

Patientveiligheid op de Spoedeisende Hulp: Exploratie en toepassing van principes uit hoogrisico industrieën

Het centrale doel van dit proefschrift is om inzicht te krijgen in hoe veiligheid op de Spoedeisende Hulp (SEH) verbeterd zou kunnen worden aan de hand van principes van de lerende organisatie, organisatie cultuur en de theorie achter *high-reliability* organisaties. Kort samengevat zijn de volgende onderzoeksvragen geformuleerd (Hoofdstuk 1):

- Kan het leren van incidenten bewerkstelligt worden door oorzaken analyses op dossiers van schadeclaims?
- Hoe is veiligheidscultuur gerelateerd aan de implementatie van organisatorische vangnetten en zelf-gerapporteerde veiligheid?
- Kan een klassikale Crew Resource Management training bijdragen aan het ontwikkelen van niet-technische vaardigheden en een veiligheids-cultuur?

Het principe van de lerende organisatie is geoperationaliseerd door middel van het analyseren van oorzaken van schadeclaims uit dossiers. We waren geïnteresseerd of dossiers van schadeclaims konden helpen bij het leren van incidenten (Hoofdstuk 2). Organisatiecultuur, met name veiligheidscultuur, is een van de factoren die veiligheid beïnvloedt. We waren geïnteresseerd in welke dimensies van veiligheidscultuur de implementatie van organisatorische vangnetten voorspelden (Hoofdstuk 3) en welke dimensies de mate van veiligheid op afdelingen voorspelden zoals gerapporteerd door medewerkers (Hoofdstuk 4).

Een ander belangrijk principe om patiëntveiligheid te bevorderen is om karakteristieken van *high-reliability* organisaties over te nemen. Deze organisatie zijn zeer gefocust op een optimale samenwerking tussen hun medewerkers. Optimaal gebruik maken van niet-technische vaardigheden is een voorwaarde voor optimale samenwerking. Beginnend met een klassikale team training om bewustzijn te vergroten, wilden we onderzoeken of ons trainingsprogramma het gebruik van niet-technische vaardigheden kon verbeteren en of het de zorg ook uiteindelijk veiliger kon maken. Voor dat doel ontwikkelden we een observatie instrument om niet-technische

vaardigheden te kunnen meten (Hoofdstuk 5) en gebruikten we dat instrument om het effect van de training te evalueren (Hoofdstuk 6). Tot slot deden we een systematisch literatuur onderzoek om te kijken of klassikale CRM-training de patiëntveiligheidscultuur kan veranderen.

Hoofdstuk 1 geeft een weergave van de contextuele setting van het onderzoek in dit proefschrift alsmede een uiteenzetting van de fundamentele inzichten uit veiligheidsonderzoek om meer betrouwbare organisaties te bewerkstelligen, te weten de lerende organisatie, organisatie cultuur en de theorie achter hoge betrouwbaarheid. Het hoofdstuk eindigt met een korte samenvatting van de onderzoeksvragen in dit proefschrift.

Hoofdstuk 2 onderzoekt of dossiers van schadeclaims geschikt waren als bron om te leren van incidenten. Dossiers van gesloten en minnelijk geschikte claims uit de jaren 2001 en 2002 werden met PRISMA-Medische versie onderzocht, een instrument om basisoorzaken van incidenten te achterhalen. . Geschiktheid werd gedefinieerd als valide, betrouwbaar en toepasbaar. Vijftig incidenten die plaatsvonden op 31 SEH afdelingen werden gevonden in 47 dossiers. In totaal werden 114 basisoorzaken gevonden, wat neerkomt op gemiddeld 2,3 basisoorzaken per incident. Geen technische, 29% organisatorische en 66,7% menselijke en 4,4% patiëntgerelateerde basisoorzaken werden gevonden. Interbeoordelaarsbetrouwbaarheid voor het classificeren van de basisoorzaken was goed ($\kappa = 0,78$). Preventieve maatregelen die volgden uit de PRISMA classificatie-actie-matrix waren: beschikbare informatiebronnen completeren, communicatiestructuren verbeteren en het training van personeel om te voorkomen dat zij taken verkeerd uitvoeren of fouten in redenering maken. Wat betreft de toepasbaarheid van de methode vonden wij dat het grootste probleem de vertraging tussen het optreden van het incident, het ontdekken ervan en de melding van het incident bij de verzekeraar was. Aanvullende informatie over de context en relevante organisatorische en technische informatie kan daardoor nauwelijks verkregen worden. Uiteindelijk werd de validiteit van het basisoorzaken profiel als middelmatig beschouwd vanwege het ontbreken van informatie over technische en organisatorische faaloorzaken in vergelijking tot wat andere studies vonden. Geschiktheid van PRISMA-medische versie voor retrospectieve analyse van gesloten schadeclaims zou verbeterd kunnen worden als systeem-gerelateerd denken door de schadeverzekeraars en ziekenhuispersoneel gestimuleerd zou worden.

