaarteklassen zijn gebaseerd.

HSMR's en SMR's worden steeds vaker gebruikt bij het verbeteren van kwaliteit van zorg in ziekenhuizen. Vooral diagnosegroepen met significant hoge sterfte worden kritisch doorgelicht door zorgprofessionals. Hierbij is van belang dat de juiste diagnosegroepen worden aangepakt. Uit ons onderzoek blijkt echter dat het huidige model voor Nederland niet adequaat corrigeert voor casemixverschillen binnen CCS-diagnosegroepen. Daardoor kan een substantieel aantal SMR's ten onrechte aangemerkt worden als statistisch significant hoger dan het gemiddelde. Omgekeerd wordt ook een substantieel aantal SMR's aangemerkt als niet-significant verhoogd, terwijl de verhoging vermoedelijk wél statistisch significant is. Ziekenhuizen hebben derhalve nog onvoldoende zicht op sterke en zwakke pun-



TABEL 2 (vervolg)

CCS-code	CCS-diagnosegroep	c-statistic* per diagnosegroep		verschillen in SMR-uitslagen per diagnosegroep; n‡			
				lage sterfte		hoge sterfte	
		model 1†	model 2†	1: niet 2: wel	1: wel 2: niet	1: niet 2: wel	1: wel 2: niet
133	overige aandoeningen van de lage luchtwegen	0,822	0,851	2	2	1	0
145	darmobstructie zonder hernia	0,796	0,814	0	0	0	1
146	diverticulose en diverticulitis	0,805	0,812	0	0	1	2
149	galwegaandoening	0,906	0,898	0	0	1	1
150	leveraandoening, alcoholgerelateerd	0,666	0,679	1	0	1	2
151	overige leveraandoeningen	0,665	0,772	2	1	3	4
153	gastro-intestinale bloeding	0,707	0,739	2	2	0	2
155	overige gastro-intestinale aandoeningen	0,862	0,910	1	1	1	2
157	acuut en niet-gespecificeerd nierfalen	0,716	0,722	0	0	2	0
158	chronisch nierfalen	0,850	0,851	0	1	1	0
159	urinewegsinfectie	0,817	0,816	0	1	0	0
226	heupfractuur	0,725	0,722	1	0	1	0
233	intracraniaal letsel	0,836	0,940	4	11	2	2
237	complicatie van apparatuur, implantaat of transplantaat	0,798	0,823	0	1	0	2
238	complicatie chirurgische procedure of medische zorg	0,779	0,830	0	3	0	2
249	shock	0,715	0,717	1	0	0	0
			<b>totaal</b>	<b>46</b>	<b>88</b>	<b>48</b>	<b>87</b>

CCS = 'Clinical classification software', een hulpmiddel om verschillende ICD-9-diagnosen in diagnosegroepen te clusteren.

\*'C-statistic' geeft aan hoe goed een regressiemodel voorspelt. Bij een waarde van 0,5 heeft het model geen voorspellende waarde, bij 1,0 is de voorspellende waarde perfect.

†Model 1 is het gebruikelijke model voor de berekening van de 'standardized mortality ratio' (SMR), gebaseerd op sterfterisico's volgens de WHO. Model 2 is een aangepast 'Nederlands' model, gebaseerd op ziekenhuismortaliteit in Nederland.

‡Per diagnosegroep werden SMR's van 84 ziekenhuizen berekend. Weergegeven zijn de aantallen ziekenhuizen waarbij de SMR volgens het ene model wees op een significant lagere of hogere sterfte dan gemiddeld, en volgens het andere model niet.

