

# Resident-sensitive quality measures

## Citation for published version (APA):

Schumacher, D. J. (2020). *Resident-sensitive quality measures: defining the future of patient-focused assessment*. ProefschriftMaken Maastricht. <https://doi.org/10.26481/dis.20200319ds>

## Document status and date:

Published: 01/01/2020

## DOI:

[10.26481/dis.20200319ds](https://doi.org/10.26481/dis.20200319ds)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

## Summary

### Chapter 1

Chapter 1 describes recent calls to link educational outcomes and patient care outcomes in graduate medical education. Achieving this goal is the foundation of competency-based medical education, but current gaps in quality improvement training and resident performance assessment practices hinder efforts to achieve it. Concerning gaps in assessment, there is both unwanted variation in assessments as well as failure to align assessment with patient care. In this chapter, we propose that patient-centered performance assessment, specifically entrustable professional activities (EPAs) and quality measures that are relevant to the work that residents complete (which we term resident-sensitive quality measures, or RSQMs), is a way to overcome these barriers. By closing these gaps, we will begin to provide the ability to better link educational outcomes with patient care outcomes. However, EPAs currently lack sufficient empirical evidence despite their international popularity over the past decade, and there are also insufficient quality measures for the work that residents complete. This situation leads to the main research questions for this dissertation:

What are the characteristics of developing and implementing quality measures, which we term resident-sensitive quality measures (RSQMs), that are largely attributable to residents and important to patient care?

What is the association between entrustment decisions made about residents and their performance on RSQMs?

The studies in this dissertation took place in the U.S. pediatric residency context. Specifically, we focused in the pediatric emergency department for our initial efforts in this novel research direction.

### Chapter 2

Chapter 2 extends the call to focus on new approaches of performance assessment to better link educational and patient care outcomes, arguing that the future of high-quality care depends on better physician performance assessment. This chapter proposes a three-pronged research agenda in this area. First, EPAs should be implemented and studied broadly. Second, the association between entrustment decisions made about residents and quality measures for the care they provide should be investigated. Finally, with only 2% of global health expenditure dedicated to health education for all professions combined, these efforts should receive funding and resource prioritization from public and private sectors.

---

### Chapter 3

The research program presented in this dissertation focuses on RSQM performance and entrustment decisions for individuals. However, given the recent focus on health care provided by teams, it is important to justify the focus on individuals. Chapter 3 argues the importance of focusing on individuals (by considering *attribution* of care) as well as teams, programs, and systems (by considering *contribution* to care). Contribution focuses on the various degrees different entities contribute to an outcome, and it plays an important role in analyzing program- and system-level outcomes that inform program evaluation and program-level improvements for the future. Equally important in health care are the role of the individual and aspects of care that can be attributed to an individual team member. Providing individuals with data related to their performance on the team has value in driving their personal improvement. Furthermore, we graduate, certify, and credential individual physicians, not teams, for medical practice. Additionally, incentive-based payment models focus on the quality of care provided by individuals, making it important to be able to determine outcomes of performance attributed to individuals. In this chapter, we explore how attribution and contribution analyses can be used in a complimentary fashion to discern what outcomes can and should be attributed to individuals as well as to teams and programs.

### Chapter 4

Chapter 4 describes efforts to begin closing the training and assessment gap with respect to quality measures available for use with residents through the development of RSQMs for use in the pediatric emergency department setting. In this study, we used a nominal group technique (NGT) and Delphi process with faculty and fellow supervisors at Cincinnati Children's Hospital Medical Center to develop (through the NGT) and prioritize (through the Delphi) possible RSQMs for acute asthma exacerbation, bronchiolitis, and closed head injury encounters. Groups were asked to generate (in the NGT) and rate (in the Delphi) possible RSQMs based on two criteria: 1) importance of the measure to the illness of interest and 2) likelihood that a resident, and not another member of the team or the team collectively, completes the measure. The NGT produced 67 measures for asthma, 46 for bronchiolitis, and 48 for closed head injury. These were used to populate the first round of the Delphi process. After two rounds, 18 measures for asthma, 21 for bronchiolitis, and 21 for closed head injury met automatic inclusion criteria. In round three, participants categorized the potential final measures by their top 10 and next 5. Prioritized measures often fell into one of three categories: 1) appropriate medication use, 2) documentation, and 3) information provided at patient discharge. In summary, this study provides a template for identifying and developing RSQMs that may help to promote high-quality care delivery during and following training.

