



Paper

Van handmatig naar automatisch coderen van doodsoorzaken

Een bridge coding study

Dr. P.P.M. Harteloh
December 2015

Inhoud

Samenvatting 3

1. Inleiding 3

2. Van handmatig naar het automatisch coderen van doodsoorzaken 4

2.1 Het handmatig coderen van doodsoorzaken 5

2.2 Het automatisch coderen van doodsoorzaken 6

2.3 Bridge coding study 7

3. Resultaten van de 'bridge coding study' 8

3.1 Het vóórkomen van doodsoorzaken bij handmatig en automatisch coderen 9

3.2 De overeenstemming tussen handmatig en automatisch coderen 19

3.3 Het handmatig en automatisch coderen van de European Shortlist (ESL) 23

3.4 Het vergelijken van automatisch en handmatig coderen 26

3.5 De kwaliteit van de doodsoorzakenstatistiek 28

4. Conclusies 28

5. Begrippen 29

6. Referenties 33

Samenvatting

Het Centraal Bureau Statistiek (CBS) is in 2013 overgegaan op het automatisch coderen van doodsoorzaken. Hiermee wordt een betere internationale vergelijkbaarheid en stabiliteit in de tijd (trends) van de doodsoorzakenstatistiek nagestreefd. De overgang op het automatisch coderen is een breuk met een meer dan honderd jaar oude traditie van het handmatig coderen van doodsoorzakenformulieren en het is de vraag welke gevolgen deze methodewijziging voor de doodsoorzakenstatistiek heeft. Om deze vraag te beantwoorden is het jaarbestand 2009 zowel handmatig als automatisch gecodeerd. De overeenstemming tussen handmatig en automatisch coderen bleek samen te hangen met het detail niveau van de ICD-10 code. Bij ruim 78 procent van de sterfgevallen is er (perfecte) overeenstemming tussen automatisch en handmatig coderen wat betreft onderliggende doodsoorzaak op vier digit niveau van de ICD-10 code, het niveau waarop CBS gegevens aanlevert aan onderzoekers, Eurostat of de WHO. Op het drie digit niveau van de ICD-10 code (CBS Statline/nationale publicaties) is de overeenstemming bijna 85 procent en op het ICD-10 hoofdstukniveau is de overstemming tussen handmatig en automatisch coderen bijna 90 procent. Deze hoge mate van overeenstemming maakt de overgang op automatisch coderen verantwoord. De introductie van het automatisch coderen brengt echter wel (eenmalige) verschuivingen in de statistiek met zich mee. Er is een significante *toename* van de infectieziekten ('niet gespecificeerde diarree' en sepsis), uitdroging, dementie, de ziekte van Alzheimer en het laat gevolg van een CVA als onderliggende doodsoorzaak, en een significante *afname* van longontsteking, COPD, ziekten van het spijsverteringsstelsel ('niet gespecificeerd diarree'), urineweginfectie en ouderdom als onderliggende doodsoorzaak. Wat betreft het vóórkomen van belangrijke doodsoorzakengroepen als de (kwaadaardige) nieuwvormingen en hart- en vaatziekten is het verschil tussen het automatisch en handmatig coderen gering. Verklaringen voor de verschuivingen worden gevonden in: een andere interpretatie van selectieprincipes voor de onderliggende doodsoorzaak, een andere codering van doodsoorzaken, het al of niet hanteren van ICD-10 updates, technische beperking van IRIS (vooral bij combinatiecodes of de volgorde van tijdsintervallen) en het al of niet gebruiken van additionele informatie bij het coderen van een doodsoorzakenformulier. Door ingrijpen in de verwerking van doodsoorzakenformulieren (bijvoorbeeld voorselectie en/of handmatige verwerking van formulieren met opmerkingen), het plaatsen van codeervlaggen voor het omgaan met tijdsduren, en nacontrole op onderliggende doodsoorzaken zijn de ongewenste effecten van de overgang op automatisch coderen deels ongedaan te maken bij de introductie van IRIS in het (routine) productieproces van de doodsoorzakenstatistiek. Daarbij kan uiteindelijk tweederde van de formulieren volledig automatisch worden verwerkt en zal eenderde van een jaarbestand nog altijd handmatig ingrijpen van een CBS codeur behoeven om het doodsoorzakenformulier te verwerken.

1. Inleiding

Het Centraal Bureau Statistiek (CBS) is met ingang van het verslagjaar 2013 overgegaan op het gebruik van IRIS, een programma voor de automatische codering van doodsoorzakenformulieren dat inmiddels als *de* Europese standaard kan worden beschouwd (Pavillon

et al., 1998). Alvorens dit programma in gebruik is genomen, is een onderzoek verricht naar de gevolgen van automatisch coderen voor de doodsoorzakenstatistiek (Harteloh, 2014; Harteloh et al., 2014; Harteloh, 2015). Een verzameling handmatig gecodeerde doodsoorzakenformulieren (verslagjaar 2009) is ingevoerd en opnieuw gecodeerd met behulp van IRIS. Voor het volautomatisch verwerkte gedeelte van het bestand beschikken we dan per sterfgeval (record) over een handmatig en een automatisch toegekende ICD-10 code voor de onderliggende doodsoorzaak. We kunnen zo de resultaten van de twee verschillende methoden vergelijken. In dit verslag worden de verschillen in vóórkomen van onderliggende doodsoorzaken beschreven en zo mogelijk verklaard. Het onderzoek beperkt zich (noodzakelijkerwijs) tot het volautomatisch gecodeerde gedeelte van het bestand ofwel de verzameling volledig gecodeerde doodsoorzakenformulieren (ongeveer tweederde). Bij ongeveer een derde gedeelte van het bestand blijft handmatige interventie bij verwerking van een doodsoorzakenformulier een rol spelen. Het gedeelte van de doodsoorzakenformulieren dat bij voorbaat al voor handmatige interventie in aanmerking komt (niet-natuurlijk overlijden, doodgeborenen, euthanasie- en correspondentiegevallen) zijn niet in dit onderzoek opgenomen. Daardoor zijn de resultaten van dit onderzoek niet zonder meer vertaalbaar naar de jaarcijfers van het CBS. Het onderzoek levert een aantal factoren op die gebruikt kunnen worden bij een kwalitatieve verklaring van de verschillen tussen een handmatig en automatisch gecodeerd jaarbestand. Het levert ook kwantitatieve gegevens met behulp waarvan de uitkomsten van het handmatig coderen naar die van het automatisch coderen kunnen worden vertaald (en omgekeerd) om bij trendanalyse en/of het opstellen van tijdreeksen codeereffecten te onderscheiden van epidemiologische factoren. Het hoofdstuk 2 van dit rapport beschrijft de methode die wordt gevolgd bij het handmatig en bij het automatisch coderen van doodsoorzakenformulieren, alsmede de manier waarop beide methoden in dit onderzoek met elkaar worden vergeleken. Hoofdstuk 3 geeft de resultaten van het onderzoek weer en gaat in op de belangrijkste verklaringen voor de gevonden verschillen tussen beide methoden. Tot slot trekken we enige conclusies (hoofdstuk 4). Een uitgebreide verklaring van de belangrijkste begrippen in dit rapport wordt aangetroffen in hoofdstuk 5. De lezer die niet bekend is met het (automatisch) coderen van doodsoorzaken wordt aangeraden dit hoofdstuk als eerste te lezen.

2. Van handmatig naar het automatisch coderen van doodsoorzaken

Dit hoofdstuk beschrijft de methode die wordt gevolgd bij het handmatig en bij het automatisch coderen van doodsoorzakenformulieren, alsmede de manier waarop we beide methoden met elkaar zullen vergelijken.

2.1 Het handmatig coderen van doodsoorzaken

Het CBS kent een meer dan honderd jaar oude traditie van het handmatig coderen van doodsoorzaken door medisch geschoold personeel (codeurs) (Sonsbeek, 2005). Uitgangspunt is de opgave van de doodsoorzaak door een arts of forensisch geneeskundige middels het invullen van een doodsoorzakenformulier. Hierop worden handgeschreven doodsoorzaken vermeld en is de invuller vrij in formulering of notatievolgorde. De verwerking van deze formulieren geschiedt (centraal) door het CBS. Daarbij worden twee verschillende stappen genomen. Ten eerste worden medische uitdrukkingen op een doodsoorzakenformulier vertaald in ICD-10 codes (codering). Ten tweede wordt uit de causale reeks vermeld op het doodsoorzakenformulier de onderliggende doodsoorzaak geselecteerd (selectie). Bij handmatig coderen neemt een codeur van het CBS deze stappen door het doodsoorzakenformulier te lezen en de onderliggende doodsoorzaak in een bestand in te voeren. Naast de onderliggende doodsoorzaak kan de CBS codeur daarbij maximaal drie doodsoorzaken in het bestand invoeren die aan het overlijden hebben bijgedragen (secundaire doodsoorzaken). Vanuit dit bestand wordt de doodsoorzakenstatistiek gemaakt (1901–2012). Sinds 1996 gebruikt het CBS bij het maken van de doodsoorzakenstatistiek de ICD-10. De ICD-10 bestaat uit een systematiek (classificatie) van ziekten en aandoeningen (deel 1), aanwijzingen voor het coderen en selecteren van de onderliggende doodsoorzaak (deel 2), en een index van termen met de bijbehorende codes (deel 3). De ICD-10 kent ook een lijst met updates, bestaande uit jaarlijkse aanvullingen op de classificatie om tegemoet te komen aan medische ontwikkelingen (nieuwe ziekten of inzichten in causale mechanismen) of om onvolkomenheden in de classificatie te corrigeren. Het CBS heeft deze updates bij het handmatig coderen niet gevolgd en hanteert sinds 1996 de in dat jaar gepubliceerde versie van de ICD-10. Daarnaast is er door de codeurs en de medisch ambtenaar van het CBS in de loop der tijd een lijst met 1 024 regels opgesteld bestaande uit (i) aanwijzingen voor het coderen van medische termen die niet in de ICD-10 index voorkomen (ongeveer 700 regels, bijvoorbeeld: Balint-Holmes syndroom = G23.8), (ii) afspraken over het hanteren van een tijdsduur bij het coderen van medische termen (ongeveer 200 regels, bijvoorbeeld dat een leverfalen (niet nader gespecificeerd) als ‘acuut’ wordt gecodeerd indien deze korter dan vier weken bestaat en anders als ‘chronisch’), en (iii) regels voor het selecteren van de onderliggende doodsoorzaak als het deel 2 van de ICD-10 daarvoor geen specifieke instructies geeft (ongeveer 100 regels, veelal over causale verbanden, bijvoorbeeld dat een hartklepaandoening een hartinfarct kan veroorzaken). De lijst van het CBS fungeert als aanvulling op de delen 1, 2 of 3 van de ICD-10. De lijst bevat (nationale) inzichten die volgens het CBS moeten worden gehanteerd bij het coderen van doodsoorzaken en is aan een continue veranderingen onderhevig geweest. Bij overgang op automatisch coderen is deze lijst gedeeltelijk in het woordenboek (i) of het vlaggensysteem van IRIS (ii) verwerkt, zodat daarmee de vigerende Nederlandse inzichten ook in de automatisch codering van doodsoorzaken tot uitdrukking komen. Als het gaat om causale verbanden (iii) zal veelal de (internationale) opvatting gevolgd gaan worden, omdat deze is vastgelegd in de programmatuur van IRIS. Evidente fouten in de programmatuur worden ondervangen door nacontrole op codes of voorselectie van doodsoorzakenformulieren, waarbij de automatisch toegekende code door de codeur zo nodig handmatig kan worden veranderd.

De voor- en nadelen van het automatisch coderen

Voordelen:

Toegenomen betrouwbaarheid/consistentie bij het coderen en selecteren van de onderliggende doodsoorzaak

Verbetering van de internationale vergelijkbaarheid van doodsoorzakenstatistieken

Selectie van een onderliggende doodsoorzaak is expliciet en gedocumenteerd (per record/sterfgeval) middels log files

Gemakkelijke implementatie van ICD-10 updates of revisies

Mogelijkheid tot 'bridge coding studies' om verschuivingen in statistiek te bestuderen en te documenteren bij veranderingen in de ICD-10 of het woordenboek (gemakkelijker en snel uit te voeren)

Verminderde werkbelasting van menselijke codeurs (twee derde bestand volautomatisch)

Verschuiving in werkzaamheden codeurs: van louter codering naar kwaliteitscontrole

Toegenomen snelheid van verwerking en/of terugkoppeling aan invullend artsen

Verbetering van data opslag (complete tekst van certificaten plus corresponderende codes)

'Multiple cause coding' (Iedere term op het formulier wordt gecodeerd)

Mogelijkheid tot gedetailleerde analyse van certificaten (op tekst of codes)/toename van kennis over causale mechanismen bij overlijden

Nadelen:

Traag mechanisme van aanpassing bij fouten of onvolkomenheden in de IRIS beslistabellen (via het internationale circuit: IRIS User Group en/of WHO Mortality Reference Group)

Verschuivingen in statistiek bij introductie automatisch coderen.

Overcodering/ondercodering van doodsoorzaken door gebruik IRIS beslistabellen

Grotere afhankelijkheid van correct ingevuld en aangeleverd doodsoorzakenformulier

Grotere afhankelijkheid van plaats medische term op doodsoorzakenformulier bij selectie onderliggende doodsoorzaak

Technische afhankelijkheid van bouwers (IRIS Core Group)

2.2 Het automatisch coderen van doodsoorzaken

Sinds de jaren zeventig van de vorige eeuw is er in de Verenigde Staten software ontwikkeld en gebruikt voor het automatisch coderen van doodsoorzaken, 'MMDS' genaamd (Birch, 1993). Het software pakket IRIS heeft deze Amerikaanse software bruikbaar gemaakt voor niet-Engelstalige landen (Pavillon & Johansson, 2001). IRIS wordt door het CBS met ingang van het statistiekjaar 2013 gebruikt voor het automatisch coderen van doodsoorzaken. Het handgeschreven doodsoorzakenformulier vormt nog altijd het uitgangspunt. Dit wordt middels data entry omgezet in elektronische vorm (overgetypt). Bij het automatisch coderen van doodsoorzaken worden vervolgens vier stappen genomen. Ten eerste worden medische uitdrukkingen op een doodsoorzakenformulier omgezet in standaardtermen. Hierbij worden reguliere expressies gebruikt, ofwel programmeerregels die spelvarianten van woorden lezen en omzetten in een standaardterm. Vervolgens worden deze standaardtermen met behulp van een woordenboek van ICD-10 codes voorzien. Deze codes worden automatisch

gecontroleerd, bijvoorbeeld naar leeftijd en geslacht, en eventueel omgezet naar gelang de duur van de ziekte of zo nodig gecombineerd met andere codes op het formulier. Tot slot bepaalt de programmatuur met behulp van beslistabellen gebaseerd op de ICD-10 selectieregels (deel 2 van de ICD-10) de onderliggende doodsoorzaak. Het automatisch coderen beoogt zo de reproduceerbaarheid en internationale vergelijkbaarheid van de doodsoorzakenstatistiek te vergroten. Voor de eerste twee stappen heeft het CBS ongeveer 600 reguliere expressies gebouwd en een woordenboek opgesteld bestaande uit bijna 3 600 standaarduitdrukkingen. Bij het opstellen van het woordenboek is gebruik gemaakt van de Nederlandstalige index van de ICD-10, de aanwezige lijst met codeeraanwijzingen voor CBS codeurs en een woordenboek dat door de collega's van de Vlaamse doodsoorzakenstatistiek is samengesteld. De lijst is nagekeken en gevalideerd door de medisch ambtenaar van het CBS. Het ontwikkelen van het woordenboek in de eigen taal is de verantwoordelijkheid van een gebruiker van IRIS. Voor de controle van de codes en de selectie van de onderliggende doodsoorzaak wordt gebruik gemaakt van de (ingebouwde) algoritmes van IRIS. Deze zijn gebaseerd op de ICD-10 codeer- en selectieregels zoals die in 1996 zijn opgesteld en bevatten ook de jaarlijkse updates van de ICD-10. Bij het werken met IRIS kunnen een volautomatische, een semiautomatische en een handmatige verwerking van doodsoorzakenformulieren worden onderscheiden. Bij volautomatische verwerking zijn alle medische uitdrukkingen op het doodsoorzakenformulier van een ICD-10 code voorzien en is er een onderliggende doodsoorzaak geselecteerd ('automatic final'). IRIS legt een formulier dat niet volautomatisch kan worden verwerkt voor aan een CBS codeur ('rejected'). Deze kan dan proberen de invoer (spelfouten/varianten, afkortingen, code van term) te corrigeren waarop het formulier alsnog door IRIS kan worden verwerkt (semiautomatische verwerking). Heeft de menselijke ingreep niet het gewenste resultaat dan kan de CBS codeur besluiten zelf handmatig de onderliggende doodsoorzaak in te voeren ('manual final').

