

R.W.M. Giard

5.1 Inleiding

In de voorgaande hoofdstukken is duidelijk gemaakt dat patiëntveiligheid continu aandacht verdient: het is zowel een onderschat als een structureel probleem, niet maar een handjevol incidenten. Het is om verschillende redenen belangrijk om onderzoek te doen naar de frequentie van fouten. Hoe vaak worden er fouten gemaakt? Welke verschillen bestaan er tussen foutfrequenties van artsen, ziekenhuizen, regio's en landen? Wat kunnen we te weten komen over de oorzaken van misslagen? Hebben maatregelen die genomen worden voor verbetering van de patiëntveiligheid ook daadwerkelijk effect in de zin van minder iatrogene schade en minder slachtoffers?^{1,2} Dat vraagt nauwgezet en systematisch onderzoek van de medische praktijk.

Nu er steeds vaker van artsen wordt gevraagd de kwaliteit van medische zorg duidelijk te maken, zullen bruikbare en betrouwbare gegevens dienen te worden gepresenteerd om zo verantwoording af te kunnen leggen.³ Bij onderzoek naar beroepsfouten stuit men op een aantal belangrijke methodologische kwesties: Hoe definieer je een medische fout? Hoe ontwikkel je meetinstrumenten om deze problematiek te onderzoeken? Hoe accuraat en reproduceerbaar zijn de uitkomsten van een dergelijk onderzoek? Hoe kunnen fouten qua ernst worden gerangschikt en wat is de klinische betekenis van onjuist medisch handelen? Hoe moet je de resultaten van onderzoek naar fouten interpreteren? Pas als deze vragen bevredigend zijn beantwoord, komen we toe aan de kwestie hoe de frequentie van medische missers kan worden verminderd.²

Het verzamelen van gegevens over medische misslagen is nimmer een doel op zich, maar steeds een middel tot het doel van een betere en vooral veiliger zorg. In dit hoofdstuk wordt een aantal van die methodologische problemen rond epidemiologisch onderzoek van fouten nader besproken.

5.2 Risicoschatting: teller en noemer

Bij het kwantificeren van medische fouten biedt het absolute aantal ervan betrekkelijk weinig informatie. Liever hanteert men een relatieve maat: het risico. Hoe groot is de kans op een ongewenste gebeurtenis per blootstelling aan een medische interventie? Die schatting is het resultaat van een breuk met in de teller het aantal accidentele gebeurtenissen en in de noemer het aantal behandelde patiënten. Die uitkomst heeft niet het karakter van een onwrikbare natuurconstante. De teller omvat ongelijksoortige uitkomstcategorieën: sterfte, reversibele en irreversibele letselschade, psychische schade en ontevredenheid over de kwaliteit van de dienstverlening.

De noemer bestaat eveneens uit verschillende componenten. De foutfrequentie tijdens ziekenhuisopname wordt onderzocht binnen uiteenlopende groepen patiënten. Die populaties kunnen van zorginstelling tot zorginstelling variëren en een dergelijke redenering gaat ook op voor het ziektespectrum per aandoening of per afdeling.

Welke betekenis men mag toekennen aan het geschatte risico in een bepaalde situatie of de vergelijking van risico's tussen twee ziekenhuizen hangt af van de eigenschappen van de teller (welk soort risico onderzocht? hoe gedefinieerd? hoe gemeten?) en de noemer (welk soort patiënten? welk soort behandeling? welk ziektespectrum? hoe gemeten?). Daardoor is die vergelijkbaarheid tussen ziekenhuizen op voorhand problematisch.

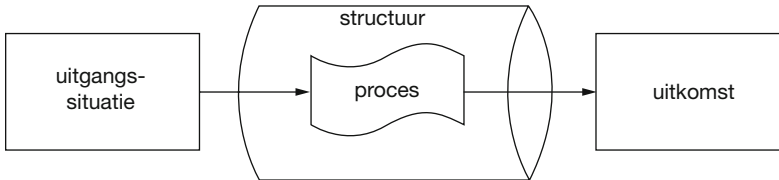
5.3 Definitieproblematiek

Het startpunt van onderzoek naar de frequentie van medische fouten is het ondubbelzinnig definiëren ervan en dat blijkt allesbehalve eenvoudig.⁴ Daaraan zijn meerdere facetten te onderscheiden zoals: 1) het omschrijven van het begrip medische fout, 2) het rubriceren van medische fouten in soorten, 3) het classificeren van gebeurtenissen naar hun aard, zoals wel of niet vermijdbaar, een cognitieve of uitvoeringstechnische fout en 4) het differentiëren naar ernst. We zullen ons hier vooral richten op het eerste punt, het definiëren van de term 'medische fout'.