## Chapter 5

Chapter 5 extends the work of chapter 4 in two ways: 1) replicating the process of developing RSQMs for asthma, bronchiolitis, and closed head injury in the pediatric emergency department setting with pediatric residents, and 2) engaging resident and supervisor stakeholders to develop and inform next steps in creating RSQMs. Like faculty and fellow supervisors, residents in the NGT and Delphi groups placed considerable focus across the three illnesses of interest on measures that pertained to 1) appropriate medication use, 2) documentation, and 3) information provided at patient discharge. Subsequent focus groups with residents as well as faculty and fellow supervisors highlighted hospital medicine and general pediatrics as priority areas for developing future RSQMs but also noted contextual variables that influence the application of similar measures in different settings (e.g., team structure of general pediatrics clinic may limit the utility of some measures). Residents and supervisors had both similar and unique insights into developing RSQMs. For example, the resident NGT produced more measures than the supervisor NGT, and both focus groups felt this may be because residents work on the front line and know more details of care provided there. However, supervisors also noted that residents may not be able to prioritize well or do not know resources available to help complete their work, and residents noted that supervisors may not know details that go into actually providing frontline care. In summary, this chapter builds on the work of chapter 4 to continue to pave the path forward in developing future RSQMs by exploring specific settings, measures, and stakeholders to consider when undertaking this work.

## Chapter 6

Chapter 6 presents the range of individual and collective performance for the RSQMs introduced in chapters 4 and 5 when these RSQMs are used to assess pediatric residents working in the pediatric emergency department. During the 2017–2018 academic year, pediatric residents in the Cincinnati Children's Hospital Medical Center pediatric emergency department were assessed using RSQMs for patient encounters focused on acute asthma exacerbation (21 RSQMs), bronchiolitis (23), and closed head injury (19). Of included encounters, 83 residents cared for 110 patients with asthma, 112 with bronchiolitis, and 77 with closed head injury. Residents had the opportunity to meet most RSQMs, but exceptions existed (e.g., patients with asthma exacerbation presenting to the medical resuscitation area had initial orders placed by the supervising physician, per protocol, precluding residents from meeting some RSQMs that pertained to initial orders). In encounters presenting the opportunity, there was a wide range of frequency for how often residents performed RSQMs. One closed head injury measure was met in all encounters. Across illnesses, another 9 RSQMs were met in almost all encounters, but most RSQMs were not met, with a frequency ranging from several times to the majority of encounters. To determine the proportion of individual RSQMs correctly performed out of total possible RSQMs that could have been performed, we generated a composite score for each encounter. RSQM composite scores demonstrated notable range and variation: asthma mean,

---

0.81 (SD: 0.11); bronchiolitis mean, 0.62 (SD: 0.12); closed head injury mean, 0.63 (SD: 0.10). In summary, chapter 6 demonstrates that individual and composite RSQMs capture and discern resident performance across patient encounters, opening the door for their use in other contexts and by other specialties.

## **Chapter 7**

Chapter 7 explores the association between RSQM composite scores for individual pediatric emergency department encounters, supervisor entrustment decisions (as measured by supervision level assigned) for those encounters, and patient acuity and complexity for those encounters. The 83 residents and 299 patient encounters for asthma, bronchiolitis, and closed head injury presented in chapter 6 were used for the study presented in this chapter. To measure the association of RSQM composite scores with the other variables of interest, we used mixed models to account for observations nested within residents. Entrustment decisions were positively associated with asthma RSQM composite scores (beta-coefficient 0.03;  $P = .0004$ ), but there was no significant association between RSQM composite scores and entrustment decisions for bronchiolitis or closed head injury. While statistically significant, the positive linear relationship for asthma is likely not educationally or clinically significant because RSQM composite scores only increased 0.03 (out of 1) for every 1 (out of 7) point increase in entrustment score. A positive relationship would be expected between entrustment decisions and RSQM composite scores. Given that RSQMs are a more objective measure of performance, the findings of this study should prompt consideration of whether RSQM performance data should help inform entrustment decisions. RSQM composite scores were significantly higher when acuity was also higher and significantly lower when acuity was also lower for both asthma ( $P = .0009$ ) and bronchiolitis ( $P = .01$ ). However, RSQM composite scores were almost identical for the different levels of acuity for closed head injury ( $P = .94$ ). There was no association between RSQM composite scores and patient complexity.