2.3 Bridge coding study

Bij vergelijking van handmatig en automatisch coderen volgen we de methodiek van de 'bridge coding study'. Een 'bridge coding study' is gedefinieerd als: een onafhankelijke codering van dezelfde verzameling doodsoorzakenformulieren door twee verschillende methoden (handmatig of geautomatiseerd coderen) (Eurostat, 2013). Gangbare maten in bridge coding onderzoek zijn de 'Comparability Ratio' (CR) en het 'Perfect Compatibility Percentage' (PCP). Deze beide maten moeten in samenhang worden beoordeeld. Zij geven respectievelijk inzicht in de verschuivingen in statistiek in termen van frequenties (CR) en op het niveau van de individuele gevallen (PCP) (zie hoofdstuk 5).

In dit onderzoek vergelijken we de code voor de onderliggende doodsoorzaak zoals die automatisch aan een sterfgeval is toegekend door IRIS met de handmatig toegekende code. Het onderzoek is verricht op het jaar 2009, een representatief jaar in de doodsoorzakenstatistiek zonder noemenswaardige epidemieën, rampen of andere uitschieters. Het bevat 134 263 sterfgevallen. Een aantal doodsoorzakenformulieren (n = 10 196) is van onderzoek uitgesloten en niet overgetypt.¹⁾ Dit betrof ten eerste de doodgeborenen en de gevallen met een expliciete vermelding van een niet natuurlijke overlijden (aangekruist op het doodsoorzakenformulier).²⁾ Vooralnog ontbreekt in IRIS een

¹⁾ Bij inzet van IRIS in de routineproductie worden dit soort formulieren (niet-natuurlijke doodsoorzaak, doodgeboorte of nul-jarige, euthanasie, mogelijk correspondentiegeval of voorzien van aanvullende opmerking) door voorselectie aan de CBS codeur aangeboden voor inspectie en/of handmatige codering.

²⁾ Soms vermeldt de invuller een niet-natuurlijke dood op deel 1 van het doodsoorzakenformulier (zie figuur 5.1)

algoritme om bij dit soort sterfgevallen de onderliggende doodsoorzaak op correcte wijze te bepalen. Ten tweede zijn ook gevallen van euthanasie niet in het onderzoek opgenomen. Deze hebben een aparte plaats en vorm van vermelding op het doodsoorzakenformulier.³⁾ Ten derde zijn sterfgevallen waarbij correspondentie plaatsvond over de doodsoorzaak niet in het onderzoek opgenomen, omdat de codering mogelijk is veranderd op basis van informatie waarover IRIS bij het coderen ervan niet kan beschikken (ongeveer 3 000 gevallen per jaar). De resterende verzameling doodsoorzakenformulieren van het jaar 2009 (ongeveer 124 duizend) is letterlijk overgetypt (tekst zoals aangetroffen op formulier, inclusief spelfouten).

Na het overtypen van de formulieren (2010–2011) en de ontwikkeling van een woordenboek, reguliere expressies voor standaardisatie van de invoer (2012) en codeervlaggen (2013) is het jaarbestand 2009 in juni 2014 voor codering aan IRIS versie 4.4.1 (beschikbaar januari 2014) aangeboden.⁴⁾ Voordat verwerking door IRIS plaatsvond zijn er 142 reguliere expressies voor het zetten van komma's tussen duidelijk aparte ziekten (begrippen) en het standaardiseren van de opgegeven tijdsduren toegepast, alsmede 632 reguliere expressies voor het vertalen van een variëteit aan termen op het B-formulier in standaarduitdrukkingen. Met behulp van een woordenboek – versie juni 2014, incl. codeervlaggen – kan IRIS vervolgens deze standaarduitdrukking voorzien van een ICD-10 code en de onderliggende doodsoorzaak selecteren als er meer dan één doodsoorzaak op het doodsoorzakenformulier wordt gerapporteerd. Ongeveer twee derde van deze formulieren kon door IRIS op een dergelijke wijze van een onderliggende doodsoorzaak worden voorzien en werd 'automatic final' verklaard. De analyse is uiteindelijk uitgevoerd op het aantal formulieren dat door IRIS volledig is gecodeerd *en* dat zowel handmatig als automatisch een natuurlijke doodsoorzaak kreeg toegekend (n = 86 893).

3. Resultaten van de 'bridge coding study'

In dit hoofdstuk beschrijven we de overeenstemming en verschillen tussen handmatig en automatisch coderen op verschillend detailniveau van een ICD-10 code om vervolgens in te gaan op de belangrijkste verklaringen voor verschillen tussen beide methoden.⁵⁾ De subparagrafen beschrijven de verschuivingen tussen ICD-10 hoofdstukken in detail en zijn wellicht alleen interessant voor degene die statistiek over doodsoorzaken in het desbetreffende hoofdstuk gebruikt. De lezer die slechts geïnteresseerd is in het globale

en is er een 'natuurlijke dood' aangekruist. De aard van het overlijden wordt dan door het CBS veranderd. Dit soort formulieren zijn wel overgetypt, maar niet in de analyse opgenomen.

³⁾ Omdat een maligniteit bij euthanasie vaak de onderliggende doodsoorzaak is, is een relatief groot aantal formulieren met maligniteiten (6,7 procent van het totaal aantal maligniteiten) niet in het onderzoek opgenomen. Bij euthanasie vindt maar opgave van één doodsoorzaak plaats. Er is geen selectie van een onderliggende doodsoorzaak nodig. Het is dus niet aannemelijk dat uitsluiting van dit soort formulieren het resultaat van het onderzoek in negatieve zin zal hebben vertekend.

⁴⁾ Voor het verwerken van het jaarbestand 2013 is aanvankelijk IRIS versie 4.2.0 gebruikt. Deze versie was nog niet voorzien van codeervlaggen en technisch minder ver ontwikkeld dan versie 4.4.1. Het jaarbestand 2013 is echter in juni 2014 volledig (opnieuw) verwerkt met behulp van IRIS versie 4.4.1, voorzien van codeervlaggen. IRIS versie 4.4.1 is vervolgens ook gebruikt bij deze bridge coding study zodat de resultaten van dit onderzoek gebruikt kunnen worden om waargenomen verschuivingen in het jaarbestand te verklaren.

⁵⁾ Een ICD-10 code is opgebouwd uit een letter en (minimaal) twee, danwel drie of vier cijfers.

effect van automatisch coderen op de doodsoorzakenstatistiek kan deze paragrafen bij lezing overslaan zonder de kernboodschap van dit rapport te missen. Bij een mogelijke meerduideligheid in de beschrijving wordt de ICD-10 code tussen haakjes achter de doodsoorzaak vermeld. Wanneer deze niet is vermeld komt de code overeen met die voor de term in deel 1 van de ICD-10.

3.1 Het vóórkomen van doodsoorzaken bij handmatig en automatisch coderen

Een ICD-10 hoofdstuk bevat ziekten of aandoeningen van overeenkomstige soort. De indeling is naar aard (bijv. infectieziekten of nieuwvormingen), orgaanstelsel (bijv. ademhaling, spijsvertering) of ontstaan (bijv. perinataal, congenitaal). Ook is er een hoofdstuk voor symptomen of onvolledig omschreven ziektebeelden.⁶⁾ *Tabel 3.1.1* toont het verschil in vóórkomen van doodsoorzaken bij automatisch en handmatig coderen van eenzelfde verzameling doodsoorzakenformulieren voor ICD-10 hoofdstukken. De CR is berekend op basis van de volautomatisch gecodeerde doodsoorzakenformulieren die zowel bij automatische als bij handmatige codering van een natuurlijke dood werden voorzien (n = 86 893).

3.1.1 Comparability Ratio (CR) voor ICD-10 hoofdstukken

	Automatisch gecodeerd		Handmatig gecodeerd		CR ¹⁾	2009 (onderzoeksbestand)
	Aantal	Percentage	Aantal	Percentage		Percentage
Infectieziekten (A00–B99)	1 805	2,1	1 225	1,4	1,47	1,6
Nieuwvormingen (C00–D48)	28 845	33,2	29 070	33,5	0,99	32,1
w.v. Maligne nieuwvormingen (C00–C97)	28 096	.	28 583	.	0,98	.
Bloedvormende organen (D50–D89)	296	0,3	258	0,3	1,15	0,4
Endocriene/voedingsziekten (E00–E90)	2 874	3,3	2 478	2,9	1,16	3,1
Psychische stoornissen (F00–F99)	7 250	8,3	5 512	6,3	1,32	5,7
Zenuwstelsel en zintuigen (G00–H95)	3 225	3,7	2 744	3,2	1,18	3,5
Hart- en vaatziekten (I00–I99)	28 217	32,5	27 870	32,1	1,01	31,3
Ademhalingsstelsel (J00–J99)	7 803	9,0	10 065	11,6	0,78	11,2
Spijsverteringsstelsel (K00–K93)	2 238	2,6	2 625	3,0	0,85	4,1
Huidziekten (L00–L99)	114	0,1	162	0,2	0,70	0,3
Spier- en botziekten (M00–M99)	327	0,4	281	0,3	1,16	0,6
Nier en urinewegen (N00–N99)	1 694	1,9	2 166	2,5	0,78	2,5
Zwangerschap en bevalling (O00–O99)	0	0,0	1	0,0		0,0
Perinatale periode (P00–P96)	32	0,0	30	0,0	1,07	0,8
Congenitale afwijkingen (Q00–Q99)	167	0,2	166	0,2	1,01	0,5
Symptomen en onvolledig omschreven ziektebeelden (R00–R99)	2 006	2,3	2 240	2,6	0,90	2,3
Totaal	86 893	100	86 893	100	1,00	100

¹⁾ CR vet: significant verschil bij tweezijdige toetsing met T-test op percentages

⁶⁾ De niet-natuurlijke doodsoorzaken worden met twee codes beschreven, één voor de oorzaak (bijv. val of vergiftiging) en één voor het letsel (bijv. wond, hersenschudding), en kennen aparte hoofdstukken in de ICD-10.

Negen van de zeventien onderzochte ICD-10 hoofdstukken vertonen een significant verschil in vóórkomen bij handmatig en automatisch coderen van eenzelfde verzameling doodsoorzakenformulieren. Ook voor de groep van de niet-maligne nieuwvormingen in het hoofdstuk van de nieuwvormingen is dit het geval. Bij automatisch coderen worden er significant meer infectieziekten, endocriene/voedingsziekten, psychische stoornissen en ziekten van zenuwstelsel of zintuigen als onderliggende doodsoorzaak waargenomen dan bij handmatig coderen. Bij automatisch coderen worden significant minder ziekten van het ademhalings- of spijsverteringstelsel, de huid, nieren of urinewegen en symptomen/onvolledig omschreven ziektebeelden als onderliggende doodsoorzaak waargenomen dan bij handmatig coderen. Het vóórkomen van de nieuwvormingen, hart- of vaatziekten en congenitale aandoeningen blijft nagenoeg gelijk.⁷⁾ Een andere codering van 'niet gespecificeerde gastroenteritis' (A09.9 in plaats van K52.9) verklaart de respectievelijk toename van infectieziekten en afname van ziekten van het spijsverteringsstelsel. Op aanwijzing van een ICD-10 update wordt deze aandoening met ingang van januari 2010 in een ander ICD-10 hoofdstuk geplaatst. Hetzelfde geldt voor 'necrotiserende fasciitis' (M72.6 in plaats van L08.8). Deze doodsoorzaak verschuift op grond van een ICD-10 update uit januari 2003 van de huidziekten naar de ziekten van het bot/spierstelsel. De respectievelijke toename van de sterfte aan psychische stoornissen en ziekten van het zenuwstelsel/zintuigen, en de afname van de sterfte aan ziekten van de luchtwegen en ziekten van de nier of urinewegen is een gevolg van een andere selectie van de onderliggende doodsoorzaak, namelijk van dementie en de ziekte van Alzheimer (toename) ten koste van longontstekingen of urineweginfecties (afname). Bij het automatisch coderen worden er andere causale verbanden gelegd dan bij handmatig coderen. De toename van endocriene/voedingsziekten is vooral toe te schrijven aan de selectie van dehydratie (E86) als onderliggende doodsoorzaak. Hoewel IRIS de selectieregels hier correct toepast en de codering van dehydratie correct is, dekt deze selectie doorgaans niet de lading van het doodsoorzakenformulier waar dehydratie vaak als verschijnsel van overlijden bij bewoners van een verpleeghuis wordt vermeld.⁸⁾ De afname van symptomen/onvolledig omschreven ziektebeelden (R-codes) als onderliggende doodsoorzaak is vooral het gevolg van een afname van ouderdom als onderliggende doodsoorzaak. IRIS beschouwt R-codes (waaronder ouderdom) als verlegenheidsdiagnoses (vuilnisbakcodes) en geeft (standaard) de voorkeur aan een code uit een ander ICD-10 hoofdstuk indien aanwezig op het doodsoorzakenformulier. Bij handmatig coderen was dit niet het geval en werd ouderdom op grond van de beschrijving op het doodsoorzakenformulier in combinatie met de leeftijd van de overledene (bijvoorbeeld bij dehydratie, cachexie of hartstilstand op hoge leeftijd) wel degelijk als valide onderliggende doodsoorzaak gezien.

De rechter kolom van *tabel 3.1.1* geeft een beeld van de representativiteit van het gedeelte van een jaarbestand dat IRIS volautomatisch codeert. De verdeling van de ICD-10 hoofdstukken in het onderzoeksbestand van 2009 (rechter kolom) verwacht men terug te zien in de verdeling van de formulieren die IRIS verwerkt (percentage handmatig gecodeerd). Het volautomatisch gecodeerde bestand vertoont echter een oververtegenwoordiging van nieuwvormingen, hart- en vaatziekten, ziekten van het ademhalingsstelsel en psychische stoornissen. Er is een ondervertegenwoordiging van sterfte aan ziekten van het spijsverteringsstelsel. De verklaring voor deze selectiviteit bij verwerking van doodsoorzakenformulieren is gelegen in het gemak waarmee IRIS formulieren volautomatisch

⁷⁾ Het aantal sterfgevallen in de perinatale periode is te klein voor een zinvolle vergelijking. Bovendien is IRIS nog niet geschikt voor het verwerken van perinatale sterfgevallen.