5.3.1 Conceptuele definitie

Een voorbeeld van een algemene definitie van het begrip 'fout' is de omschrijving van de Engelse psycholoog Reason: '“Fout” wordt gebruikt als een generieke aanduiding voor al die situaties waarin een voorgenomen opeenvolging van mentale of fysieke activiteiten niet leidt tot de bedoelde uitkomst en waarbij deze mislukkingen niet kunnen worden toegeschreven aan een vorm van toeval.⁴ Een ongewenste uitkomst als resultaat van een verkeerd proces. Er zijn bijgevolg twee materiële elementen bij het definiëren, namelijk de uit te voeren taak en het resultaat daarvan en vervolgens een normatief oordeel over dit geheel. In de medische praktijk hoeft een ongewilde afloop – een complicatie of zelfs sterven – niet steeds te betekenen dat

de arts dit had kunnen vermijden. Als conceptueel raamwerk voor het beoordelen van medisch handelen kan de trias structuur – proces – uitkomst van Donabedian zijn dienst bewijzen.^{5,6} De samenhang tussen structuur, proces en uitkomst impliceert dat een medische fout niet uitsluitend op grond van de afloop wordt gedefinieerd (figuur 5.1).⁶ Er is bij het definiëren dus ook een belangrijke rol weggelegd voor het proces.



Figuur 5.1

De trias structuur, proces en uitkomst van Donabedian (figuur naar Battles & Lilford).⁶

Een complicatie is iedere ongewenste uitkomst van medisch handelen ongeacht de kwaliteit van het proces, een fout is een tekortschietende kwaliteit van het proces ongeacht de uitkomst.^{5,7} De samenhang tussen deze twee wordt weergegeven in het Venn-diagram (figuur 1.1). Er kan sprake zijn van fouten in het proces zonder complicaties bij de patiënt en van complicaties bij de patiënt zonder fouten in het proces. Er zijn echter ook complicaties die het gevolg zijn van fouten en juist die schade zou – althans in theorie – te voorkomen moeten zijn. Een antwoord op de vraag *hoe* een medische fout dient te worden gedefinieerd, is hiermee nog niet gegeven: omschrijft men een fout primair vanuit het proces of vanuit de uitkomst? Kijkt men alleen naar de uitkomst, dan mist men de fouten die niet gepaard gaan met een complicatie. Kiest men het proces, dan betekent dat niet dat al deze mislagen ook bij de patiënt tot schade zullen leiden. Er zijn dus verschillende definitiemogelijkheden, die elk een andere patiëntenpopulatie zullen identificeren waarbij sprake is van fout medisch handelen en/of schade.

5.3.2 Operationaliseren van de conceptuele definitie

Voor het opsporen van medische fouten zijn twee richtingen in de tijd mogelijk: een retrospectieve of een prospectieve analyse. Vervolgens rijst de vraag welke invalshoek daarbij gekozen wordt: proces- of uitkomsttoetsing of de combinatie daarvan?^{6,7} De causale relatie tussen proces en uitkomst is onzeker. Welke benadering men ook kiest, er zal praktisch onderscheid gemaakt moeten worden tussen juist en onjuist medisch handelen. Maar het formuleren van een eenduidige en toetsbare norm is minder eenvoudig dan het lijkt. Bijvoorbeeld wat betreft de uitkomsten. Wanneer is een liesbreukoperatie adequaat uitgevoerd? Wat is een geoorloofd percentage recidief liesbreuken? Hoeveel mammacarcinomen mag een screeningsradioloog missen? Welke postoperatieve sterfte is aanvaardbaar bij coronaire bypass-

operaties? Hoe definieert men een juiste procesvoering en welk empirisch onderzoek is daarvoor nodig? We krijgen te maken met een complex geheel van biologische, organisatorische en psychologische factoren. Wie ondubbelzinnige begripsomschrijvingen van juist medisch handelen probeert te maken, wordt geconfronteerd met het veelal nog ontbreken van standaarden voor evidence-based medicine.