Hoofdstuk 3 onderzoekt de associatie tussen elf patiëntveiligheidscultuur dimensies en de implementatie van zeven organisatorische patiëntveiligheidsvangnetten op de SEH. De vangnetten werden geformuleerd door de grootste medische aansprakelijkheidsverzekeraar van Nederland. De vangnetten waren: 1. Een inwerkprogramma voor alle nieuwe arts-assistenten, 2. Onderwijs aan arts-assistenten over diagnose en behandeling van vaak gemiste diagnoses of verkeerd behandelde aandoeningen, 3. Evaluatie van de vaardigheden van arts-assistenten direct na de training en vervolgens twee keer per jaar wat betreft hun medische kennis, technische vaardigheden, dossiervorming en communicatieve/sociale vaardigheden, 4. Behandelprotocollen over fracturen en peesletsels moeten op de SEH aanwezig zijn en elke drie jaar geüpdatet worden, 5. Dagelijks controle van patiëntendossiers tijdens daarvoor gearrangeerde bijeenkomsten, 6. Supervisieafspraken en regels moeten vastgelegd worden, en 7. Radiologen, chirurgen/supervisors en arts-assistenten moeten deelnemen aan dagelijkse röntgenbesprekingen.

Gegevens over de mate van implementatie van deze vangnetten werden verzameld met een vragenlijst aan het hoofd van de afdeling. Gegevens over de veiligheidscultuur werden verzameld met de COMPaZ vragenlijst (CultuurOnderzoek onder Medewerkers over Patiëntveiligheid in Ziekenhuizen), de Nederlandse versie van de Hospitals Survey on Patient Safety Culture, die elf dimensie van patiëntveiligheidscultuur meet.

Resultaten lieten zien dat verschillende cultuurdimensies positief en/of negatief geassocieerd waren met de implementatie van de veiligheidsvangnetten, waarvan twee het meest vaak in de modellen terugkwamen. Steun vanuit het management bleek faciliterend te werken voor de implementatie van de vangnetten 'inwerkprogramma voor arts-assistenten (vangnet 1)', 'aanwezigheid van protocollen (vangnet 4)', dossierbesprekingen (vangnet 5) en 'röntgenbesprekingen (vangnet 7)'. Een goede wisseling van diensten lijkt belemmerend te werken voor de invoering van 'een inwerkprogramma voor arts-assistenten (vangnet 1)', 'dossierbesprekingen (vangnet 5)', 'gestructureerde supervisie (vangnet 6)' en 'röntgenbesprekingen (vangnet 7)'. Deze resultaten suggereren dat dimensies die ogenschijnlijk goed zijn, implementatie van verbetermaatregelen in de weg kan staan door een mogelijke gebrek aan het beseffen van de noodzakelijkheid ervan. Aan de andere kant kunnen ze faciliterend werken, mogelijk doordat ze het gevoel van steun geven voor organisatorische veranderingen. De invloed van patiëntveiligheidscultuur blijkt niet altijd zo rechtlijnig te zijn als men van te voren wellicht zou verwachten.

Hoofdstuk 4 beschrijft welke patiëntveiligheidsdimensies de zelf-gerapporteerde veiligheid op de afdeling voorspelt. Daarnaast worden de verschillen tussen verpleegkundigen en artsen beschreven en wordt een overzicht van de patiëntveiligheidscultuur op Nederlandse SEH afdelingen gepresenteerd. Voor de analyses werd de data van hoofdstuk 3 gebruikt van 480 verpleegkundigen, 159 artsen en 91 medewerkers in andere functies.