⁸⁾ Bij ingebruikname van IRIS zal hiervoor handmatig worden gecorrigeerd. De codeur is beter in staat tot het interpreteren van de context dan de software en kan hier eventueel met de invuller over corresponderen.

codeert. Daarbij is een sterke samenhang met het aantal codes enerzijds en met de aard van de medische terminologie anderzijds. Het aantal formulieren dat niet volautomatisch kan worden verwerkt neemt (sterk) toe met het aantal codes op een formulier. Daarnaast verschilt het aantal termen in het woordenboek van IRIS per ICD-10 hoofdstuk en daarmee ook de trefkans om een term succesvol te coderen.

Infectie ziekten en parasitaire aandoeningen (A00–B99)

Bij het automatisch coderen worden 47 procent (CR=1,47) meer infectieziekten als onderliggende doodsoorzaak aangetroffen dan bij handmatig coderen van dezelfde verzameling doodsoorzakenformulieren. In dit onderzoek is deze toename de resultante van een instroom van 800 en een uitstroom van 220 sterfgevallen. De instroom van nieuwe sterfgevallen in dit hoofdstuk is vooral afkomstig van 'niet gespecificeerde diarree' (K52.9, 25 procent), 'longontsteking' (25 procent) en een groot aantal doodsoorzaken in kleine aantallen gespreid over andere ICD-10 hoofdstukken.⁹⁾ De 'niet nader gespecificeerde diarree' verhuist bij het automatisch coderen van code K52.9 naar A09.9 ofwel van een ziekte van het spijsverteringsstelsel naar een infectieziekte. Deze verschuiving is het gevolg van een update van de ICD-10 die in januari 2010 van kracht is geworden. De code A09 kreeg toen een vierde digit, waarbij de '0' werd gereserveerd voor (expliciet) als infectieus aangeduide gastro-enteritis /diarree en de '9' voor een op het doodsoorzakenformulier niet nader gespecificeerde gastro-enteritis /diarree. De code K52.9 is sinds januari 2010 uitsluitend gereserveerd voor (expliciet) als 'niet infectieus' aangeduide gastro-enteritis. Aangezien een gastro-enteritis op het doodsoorzakenformulier zelden gespecificeerd is als 'niet infectieus' blijkt bij het automatisch coderen 82 procent van de gevallen in code A09 afkomstig van K52.9. Bij de instroom van 'longontsteking' in het hoofdstuk 'infectieziekten' blijkt het merendeel van de gevallen (ongeveer 88 procent) die handmatig als longontsteking werden gecodeerd (J18) terecht te komen in de code A41.9, 'niet nader gespecificeerde sepsis'. Een en ander is het gevolg van een andere opvatting over de selectie van de onderliggende doodsoorzaak. Als IRIS longontsteking en sepsis op een doodsoorzakenformulier aantreft, wordt de volgorde van invullen gevolgd en kan sepsis als oorzaak van longontsteking worden aangemerkt. De CBS codeur gaf in een dergelijke situatie de voorkeur aan longontsteking als onderliggende doodsoorzaak. De uitstroom van 220 sterfgevallen uit het hoofdstuk 'infectieziekten' wordt voornamelijk gevormd door sepsis (58 procent) en de groep van overige infectieziekten (B99, 16 procent). Sterfgevallen met een sepsis als onderliggende doodsoorzaak verschuiven naar hartfalen (I50) of nierfalen (N17-N19). Een andere opvatting over de selectie van de onderliggende doodsoorzaak is hiervoor de verklaring. Anders dan bij handmatig coderen accepteert IRIS hart- of nierfalen als oorzaak van sepsis indien de invuller van het doodsoorzakenformulier dat zo aangeeft.

De toename van infectieziekten is bij de overgang op automatisch coderen vooral een gevolg van een ICD-10 update die een veranderende opvatting over de aard van de niet-gespecificeerde gastro-enteritis weergeeft. Aangezien het CBS de ICD-10 updates bij handmatig coderen niet heeft gevolgd, wordt deze update bij automatisch coderen van eenzelfde verzameling formulieren in één klap zichtbaar. Bij de selectie van de onderliggende doodsoorzaak blijkt sepsis de plaats in te nemen van de longontsteking. Hier treedt een uitwisseling van infectieziekten tussen ICD-10 hoofdstukken op door een verschil van opvatting over causale relaties tussen ziektebeelden zoals vermeld op het doodsoorzakenformulier.

⁹⁾ Instroom = sterfgevallen, bij handmatig coderen niet en bij automatisch coderen wel in desbetreffende ICD-10 hoofdstuk. Uitstroom = sterfgevallen, bij handmatig coderen wel en bij automatisch coderen niet in het desbetreffende ICD-10 hoofdstuk.

Maligne (kwaadaardige) nieuwvormingen (C00–C97)

Bij automatisch coderen worden (2 procent) minder maligne nieuwvormingen als onderliggende doodsoorzaak (CR = 0, 98) aangetroffen dan bij handmatig coderen van dezelfde verzameling doodsoorzakenformulieren. In dit onderzoek is deze afname de resultante van een uitstroom van 956 en een instroom van 469 sterfgevallen. Het verschil tussen automatisch en handmatig coderen is (ook binnen dit hoofdstuk) niet significant. De uitstroom van sterfgevallen wordt voornamelijk gevormd door longkanker (23 procent), borstkanker (10 procent), prostaat­kanker (9 procent), maligne hersentumoren (8 procent) en dikke darmkanker (6 procent).

De uitstroom van sterfgevallen met longkanker gaat in bijna 23 procent van de gevallen naar COPD/ademhalingsstilstand. IRIS laat (terecht) deze aandoeningen niet uit de longkanker ontstaan. Het is echter de vraag hoe nauwkeurig de invullend arts de termen hier heeft gebruikt. Strikt genomen veroorzaakt longkanker geen COPD. Vaak lijkt de arts echter met de COPD een ernstige vorm van benauwdheid aan te duiden die past bij de longkanker of bij de behandeling ervan (operatie en/of bestraling). Ongeveer 10 procent van de uitstroom gaat naar de longtumor. Bij vermelding van verschijnselen die op een maligniteit wijzen (bijvoorbeeld: cachexie of embolie) werd bij handmatig coderen een tumor als maligne beschouwd. IRIS doet/kan dit niet. Verder verschuift nog 10 procent van de gevallen naar het acuut hartinfarct en 10 procent naar hartfalen. Hier ligt een andere opvatting over de selectie van de onderliggende doodsoorzaak aan ten grondslag.

De uitstroom van de sterfgevallen met borstkanker gaat voornamelijk naar hartfalen (15 procent), hersenbloeding (13 procent), sepsis (11 procent) en hartinfarct (9 procent). IRIS laat deze ziekten niet ontstaan uit de maligniteit van de borst, vermeld op deel 2 van het doodsoorzakenformulier (*zie figuur 5.1*). De uitstroom van de sterfgevallen met prostaat­kanker gaat (net als bij borstkanker) voornamelijk naar hartfalen (12 procent), sepsis (8 procent), hartinfarct (7 procent) en de hersenbloeding (6 procent). IRIS laat deze ziekten niet ontstaan uit de maligniteit van de prostaat, vermeld op deel 2 van het doodsoorzakenformulier (*zie figuur 5.1*). Bij handmatig coderen werden in deze situatie andere regels gehanteerd voor het aanwijzen van een maligniteit als onderliggende doodsoorzaak.

De uitstroom van de sterfgevallen met een maligne hersentumor gaat voor bijna 50 procent naar een maligniteit van het centraal zenuwstelsel (C72.9). Hier ligt een verandering in codeerpraktijk (het woordenboek) aan ten grondslag. Het glioom werd handmatig als maligniteit van de hersenen gecodeerd (C71.9), omdat het daar het meeste voorkomt. Strikt genomen kan het echter ook buiten de hersenen voorkomen en is de code C72.9 van toepassing als de plaats van het glioom niet is gespecificeerd. Bijna 45 procent verschuift naar de code voor hersentumor. Bij handmatig coderen werd een hersentumor vaak als maligne beschouwd. IRIS doet/kan dit niet.

De uitstroom van de sterfgevallen met dikkerdarmkanker gaat voornamelijk naar een ongespecificeerde maligniteit van de darmen (16 procent) in gevallen van dubbeltumoren, de darmtumor (7 procent) door ontbreken van de interpretatie van de context bij automatisch coderen, en de ongespecificeerde darmaandoening (7 procent) die IRIS niet laat ontstaan uit een maligniteit van de darmen. Net als bij COPD/longkanker is het automatisch coderen hier meer dan bij handmatig coderen het geval is, gevoelig voor de onderlinge positie van termen op het doodsoorzakenformulier.

Er is een instroom van 469 nieuwe sterfgevallen in dit hoofdstuk die bij handmatig coderen longontsteking (35 procent), COPD (5 procent), longembolie (4 procent), urineweginfectie (4 procent), sepsis (3 procent), acuut hartinfarct (3 procent) en aspiratiepneumonie (3 procent) als onderliggende doodsoorzaak toegewezen kregen. Deze komen vooral terecht bij prostaatkanker, longkanker en borstkanker. Een verschil van opvatting over de selectie van de onderliggende doodsoorzaak ligt hieraan ten grondslag. De bij instroom genoemde ziekten lezen als een logisch gevolg van kwaadaardige nieuwvormingen, maar kunnen op grond van tijdsduur (bijvoorbeeld borstkanker langer dan 10 jaar geleden) of plaats op het doodsoorzakenformulier (bijvoorbeeld prostaatkanker zonder metastasen op deel 2 van het doodsoorzakenformulier) bij handmatig coderen toch als onderliggende doodsoorzaak zijn aangewezen.

Niet-kwaadaardige nieuwvormingen (D00–D48)

Bij het automatisch coderen worden er 54 procent (CR=1,54) meer niet-kwaadaardige nieuwvormingen (goedaardige of van onbestemde aard) als onderliggende doodsoorzaak aangetroffen dan bij handmatig coderen van dezelfde verzameling doodsoorzakenformulieren. In dit onderzoek is deze toename de resultante van een instroom van 285 en een uitstroom van 23 sterfgevallen. De instroom van nieuwe sterfgevallen in dit hoofdstuk wordt vooral gevormd door sterfgevallen die bij handmatig coderen maligniteiten van hersenen (C71.9, 24 procent) en longen (C34.9, 9 procent) als onderliggende doodsoorzaak toegewezen kregen. Bij automatisch coderen krijgen deze sterfgevallen hersentumoren (D43.2, 33 procent), longtumoren (D38.1, 11 procent) en darmtumoren (D37.7, 9 procent) als onderliggende doodsoorzaak toegewezen. Het effect van een interpretatie van de term ‘tumor’ op basis van context (cachexie, embolie, tijdsduur, etc.) wordt zichtbaar in deze verschuiving van een C (lees maligniteit) naar een D code (lees tumor).¹⁰⁾ Er is een specifieke uitstroom van sterfgevallen uit dit hoofdstuk voornamelijk naar het hoofdstuk kwaadaardige nieuwvormingen (C00-C97).

Ziekten van bloed en bloedvormende organen (D50–D89)

Bij het automatisch coderen worden er 15 procent (CR=1,15) meer ziekten van het bloed of de bloedvormende organen als onderliggende doodsoorzaak aangetroffen dan bij handmatig coderen van dezelfde verzameling doodsoorzakenformulieren. In dit onderzoek is deze toename de resultante van een instroom van 122 en een uitstroom van 84 sterfgevallen. De instroom van nieuwe sterfgevallen in dit hoofdstuk wordt vooral gevormd door hartfalen (I50, 29 procent), nierfalen (N18.9, 7 procent), maagdarmbloedingen (K92.2, 6 procent), longkanker (C34.9, 5 procent) en longontsteking (J18.9, 5 procent). Bij automatisch coderen krijgen deze sterfgevallen een of andere vorm van anemie (D64.9, 53 procent of D53.9, 16 procent, of D61.9, 10 procent) als onderliggende doodsoorzaak toegewezen. De uitstroom wordt voornamelijk gevormd door sterfgevallen met anemie (D64.9, 66 procent; D53.9, 12 procent) als onderliggende doodsoorzaak. Deze krijgen bij automatisch coderen dementie (24 procent), CVA (7 procent) of hartfalen (7 procent) als onderliggende doodsoorzaak. Een andere selectie van de onderliggende doodsoorzaak en een andere codeerpraktijk liggen aan deze verschuivingen ten grondslag. IRIS beschouwt een anemie als oorzaak van hartfalen. Dit deden de CBS codeurs alleen als er een duidelijk verband tussen beide doodsoorzaken was aangegeven. Ook werd bij handmatig coderen de anemie gecombineerd met nierfalen, een verband dat IRIS niet kan leggen, omdat het in de beslistabellen ontbreekt.

¹⁰⁾ Bij het aantreffen van de term ‘hersentumor’ op het doodsoorzakenformulier vraagt de CBS codeur de invuller doorgaans om een nadere specificatie (kwaad- of goedaardig?). In dit onderzoek zijn correspondentiegevallen buiten beschouwing gelaten.

Endocriene, voedings- en metabole ziekten (E00–E90)

Bij het automatisch coderen worden er 16 procent (CR= 1,16) meer onderliggende doodsoorzaken in dit hoofdstuk aangetroffen dan bij handmatig coderen van dezelfde verzameling doodsoorzakenformulieren. In dit onderzoek is deze toename de resultante van een instroom van 761 en een uitstroom van 365 sterfgevallen. De instroom van sterfgevallen is gespreid over een groot aantal gevallen uit andere hoofdstukken, waaronder ouderdom (11 procent), hartinfarct (8 procent) en longontsteking (8 procent). De instroom vanuit ouderdom kan worden verklaard doordat CBS codeurs dehydratie (E86) in combinatie met cachexie interpreteren als ouderdom (R54). IRIS doet dit niet. Bij het automatisch coderen zien we dan ook een sterke toename van dehydratie (E86) als onderliggende doodsoorzaak ten koste van ouderdom. De uitstroom uit dit hoofdstuk wordt vooral gevormd door sterfgevallen met verschillende vormen van diabetes (ruim 70 procent) als onderliggende doodsoorzaak. Bij automatisch coderen krijgen deze sterfgevallen hart- en vaatziekten als hartinfarct, hartfalen of CVA als onderliggende doodsoorzaak toegewezen. Een verschil in opvatting over het vierde digit bij diabetes ligt hieraan ten grondslag. Hartinfarct of CVA wordt door CBS codeurs gezien als complicatie van diabetes en handmatig als E14.6 gecodeerd. IRIS doet dit, gebruik makende van een tabel bij de ICD-10 updates van januari 2010 die complicaties van diabetes definieert, niet. De verschuivingen van of naar dit hoofdstuk van de ICD-10 kunnen dan ook grotendeels aan een verschil in nationale of internationale codeergewoonten worden toegeschreven.

Psychische ziekten en gedragsstoornissen (F00–F99)

Bij het automatisch coderen worden er 31 procent (CR = 1,31) meer psychische ziekten als onderliggende doodsoorzaak aangetroffen dan bij handmatig coderen van dezelfde verzameling doodsoorzakenformulieren. In dit onderzoek is deze toename de resultante van een instroom van 2 001 en een uitstroom van 263 sterfgevallen. De instroom van nieuwe sterfgevallen in dit hoofdstuk is grotendeels afkomstig van longontsteking/luchtweginfecties (47 procent) en urineweginfecties (12 procent). Ruim 82 procent van deze gevallen krijgt bij automatisch coderen dementie als onderliggende doodsoorzaak. Een andere selectie van de onderliggende doodsoorzaak is hiervan de oorzaak. IRIS merkt een longontsteking of urineweginfectie aan als een gevolg van (vasculaire) dementie, ook wanneer dementie op deel 2 van het doodsoorzakenformulier staat (zie *figuur 5.1*). Hierbij wordt een beroep gedaan op een update van de selectieregel 3 uit deel 2 van de ICD-10 (januari 2005) die stelt dat :

‘Rule 3 . . . Assumed direct consequences of another condition

Secondary or unspecified anaemia, malnutrition, marasmus or cachexia may be assumed to be a consequence of any malignant neoplasm, paralytic disease, or disease which limits the ability to care for oneself, including dementia and degenerative diseases of the nervous system.’