5.4 Retrospectief statusonderzoek

De meeste onderzoeken naar de frequentie van medische fouten hanteren de retrospectieve methode door statusonderzoek. Dat is tijdrovend en dus kostbaar. Eerst formuleert men een waaier van criteria waarmee gescreend wordt op gebeurtenissen die verband kunnen houden met medische fouten. Voorbeelden daarvan zijn overlijden, heroperaties na een eerdere ingreep, heropnames na behandelingen, overplaatsing naar afdelingen voor intensieve zorg, bloedtransfusies, in het ziekenhuis opgelopen infecties, andere ziektebeelden of letsels ontstaan gedurende de opname, geneesmiddelreacties, klachten van de patiënt of diens familie of juridische acties. Opsommingen van dergelijke criteria worden vermeld bij onderzoeken.⁸⁻¹² Deze methode is vrij grofmazig.¹³ Een praktisch punt is dat men geheel afhankelijk is van de kwaliteit van vastleggen van gedachten en gegevens. Het paradoxale doet zich voor dat hoe groter de ijver om alle informatie in het dossier vast te leggen, des te groter de kans dat fouten zichtbaar zijn. Om dit te voorkomen zal men zich alleen op van tevoren omschreven primaire data moeten richten.¹⁴

Als het startpunt van onderzoek de uitkomst is, spoort men fouten in het proces die niet gepaard gaan met ‘verklikkende’ gebeurtenissen zo dus niet op. Dat kan zijn omdat de signalerende gebeurtenis niet plaatsvond, niet werd herkend of niet werd vastgelegd in de status.

Een andere reden waarom fouten gemist kunnen worden, is als de uiting van de fout pas merkbaar wordt na ontslag uit het ziekenhuis. Door de alsmaar kortere opnameduur en omdat steeds meer ingrepen in dagbehandeling worden gedaan, onttrekt een kwart van de complicaties – al dan niet het gevolg van een medische fout – zich aan waarneming bij statusonderzoek.¹⁵

Omdat statusonderzoek zo tijdrovend is en vrijwel alle ziekenhuizen beschikken over een geautomatiseerd informatiesysteem, wordt gezocht naar alternatieve mogelijkheden om op indirecte wijze gevallen van medische fouten op te sporen.¹⁶ Dat kan bijvoorbeeld aan de hand van de opnameduur, de gemaakte kosten, gebruik van farmaca en sterfte.

Op grond van het bovenstaande is het aannemelijk dat de retrospectieve methode niet erg gevoelig zal zijn. Ook is het de vraag welke waarde men mag hechten aan totaalpercentages fouten voor een ziekenhuis, want binnen de instelling verschillen de risico's van afdeling tot afdeling en fouten zijn er qua ernst en gevolgen in soorten.

5.5 Prospectief observationeel onderzoek naar medische fouten

Waar retrospectief statusonderzoek begint bij beoordeling van de uitkomst, observeren bij een prospectief onderzoek getrainde waarnemers in 'real time' eerst het proces, om vervolgens vast te stellen wat de uitkomst daarvan is en hoe die zich verhoudt tot het gewenste resultaat. Deze benadering wordt wel aangeduid als participerende observatie of etnografische methode. Zo observeerde een Amerikaans team het werk op een afdeling algemene heelkunde en constateerde ernstige fouten bij 18% van de opgenomen patiënten.¹⁷ Ook afdelingen voor intensieve zorg zijn op die manier beoordeeld¹⁸ en tegenwoordig wordt ook dergelijk onderzoek verricht waarbij tv-camera's de werkzaamheden registreren op bijvoorbeeld een ok, de intensive care of de afdeling voor spoedeisende hulp.¹⁹ De waarnemingen van de observatoren kunnen dan achteraf worden getoetst aan de hand van de televisiebeelden.