In de ongecorrigeerde analyses waren alle cultuurdimensies positief geassocieerd met de zelf-gerapporteerde patiëntveiligheid. Zes van deze dimensies bleven significant voorspellend voor de patiëntveiligheid in multivariate analyse. Deze dimensies waren 'samenwerking tussen afdelingen', 'meldingsfrequentie', 'open communicatie', 'feedback over en leren van fouten' en 'algemene perceptie van veiligheid'. Verschillen tussen verpleegkundigen en artsen werden gevonden op twee dimensies ('meldingsfrequentie' en 'steun vanuit het management'). Artsen waren geneigd om de veiligheid hoger te beoordelen dan verpleegkundigen, wanneer zij deze dimensies hetzelfde beoordeelden.

We kunnen concluderen dat het personeel diverse dimensies van veiligheidscultuur identificeerden die gerelateerd waren aan de zelf-gerapporteerde veiligheid op de SEH, waarbij artsen sommige dimensies als minder belangrijk beschouwden als verpleegkundigen.

Hoofdstuk 5 omschrijft de ontwikkeling van het EPOC (*Explicit Professional Oral Communication*) observatiemodel, dat het gebruik van niet-technische vaardigheden meet. Daartoe vertaalden, testten en pasten we een bestaand observatiemodel, genaamd SHAPE, aan. Dit model stamt uit de luchtvaart en is vertaald naar de zorg, in eerste instantie specifiek de SEH en de Intensive Care (IC).

Het uiteindelijke EPOC instrument bestaat uit drie hoofddimensies en zes sub dimensies; assertiviteit vormt de categorie *Self* (zelf), samenwerken met anderen, taakgericht en mensgericht leiderschap vormen de categorie *Human Interaction* (HI) (menselijke interactie) en situationeel bewustzijn en plannen en anticiperen vormen de dimensie *Anticipation on Environment* (AoE) (Anticiperen op de omgeving). Elke dimensies omvat een aantal concrete items die verbale communicatie reflecteren.

We evalueerden de interbeoordelaarsbetrouwbaarheid (*Intraclass Correlation* (ICC)) voor elke dimensie en voor een EPOC totaal score. Het model werd geëvalueerd in vier SEH afdelingen en zes IC afdelingen.

Er werden respectievelijk 378 en 1144 individuele en 51 en 68 dubbele observaties op de SEH en IC uitgevoerd. Alle EPOC dimensies werden geregeld geobserveerd, behalve assertiviteit, dat nauwelijks zichtbaar was. De

interbeoordelaarsbetrouwbaarheid was hoog met een ICC van 0,91 (95% betrouwbaarheidsinterval (BI) 0,84-0,95) voor de totale EPOC score, 0,90 (95% BI 0,80-0,95) voor HI en 0,85 (95% BI -,75-0,91) voor AoE. Voor de sub dimensies varieerden de ICCs tussen 0,53 en 0,95.

We concludeerden dat de EPOC betrouwbaar was in twee zeer verschillende *settings*. Door professioneel gedrag te kwantificeren kan dit instrument het meten van gedragsverandering over de tijd meten. De resultaten suggereren dat het instrument ook vertaald kan worden naar andere *settings*.

Hoofdstuk 6 evalueert het effect van een klassikale CRM training op de SEH op de niet-technische vaardigheden van medewerkers. Hiervoor werd het in hoofdstuk 5 genoemde observatie instrument gebruikt (EPOC). We voerden een pragmatisch, gecontroleerd onderzoek uit met twee SEH afdelingen in de interventiegroep en twee in de controlegroep.

SEH verpleegkundigen en artsen van de interventieafdelingen ontvingen een tweedaagse interactieve, klassikale training. Directe observaties werden gemaakt met het EPOC instrument om de niet-technische vaardigheden van de verpleegkundigen en artsen voor en na de training te meten. Drie dimensies van EPOC werden geanalyseerd; HI, AoE en een EPOC totaal score. Lineaire en logische *multilevel* analyses werden uitgevoerd. Modellen werden gecorrigeerd voor de uitkomst maat op baseline, het aantal dagen tussen training en observatie, en patiëntveiligheidscultuur en foutenmanagement cultuur op baseline.