Dementie wordt hier genoemd als ziekte die de mogelijkheid tot zelfzorg beperkt. Daardoor kan een urineweginfectie of longontsteking minder goed worden behandeld en tot de dood leiden. Bij handmatig coderen werd deze veronderstelling niet gehanteerd en slechts een verband gelegd tussen een doodsoorzaak op deel 1 en dementie op deel 2 van het doodsoorzakenformulier als dit expliciet door de invuller was beschreven. De uitstroom uit

dit hoofdstuk betreft vooral sterfgevallen met dementie (64 procent) en vasculaire dementie (13 procent) als onderliggende doodsoorzaak. Bij automatisch coderen komen deze sterfgevallen terecht bij een groot aantal andere codes waaronder het laat gevolg van een CVA (10 procent) en de Fronto-temporale dementie (8 procent).

Het significante verschil tussen automatisch en handmatig coderen bij de overige psychische stoornissen wordt veroorzaakt doordat IRIS delier (F05.9) als onderliggende doodsoorzaak codeert, terwijl dit veelal in Nederland als een verschijnsel van het overlijden wordt gerapporteerd. Hoewel de selectie van IRIS hier strikt genomen correct is, zal bij implementatie van het automatisch coderen in de praktijk correctie middels handmatige nacontrole plaats moeten vinden om de inhoudelijke juistheid van de codering vast te stellen.

Ziekten van het zenuwstelsel en zintuigen (G00-H95)

Bij het automatisch coderen worden er 17 procent (CR=1,17) meer ziekten van het zenuwstelsel als onderliggende doodsoorzaak aangetroffen dan bij handmatig coderen van dezelfde verzameling doodsoorzaken. In dit onderzoek is deze toename de resultante van een instroom van 698 en een uitstroom van 217 sterfgevallen. De instroom van nieuwe sterfgevallen in dit hoofdstuk is grotendeels afkomstig van sterfgevallen met ontsteking/luchtweginfecties (32 procent) en urineweginfecties (10 procent) als onderliggende doodsoorzaak. Bij automatisch coderen krijgen deze sterfgevallen vooral de ziekte van Alzheimer (51 procent) als onderliggende doodsoorzaak toegewezen. De verklaring hiervoor is dezelfde als die voor dementie (te weten de update van selectieregel 3 uit deel twee van de ICD-10). De uitstroom uit dit hoofdstuk betreft vooral sterfgevallen met de ziekte van Alzheimer (G30.1, 30 procent), de ziekte van Parkinson (22 procent), epilepsie (20 procent) en TIA (11 procent) als onderliggende doodsoorzaak. Bij automatisch coderen krijgen deze sterfgevallen vooral dementie (29 procent), dehydratie (8 procent) en het laat gevolg van een CVA (6 procent) als onderliggende doodsoorzaak toegewezen.

Hart- en vaatziekten (I00-I99)

Bij het automatisch coderen wordt er 1 procent (CR = 1,01) meer hart- en vaatziekten als onderliggende doodsoorzaak aangetroffen dan bij handmatig coderen van dezelfde verzameling doodsoorzakenformulieren. In dit onderzoek is deze toename de resultante van een instroom van 2 171 en een uitstroom van 1 824 sterfgevallen. De instroom van nieuwe sterfgevallen wordt vooral gevormd door longontsteking (17 procent), COPD (14 procent), diabetes (8 procent) en urineweginfecties (4 procent) en een groot aantal andere ziekten of aandoeningen in kleine aantallen. Bij automatisch coderen komen deze sterfgevallen terecht bij het hartfalen (27 procent), de late gevolgen van het CVA (23 procent), het acuut hartinfarct (8 procent) en het (recent) CVA (6 procent). De uitstroom uit dit hoofdstuk betreft het acuut hartinfarct (20 procent), hartfalen (15 procent), hartstilstand (10 procent) en het CVA (10 procent). Bij automatisch coderen komen deze sterfgevallen voornamelijk terecht bij dementie (10 procent), COPD (7 procent), longontsteking (6 procent), diabetes (5 procent) en in kleine aantallen bij een groot aantal andere ziekten of aandoeningen. De uitstroom van sterfgevallen uit dit hoofdstuk wordt overtroffen door de instroom van nieuwe sterfgevallen in dit hoofdstuk. Een andere selectie van de onderliggende doodsoorzaak is verantwoordelijk voor de waargenomen verschuivingen doordat IRIS strikter vasthoudt aan de volgorde van vermelding op het doodsoorzakenformulier dan bij handmatig coderen het geval was. Wanneer bijvoorbeeld een longontsteking of COPD op het doodsoorzakenformulier onder het hartinfarct vermeld staat, kiest IRIS respectievelijk de longontsteking of COPD als onderliggende doodsoorzaak (uitstroom uit hoofdstuk). Bij handmatig coderen werd dan

het hartinfarct als onderliggende doodsoorzaak gekozen vanuit het oogpunt van statistische relevantie. Een verandering van codeerpraktijk is verantwoordelijk voor de toename van het laat gevolg van een CVA als onderliggende doodsoorzaak. Bij handmatig coderen werd een CVA op deel 2 van het doodsoorzakenformulier niet gecodeerd, tenzij een neuroloog het formulier had ingevuld. IRIS codeert alle termen op het doodsoorzakenformulier, waarbij het (ongespecificeerd) CVA op deel 2 van het doodsoorzakenformulier de code voor een laat gevolg van een CVA (I69) krijgt, en trekt de causale keten zo nodig door naar dit laat gevolg van het CVA.

Ziekten van ademhalingsstelsel en luchtwegen (J00–J99)

Bij het automatisch coderen worden er 22 procent (CR = 0,78) minder ziekten van ademhalingsstelsel of luchtwegen als onderliggende doodsoorzaak aangetroffen dan bij handmatig coderen van dezelfde verzameling doodsoorzakenformulieren het geval is. In dit onderzoek is deze afname de resultante van een uitstroom van 2 859 en een instroom van 597 sterfgevallen. De uitstroom wordt vooral gevormd door sterfgevallen die bij handmatig coderen longontsteking (68 procent), COPD (18 procent) of luchtweginfectie (6 procent) als onderliggende doodsoorzaak hadden. Bij automatisch coderen krijgen deze sterfgevallen dementie (41 procent), hartfalen (13 procent), laat gevolg van het CVA (10 procent) of sepsis (8 procent) als onderliggende doodsoorzaak toegewezen.

De grote afname van de sterfte aan longontsteking is gebaseerd op een aanpassing van de selectieregel 3 uit deel 2 van de ICD-10 waardoor de longontsteking bijna altijd wordt veroorzaakt door een andere doodsoorzaak op het formulier. Zo lezen we in de ICD-10 update van januari 2008:

'Rule 3 . . . Assumed direct consequences of another condition

Any pneumonia in J12-J18 should be considered an obvious consequence of conditions that impair the immune system. Pneumonia in J18.0 and J18.2-J18.9 should be considered an obvious consequence of wasting diseases (such as malignant neoplasm and malnutrition) and diseases causing paralysis (such as cerebral haemorrhage or thrombosis), as well as serious respiratory conditions, communicable diseases, and serious injuries. Pneumonia in J18.0 and J18.2-J18.9, J69.0, and J69.8 should also be considered an obvious consequence of conditions that affect the process of swallowing.'

Deze wijziging van codeerregel heeft een groot gevolg. Slecht 49 procent van de handmatig als longontsteking gecodeerde formulieren krijgt door IRIS ook deze code als onderliggende doodsoorzaak toegekend. Met de automatisch gecodeerde sterfgevallen als noemer is dit percentage 91 (*zie tabel 3.3.1*). Dit betekent dat de CBS codeurs het veelal eens zijn met de automatisch toegekende code, maar omgekeerd dat IRIS veel sterfgevallen die handmatig de code longontsteking kregen anders codeert.

De instroom van 597 nieuwe sterfgevallen wordt vooral gevormd door sterfgevallen die bij handmatig coderen het hartinfarct (23 procent), longkanker (11 procent) en het CVA (7 procent) als onderliggende doodsoorzaak hadden. Bij automatisch coderen komen deze terecht bij COPD (43 procent), longontsteking (24 procent), en overige longziekten (6 procent)

als onderliggende doodsoorzaak. Aan de uitwisseling tussen longkanker en COPD ligt een positieprobleem ten grondslag. Wanneer COPD boven of als eerste in het zelfde vakje als longkanker vermeld wordt, kiest IRIS COPD als onderliggende doodsoorzaak in plaats van de longkanker. Bij handmatig coderen werd dit niet gedaan.

Ziekten van spijsverteringsorganen (K00–K93)

Bij het automatisch coderen worden er 15 procent (CR = 0,85) minder ziekten van de spijsverteringsorganen als onderliggende doodsoorzaak aangetroffen dan bij handmatig coderen van dezelfde verzameling doodsoorzakenformulieren het geval was. De ICD-10 update uit januari 2010 die aangeeft 'niet-infectieuze diarree' als A09.9 te coderen in plaats van K25.9 is grotendeels verantwoordelijk voor deze afname. In dit onderzoek is de afname van spijsverteringsziekten als onderliggende doodsoorzaak de resultante van een uitstroom van 612 en een instroom van 225 sterfgevallen. De uitstroom uit dit hoofdstuk betreft vooral sterfgevallen die bij handmatig coderen 'niet-infectieuze diarree' met als code K52.9 (39 procent), maagdarmbloedingen (24 procent), ileus (6 procent), levercirrose (4 procent) en alcoholische leverziekten (3 procent) als onderliggende doodsoorzaak toegewezen kregen. Bij automatisch coderen krijgen deze sterfgevallen vooral 'niet infectieuze diarree' met als code A09.9 (33 procent), alcoholabusus (F10.1, 7 procent) en sepsis (A14.9, 6 procent) als onderliggende doodsoorzaak toegewezen. De instroom wordt gevormd door het acuut hartinfarct (8 procent), longontsteking (8 procent), dikke darmkanker (6 procent), longkanker (4 procent) en prostaatkanker (4 procent). Bij automatisch coderen krijgen deze sterfgevallen de ileus (12 procent), levercirrose (11 procent) en de maagdarmbloeding (9 procent) als onderliggende doodsoorzaak toegewezen. Vooral de instroom van maligniteiten in dit hoofdstuk is opmerkelijk. Dit is het gevolg van een selectie die hoewel correct uitgevoerd (IRIS laat een niet nader gespecificeerde darmafwijking niet ontstaan uit een maligniteit), niet in overeenstemming lijkt te zijn met de bedoeling van de invuller van het doodsoorzakenformulier. Deze noteert een darmafwijking als fase van het overlijden met de maligniteit als onderliggende doodsoorzaak.

Ziekten van huid en onderhuids bindweefsel (L00–L99)

Bij het automatisch coderen worden er 30 procent (CR = 0,70) minder ziekten van huid of onderhuids bindweefsel als onderliggende doodsoorzaak aangetroffen dan bij handmatig coderen van dezelfde verzameling formulieren het geval is. In dit onderzoek is deze afname de resultante van een uitstroom van 82 en een instroom van 34 sterfgevallen. De uitstroom uit dit hoofdstuk betreft vooral sterfgevallen met decubitus (77 procent), huidinfecties (11 procent) en ulcera van de onderbenen (5 procent) als onderliggende doodsoorzaak. Deze sterfgevallen krijgen bij automatisch coderen dementie (35 procent), laat gevolg van een CVA (12 procent), sepsis (10 procent) of fasciitis necroticans (9 procent) als onderliggende doodsoorzaak toegewezen. De instroom van nieuwe sterfgevallen betreft een klein aantal doodsoorzaken gespreid over de verschillende ICD-10 hoofdstukken. De maagulcus valt hierbij op (15 procent). Mogelijk is er hier sprake van een codeerfout. 'Maag spatie ulcus' wordt als twee termen gelezen en gecodeerd. Hier wordt een technische beperking van het systeem zichtbaar die zo mogelijk al bij de invoer van de gegevens zal moeten worden gecorrigeerd. De verschuiving van sterfgevallen naar fasciitis necroticans is het gevolg van een ICD-10 update uit januari 2003 die de code M72.6 toewijst aan necrotiserende fasciitis. Op grond van onderlinge afspraak werden sterfgevallen met deze doodsoorzaak (bij gebrek aan betere code) door CBS codeurs aan L08.8 toegewezen (gespecificeerde infectie van huid en onderhuids bindweefsel). Zij zijn daarbij niet onderscheiden van andere sterfgevallen in L08.8. De onderhavige update maakt een specifieke codering mogelijk en is in het bridge coding onderzoek als verschuiving van een L naar een M code zichtbaar.

Ziekten van spierstelsel, bindweefsel en botten (M00–M99)

Bij het automatisch coderen worden er 11 procent (CR = 1,11) meer ziekten van spierstelsel, bindweefsel en botten als onderliggende doodsoorzaak aangetroffen dan bij handmatig coderen van dezelfde verzameling formulieren het geval is. In dit onderzoek is deze toename de resultante van een instroom van 102 en een uitstroom van 56 sterfgevallen. De instroom wordt vooral gevormd door sterfgevallen die bij handmatig coderen longontsteking (23 procent), COPD (7 procent) en huidinfecties (L08.8, 7 procent) als onderliggende doodsoorzaak toegewezen kregen. Bij automatisch coderen komen deze sterfgevallen vooral ten gunste van reumatoïde artritis (M06.9, 37 procent) als onderliggende doodsoorzaak. De uitwisseling met huidinfecties berust op een ICD-10 update uit januari 2003. De uitstroom uit dit hoofdstuk betreft vooral sterfgevallen die bij handmatig coderen de ongespecificeerde gewrichtsaandoening (M25.9, 11 procent) als onderliggende doodsoorzaak toegewezen kregen. Bij automatisch coderen krijgen deze sterfgevallen een groot aantal andere doodsoorzaken toegewezen in kleine aantallen verspreid over andere ICD-10 hoofdstukken.

Ziekten van nier of urinewegen (N00–N99)

Bij het automatisch coderen worden er 22 procent (CR=0,78) minder ziekten van nier en urinewegen als onderliggende doodsoorzaak aangetroffen dan bij handmatig coderen van dezelfde verzameling formulieren het geval is. In dit onderzoek is deze afname de resultante van een uitstroom van 754 en een instroom van 282 sterfgevallen. De uitstroom uit dit hoofdstuk betreft vooral sterfgevallen met urineweginfecties (N39.0, 63 procent) en nierfalen (N17-N19, 30 procent) als onderliggende doodsoorzaak. Bij urineweginfecties hanteert IRIS een soortgelijke interpretatie van ICD-10 selectieregel 3 als bij longontsteking of sepsis. Dit betekent dat urineweginfecties veel vaker dan bij handmatig coderen het geval was als het gevolg van andere ziekten of aandoeningen wordt gezien. Bij automatisch coderen krijgen deze sterfgevallen dementie/ziekte van Alzheimer (44 procent) en laat gevolg van een CVA (9 procent) als onderliggende doodsoorzaak toegewezen. Deze uitstroom overtreft de instroom van nieuwe sterfgevallen die bij handmatig coderen het acuut hartinfarct (8 procent), longontsteking (8 procent), hartfalen (I50.9, 7 procent), sepsis (6 procent) en een groot aantal doodsoorzaken in kleine aantallen verdeeld over verschillende ICD-10 hoofdstukken als onderliggende doodsoorzaak kregen toegewezen. Bij automatisch coderen krijgen deze sterfgevallen vooral nierfalen (72 procent) en de urineweginfectie (21 procent) als onderliggende doodsoorzaak toegewezen.