## **Chapter 8**

Having established initial implementation evidence for using RSQMs in the pediatric emergency department, chapter 8 sought to consider how RSQM data might be used by those who make assessment decisions outside the pediatric emergency department. The study presented in this chapter achieved this goal by exploring how individual clinical competency committee (CCC) members interpret, use, and prioritize RSQM data inserted into their usual review processes. In this constructivist grounded theory study, we purposively and theoretically sampled 19 pediatric residency CCC members from Cincinnati Children's Hospital Medical Center. Participants were provided a resident assessment portfolio consisting of performance ratings and narrative comments for five rotations as well as RSQM data for one of these rotations. They were asked to review with the purpose of making a decision about the residents' ability to care for patients presenting with common, acute problems (a general pediatric entrustable

professional activity). Data collection consisted of two phases: 1) observation and think aloud in which participants reviewed the resident portfolio, and 2) semistructured interviews to probe participant reviews. Analysis moved from close reading to coding and theme development followed by the development of a model illustrating how themes interact with one another. We identified five dimensions for how participants view and use RSQMs: 1) ability to orient to RSQMs: confusing to self-explanatory; 2) propensity to use RSQMs: reluctant to enthusiastic; 3) RSQM interpretation: requires contextualization to self-evident; 4) RSQMs for assessment decisions: not sticky to sticky; and 5) expectations for residents: potentially unfair to fair use of RSQMs. These dimensions interact with one another to generate three profiles regarding how participants used RSQM data: eager incorporation, willing incorporation, and disinclined incorporation. The former two profiles were most common, with most participants using RSQMs to some extent, supporting the inclusion of RSQMs as resident assessment data for CCC review.

## **Chapter 9**

As RSQMs continue to be developed, it will be important to optimize ways to extract them from the electronic health record (EHR) and to automatically identify which residents they should be attributed to in settings where multiple residents provide care to a given patient, such as the inpatient environment. Chapter 9 sought a method for attributing care of individual patients to individual interns based on “footprints” in the EHR. Primary interns caring for patients on an internal medicine inpatient service were recorded daily by five attending physicians of record at University of Cincinnati Medical Center in August 2017 and January 2018. These records were considered gold standard identification of primary interns. The following EHR variables were explored to determine representation of primary intern involvement in care: postgraduate year, progress note author, discharge summary author, physician order placement, and logging clicks in the patient record. These variables were turned into categorical attributes (e.g., progress note author: yes/no), and informative attributes were selected and modeled using a decision tree algorithm. A total of 1,511 access records were generated; 116 were marked as having a primary intern assigned. All variables except discharge summary author displayed at least some level of importance in the models. The best model achieved 78.95% sensitivity, 97.61% specificity, and an area under the receiver-operator curve of approximately 91%. Thus, this study successfully predicted primary interns caring for patients on inpatient teams using EHR data with excellent model performance. These results provide a foundation for attributing patients to primary interns for the purposes of determining profiles of patient diagnoses by trainee, supporting continuous quality improvement efforts in graduate medical education, and ultimately attributing RSQMs to residents.

## **Chapter 10**

Chapter 10 details calls for ensuring high value health care and posits that medical education can and should drive progress toward this goal. We argue that getting medical education to

---

the point where it can do this work will require a clear, sustained focus on the patient, which is not traditionally a notable focus of medical education research or practice. Patient-focused assessment efforts, such as RSQMs, can help achieve this goal. The summary of our research is that RSQMs developed through consensus group methodology largely do reflect the performance of residents and that performance varies widely across residents and encounters. Future development of RSQMs will need to consider engaging other stakeholders and expanding to other settings and specialties. Our research suggests future settings for pediatrics, including hospital medicine and primary care outpatient pediatrics, but these efforts should also be undertaken in other medical specialties as well as surgical and hospital-based specialties. Completing this work should include engaging with stakeholders we have not yet involved in our work to date, including nurses and patients/families.