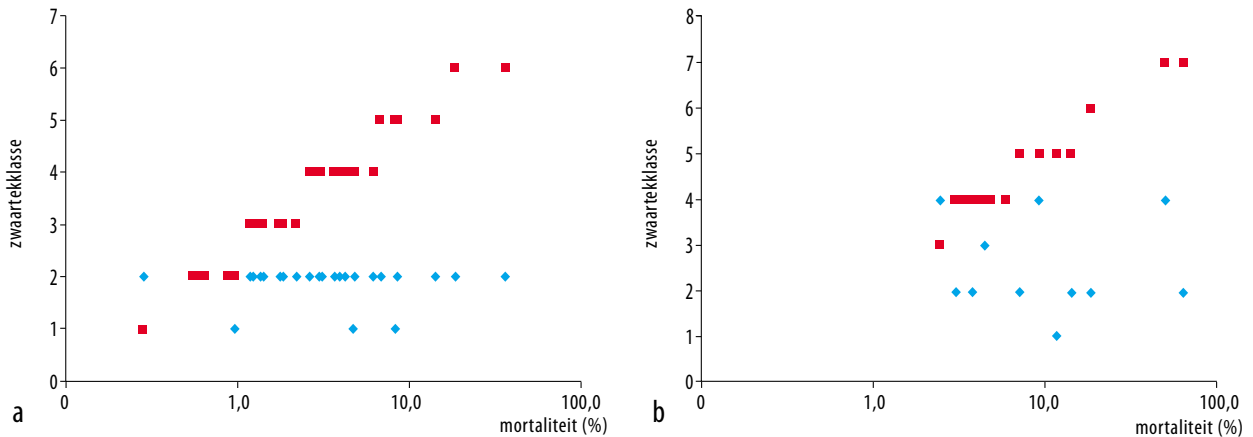
ten van de verleende zorg. Bij de analyse van gestandaardiseerde ziekenhuissterftecijfers in de rapportages over 2007-2009, dient men hierop bedacht te zijn; tabel 2 kan hierbij helpen.

In dit onderzoek bleek één bron van ongecorrigeerde variatie de uitkomsten van SMR- en HSMR-berekeningen al te kunnen vertekenen. Eerder onderzoek heeft laten zien dat variaties in codering,<sup>5</sup> het effect van bijzondere medische verrichtingen,<sup>7</sup> de onvergelykbaarheid van context,<sup>8</sup> variaties in ziekenhuissterfte van terminaal geïndiceerde patiënten en het effect van heropnamen eveneens van invloed kunnen zijn op de HSMR van een ziekenhuis.<sup>13,14</sup> Daarom is er behoefte aan een beter inzicht in de aard en impact van ongecorrigeerde variatiebronnen voordat men de HSMR-

methodiek betekenisvol kan inzetten, zoals bij het werken aan verbetering van kwaliteit en patiëntveiligheid in ziekenhuizen. Deze behoefte wordt steeds groter naarmate het HSMR-model verder wordt ontwikkeld en de besluitvorming rond openbaarmaking van het gestandaardiseerde ziekenhuissterftecijfer voortschrijdt.

## CONCLUSIES

Het huidige HSMR-model voor Nederland corrigeert onvoldoende voor casemixverschillen op het niveau van ICD-9-hoofddiagnosen. De oorzaak hangt samen met een inadequate indeling van ICD-9-codes in zwaarteklassen. Een betere correctie wordt mogelijk na



**FIGUUR 4** Verband tussen de mortaliteit en de zwaarteklasse volgens model 1 (blauwe ruit) en model 2 (rood vierkant) van ICD-9-diagnosen in (a) de diagnosegroep 'Overige circulatoire aandoeningen' en (b) de diagnosegroep 'Overige leveraandoeningen'. De X-as heeft een logaritmische schaal.