Aangeboren afwijkingen en chromosoom stoornissen (Q00–Q99)

Bij het automatisch coderen worden er 1 procent (CR = 1,01) meer aangeboren afwijkingen en chromosoom stoornissen als onderliggende doodsoorzaak aangetroffen dan bij handmatig coderen van dezelfde verzameling formulieren het geval is. In dit onderzoek is deze toename de resultante van een instroom van 26 en een uitstroom van 25 sterfgevallen. De instroom wordt vooral gevormd door sterfgevallen die bij handmatig coderen longontsteking (46 procent) als onderliggende doodsoorzaak toegewezen kregen. Bij automatisch coderen krijgen deze vooral het syndroom van Down (77 procent) als onderliggende doodsoorzaak. De uitstroom van 25 sterfgevallen uit dit hoofdstuk betreft eveneens sterfgevallen met het syndroom van Down (Q909, 28 procent) als onderliggende doodsoorzaak. Bij automatisch coderen krijgen deze sterfgevallen een groot aantal doodsoorzaken in kleine aantallen verdeeld over verschillende ICD-10 hoofdstukken als onderliggende doodsoorzaak toegewezen. Een andere selectie van de onderliggende doodsoorzaak is verantwoordelijk voor deze verschuivingen.

Symptomen en onvolledig omschreven ziektebeelden (R00–R99)

Bij automatisch coderen worden er 11 procent (CR = 0,89) minder symptomen, onvolledig omschreven ziektebeelden, tekenen en afwijkende laboratorium uitslagen als onderliggende doodsoorzaak aangetroffen dan bij handmatig coderen van dezelfde verzameling formulieren het geval is. In dit onderzoek is deze afname de resultante van een uitstroom van 497 en een instroom van 263 sterfgevallen. De uitstroom uit dit hoofdstuk betreft vooral sterfgevallen met ouderdom (32 procent), plotselinge dood met onbekende oorzaak (R96, 10 procent), koorts (10 procent), bloedingen (6 procent), acute buik (5 procent), misselijkheid en braken (5 procent), cachexie (5 procent), malaise & moeheid (5 procent) als onderliggende doodsoorzaak. Bij automatisch coderen krijgen deze vooral dehydratie (19 procent), laat gevolg van een CVA (11 procent) en dementie (7 procent) als onderliggende doodsoorzaak toegewezen. De uitstroom overtreft de instroom van sterfgevallen met hartfalen (150, 15 procent), cor pulmonale (127,9, 13 procent) en acute hartstilstand (146, 13 procent) als onderliggende doodsoorzaak. Bij automatisch coderen krijgen deze sterfgevallen vooral een plotselinge dood met onbekende oorzaak (R96, 24 procent), ouderdom (22 procent) en ademhalingsstilstand (17 procent) als onderliggende doodsoorzaak toegewezen.

3.2 De overeenstemming tussen handmatig en automatisch coderen

De PCP geeft een beeld van de overeenstemming tussen automatisch en handmatig coderen met betrekking tot het coderen en selecteren van de onderliggende doodsoorzaak bij hetzelfde sterfgeval (tabel 3.2.1). In 78,3 procent van de doodsoorzakenformulieren, die IRIS volautomatisch kan verwerken, wordt er bij automatisch en handmatig coderen precies

3.2.1 Perfect Compatibility Percentage (PCP) voor ICD-10 hoofdstukken

	Automatisch gecodeerd			Handmatig gecodeerd
	ICD-10 code op 4-digit niveau	ICD-10 code op 3-digit niveau	ICD-10 code op 1-digit niveau	
	PCP			Aantal
Infectieziekten (A00–B99)	73,0	77,8	80,8	1 225
Maligne nieuwvormingen (C00–C97)	91,4	94,2	96,7	28 583
Niet maligne Nieuwvormingen (D00–D48)	89,7	92,8	95,3	487
Bloedvormende organen (D50–D89)	61,6	64,0	70,2	258
Endocriene/voedingsziekten (E00–E90)	43,1	81,8	85,3	2 478
Psychische stoornissen (F00–F99)	90,3	92,1	95,2	5 512
Zenuwstelsel en zintuigen (G00–H95)	73,2	86,8	92,1	2 744
Hart- en vaatziekten (I00–I99)	79,3	83,8	93,5	27 870
Ademhalingsstelsel (J00–J99)	56,4	69,1	71,6	10 065
Spijverteringsstelsel (K00–K93)	62,1	72,1	76,7	2 625
Huidziekten (L00–L99)	14,8	48,8	49,4	162
Spier- en botziekten (M00–M99)	21,1	72,6	80,1	281
Nier en urinewegen (N00–N99)	52,4	61,9	65,2	2 166
Congenitale afwijkingen (Q00–Q99)	83,7	84,3	84,9	166
Symptomen en onvolledig omschreven ziektebeelden (R00–R99)	71,4	71,5	77,8	2 240
Totaal ¹⁾	78,3	84,7	89,8	86 893
Totaal (formulieren zonder vrije tekst)	79,1	85,4	90,2	81 120
Totaal (formulieren met vrije tekst)	66,8	74,3	84,8	5 773
Formulieren ACME Maybe ²⁾	28,9	35,0	42,2	1 812

¹⁾ Inclusief O00-P99.

²⁾ Onderdeel van het portie rejected formulieren.

dezelfde ICD-10 code voor de onderliggende doodsoorzaak aangetroffen. Deze overeenstemming neemt toe met een afnemend detailniveau van de code: 84,7 procent op drie digit niveau (het niveau van de CBS Statline publicaties) en 89,8 procent op ICD-10 hoofdstukniveau. De overeenstemming tussen handmatig en automatisch coderen is hoger voor formulieren zonder dan voor formulieren met een vrije tekst in het opmerkingenveld. Gelet op de overeenstemming voor het totale bestand (78,3 procent) is de overeenkomst tussen automatisch en handmatig coderen groot voor (maligne) nieuwvormingen en psychische stoornissen. De overeenkomst is klein voor endocriene/voedingsziekten, ziekten van het ademhalingsstelsel, ziekten van spijsverteringsstelsel, huidziekten, ziekten van spieren of botten en aandoeningen van nieren of urinewegen.¹¹⁾ Een groot verschil in PCP tussen het 3- en 4-digit-niveau van de ICD-10 code zoals bij endocriene/voedingsziekten, huidziekten en ziekten van spieren of botten duidt op een verschillende interpretatie van het vierde digit van de ICD-10 code. Dit is onder meer het geval voor diabetes mellitus waar het vierde digit een complicatie aanduidt. Handmatig werden cardiovasculaire aandoeningen als een dergelijke complicatie aangeduid (E14.6). IRIS doet dat niet. Bij de huidziekten is decubitus bij een ICD-10 update in januari 2010 van een vierde digit voorzien. Bij handmatig coderen werden geen updates gehanteerd. Het effect van een update wordt hier dus zichtbaar in een lage overeenstemming op 4-digit niveau.

Ongeveer 10 procent van de formulieren uit 2009 blijkt additionele tekst in het veld opmerkingen (g) van het doodsoorzakenformulier te bevatten. Van deze formulieren wordt slechts 43 procent door IRIS volautomatisch gecodeerd (n=5 773). IRIS kan daarbij de opmerkingen niet gebruiken. De CBS codeur deed dat bij handmatig coderen wel. De PCP van de formulieren met additionele informatie bedraagt 84,8 procent op ICD10 hoofdstukniveau, 74,3 procent op drie digit niveau en 66,8 procent op vier digit niveau (*tabel 3.2.1*). Deze PCP-s zijn vooral op drie- en vier digit niveau beduidend lager dan die voor de formulieren zonder opmerkingen, hetgeen duidt op een specificering van de code door additionele informatie die de CBS codeur bij handmatig coderen gebruikt. In productie is het daarom raadzaam formulieren met opmerkingen aan de CBS codeur voor te leggen alvorens deze automatisch te verwerken. De CBS codeur kan dan besluiten het formulier door IRIS te laten coderen of de onderliggende doodsoorzaak zelf te bepalen (manual final).

Ongeveer 2 procent van de doodsoorzakenformulieren wordt door IRIS volledig gecodeerd, maar voorzien van een twijfel bij de gevonden causale relatie (AcmeMaybe) en aan een CBS codeur voorgelegd (moeilijke formulieren). De overeenstemming tussen de door IRIS gesuggereerde code en de handmatig toegekende code ligt voor deze formulieren beduidend lager op elk niveau van de ICD-10 code (*tabel 3.2.1*), hetgeen aangeeft dat de CBS codeur nodig is om de code te herzien.

Tabel 3.2.2 geeft een overzicht van de tien meest frequent voorkomende doodsoorzaken binnen respectievelijk de verzameling doodsoorzakenformulieren die *geen* (n = 68 057) en de verzameling doodsoorzakenformulieren die *wel* (n = 18 836) een verschil vertonen bij vergelijking van automatisch en handmatig coderen op onderliggende doodsoorzaak (4-digit niveau) alsmede de formulieren die IRIS niet volledig kan verwerken (n = 35 674).

Voor de doodsoorzaken die het meest frequent voorkomen bij sterfgevallen met een verschil in onderliggende doodsoorzaak bij handmatig en automatisch coderen is middels kruistabellen naar een verklaring gezocht.

¹¹⁾ De overeenstemming tussen CBS codeurs (intercodeur variabiliteit) is bij herhaald coderen van een handmatig gecodeerd bestand overigens ook ongeveer 75 procent, zodat deze overall reproductie van codes niet slecht genoemd kan worden (Harteloh et al., 2010).

3.2.2 Onderliggende doodsoorzaak (ODO) in jaarbestand 2009 naar records die IRIS niet (rejected) en records die IRIS wel (automatic final) volledig kon coderen

Rang-nummer	Rejected (n=35 674)	Automatic Final		verklaring verschil			
		zelfde onderliggende doodsoorzaak (n=68 057)	verschil in onderliggende doodsoorzaak (n=18 836)				
		%	%	%			
1.	Longkanker (C34.9)	4,7	Longkanker (C34.9)	10,7	Longontsteking (J18.9)	10,2	Selectie ODO
2.	Acuut hartinfarct (I21.9)	4,2	Acuut hartinfarct (I21.9)	6,8	COPD (J44.9)	8,3	Selectie ODO/codering vierde digit (update)
3.	Longontsteking (J18.9)	4,0	Dementie (F03)	6,4	Atrium fibrilleren (I48)	4,3	Codering: vierde digit (update)
4.	CVA (I64)	2,9	Hartfalen (I50.9)	5,0	Diabetes (E14.6)	4,2	Codering: vierde digit
5.	Dementie (F03)	2,8	CVA(I64)	4,8	Hartfalen (I50.9)	3,6	Codering: vierde digit
6.	COPD (J44.9)	2,6	Borstkanker (C50.9)	3,4	Primaire maligniteit nos (C80)	3,2	Codering vierde digit/andere definitie van code (update)
7.	Hartfalen (I50.9)	2,2	COPD (J44.9)	3,2	Acuut hartinfarct (I21.9)	2,8	Selectie ODO
8.	Borstkanker (C50.9)	1,8	Dikgedarm kanker (C18.9)	3,0	Urineweginfectie (N39.0)	2,6	Selectie ODO
9.	Primaire maligniteit nos (C80)	1,6	Longontsteking (J18.9)	2,8	Alzheimer voor 65e (G30.1)	2,3	Andere coderingspraktijk
10.	Dikgedarm kanker (C18.9)	1,5	Nier/blaaskanker(C61)	2,7	CVA (I64)	2,0	Selectie ODO

Van alle sterfgevallen in het jaarbestand 2009 met longontsteking (J18.9) als onderliggende doodsoorzaak komt slecht 49 procent bij automatisch coderen op dezelfde doodsoorzaak uit (PCP=49 procent). Ongeveer 25 procent verschuift bij automatische codering naar dementie/ Alzheimer (F01.9, F03, G30.9), 6 procent naar het laat gevolg van hersenbloedingen (I69.4) en 5 procent naar sepsis (A41.9). Dit is een gevolg van een verandering van selectieregels voor de onderliggende doodsoorzaak. IRIS volgt de internationaal overeengekomen regel die stelt dat (broncho)pneumonie als complicatie van elke andere ziekte op het doodsoorzakenformulier mag worden beschouwd tenzij dit in de ICD-10 handleiding anders is aangegeven (WHO,1993, p39). Deze regel werd bij handmatig coderen niet gehanteerd. Een longontsteking werd alleen toegeschreven aan een andere ziekte als er expliciet een causaal verband was beschreven.

Van alle sterfgevallen in het jaarbestand 2009 met COPD (J44.9) als onderliggende doodsoorzaak komt slecht 58 procent bij automatisch coderen op dezelfde doodsoorzaak uit (PCP = 58 procent). Ongeveer 29 procent verschuift bij automatische codering naar COPD met luchtweginfectie: J44.0. Hier is sprake van een verschil op niveau van het vierde digit doordat IRIS COPD en longontsteking combineert tot J44.0. Deze combinatie wordt door een ICD-10 update uit januari 2010 voorgeschreven en werd bij handmatig coderen in 2009 nog niet gehanteerd. In deze verschuiving wordt het effect van een update zichtbaar. Ruim 5 procent verschuift naar hartfalen (I50.1 of I50.9). Hier wordt een andere selectie van de onderliggende doodsoorzaak zichtbaar bij een veel voorkomende vermelding van doodsoorzaken op eenzelfde formulier: COPD en hartfalen.

De code I48 (atriumfibrilleren) is bij een update van januari 2013 van een vierde cijfer voorzien. Van sterfgevallen met de code I48 in het jaarbestand 2009 krijgt bij automatische codering 76 procent de code I48.9. Het gevolg van de update wordt in het vierde digit van de ICD-10 code zichtbaar. Ongeveer 3 procent van de gevallen krijgt hypertensie (I10, I11) en 2 procent een hersenbloeding (I64) als onderliggende doodsoorzaak. Hier wordt een andere selectie van de onderliggende doodsoorzaak zichtbaar bij een veel voorkomende vermelding van doodsoorzaken op eenzelfde formulier: atriumfibrilleren en hersenbloeding of hypertensie.

Van sterfgevallen met de code diabetes mellitus met andere gespecificeerde complicaties (E14.6) is de PCP slechts 2 procent. In het jaarbestand 2009 verschuift bij automatische codering 73 procent naar de code voor een ongespecificeerde diabetes mellitus (E14.9), 6 procent naar de code voor diabetes mellitus met nier complicaties (E14.2), 5 procent naar het acuut hartinfarct (I21.9) en 2 procent naar hartfalen (I50). Er is hier dus voornamelijk een verschil in codering van het vierde digit dat complicaties bij diabetes aangeeft. Cardiovasculaire verschijnselen werden door de CBS codeurs als complicaties van diabetes aangewezen en in het vierde digit tot uitdrukking gebracht. Dit is niet in overeenstemming met de ICD-10 codeerregels (updates ICD-10 januari 2010). Hier wordt een verschil van opvatting in het coderen (niet in het selecteren) van de onderliggende doodsoorzaak zichtbaar.

Door de relatief hoge frequentie van voorkomen heeft hartfalen (I50.9) ook een groot aandeel in het verschil tussen handmatig en automatisch coderen, hoewel de PCP relatief hoog is: 83,5 procent. In het jaarbestand 2009 verschuift bij automatische codering ruim 5 procent naar linker hartfalen (I50.1), 1 procent naar hypertensief hartfalen (I11.0) en 1 procent naar atriumfibrilleren (I48.9). Een andere selectie van de onderliggende doodsoorzaak is verantwoordelijke voor deze verschuiving.