De prospectieve benadering heeft verschillende voordelen ten opzichte van het retrospectieve statusonderzoek. Men is niet langer afhankelijk van wat er is geregistreerd en evenmin is men aangewezen op uitkomsten als mogelijk signaal voor een verkeerd proces. Daarom ook is deze methode veel gevoeliger. Waarnemingen kunnen direct worden teruggekoppeld naar de uitvoerende hulpverleners, wat een betere interpretatie ervan mogelijk maakt, en dat maakt deze benadering eventueel ook specifiek.

5.6 Meetfouten bij het meten van medische fouten

Hoe nauwkeurig en reproduceerbaar zijn de methoden van onderzoek naar medische fouten?¹⁴ Bepaling van de nauwkeurigheid vraagt een onafhankelijke gouden standaard, maar die is er helaas niet. Dus zal men moeten proberen een enigszins betrouwbare schatting te doen voor de verschillende methoden van onderzoek, rekening houdend met factoren als ziektespectrum en bias. De gevoeligheid – welk percentage van de fouten kunnen we achterhalen? – wordt geschat op ongeveer 70-80%.^{13,14} Bij vergelijking van de retro- en prospectieve methode bleek de prospectieve werkwijze meer vermijdbare foutgevallen te identificeren.²⁰ De specificiteit – is een 'fout' ook werkelijk een fout? – is bij onderzoek naar mislagen van groot belang, want dit etiket kan grote gevolgen hebben, ook juridische. Er is over de specificiteit minder bekend dan over de sensitiviteit, maar deze wordt op ongeveer 95% geschat.²⁰

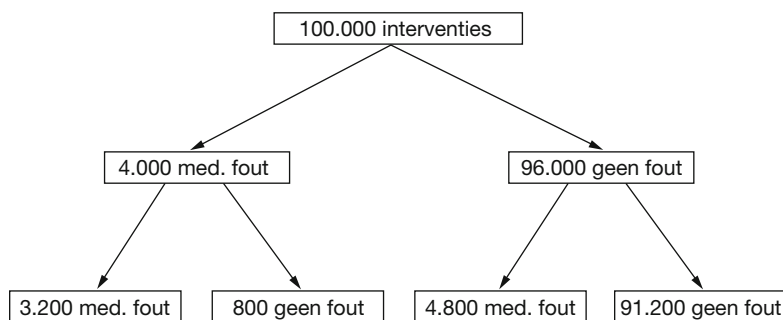
Een belangrijke kwestie is dat de overeenstemming tussen verschillende waarnemers of iets als 'fout' moet worden geclassificeerd, de reproduceerbaarheid, veelal bedroevend laag blijkt.¹⁴ Over de vraag of een ongewenste afloop door anders te handelen te voorkomen zou zijn geweest, bestaat al evenmin eenstemmigheid.²¹ Juist deze gebrekkige reproduceerbaarheid is de achilleshiel van onderzoek naar medische fouten: de willekeur van het oordeel erover is te groot, vooral over de vermijdbaarheid ervan. In welke mate training en/of het gebruik van meerdere waarnemers een verbetering betekent, is nog onvoldoende duidelijk.

Een ander probleem betreft bias. Als men een complicatie kiest als uitgangspunt voor verder onderzoek van het proces, krijgt men te maken met *hindsight bias*: men oordeelt meestal anders – strenger – als de ongelukkige afloop bekend is.²² Een tweede kwestie is waarnemerbias. Bij het onderzoek weten waarnemers wie er handelden en dat kan hun oordeel beïnvloeden. Anonimiseren is daarvoor een oplossing, maar blijkt praktisch lastig.

De methodologie van onderzoek en beoordeling van medisch handelen toont dus veel onvolkomenheden en dat heeft gevolgen voor de duiding van de resultaten van onderzoek.

5.7 De interpretatie van onderzoek naar medische fouten

Bij het interpreteren van onderzoek kan het gaan om de uitkomst van één onderzoek of om het toekennen van betekenis aan verschillen tussen uitkomsten van meerdere onderzoeken. Als we 100.000 statussen onderzoeken op het vóórkomen van medische fouten, met een testmethode die een gevoeligheid heeft van 80% en een specificiteit van 95%, terwijl er bij 1 op de 25 medische interventies in werkelijkheid van een medische fout sprake is (foutprevalentie: 4%), kunnen we de resultaten weergeven in een boomstructuur (figuur 5.2). De voorspellende waarde van de testuitkomst ‘medische fout’ is maar 40%! Gezien de relatief lage prevalentie van misslagen zal er bij de meeste interventies geen sprake zijn van een fout; maar 5% daarvan wordt als ‘fout’ beoordeeld. Het gevolg is dat het probleem van medische fouten dus kan worden overschat (gemeten foutfrequentie: 8%).