#### LEERPUNTEN

- Voor de berekening van gestandaardiseerde ziekenhuissterftecijfers (HSMR en SMR) wordt een model gebruikt dat corrigeert voor vertekende factoren.
- Een van de vertekende factoren is de ernst van de diagnose bij opname; correctie vindt plaats door diagnoses in te delen in zwaarteklassen.
- De indeling van diagnoses in zwaarteklassen is in het huidige HSMR-model gebaseerd op sterfterisico's volgens de WHO.
- Een betere correctie is mogelijk met een HSRM-model met zwaarteklassen die zijn gebaseerd op de Nederlandse ziekenhuismortaliteit.
- Bij gebruik van de gestandaardiseerde sterftecijfers die in 2010 in Nederland zijn verspreid, dient men er rekening mee te houden dat mogelijk niet goed gecorrigeerd is voor casemixverschillen binnen diagnosegroepen.

aanpassing van deze indeling op geleide van de ziekenhuismortaliteit in Nederland. Bij vergelijking van SMR-waarden volgens het oorspronkelijke model en volgens een nieuwe berekening springen substantiële verschillen op CCS-diagnosegroepsniveau in het oog. Op ziekenhuisniveau (HSMR-waarden) zijn de verschillen beperkt. Ter verbetering van het HSMR-model bevelen wij aan de zwaarteklassenindeling te baseren op de Nederlandse ziekenhuismortaliteit.

Belangenconflict: geen gemeld. Financiële ondersteuning: geen gemeld.

Aanvaard op 11 mei 2011

Citeer als: Ned Tijdschr Geneeskd. 2011;155:A3299

[Meer op www.ntvg.nl/onderzoek](http://www.ntvg.nl/onderzoek)

#### LITERATUUR

- 1 Inspectie voor de Gezondheidszorg. Veiligheidsindicatoren ziekenhuizen 2010 t/m 2012. Utrecht: IGZ; 2010.
- 2 Tweede Kamer der Staten-Generaal. Antwoorden op kamervragen van Koser Kaya over het bericht dat ziekenhuizen zwijgen over hun sterftecijfers. Aanhangsel der Handelingen. 2008-2009, 1422. Den Haag: SDU, 2009.link
- 3 Sterftecijfer nog niet publiek in 2010. Med Contact. 2009;64:1320.
- 4 Sterftecijfers: ja of nee? Elsevier, 24 oktober 2009.
- 5 Van den Bosch WF, Rozenendaal KJ, Silberbusch J, Wagner J. Variatie in codering patiëntgegevens beïnvloedt gestandaardiseerd ziekenhuissterftecijfer (HSMR). Ned Tijdschr Geneeskd. 2010;154:A1189.
- 6 Pieter D, Kool RB, Westert GP. Beperkte invloed gegevensregistratie op gestandaardiseerd ziekenhuissterftecijfer (HSMR). Ned Tijdschr Geneeskd. 2010;154:A2186.
- 7 Van den Bosch WF, Graafmans WC, Pieter D, Westert GP. Hartcentra en het effect van bijzondere medische verrichtingen op de gestandaardiseerde ziekenhuissterfte. Ned Tijdschr Geneeskd. 2008;152:1221-7.
- 8 Mohammed MA, Deeks JJ, Girling A, et al. Evidence of methodological bias in hospital standardised mortality ratios: retrospective database study of English hospitals. BMJ. 2009;338:b780.
- 9 Lilford R, Mohammed MA, Spiegelhalter D, Thomson R. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. Lancet. 2004;363:1147-54.



- 10 Penfold RB, Dean S, Flemons W, Moffatt M. Do hospital standardized mortality ratios measure patient safety? HSMRs in the Winnipeg Regional Health Authority. *Healthc Pap.* 2008;8:8-24.
- 11 Pleizier CM, Geerlings W, Pieter D, Boiten J. Patiëntenmix beïnvloedt HSMR. *Medisch Contact.* 2010;65:1777-9.
- 12 Jarman B, Pieter D, van der Veen AA, et al. The Hospital Standardised Mortality Ratio: a powerful tool for Dutch hospitals to assess their quality of care? (2009). *Qual Saf Health Care.* 2010;19:9-13.
- 13 Black N. Assessing the quality of hospitals. *BMJ.* 2010;340:c2066.
- 14 Van den Bosch WF, Kelder HC, Wagner C. Predicting hospital mortality among frequently readmitted patients: HSMR biased by readmission. *BMC Health Serv Res.* 2011;11:57.