De code C80 (maligniteit n.o.s.) kreeg met ingang van januari 2010 een vierde digit waarbij de '0' wordt gereserveerd voor een maligniteit met een (expliciet) als onbekend gespecificeerde locatie en de '9' voor een maligniteit zonder (expliciet) gespecificeerde locatie. Van sterfgevallen met de code C80 in het jaarbestand 2009 verschuift bij automatische codering 41 procent naar de code C80.0 en 36 procent naar de code C80.9. Het effect van een ICD-10 update wordt hier zichtbaar. Ongeveer 14 procent verschuift bij automatisch coderen naar codes voor metastasen. Hier wordt een verschil van opvatting in het coderen zichtbaar. Metastasen werden door de CBS codeurs als C80 gecodeerd. De daarvoor bestemde ICD-10 codes (C78.0-C79.9) werden niet gebruikt. Bij automatisch coderen gebeurt dit wel.

Door de relatief hoge frequentie van voorkomen heeft het hartinfarct ook een groot aandeel in het verschil tussen handmatig en automatisch coderen. Van sterfgevallen met de code acuut hartinfarct (I21.9) in het jaarbestand 2009 komt bijna 90 procent bij automatisch coderen op dezelfde doodsoorzaak uit (PCP = 89,7). Ruim 1 procent verschuift bij automatische codering naar het oud infarct (I25.8), 1 procent naar longontsteking (J18.9) en 1 procent naar COPD (J44.9). Een andere selectie van de onderliggende doodsoorzaak is verantwoordelijk voor deze verschuivingen.

Van alle sterfgevallen in het jaarbestand 2009 met urineweginfectie (N39.0) als onderliggende doodsoorzaak komt slecht 50 procent bij automatisch coderen op dezelfde doodsoorzaak uit (PCP = 50 procent). In het jaarbestand 2009 verschuift bij automatische codering 29 procent naar dementie/Alzheimer, 6 procent naar de late gevolgen van een hersenbloeding (I69) en 1 procent naar hartfalen (I50.9). Deze verschuiving is een gevolg van de inrichting van de beslistabellen van IRIS. Een urineweginfectie wordt als complicatie van bijna elke andere ziekte op het doodsoorzakenformulier beschouwd. Dit werd bij handmatig coderen niet gedaan. Een andere selectie van de onderliggende doodsoorzaak wordt in deze verschuiving zichtbaar.

Van sterfgevallen met de code G30.1: 'ziekte van Alzheimer ontstaan boven de leeftijd van 65 jaar' komt bij automatisch coderen ruim 63 procent op dezelfde doodsoorzaak uit. Bijna

30 procent verschuift naar de code G30.9 'Alzheimer niet nader gespecificeerd'. In deze verschuiving wordt een andere opvatting van coderen zichtbaar. Bij handmatig coderen werd bij ontbreken van informatie over de ziekteduur gebruik gemaakt van de leeftijd om de code toe te kennen. Deze kwam daardoor altijd op de G30.0 (onder de 65 jaar) of G30.1 (boven de 65 jaar) uit. De code G30.9 werd bij handmatig coderen niet gebruikt. IRIS doet dit niet. IRIS werkt met het verschil tussen leeftijd en ziekteduur. Bij ontbreken van gegevens over de ziekteduur wordt de ziekte van Alzheimer als G30.9 gecodeerd.

Van sterfgevallen met de code hersenbloeding I64 (PCP = 90 procent) in het jaarbestand 2009 verschuift bij automatische codering 4 procent naar de code voor de late gevolgen van hersenbloeding (I69.4). Dit is een gevolg van een andere selectie van de onderliggende doodsoorzaak. IRIS verleent het late gevolg voorrang boven een recente hersenbloeding. Bij handmatig coderen was dit niet het geval.

3.3 Het handmatig en automatisch coderen van de European Shortlist (ESL)

Tabel 3.3.1 toont een vergelijking van het handmatig en automatisch coderen voor de 'European Shortlist for Causes of Death' (ESL), versie 2012 (Eurostat, 2015). De lijst met belangrijke doodsoorzaken (BELDO) zoals deze door het CBS op Statline wordt geplaatst is hiervan afgeleid. De ESL-2012 biedt meer detail. Voor het gemak van de lezer geven we de

3.3.1 Comparability Ratio (CR) en Perfect Compatibility Percentage (PCP) voor de ESL-2012

		Auto- matisch gecodeerd	Hand- matig gecodeerd	CR ¹⁾	PCP ten	PCP ten	Totaal 2009
					opzichte van hand- matig gecodeerd	opzichte van auto- matisch gecodeerd	
		Aantal		Ratio	Percentage		Aantal
1	Infectieziekten (A00–B99) ²⁾	1 805	1 225	1,47	82,0	55,7	2 022
1.1	Tuberculose (A15–A19, B90)	21	22	0,96	63,6	66,7	62
1.2	AIDS (B20–B24)	15	25	0,60	52,0	86,7	70
1.3	Virale hepatitis (B15–B19)	27	23	1,17	73,9	63,0	46
1.4	Overige infectieziekten	1 742	1 155	1,50	82,6	54,7	1 838
2	Nieuwvormingen (C00–D48)	28 845	29 070	0,99	97,5	98,2	42 396
2.1	Kwaadaardige Nieuwvormingen (KNV) (C00–C97)	28 096	28 583	0,98	96,7	98,3	41 322
2.1.1	KNV lip, mond en keel (C00–C14)	280	308	0,91	88,6	97,5	630
2.1.2	KNV slokdarm (C15)	1 081	1 100	0,98	96,8	98,5	1 569
2.1.3	KNV maag (C16)	1 047	1 059	0,99	97,0	98,1	1 492
2.1.4	KNV dikke darm, endeldarm en anus (C18–C21)	3 260	3 356	0,97	95,3	98,1	.
2.1.5	KNV lever en galwegen (C22)	448	465	0,96	91,8	95,3	522
2.1.6	KNV pancreas (C25)	1 754	1 791	0,98	96,9	99,0	2 379
2.1.7	KNV larynx (C32)	118	117	1,01	93,2	92,4	200
2.1.8	KNV trachea en long (C33–C34)	7 470	7 632	0,98	96,5	98,9	9 960
2.1.9	Melanoom (C43)	495	502	0,99	98,0	99,4	794
2.1.10	KNV borst (C50)	2 366	2 397	0,99	95,6	96,9	3 203
2.1.11	KNV cervix (C53)	148	146	1,01	97,3	95,9	209
2.1.12	KNV baarmoeder (C54–C55)	241	246	0,98	93,9	95,9	422
2.1.13	KNV eierstok (C56)	709	729	0,97	96,7	99,4	1 006
2.1.14	KNV prostaat (C61)	1 895	1 911	0,99	94,9	95,7	2 492
2.1.15	KNV nier/urineewegen (C64)	584	593	0,99	96,1	97,6	915
2.1.16	KNV blaas (C67)	848	874	0,97	93,9	96,8	1 204
2.1.17	KNV hersenen en centraal zenuwstelsel (C70–C72)	426	496	0,86	82,5	96,0	.

3.3.1 Comparability Ratio (CR) en Perfect Compatibility Percentage (PCP) voor de ESL-2012 (slot)

		Auto- matisch gecodeerd	Hand- matig gecodeerd	PCP ten opzichte van hand- matig gecodeerd CR ¹⁾	PCP ten opzichte van auto- matisch gecodeerd	Totaal 2009	
		Aantal	Ratio	Percentage	Aantal		
2.1.18	KNV schildklier (C73)	60	62	0,97	96,8	100,0	.
2.1.19	Ziekte van hodgkin en lymfomen (C81–C86)	701	679	1,03	96,5	93,4	.
2.1.20	Leukemie (C91–95)	821	791	1,04	96,2	92,7	.
2.1.21	KNV bloed/lymfe (C88, C90, C96)	488	459	1,06	96,7	91,0	3 086
2.1.22	Overige kwaadaardige nieuwvormigen	2 856	2 870	0,99	93,0	93,6	.
2.2	Niet kwaadaardige nieuwvorming (D00–D48)	749	487	1,54	95,3	61,9	1 074
3	Ziekten van bloed (D50–D89)	296	258	1,15	67,4	58,8	466
4	Endocriene/voedingsziekten (E00–E90)	2 874	2 478	1,16	85,3	73,5	3 826
4.1	Diabetes (E10–E14)	2 072	2 029	1,02	85,6	83,8	3 052
4.2	Overige endocriene ziekten	802	449	1,79	78,0	43,6	774
5	Psychische stoornissen (F00–F99)	7 250	5 512	1,32	95,2	72,4	7 085
5.1	Dementie (F01, F03, F05.1)	6 697	5 260	1,27	95,0	74,6	.
5.2	Alcohol (F10)	265	132	2,01	90,9	45,3	227
5.3	Drugs (F11–F16, F18–F19)	6	5	1,20	20,0	16,7	18
5.4	Overige psychische stoornissen	282	115	2,45	61,7	25,2	.
6	Ziekten van zenuwstelsel/zintuigen (G00–H95)	3 225	2 744	1,18	92,1	78,4	4 481
6.1	Parkinson (G20–G21)	750	828	0,91	83,0	91,6	1 170
6.2	Alzheimer (G30)	1 480	1 181	1,25	94,2	75,1	.
6.3	Overige ziekten zenuwstelsel/zintuigen	995	735	1,35	85,9	63,4	.
7	Ziekten van het hart en vaatstelsel (I00–I99)	28 271	27 870	1,01	93,5	92,3	38 897
7.1	Ischemische hartziekten (I20–I25)	7 582	7 713	0,98	90,8	92,4	10 665
7.1.1	Acuut hartinfarct (I21–I22)	5 190	5 351	0,97	89,6	92,4	7 104
7.1.2	Overige ischemische hartziekten (I20, I23–I25)	2 392	2 362	1,01	86,2	85,1	3 561
7.2	Andere hartziekten w.o. klepafwijkingen (I30–I52)	10 606	10 714	0,99	86,5	87,4	12 009
7.3	CVA. Cerebrovasulair Accident (I60–I69)	7 173	6 486	1,11	94,5	85,4	8 983
7.4	Overige ziekten van het hart en vaatstelsel	2 856	2 957	0,97	75,5	78,2	7 240
8	Ziekten van het ademhalingsstelsel (J00–J99)	7 803	10 065	0,78	71,6	92,3	13 962
8.1	Griep/influenza (J09–J11)	86	72	1,19	95,8	80,2	127
8.2	Longontsteking (J12–J18)	2 131	3 931	0,54	49,3	90,9	5 556
8.3	Chronische Luchtwegaandoeningen (J40–J47)	4 550	4 801	0,95	87,9	92,8	6 284
8.3.1	Astma (J45–J46)	70	39	1,80	89,7	50,0	58
8.3.2	COPD. Chronisch obstructieve Longaandoening (J40–J44, J47)	4 480	4 762	0,94	87,4	92,9	6 226
8.4	Overige ziekten van het ademhalingsstelsel	1 036	1 261	0,82	66,1	80,4	1 995
9	Ziekten van het spijsverteringsstelsel (K00–K93)	2 238	2 625	0,85	76,7	89,9	5 146
9.1	Zweer van maag en/of duodenum (K25–K28)	94	112	0,84	76,8	91,5	229
9.2	Chronische Leverziekten (K70, K73–K74)	343	362	0,95	83,7	88,3	819
9.3	Overige ziekten van het spijsverteringsstelsel	1 801	2 151	0,84	74,4	88,9	4 098
10	Ziekten van huid & bindweefsel (L00–L99)	114	162	0,70	49,4	70,2	350
11	Ziekten van spieren/bindweefsel (M00–M99)	327	281	1,16	80,1	68,8	789
11.1	Reumatoïde artritis, artrose (M05–06, M15–19)	118	67	1,76	86,6	49,2	236
11.2	Overige ziekten spieren/bindweefsel	209	214	0,98	72,0	73,7	553
12	Ziekten van nier/urinewegen & geslachtsorganen (N00–N99)	1 694	2 166	0,78	65,2	83,4	3 125
12.1	Ziekten van nier/urinewegen (N00–N20, N22–N29)	1 093	1 125	0,97	76,8	79,0	1 665
12.2	Overige ziekten nier/urinewegen & geslachtsorganen (N21, N30–99)	601	1 041	0,58	50,4	87,4	1 460
13	Complicaties van zwangerschap en bevalling (O00–O99)	0	1	0,00		0,0	9
14	Aandoeningen van perinatale periode (P00–P99)	32	30	1,07	83,3	78,1	384
15	Aangeboren afwijkingen (Q00–Q99)	167	166	1,01	84,9	84,4	460
16	Symptomen en onvolledig omschreven ziektebeelden (R00–R99)	2 006	2 240	0,90	77,8	86,9	5 219
16.1	Wiegendood (SIDS) R95	2	2	1,00	100,0	100,0	19
16.2	Onvolledig omschreven ziektebeelden en onbekende oorzaken (R96–R99)	393	399	0,98	79,7	80,9	2 825
16.3	Overige symptomen en onvolledig omschreven ziektebeelden (R00–R94)	1 611	1 839	0,88	76,8	87,6	2 375

¹⁾ CR vet = significant verschil tussen automatisch en handmatig coderen bij tweezijdige toetsing met T-test op percentages.

²⁾ Er is een klein verschil met de PCP in tabel 3.2.1, waar alleen de sterfgevallen met A=A en B=B geteld zijn bij de berekening van de PCP op hoofdstukniveau. Hier zijn alle gevallen A-B = A-B geteld.

PCP zowel met het aantal handmatig als met het aantal automatisch gecodeerde sterfgevallen (records) als noemer. Vergelijken van beide maten geeft een beeld van de gevoeligheid van IRIS enerzijds en van de specificiteit van IRIS anderzijds. Zo zien we bijvoorbeeld bij de doodsoorzaak dementie een hoge PCP ten opzicht van handmatig coderen, hetgeen aangeeft dat IRIS bijna alle gevallen (95 procent) die handmatig aan deze doodsoorzaak worden toegekend op dezelfde manier codeert. De PCP met het automatisch aantal gecodeerde gevallen als noemer is echter een stuk lager (74,6 procent). Dit betekent dat IRIS een aanzienlijk deel van de sterfgevallen die handmatig niet dementie als onderliggende doodsoorzaak kregen toegewezen bij automatisch coderen wel aan deze doodsoorzaak toewijst. Dit komt ook in de CR tot uitdrukking. Deze is (veel) groter dan 1.

Er zijn significante verschillen tussen automatisch en handmatig coderen voor 26 van de 75 groepen van doodsoorzaken uit de meest recente versie van de ESL. Ceteris paribus verwachten we deze verschillen terug te zien bij de implementatie van IRIS in de praktijk. Bij automatisch coderen worden significant meer infectieziekten, niet-kwaadaardige nieuwvormingen, endocriene/voedingsziekten, psychische stoornissen (waaronder dementie en alcoholgebruik), ziekten van het zenuwstelsel (waaronder de ziekte van Alzheimer), het CVA en reumatoïde artritis als onderliggende doodsoorzaak aangetroffen dan bij handmatig coderen. Er worden significant minder ziekten van de ademhalingsorganen (waaronder longontsteking, astma en COPD), ziekten van het spijsverteringsstelsel, ziekten van huid en bindweefsel, ziekten van de nier en urinewegen en symptomen of onvolledig omschreven ziektebeelden als onderliggende doodsoorzaak aangetroffen dan bij handmatig coderen.