Figuur 5.2

Boomstructuur volgens Bayes. Onderzoek naar fouten toegepast op 100.000 interventies. De middelste rij toont de werkelijke situatie, de onderste de uitkomsten van het onderzoek naar medische fouten. De testmethode heeft een sensitiviteit van 80% en een specificiteit van 95%.

Iets anders is de vergelijking van uitkomsten van meerdere onderzoeken, zoals in eerdere hoofdstukken gepresenteerd. Gaat het bij gesignaleerde diversiteit van foutfrequenties ook om verschillen in kwaliteit van zorg? Voor dat uiteenlopen van resultaten bestaan meerdere verklaringen.²³ Het

kan gaan om verschillen in definities en meetmethoden. Uitkomsten van interventies worden echter niet alleen bepaald door de kwaliteit van zorg maar door tal van andere patiënt- en omgevingsgebonden factoren. Daarnaast kan de case mix van ziekenhuis tot ziekenhuis verschillen en ook toeval kan een belangrijke rol spelen. Alvorens een bepaald foutpercentage gezien kan worden als uiting van gebrekkige kwaliteit en vóór men kan stellen dat in het ene ziekenhuis meer fouten worden gemaakt dan in het andere, zal men alle mogelijke factoren van invloed op die uitkomsten en alternatieve verklaringen moeten nalopen.²⁴

5.8 Conclusie

De neiging is groot om elke medische fout als onaanvaardbaar te beschouwen, maar dat is niet erg realistisch. Is ultraveilige medische zorg wel haalbaar?²⁵ Bij epidemiologisch onderzoek naar fouten is het doel niet alleen descriptief maar uiteindelijk ook normatief. Om te kunnen beoordelen of de zorg wel voldoet aan de daaraan gestelde eisen zal een norm moeten worden geformuleerd. Maar zijn die eisen wel zo eenduidig en toetsbaar? Wat is de realiseerbare bovengrens van medische prestaties? Daarvoor is zowel observationeel als experimenteel onderzoek nodig en onze normering van medisch handelen zal gestuurd moeten worden door wetenschappelijk bewijs.

De brandende vraag is: wat worden we nu wijzer van de epidemiologische gegevens over medische fouten en hoe kunnen we die gegevens gebruiken om een veiliger zorg te realiseren? Hier speelt weer de vraag of we prioriteit geven aan uitkomst- of aan procestoetsing. De onzekere relatie tussen processen en uitkomsten maakt uitkomsttoetsing als kwaliteitsinstrument minder bruikbaar. De grootste winst lijkt te kunnen worden geboekt via de weg van procestoetsing, maar dat vergt niet alleen een medisch-technische analyse, maar ook een brede organisatorische aanpak.^{14,25} Toch kunnen we niet zonder systematisch verzamelde gegevens over de afloop van medische interventies wanneer we het effect van verbeteringen van processen willen evalueren. De methodiek daarvan staat nog in de kinderschoenen.^{13,14}

De media tonen grote gretigheid om percentages medische fouten als 'schokkend' te presenteren, maar dergelijke publicaties roepen ongedult op wantrouwen op jegens artsen, met irritatie en frustratie van professionals als gevolg.³ Verantwoording afleggen betekent dat de prestaties worden afgemeten aan een norm, dus de eerste stap in dit proces van rekenschap is de normstelling langs de weg van evidence-based medicine. Die inspanning mag worden verwacht, maar juist de complexiteit van de hedendaagse geneeskunde maakt dat een lastige opgave. Worden er minder fouten gemaakt als de aanpak van patiëntveiligheid over een andere boeg wordt gegooid? Dat zal een rationele keuze moeten zijn en dat vergt uiteenlopend empirisch onderzoek en verkenning van de huidige barrières op weg naar maximale veiligheid.^{1,2}