Our findings for the association between RSQMs and entrustment decisions made about residents varied. RSQM composite scores were significantly higher for residents entrusted with unsupervised practice or the ability to supervise others than for residents entrusted to lower supervision levels for all conditions (asthma, closed head injury, and all conditions combined) other than bronchiolitis. This finding offers validity evidence for RSQM composite scores; however, the absolute difference was modest. Additionally, when considering each supervision level and not combining them into two categories (at least entrusted with unsupervised practice or not entrusted to that level), a significant positive linear relationship between RSQM composite scores and entrustment decisions was only seen for asthma. However, that observation for asthma is likely not educationally or clinically significant. We believe this finding may be more problematic for the consideration of entrustment as an assessment framework than for using RSQMs to inform resident assessment. Specifically, we posit that RSQMs may help to improve the veracity of entrustment decisions.

Despite the need for further understanding and validity evidence, we believe patient-focused assessment efforts, such as RSQMs and entrustment, should be embraced given the promise to ensure high value care embodied in their attention to the patient. However, hazards to their use exist. There can be reluctance to adopt new ways of approaching assessment. Nonetheless, recent conversations in medical education about moving away from psychometrics and toward welcoming the subjective in performance assessment suggest change is possible. With calls for ensuring high value health care as well as uniting the missions of medical education and health care delivery, medical education may be at a crossroads. Efforts such as RSQMs can help the community take a clear step toward focusing on the patient in our assessment efforts.







## Samenvatting

### Hoofdstuk 1

Hoofdstuk 1 beschrijft de recente vraag naar een verbeterde link tussen de onderwijsresultaten uit de medische vervolgopleiding en de kwaliteit van de patiëntenzorg. Hoewel het bereiken van dit doel aan de basis ligt van competentiegericht onderwijs, belemmeren huidige hiaten in toetspraktijken en trainingen gericht op kwaliteitsverbetering pogingen om dit te bereiken. Met hiaten in toetspraktijken refereren we aan de ongewenste variatie in toetsen alsook aan de slechte afstemming van deze toetsing op de patiëntenzorg. In dit hoofdstuk stellen we dat patiëntgerichte toetsing van prestaties, en in het bijzonder *entrustable professional activities* (EPA's) en kwaliteitsmaten die betrekking hebben op het werk dat aiossen verrichten (welke we in het navolgende "aios-specifieke kwaliteitsmaten", oftewel ASKM's zullen noemen), een manier is om deze belemmeringen weg te nemen. Met het nemen van deze hordes zetten we een eerste stap richting een betere afstemming van onderwijsresultaten op de patiëntenzorg. Hoewel EPA's de afgelopen 10 jaar veel internationale populariteit hebben genoten, is het empirische bewijs daarvoor momenteel onvoldoende; daarnaast zijn er onvoldoende kwaliteitsmaten voor het werk dat aiossen verrichten. Dit bracht ons tot de volgende centrale onderzoeksvragen voor dit proefschrift:

Welke aspecten komen er kijken bij de ontwikkeling en toepassing van kwaliteitsmaten, welke we "aios-specifieke kwaliteitsmaten" (ASKM's) noemen, die grotendeels aan aiossen zijn toe te schrijven en die van belang zijn voor de patiëntenzorg?

Wat is het verband tussen *entrustment decisions*, oftewel besluiten om een kritische activiteit aan aiossen toe te vertrouwen, en hun prestaties op de ASKM's?

De studies in dit proefschrift vinden plaats in de context van een medische vervolgopleiding Kindergeneeskunde in de VS. Bij onze eerste stappen in deze baanbrekende onderzoeksrichting richtten we ons meer specifiek op de afdeling spoedeisende hulp (SEH) van dit specialisme.

### Hoofdstuk 2

In Hoofdstuk 2 wordt de vraag naar nieuwe benaderingen van prestatietoetsing teneinde onderwijsresultaten en de kwaliteit van de patiëntenzorg beter op elkaar te laten aansluiten, verder aangevuld door te stellen dat de toekomst van kwalitatief hoogwaardige zorg afhangt van een verbeterde toetsing van artsprestaties. Er wordt een drieledige onderzoeksagenda op dit gebied voorgesteld. Ten eerste moeten EPA's ingevoerd en uitgebreid onderzocht worden. Ten tweede moet het verband tussen *entrustment decisions* over aiossen en kwaliteitsmaten voor de zorg die zij verlenen onderzocht worden. Ten slotte zouden de publieke en private sector deze inspanningen moeten financieren en daar voorrang aan moeten geven, aangezien slechts

---

2% van de totale uitgaven voor de gezondheidszorg naar gezondheidszorgonderwijs voor alle beroepen samen gaat.