Soms worden verschuivingen niet specifiek zichtbaar in de ESL. Zo vindt bijvoorbeeld in het hoofdstuk van de infectieuze ziekten de belangrijkste verandering plaats bij de sterfte aan diarree en sepsis. Deze doodsoorzaken zijn niet onderscheiden in de ESL. Zij zitten in de klasse 'overige infectieziekten' die dan ook het patroon van de hoofdgroep volgt, terwijl de binnen het hoofdstuk onderscheiden doodsoorzaken dat niet doen. Hetzelfde doet zich ook voor bij de sterfte aan niet-kwaadaardige nieuwvormingen, endocriene/voedingsziekten, psychische stoornissen, ziekten van het zenuwstelsel/zintuigen.

Bij de sterfte aan niet kwaadaardige nieuwvormingen zijn geen klassen onderscheiden. De belangrijkste klassen waar zich hier de verschuivingen voordoen worden gevormd door de sterfte aan niet kwaadaardige nieuwvormingen van hersenen/meningen, longen of darmen, en het myelodysplastisch syndroom (D46).

Bij de sterfte aan overige endocriene/voedingsziekten wordt het verschil tussen handmatig en automatisch coderen vooral veroorzaakt door de sterfte aan dehydratie, cachexie en de risicofactoren adipositas en hypercholesterolemie. Hoewel correct gecodeerd past de selectie van dehydratie of cachexie bij automatisch coderen doorgaans niet bij de bedoeling van de invuller van het doodsoorzakenformulier. Het betreft verschijnselen van het overlijden die IRIS selecteert als onderliggende doodsoorzaak. De selectie van de onderliggende doodsoorzaak is hier niet valide. Door een voorselectie en/of controle van de codes voor cachexie, dehydratie zal foutieve selectie als onderliggende doodsoorzaak in het productieproces van de doodsoorzakenstatistiek ondervangen moeten worden. Bij vermelding van adipositas of hypercholesterolemie op het doodsoorzakenformulier zal door interpretatie van de context moeten worden beoordeeld of het gaat om een (epidemiologische) risicofactor dan wel een feitelijke doodsoorzaak.

Ook bij de sterfte aan de overige psychische stoornissen treedt een aanzienlijk verschil op tussen handmatig en automatisch coderen. Dit wordt vooral veroorzaakt door codering

van het delier als onderliggende doodsoorzaak bij automatisch coderen. Hoewel correct gecodeerd past de selectie van het delier doorgaans niet bij de bedoeling van de invuller van het doodsoorzakenformulier. Het is een verschijnsel van overlijden en niet een psychiatrische aandoening. De selectie van de onderliggende doodsoorzaak is hier niet valide. Door controle van de codering zal foutieve selectie van delier als onderliggende doodsoorzaak in het productieproces van de doodsoorzakenstatistiek ondervangen moeten worden. Het significant verschil tussen handmatig en automatisch coderen bij de sterfte aan overige ziekten van zenuwstelsel/zintuigen wordt vooral veroorzaakt door een verschil in codering van de status epilepticus (G40.9) en de TIA (G45.9). Bij automatisch coderen worden deze ziekten anders dan bij handmatig coderen als onderliggende doodsoorzaak beschouwd, hetgeen op eerste gezicht niet voor de hand ligt. De status epilepticus is onderdeel van de ziekte epilepsie en de TIA per definitie van voorbijgaande aard. Daarbij wordt de term 'status epilepticus' soms ook gebruikt om verschijnselen van het overlijden aan te duiden, te weten stuipen die de persoon bij overlijden heeft als gevolg van een zuurstoftekort in de hersenen. Ook hier zal door controle van de codering foutieve selectie van 'status epilepticus' als onderliggende doodsoorzaak in het productieproces van de doodsoorzakenstatistiek ondervangen moeten worden. De selectie van de TIA berust op de ICD-10 selectieregel die stelt dat een triviale aandoening onderliggende doodsoorzaak kan zijn als deze de oorzaak is van een andere (triviale) aandoening. Ook hier toont IRIS zich niet sterk in de interpretatie van de context en worden relevante doodsoorzaken (onterecht) vermeld op deel 2 van het doodsoorzakenformulier over het hoofd zien.

De afname van sterfte aan de overige ziekten van het ademhalingsstelsel bij automatisch coderen wordt vooral veroorzaakt door een afname van luchtweginfecties (J20–22). Deze volgen bij selectie als onderliggende doodsoorzaak hetzelfde patroon als de longontsteking en COPD en krijgen dementie of de ziekte van Alzheimer als onderliggende doodsoorzaak toegewezen.

De afname van de sterfte aan overige ziekten van nier of urinewegen bij automatisch coderen wordt vooral veroorzaakt door een afname van de urineweginfectie die net als longontsteking door IRIS bij voorkeur als gevolg van een andere ziekte of aandoening wordt gezien.

3.4 Het vergelijken van automatisch en handmatig coderen

Het handmatig coderen kent in Nederland een meer dan honderd jaar oude traditie waarbij interpretaties van het doodsoorzakenformulier ingekleurd zijn door codeergewoonten of inzichten die kunnen verschillen van die in andere landen. Zo heeft het CBS sinds de ingebruikname van de ICD-10 in 1996 de ICD-10 updates niet gevolgd en de voorgeschreven wijzigingen niet in de classificatie van doodsoorzaken aangebracht.¹²⁾ Het volgen van de updates bleek te ingewikkeld voor handmatige codering en de verschuivingen in statistiek niet te onderzoeken of te documenteren. IRIS werkt echter met codes en tabellen waarin de ICD-10 updates zijn verwerkt zodat deze onvermijdelijk deel uitmaken van het automatisch coderen. Dit heeft onder meer invloed op de selectie van de onderliggende doodsoorzaak en op het gebruik van ICD-10 codes. Een wijziging van selectieregels heeft grote invloed op

¹²⁾ Sinds 1999 heeft er een groot aantal jaarlijkse wijzigingen van codes (verwijderen, toevoegen) en bijstellingen van definities plaatsgevonden. Updates komen tot stand via een procedure waarbij gebruikers van de ICD-10 hiaten of fouten in de classificatie aangeven, die vervolgens in een internationaal forum (Mortality Reference Group) worden besproken en na bekrachtiging door de WHO jaarlijks als officiële aanvullingen of wijzigingen van de ICD-10 worden afgekondigd (zie: <http://www.who.int/classifications/icd/icd10updates/en/>).

de verschuivingen in statistiek, maar is moeilijk te kwantificeren. Gelet op de uitstroom uit de ICD-10 hoofdstukken 8 en 12 kan de verandering van de ICD-10 selectieregel 3 alleen al verantwoordelijk worden gehouden voor een verschuiving van ongeveer 2 600 sterfgevallen, ofwel 14 procent van het waargenomen verschil tussen handmatig en automatisch coderen (2 600/18 863).

3.4.1 Het effect van ICD-10 updates op het coderen van doodsoorzaken

ICD-10 code gebruikt bij handmatig coderen, maar niet bij automatisch coderen	Aantal	ICD-10 code (na update) gebruikt bij automatisch coderen, maar niet bij handmatig coderen	Aantal	ICD-10 code (voor update) bij handmatig coderen
B34.9	65	A09.0	13	A09, K52.9
C85.0	2	A09.9	230	K52.9
C94.5	4	C80.9	431	C80
C96.1	1	C84.9	26	C84.5, C85
D75.2	8	C96.8	3	C44.9
D76.0	1	D47.4	13	D74.1
I84.8	1	G21.4	2	G20
K35.0	15	J09	10	.
K35.1	3	K12.3	4	.
K35.9	2	K22.7	2	.
N18.0	325	K35.3	2	.
N18.8	2	K75.4	7	M35.9, K73.8
		K85.1	8	K85
		K85.2	3	K86
		K85.9	73	K85, K86.1
		L89.9	65	L89
		M72.6	10	L08.8
		N18.5	213	N18.0, N18.9
		R26.3	1	.
		R63.6	6	.
Totaal	429 (waarvan 286 automatic final)	Totaal	1 122	

Het verschil in geldige ICD-10 codes is weergegeven in *tabel 3.4.1*. In 2009 werden bij handmatig coderen 429 sterfgevallen gecodeerd met behulp van 12 codes die volgens de ICD-10 updates nu niet meer geldig zijn. Het merendeel daarvan betreft N18.0, het eindstadium van een nierziekte. Door een specificatie van de code N18 op 4-digit niveau (januari 2010) verschuiven deze gevallen naar N18.5. Ook werd de code B34.9 (overige virus infecties/viraemie) geschrapt. De sterfgevallen met deze code als onderliggende doodsoorzaak verschuiven naar B34.8. IRIS gebruikt 20 (nieuwe) ICD-10 codes die door het negeren van de updates bij handmatig coderen niet werden gebruikt. De vijfde kolom van *tabel 3.4.1* toont waar de desbetreffende sterfgevallen bij handmatig coderen worden aangetroffen. De geschrapte en de nieuwe ICD-10 codes genereren per definitie een verschil tussen automatisch en handmatig coderen. Uit de *tabel 3.4.1* is af te lezen dat 1 408 sterfgevallen een onderliggende doodsoorzaak hadden die bij automatisch coderen afweek van de onderliggende doodsoorzaak bij handmatig coderen. Naar schatting is dan ook minstens 7,5 procent van het waargenomen verschil tussen handmatig en automatisch coderen toe te schrijven aan een verandering van codes door de ICD-10 updates (1 408/18 836).

3.5 De kwaliteit van de doodsoorzakenstatistiek

In dit onderzoek beschrijven we de verschillen tussen handmatig en automatisch coderen. De resultaten van de ‘bridge coding study’ overziend, valt op dat de CBS codeurs en IRIS vooral van mening verschillen over de rol van longontsteking, dementie, hartfalen, sepsis en COPD bij selectie van de onderliggende doodsoorzaak. Een beoordeling van de waargenomen verschillen is lastiger. Bij het vergelijken van handmatig en automatisch coderen is er hiervoor geen gouden standaard. Een doodsoorzakenformulier laat soms meerdere interpretaties bij de selectie van een onderliggende doodsoorzaak toe, bijvoorbeeld: longontsteking op 1a en dementie op deel 2 van het doodsoorzakenformulier (zie figuur 5.1). Handmatig werd longontsteking als onderliggende doodsoorzaak gekozen, omdat er geen expliciet causaal verband tussen de longontsteking en dementie wordt beschreven. IRIS selecteert in deze gevallen ‘dementie’ als onderliggende doodsoorzaak onder de veronderstelling dat de persoon zonder de dementie niet aan de longontsteking zou zijn overleden. Selectie bij handmatig coderen was niet ‘fout’.¹³⁾ IRIS hanteert echter een andere interpretatie van het doodsoorzakenformulier en is daarin (onvermijdelijk) consistent. De consistentie komt ten goede aan de stabiliteit van de doodsoorzakenstatistiek. Daarbij is het zaak evidente fouten in het coderen zoveel mogelijk te vermijden. Hiertoe vindt voorselectie van formulieren en nacontrole op codes plaats. Ook bestaat er de mogelijkheid de IRIS beslistabellen aan te passen door participatie in de IRIS Core Group en WHO Mortality Reference Group. Bij gebruik van IRIS is er dan ook altijd de afweging van klakkeloos volgen tegen aanpassen van het systeem, ofwel het afwegen van de betrouwbaarheid (consistentie in tijd of plaats/land) tegen de inhoudelijke validiteit van de codering die aan een couleure locale onderhefing kan zijn.

4. Conclusies

Dit rapport doet verslag van een onderzoek naar overeenkomsten en verschillen tussen het automatisch en handmatig coderen van doodsoorzaken. Eenzelfde verzameling doodsoorzakenformulieren is zowel handmatig als automatisch gecodeerd. De relatief grote mate van overeenstemming (78 procent) tussen handmatig en automatisch coderen maakt de overgang op het automatisch coderen van doodsoorzaken verantwoord. Daarbij treden er eenmalige verschuivingen in de statistiek op. Bij automatisch coderen worden er significant *meer* infectieziekten (‘niet gespecificeerde diarree’ als infectieziekte in plaats van een spijsverteringsziekte, sepsis), uitdroging, dementie, de ziekte van Alzheimer en het laat gevolg van een CVA, en significant *minder* longontsteking, COPD, ziekten van het spijsverteringsstelsel (‘niet gespecificeerd diarree’ als spijsverteringsziekte), urineweginfecties en ouderdom als onderliggende doodsoorzaak waargenomen dan bij handmatig coderen. Wat betreft het vóórkomen van belangrijke doodsoorzakengroepen als kwaadaardige nieuwvormingen en hart- en vaatziekten is het verschil tussen het automatisch en handmatig coderen gering. Er is een aantal verklaringen voor de waargenomen verschillen. Ten eerste is er een *andere opvatting over de selectie van de onderliggende doodsoorzaak* bij vermelding van meer dan één doodsoorzaak op het doodsoorzakenformulier (bijvoorbeeld: bij de verschuiving van longontsteking, lucht- of urineweginfecties naar dementie/Alzheimer en de selectie van het

¹³⁾ Zo zijn sommige (major) updates pas na 2010 van kracht geworden, zodat opnieuw coderen van het jaarbestand 2009 met een IRIS versie uit 2014 eerder het effect van de ICD-10 updates op de statistiek weergeeft dan een discrepantie bij coderen met behulp van dezelfde regels.

laat gevolg van een CVA). Ten tweede is er een verschil in codering van medisch begrippen: a. door wijzigingen in het *woordenboek* dat door IRIS bij automatisch coderen wordt gebruikt (bijvoorbeeld bij de codering van het buikaneurysma of het glioom), en b. door het gebruik van de *ICD-10 updates*. Deze updates werden bij handmatig coderen niet gebruikt, maar zijn integraal onderdeel van het automatisch coderen door IRIS en verantwoordelijk voor ongeveer 21 procent van het waargenomen verschil tussen handmatig en automatisch coderen. Ten derde is er bij automatisch coderen een grotere afhankelijkheid van de *positie van een term* op het doodsoorzakenformulier dan bij handmatig coderen. Zo is er in het onderzoek een verschuiving van longkanker naar COPD waargenomen, doordat IRIS (strikt genomen terecht) de causale keten afbreekt als COPD boven een longcarcinoom vermeld staat, of COPD als doodsoorzaak selecteert als dit samen met een longcarcinoom als eerste in eenzelfde vakje op het doodsoorzakenformulier staat. Bij handmatig coderen werd dit niet gedaan. COPD werd dan als 'symptoom' (ernstige benauwdheid) ofwel als een obstructie van de luchtwegen bij longcarcinoom geduid, tenzij tijdsduur of volgorde andere aanwijzingen gaven. Hetzelfde wordt waargenomen bij de verschuiving van darmmaligniteiten naar darmziekten. Ook bij de interpretatie van het begrip 'tumor' werd er bij handmatig coderen gebruik gemaakt van andere informatie op het doodsoorzakenformulier. Wanneer er bijvoorbeeld ook cachexie of embolie op het doodsoorzakenformulier stond, werd 'de tumor' als maligne aangemerkt. IRIS doet/kan dit niet. Bij handmatig coderen wordt de context van een term zo meer in beschouwing genomen dan bij automatisch coderen. Ten vierde is er bij automatisch coderen vooralsnog geen *interpretatie van aanvullende tekst* op het doodsoorzakenformulier mogelijk. Tot slot is er een *codeertechische beperking* (Bijvoorbeeld bij de niet gespecificeerde luchtweginfectie. Deze wordt (correct) als J98 gecodeerd, maar dan door IRIS niet als zodanig herkend (J98 is een overige longaandoening) en dientengevolge bij de selectie van de onderliggende doodsoorzaak anders behandeld dan luchtweginfecties die wel in een code voor een infectieziekte van de luchtwegen terecht komen). Deze factoren kunnen worden gebruikt om het verschil tussen handmatig en automatisch coderen (lees de jaarbestanden 2012 en 2013) te duiden bij in gebruik name van IRIS. Verschillen tussen handmatig en automatisch coderen kunnen (gedeeltelijk) worden ondervangen bij gebruik van IRIS in de routine productie van de doodsoorzakenstatistiek door een voorselectie van doodsoorzakenformulieren (bijvoorbeeld van formulieren met vrije tekst), nacontrole op codes (bijvoorbeeld bij de selectie van delier, status epilepticus of TIA), het plaatsen van vlaggen (bijvoorbeeld bij codering van de ziekte van Alzheimer of het CVA op deel 2 van het doodsoorzakenformulier) of het aanpassen van de tabellen waarmee IRIS werkt.