Literatuur

- 1 Brennan TA, Gowanda A, Thomas E, Studdert D. Accidental deaths, saved lives, and improved quality. *N Engl J Med* 2005;353:1405-9.
- 2 Forster AJ, Shojania KG, Walraven C van. Improving patient safety: moving beyond the 'hype' of medical errors. *CMAJ* 2005;173:893-4.
- 3 Wetenschappelijke Raad voor het Regeringsbeleid. Bewijzen van goede dienstverlening. Hoofdstuk 7: Leren en controleren. Amsterdam: Amsterdam University Press, 2004, p. 185-216.
- 4 Reason J. Human error. Cambridge: Cambridge University Press 1990, p. 9.
- 5 Donabedian A. Evaluating the quality of medical care. *Milbank Mem Fund Q* 1966;44:166-203.
- 6 Battles JB, Lilford RJ. Organizing patient safety research to identify risks and hazards. *Qual Saf Health Care* 2003;12(Suppl 2):ii2-7.
- 7 Hofer TP, Kerr EA, Hayward RA. What is an error? *Effect Clin Pract* 2000;3:261-9.
- 8 Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *N Engl J Med* 1991;324:370-6.
- 9 Wilson RM, Runciman WB, Gibberd RW, Harrison BT, Newby L, Hamilton JD. The Quality in Australian Health Care Study. *Med J Aust* 1995;163:458-71.
- 10 Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ* 2001;322:517-9.
- 11 Baker GR, Norton PG, Flintoft V, Blais R, Brown A, Cox J, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *CMAJ* 2004;170:1678-86.
- 12 Michel P, Quenon JL, Djihoud A, Tricaud-Vialle S, Sarasqueta AM de, Domecq S. Les événements indésirables graves liés aux soins observés dans les établissements de santé: premiers résultats d'une étude nationale. *Études et résultats* 2005;No. 398:1-15.
- 13 Neale G, Woloshynowych M. Retrospective case record review: a blunt instrument that needs sharpening. *Qual Saf Health Care* 2003 Feb;12(1):2-3.
- 14 Lilford RJ, Mohammed MA, Braunholtz D, Hofer TP. The measurement of active errors: methodological issues. *Qual Saf Health Care* 2003;12(Suppl 2):ii8-12.
- 15 Forster AJ, Clark HD, Menard A, Dupuis N, Chernish R, Chandok N, et al. Adverse events among medical patients after discharge from hospital. *CMAJ* 2004;170:345-9.
- 16 Zhan C, Miller MR. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. *JAMA* 2003;290:1868-74.
- 17 Andrews LB, Stocking C, Krizek T, Gottlieb L, Krizek C, Vargish T, et al. An alternative strategy for studying adverse events in medical care. *Lancet* 1997;349:309-13.
- 18 Donchin Y, Gopher D, Olin M, Badihi Y, Biesky M, Sprung CL, et al. A look into the nature and causes of human errors in the intensive care unit. *Crit Care Med* 1995;23:294-300.
- 19 Rogers DA, Regehr G, MacDonald J. A role for error training in surgical technical skill instruction and evaluation. *Am J Surg* 2002 Mar;183(3):242-5.
- 20 Michel P, Quenon JL, Sarasqueta AM de, Scemama O. Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals. *BMJ* 2004;328: 199-202.

- 21 Hayward RA, Hofer TP. Estimating hospital deaths due to medical errors: preventability is in the eye of the reviewer. *JAMA* 2001;286:415-20.
- 22 Henriksen K, Kaplan H. Hindsight bias, outcome knowledge and adaptive learning. *Qual Saf Health Care*. 2003;12(Suppl 2):ii46-50.
- 23 Mant J. Process versus outcome indicators in the assessment of quality of health care. *Int J Qual Health Care*. 2001(Dec);13(6):475-80.
- 24 Bird SM, Cox D, Farewell VT, Goldstein H, Holt T, Smith PC. Performance indicators: good, bad, and ugly. *J R Stat Soc (Series A)* 2005;168:1-27.
- 25 Amalberti R, Auroy Y, Berwick D, Barach P. Five system barriers to achieving ultrasafe health care. *Ann Intern Med* 2005;142:756-64.
- 26 Wilson T, Holt T, Greenhalgh T. Complexity science: complexity and clinical care. *BMJ* 2001;323:685-8.