### Hoofdstuk 3

Het in dit proefschrift voorgestelde onderzoeksprogramma richt zich op individuele ASKM-prestaties en op besluiten om een kritische activiteit aan individuen toe te vertrouwen. Gezien de recente focus op zorgverlening door teams is het echter van belang deze toespitsing op individuen te verantwoorden. Hoofdstuk 3 licht toe waarom het belangrijk is te focussen op zowel individuen (door de *toeschrijving* van zorg aan een individu te overwegen) als teams, programma's en systemen (door de *bijdrage* aan zorg te overwegen). 'Bijdrage' richt zich op de diverse maten waarin verschillende eenheden aan een resultaat bijdragen. Daarbij speelt het een belangrijke rol bij het analyseren van resultaten op programma- en systeemniveau die als input dienen voor programma-evaluatie en -verbeteringen voor de toekomst. In de zorg is de rol van een individu en de aspecten van zorg die aan een individueel teamlid kunnen worden toegeschreven echter net zo belangrijk. Het is waardevol om individuen informatie te verschaffen over hun functioneren in het team, omdat dit hen kan aanzetten tot persoonlijke verbetering. Verder is het zo dat we individuele artsen, niet teams, laten afstuderen, hen diplomeren en hen kwalificaties toekennen voor de artsenpraktijk. Bovendien zijn prestatiebekostigingsmodellen ook gericht op de kwaliteit van de door individuen verleende zorg, waardoor het belangrijk is om de resultaten van individuele prestaties te kunnen meten. In dit hoofdstuk onderzoeken we hoe analyses van 'toeschrijving' en 'bijdrage' aanvullend kunnen worden toegepast om te achterhalen welke resultaten aan individuen en welke aan teams en programma's kunnen en moeten worden toegeschreven.

### Hoofdstuk 4

Hoofdstuk 4 beschrijft pogingen om de inhaalslag met betrekking tot kwaliteitsmaten voor gebruik onder aiossen te maken, door ASKM's te ontwikkelen voor gebruik op de afdeling kinder-SEH. In deze studie gebruikten we een *nominal group technique* (NGT) en Delphi-methode met stafleden en collega-supervisors van het Cincinnati Children's Hospital Medical Center (CCHMC) om mogelijke ASKM's te ontwikkelen (d.m.v. de NGT) en te prioriteren (d.m.v. de Delphi) voor contacten met patiënten met acute verergering van astma, bronchiolitis en licht schedeltrauma. We vroegen de groepen om mogelijke ASKM's te bedenken (in de NGT) en deze te scoren (in de Delphi) op basis van de volgende twee criteria: 1) het belang van de maat voor de betreffende ziekte; en 2) de kans dat een aios, en geen ander teamlid of het team gezamenlijk, de handeling waarop de maat van toepassing is, verricht. De NGT leverde 67 maten op voor astma, 46 voor bronchiolitis en 48 voor licht schedeltrauma. Deze vormden de input voor de eerste ronde van de Delphi-methode. Na twee rondes voldeden 18 maten voor astma, 21 voor bronchiolitis en 21 voor licht schedeltrauma aan de criteria voor automatische inclusie. Uit deze potentiële eindmaten stelden de deelnemers in ronde drie hun top 10 en de

volgende 5 op. De maten die geprioriteerd werden, vielen vaak in een van de volgende drie categorieën: 1) passend medicatiegebruik; 2) documentatie; en 3) informatieverstrekking bij ontslag van de patiënt. Samenvattend biedt deze studie een richtsnoer voor het identificeren en ontwikkelen van ASKM's die kwalitatief hoogwaardige zorgverlening tijdens en na de opleiding kunnen helpen bevorderen.