Een statistisch significante toename werd gevonden voor HI ($\beta = 0.27$, 95% CI 0.08 - 0.49) en de totale EPOC score ($\beta = 0.25$, 95% CI 0.06 -0.43), maar niet voor AoE (OR = 1.19, 95% CI .45 - 3.15). Dit betekent dat ongeveer 25% meer expliciete communicatie werd gezien na de CRM training. Het impliceert dat CRM training de potentie heeft om patiëntveiligheid te verbeteren door het verminderen van fouten in de communicatie, die een grote rol spelen bij het ontstaan van onbedoelde gebeurtenissen in de zorg.

Hoofdstuk 7 beschrijft een systematisch literatuur onderzoek naar het effect van klassikale CRM training op veiligheidscultuur. Studies werden geïdentificeerd in PubMed, Cochrane Library, PsycINOF en ERIC tot en met 19 december 2012. De *methods guide for comparative effectiveness reviews* werd gebruikt om het risico op vertekening (*bias*) in de individuele studies te meten.

Tweeëntwintig studies werden geïncludeerd voor review. De *settings* waarin de training werd gegeven, studie designs en de evaluatiemethodes verschilden zeer. De meeste studies die alleen een selectie van veiligheidscultuurdimensies

rapporteerden vonden overwegend positieve resultaten, terwijl studies die alle gemeten veiligheidscultuurdimensies vermelden wisselende resultaten vonden. Gemiddeld bestond er in de studies een gemiddeld hoog risico op vertekening.

Bewijs dat CRM training in de zorg een effect hebben op veiligheidscultuur is niet overtuigend. De validiteit van de meeste studies is beperkt. De resultaten ondersteunen de noodzaak tot meer valide studie designs, bij voorkeur een design waarin triangulatie wordt gebruikt.

Hoofdstuk 8 vat de hoofdbevindingen en de methodologische overwegingen van de studies in dit proefschrift samen. Daarnaast geeft het aanbevelingen voor de praktijk en voor toekomstig onderzoek.

Wat betreft de vraag of het leren van schadeclaims mogelijk was, vonden we dat het aansprakelijkheidsproces te gefocust is op de aansprakelijkheid van de individuele zorgverlener. Dit maakt het moeilijk om inzicht te krijgen in latente fouten in de organisatie die hebben bijgedragen tot het incident. Dit wordt extra bemoeilijkt door de vertraging tussen het ontstaan van het incident en de melding bij de aansprakelijkheidsverzekeraar.

Inzicht verkrijgen in de context waarin het incident plaatsvindt wordt gezien als een voorwaarde om te begrijpen waarom dingen misgaan. Een andere belangrijke methodologische overweging is dat schadeclaims uit 2001 en 2002 het gedachtegoed van die tijd reflecteren, die mogelijk nu anders is dan toen gezien de ontwikkelingen op het gebied van patiëntveiligheid in Nederland.

De tweede hoofdvraag hield zich bezig met veiligheidscultuur en de voorspellende waarde van de bijbehorende dimensies op de implementatie van organisatorische vangnetten en zelf-gerapporteerde veiligheid. Analyses lieten zien dat met name twee dimensies gerelateerd waren aan de implementatie van vangnetten, namelijk 'steun vanuit het management' en 'goede wisseling van diensten'. Wat betreft de laatste, speelt mogelijk *complacency* een rol, aangezien men de mogelijkheden voor verbetering wellicht over het hoofd ziet als men denkt dat veel al goed gaat.

Zes uit elf patiëntveiligheidscultuurdimensies voorspelden de zelf-gerapporteerde veiligheid op de afdeling: 'samenwerking tussen afdelingen', 'meldingsfrequentie', 'open communicatie', 'feedback op en leren van fouten', 'steun van uit het management' en de 'algemene perceptie van patiëntveiligheid'. Analyses lieten zien dat artsen geneigd zijn om de veiligheid hoger te waarderen dan verpleegkundigen terwijl zij dezelfde score hadden voor 'meldingsfrequentie' en 'steun vanuit het management'.