5. Begrippen

Codeervlaggen

Voorafgaande aan het automatisch coderen door IRIS wordt er software toegepast die een term kan markeren, dat wil zeggen, van een vlag kan voorzien. Door deze markering van begrippen kunnen zij afhankelijk van de context van de juiste ICD-10 code worden voorzien. Zo zijn er vlaggen voor leeftijd, geslacht, duur, positie op het doodsoorzakenformulier (deel 1 of 2) dan wel het al of niet samengaan met andere begrippen op het doodsoorzakenformulier. Bijvoorbeeld: als een CVA op deel 2 staat van het doodsoorzakenformulier zonder tijdsinterval wordt dit begrip van een vlag voorzien, waardoor IRIS het zal coderen als laat gevolg van een CVA (I69.4) in plaats van vers/recent CVA (I64). Zonder markering door een vlag kan IRIS dit niet en zal een dergelijk CVA als I64 worden gecodeerd.

Comparability Ratio (CR)

Gedefinieerd als de frequentie van een ICD-10 code bij automatisch coderen (IRIS) gedeeld door de frequentie van een ICD-10 code bij handmatig coderen in hetzelfde bestand. De CR is een maat voor de betrouwbaarheid van twee methoden ten opzichte van elkaar.

$$CR = \text{Freq ICD-10 code } X_{\text{IRIS}} / \text{Freq ICD-10 code } X_{\text{handmatig}}$$

De CR vergelijkt twee verschillende methoden van coderen (handmatig versus automatisch) en geeft daarbij een beeld van de te verwachten verschuivingen in statistiek. We toetsen de CR-s op significantie met behulp van een tweezijdige T-toets.¹⁴⁾

ICD-10 Grondregel

Wanneer meer dan één ziekte of aandoening op het doodsoorzakenformulier is vermeld (80 procent van de formulieren), mag de ziekte of aandoening die als enige op de onderste gebruikte regel van deel 1 staat, alleen dan geselecteerd worden als onderliggende doodsoorzaak wanneer deze de oorzaak geweest kan zijn van alle daarboven vermelde ziekten of aandoeningen (WHO, 1993, p. 36). De grondregel wordt aangevuld met een drietal selectieregels die aangeven hoe te handelen als de grondregel niet opgaat. Vooral de derde regel leidt tot interpretatieverschillen bij het selecteren van de onderliggende doodsoorzaak.

ICD-10 Selectieregel Drie

Indien de ziekte of aandoening die door de grondregel is geselecteerd duidelijk een direct gevolg is van een andere onder deel 1 of deel 2 op het doodsoorzakenformulier vermelde ziekte of aandoening, selecteer deze dan als onderliggende doodsoorzaak (WHO, 1993, pp. 39–41). Op basis van deze regel wordt door IRIS bijvoorbeeld een longontsteking of urineweginfectie op 1a door IRIS toegeschreven aan een dementie vermeld op deel 2 (zie *figuur 5.1*). Bij handmatig coderen werd dit niet gedaan. Er is hier sprake van een verschil van inzicht in het 'direct gevolg' van een ziekte of aandoening.

IRIS

Computerprogramma voor het automatisch coderen van doodsoorzaken (Internationale standaard). Door middel van een woordenboek dat medische uitdrukkingen vertaalt in ICD-10 codes is het programma in staat doodsoorzakenformulieren automatisch te coderen en vervolgens de onderliggende doodsoorzaak te bepalen met behulp van de beslistabellen van het Amerikaanse systeem MMDS.

MMDS

Medical Mortality Data System. Softwarepakket voor het coderen van doodsoorzaken en het bepalen van de onderliggende doodsoorzaak (Israel et al., 1972; Israel, 1990). Het bestaat uit verschillende modules: tabellen voor het selecteren van de onderliggende doodsoorzaak (ACME), lijsten met controles (bijvoorbeeld op leeftijd of geslacht) van codes die in ACME worden ingevoerd (MICAR200), een omzetting van ICD codes in willekeurig gekozen getallen (entity reference numbers, ERN) om zo vertaling van medische uitdrukkingen onafhankelijk te maken van verschillende ICD versies (MICAR100) en een Engelstalige (tekst) invoer module (SUPERMICAR) (Lu, 2003; Lu et al., 2005).

¹⁴⁾ Voor een algemene definitie van de CR en de berekening van het bijbehorende betrouwbaarheidsinterval: zie Statistics Canada, 2005.

5.1 De delen 1 en 2 van het doodsoorzakenformulier

d Natuurlijke dood <i>(zie toelichting ommezijde)</i>	
1a Ziekte die rechtstreeks de dood tot gevolg had	duur tussen begin ziekte en overlijden
<input type="text"/>	<input type="text"/>
1b Ziekten die hebben geleid tot de doodsoorzaak onder a. + Bij vermelding van meerdere ziekten de aan het overlijden ten grondslag liggende ziekte onder c opgeven.	
b veroorzaakt door/gevolg van	duur tussen begin ziekte en overlijden
<input type="text"/>	<input type="text"/>
c veroorzaakt door/gevolg van	duur tussen begin ziekte en overlijden
<input type="text"/>	<input type="text"/>
2 Bijkomstige, bij overlijden nog bestaande ziekten en bijzonderheden die tot de dood hebben bijgedragen, doch niet met de onder 1 genoemde ziekten in causaal verband staan.	duur tussen begin ziekte en overlijden
<input type="text"/>	<input type="text"/>

Onderliggende doodsoorzaak

(a) De ziekte of aandoening waarmee de causale reeks van ziekten of aandoeningen die rechtstreeks tot de dood leidt, een aanvang neemt (natuurlijke dood); (b) de omstandigheden van het ongeval of geweld die het letsel veroorzaakt waarmee de causale reeks van gebeurtenissen die rechtstreeks tot de dood leidt, een aanvang neemt (niet-natuurlijke dood) (WHO, 1993, p. 31). De doodsoorzakenstatistiek is een tabellering van onderliggende doodsoorzaken. Hierbij wordt de grondregel gebruikt. De grondregel is het selectieprincipe voor de onderliggende doodsoorzaak.

Perfect Compatibility Percentage (PCP)

Gedefinieerd als het aantal sterfgevallen dat bij automatisch coderen (IRIS) en bij handmatig coderen precies dezelfde code voor de onderliggende doodsoorzaak krijgt toegekend, gedeeld door het aantal sterfgevallen dat aan die code werd toegeschreven bij handmatig coderen, in hetzelfde bestand.¹⁵⁾ De PCP is een maat voor de validiteit van twee methoden ten opzichte van elkaar.

$$PCP = \frac{\text{Aantal (ICD-code}_{IRIS} = \text{ICD-code}_{handmatig})}{\text{Aantal ICD-code}_{handmatig}}$$

Aangezien we overgaan van handmatig naar automatisch coderen is het aantal handmatig gecodeerde gevallen hier als noemer gekozen. De PCP moet met de nodige omzichtigheid worden geïnterpreteerd. Een hoge PCP kan immers het gevolg zijn van een 'overcodering' door IRIS. Stel IRIS zet alle aangeboden sterfgevallen op eenzelfde code, dan is de PCP voor

¹⁵⁾ Voor een algemene definitie van de PCP: zie Gjertsen et al., 2013.

de desbetreffende code 100 procent, omdat alle gevallen die handmatig aan die code zijn toegeschreven daarbij zullen zitten (teller = noemer). Dit is niet de bedoeling. IRIS is dan weinig specifiek. De PCP moet daarom *ook* met de andere noemer worden beschouwd, namelijk als het aantal sterfgevallen dat bij automatisch coderen (IRIS) en bij handmatig coderen precies dezelfde code voor de onderliggende doodsoorzaak krijgt toegekend, gedeeld door het aantal sterfgevallen dat bij automatische codering aan die code werd toegeschreven. In het fictieve voorbeeld is deze PCP (zeer) laag aangezien de teller nu altijd (veel) kleiner zal zijn dan de noemer. Daarbij is er een samenhang tussen PCP en CR. Een hoge PCP en een lage CR wijzen op een gebrek aan specificiteit. Als voorbeeld geven we een kruistabel voor de code I21.9 'het acuut hartinfarct'.¹⁶⁾

5.2 Voorbeeld van een kruistabel voor de onderliggende doodsoorzaak I21.9 (acuut hartinfarct) met aantallen a-f:

		Automatisch gecodeerd (IRIS)		
		Niet I21.9	I21.9	
Handmatig gecodeerd	Niet I21.9	a	b	e (=c +d)
	I21.9	c	d	
			f (=b +d)	

$$CR_{I21.9} = f/e$$

$$PCP_{I21.9} = d/e \text{ (sensitiviteit van IRIS) of } PCP_{I21.9} = d/f \text{ (specificiteit van IRIS)}$$

$$\text{Specificiteit} = d/f \Rightarrow d/e \times 1/CR$$

Reguliere expressies

Voorafgaande aan het automatisch coderen door IRIS wordt er software toegepast (reguliere expressies) die een variëteit aan uitdrukkingen (w.o. synoniemen maar ook spelfouten) van eenzelfde begrip terugbrengt tot één (standaard)term. Een reguliere expressie (afgekort: regex) is een uit de informatica afkomstige methode om in tekenreeksen (waaronder lettertekens) gemakkelijk patronen te specificeren voor zowel patroonherkenning als patroonsubstitutie. In de context van IRIS en automatisch coderen van doodsoorzaken slaat het op het gebruik van deze methode om 'ruwe' tekst, zoals geschreven op het doodsoorzakenformulier om te zetten in een gestandaardiseerde vorm. Daarbij worden vormvariaties teruggebracht tot voorkeursvormen die dezelfde betekenis hebben. Bijvoorbeeld: 'myocardinfarct' of 'hartinfarkt' worden omgezet in één en dezelfde term: 'hartinfarct'. Alleen deze standaardterm wordt vervolgens ter codering aan IRIS aangeboden. Hierdoor kan de omvang van het woordenboek dat IRIS gebruikt om termen te coderen sterk worden beperkt.

¹⁶⁾ Alle codes in dit rapport zijn volgens de *International Statistical Classification of Diseases and Related Health Problems. Tenth Revision* (WHO, 1992).

6. Referenties

Birch, D. (1993). Automatic coding of causes of death. *Population Trends*, (15), 53–56.

Eurostat (2013). Statistics explained. Geraadpleegd van: http://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Bridge_coding&oldid=120312.

Eurostat (2015). Metadata. European Shortlist for causes of death, 2012. Geraadpleegd van: http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfmTargetUrl=LST_NOM_DTL&StrNom=COD_2012&Str op 12-11-2015.

Gjertsen, F., Bruzzone, S., Vollrath, M.E., Pace, M., & Ekeberg, Ø. (2013). Comparing ICD-9 and ICD-10: The impact on intentional and unintentional injury mortality statistics in Italy and Norway. *Int. J. Care Injured*, (44), 132–138.

Harteloh, P.P.M. (2014). *Verschuivingen in de doodsoorzakenstatistiek bij de introductie van het automatisch coderen*. Den Haag: Centraal Bureau voor de Statistiek (webartikel, www.cbs.nl).

Harteloh, P., Hilten O., & Kardaun, J. (2014). *Het automatische coderen van doodsoorzaken. Een nieuwe werkwijze bij de doodsoorzakenstatistiek*. Den Haag: Centraal Bureau voor de Statistiek (webartikel, www.cbs.nl).

Harteloh, P. (2015). Automated Coding of Causes of Death in The Netherlands. In: *Newsletter on the WHO-FIC*. Bilthoven: RIVM, (13), 13–16.

Harteloh, P.P.M., Bruin, K. de., & Kardaun, J. (2010). The reliability of cause-of-death coding in The Netherlands. *European J of Epidemiology*, (25), 531–53.

Israel, R.A. (1990). Automation of mortality data coding and processing in the USA. *World Health Statistics Quarterly*, (43), 259–262.

Israel, R.A., Templeton, M.C., & Evans, M.C. (1972). New approaches to coding and analysing mortality data. *American Statistical Association*, 20–24.

Lu, T.H. (2003). Using ACME software to monitor and improve the quality of death statistics. *J Epidemiology Community Health*, (57), 470–473.

Lu, T.H., Tsau, S.M., & Wu, T.C. (2005). The automated classification of medical entities (ACME) system objectively assessed the appropriateness of underlying cause-of-death certification and assignment. *J Clin Epidemiol*, (58), 1277–1281.

McKenzie, K., Walker, S., & Tong, S. (2001). Assessment of the impact of the change from manual to automated coding on mortality statistics in Australia. *Health Information Management Journal*, (15), 1–13.

Pavillon, G., Coleman, M., Johansson, L.A., Jouglà, E., & Kardaun, J. (1998). *Final report on automated coding in member states*. Luxembourg: Eurostat.

Pavillon, G., Johansson, L.A. (2001). *Production of Methods and Tools for Improving Causes of Death statistics at Codification Level*. Luxembourg (Eurostat Secretariat – Unit E3): European Commission.

Sonsbeek, J.L.A. van (2005). *Van schaduw des doods tot een licht ten leven. De historie van de methodiek van de doodsoorzakenstatistiek in Nederland 1865–2005*. Voorburg, Heerlen: Centraal Bureau voor de Statistiek (CBS).

Statistics Canada (2005). Health Statistics Division. *Comparability of ICD-10 and ICD-9 for Mortality Statistics in Canada*. Ottawa: Ministry of Industry.

Stegen, R. van der, Koren, W., Harteloh, P., Kardaun, J., & Jansen F. (2014). A Novel Time Series Approach to Bridge Coding Changes with a Consistent Solution across Causes of Death. *European J Population*, (30), 317–335.

WHO. World Health Organisation (1992). *International Statistical Classification of Diseases and Related Health Problems*. Tenth Revision. Volume 1. Geneva: WHO.

WHO. World Health Organisation (1993). *International Statistical Classification of Diseases and Related Health Problems*. Tenth Revision. Volume 2. Geneva: WHO.

Verklaring van tekens

Niets (blanco)	Een cijfer kan op logische gronden niet voorkomen
.	Het cijfer is onbekend, onvoldoende betrouwbaar of geheim
*	Voorlopige cijfers
**	Nader voorlopige cijfers
2014–2015	2014 tot en met 2015
2014/2015	Het gemiddelde over de jaren 2014 tot en met 2015
2014/'15	Oogstjaar, boekjaar, schooljaar enz., beginnend in 2014 en eindigend in 2015
2012/'13–2014/'15	Oogstjaar, boekjaar, enz., 2012/'13 tot en met 2014/'15

In geval van afronding kan het voorkomen dat het weergegeven totaal niet overeenstemt met de som van de getallen.

Colofon

Uitgever

Centraal Bureau voor de Statistiek
Henri Faasdreef 312, 2492 JP Den Haag
www.cbs.nl

Prepress

Studio BCO, Den Haag

Ontwerp

Edenspiekermann

Inlichtingen

Tel. 088 570 70 70
Via contactformulier: www.cbs.nl/infoservice

© Centraal Bureau voor de Statistiek, Den Haag/Heerlen/Bonaire, 2015.
Verveelvoudigen is toegestaan, mits CBS als bron wordt vermeld.