Hoofdzakelijk zullen de methodologische overwegingen voor deze twee studies zijn dat de uitkomsten van veiligheidscultuur zelf-gerapporteerd zijn. Er is altijd een kans dat antwoorden een over- of onderschatting van de werkelijkheid zijn en dat de ene respondent kritischer is dan de andere. Een oplossing zou zijn om triangulatie methodieken te gebruiken om veiligheidscultuur te omschrijven. Dit geeft de mogelijkheid om de onderliggende waardes en aannames van de respondenten te meten in plaats van enkel observeerbare artefacten. Daarnaast geeft het de mogelijkheid om naar meer dimensies te kijken dan naar welke door de vragenlijst gedefinieerd zijn. Triangulatie zal ook een idee geven over welke processen en structuren een organisatie zou moeten implementeren om een hoger niveau van veiligheid te creëren. Een andere belangrijke overweging is dat, gebaseerd op onze observatie dat veiligheidscultuur verschilt tussen afdelingen, ziekenhuizen en landen, waar mogelijk een correctie gemaakt moet worden voor de clustering van de antwoorden binnen afdelingen, ziekenhuizen en/of landen.

Een aanbeveling die resulteert uit bovenstaande is dat veiligheidsonderzoekers hun focus moeten leggen op de context waarin het incident plaatsvond. Deze aanbeveling geldt zowel voor de praktijk als voor wetenschappelijk onderzoek. Onderzoekers moeten ook triangulatie gebruiken waar dat gepast is en wanneer dat mogelijk is.

Wat betreft de derde hoofdvraag naar het effect van CRM, vonden we met het EPOC observatie instrument een toename van niet-technische vaardigheden van 25% wat betreft de Human Interactie categorie. Onze review naar de effecten van CRM op de veiligheidscultuur gaf echter geen eenduidige resultaten. *Singlecentre* en *multicentre* ongecontroleerde studies vonden allemaal in meer of mindere mate een effect, maar de gecontroleerde studies vonden geen interventie effect. Of dit het resultaat is van meer aan de situatie aangepaste trainingen in de ongecontroleerde studies, publicatie *bias* of een echt effect van de training zou verder onderzocht moeten worden. Het belang van steun vanuit het management vonden wij in de review. In ziekenhuizen waarin een ongecontroleerde CRM evaluatie werd uitgevoerd die uiteindelijk een groot effect liet zien, bleek of het ziekenhuismanagement achter de invoering van CRM te staan of een lokaal team met sleutelfiguren actief te zijn bij de implementatie, of beiden waren aanwezig. Een aanbeveling is daarom dat het ziekenhuismanagement patiëntveiligheid als topprioriteit moet bestempelen en dienovereenkomstig ook de nodige middelen ter beschikking moet stellen.

Methodologische overwegingen aangaande de CRM studie behelzen de noodzakelijkheid om een training te evalueren op meer dan één effect level.

We hebben geprobeerd om attitudes (level II van het model van Kirkpatrick) te meten, maar de vragenlijst die we hadden ontwikkeld bleek geen bruikbare dimensies te bevatten. Level III evalueerden we door het observeren van gedrag. Echter, we vonden dat cognitieve processen moeilijk te vangen waren met het observatie instrument. Wellicht dat vragenlijsten hiervoor beter geschikt zijn. Daarbij weten we ook niet of de gevonden effecten op de lange termijn nog zichtbaar zullen zijn. Een implementatieplan kan hier heel nuttig zijn. Level IV, uitkomsten in de organisatie, zijn moeilijk te meten. Wij richtten ons op veiligheidscultuur, als een proxy, maar optimaal zou zijn om patiëntuitkomsten ook mee te nemen.

Al met al bevelen wij aan om het gebruik van niet-technische vaardigheden en de training daarvan in de praktijk en in onderwijscurricula in te bedden om het gebruik ervan te waarborgen. Ook zou geïnvesteerd moeten worden in methodes die deze inbedding bevorderen. Voor toekomstig onderzoek is het belangrijk om te realiseren dat de evaluatie van complexe interventies een benadering van diverse invalshoeken nodig heeft, een context omschrijving en een vergelijkingsgroep, alsmede een gecombineerd gebruik van kwalitatieve en kwantitatieve metingen.

We concluderen dat een focus op de context en het gebruik van triangulatie ons inzicht in de bestaande veiligheidscultuur en ons begrip over hoe dingen mis kunnen gaan, en onze inzichten in welke interventies zouden kunnen werken, hoe zij werken en hoe ze geïmplementeerd moeten worden, kan vergroten. Met een duidelijke focus op context en veiligheidscultuur zal waarschijnlijk maximale winst te behalen zijn in het streven naar organisatorisch leren en *high reliability* en zullen de kansen groter zijn dat de behaalde effecten gewaarborgd worden.