## Hoofdstuk 5

Hoofdstuk 5 vult het werk van hoofdstuk 4 op twee manieren aan: 1) het herhaalt het proces voor het ontwikkelen van ASKM's voor astma, bronchiolitis en licht schedeltrauma op de kinder-SEH met aiossen Kindergeneeskunde; en 2) het betreft belanghebbenden als aiossen en supervisors ten behoeve van de ontwikkeling van ASKM's en vervolgstappen daarin. Net als de stafleden en collega-supervisors richtten aiossen in de NGT en Delphi-rondes zich bij de drie betreffende ziektes voornamelijk op maten die betrekking hadden op: 1) passend medicatiegebruik; 2) documentatie; en 3) informatieverstrekking bij ontslag van de patiënt. De daaropvolgende focusgroepen met zowel aiossen als stafleden en collega-supervisors merkten ziekenhuisgeneeskunde en Algemene Kindergeneeskunde aan als prioritaire gebieden voor het ontwikkelen van toekomstige ASKM's, maar wezen ook op omgevingsvariabelen die de toepassing van vergelijkbare maten in andere settings beïnvloeden (de teamsamenstelling van de afdeling Algemene Kindergeneeskunde kan bijvoorbeeld de bruikbaarheid van enkele maten beperken). Aiossen en supervisors hadden zowel vergelijkbare als unieke inzichten in het ontwikkelen van ASKM's. Zo leverde de NGT met aiossen meer maten op dan de NGT met supervisors en beide focusgroepen meenden dit te kunnen toeschrijven aan het feit dat aiossen volop in de praktijk bezig waren en daardoor meer details kenden van de zorg die daar verleend werd. Supervisors merkten echter op dat aiossen mogelijk niet voldoende in staat zijn de juiste prioriteiten te stellen of niet weten welke middelen tot hun beschikking staan om hun werk goed te kunnen verrichten. Aiossen daarentegen, gaven aan dat supervisors mogelijk zelfs geen weet hebben van de details die bij het verlenen van zorg aan de frontlinie komen kijken. Samenvattend bouwt dit hoofdstuk voort op hoofdstuk 4 om de weg vooruit te blijven effenen voor de ontwikkeling van toekomstige ASKM's, door specifieke settings, maten en belanghebbenden waarmee rekening gehouden moet worden bij het ondernemen van dit werk te onderzoeken.

## Hoofdstuk 6

Hoofdstuk 6 presenteert de verschillende individuele en collectieve resultaten voor de in hoofdstuk 4 en 5 besproken ASKM's, nadat deze waren gebruikt om aiossen Kindergeneeskunde werkzaam op de kinder-SEH te beoordelen. Tijdens het academisch jaar 2017-2018 werden aiossen Kindergeneeskunde op de afdeling kinder-SEH van het CCHMC beoordeeld aan de hand van ASKM's voor acute verergering van astma (21 ASKM's), bronchiolitis (23) en licht schedeltrauma (19). Van de geïncludeerde patiëntcontacten zorgden 83 aiossen voor 110 patiënten met astma,

---

112 met bronchiolitis en 77 met licht schedeltrauma. Op enkele uitzonderingen na (bijvoorbeeld wanneer patiënten met verergering van astma zich meldden bij de medische reanimatie-afdeling, waardoor het zorgproces, overeenkomstig het protocol, in gang gezet werd door de toeziend arts, konden de aiossen niet voldoen aan enkele ASKM's die op deze handelingen betrekking hadden), waren aiossen in de gelegenheid om de meeste ASKM-relevante activiteiten te verrichten. De frequentie waarmee ASKM-gerelateerde handelingen zich voordeden was hoog en dit bood aiossen ruim de kans om aan de kwaliteitsmaten te voldoen. Aan één kwaliteitsmaat voor licht schedeltrauma werd in alle patiëntcontacten voldaan. Over de drie ziekten werd in bijna alle patiëntcontacten nog eens aan 9 ASKM's voldaan, hoewel in een aantal tot het merendeel van de patiëntcontacten niet werd voldaan aan de meeste ASKM's. Om vast te stellen hoeveel individuele ASKM's van het totale aantal ASKM's dat had kunnen worden uitgevoerd correct werden uitgevoerd, werd er voor elk patiëntcontact een samengestelde score uitgerekend. De samengestelde scores voor de ASKM's liepen opvallend uiteen en vertoonden opmerkelijk veel variatie: gemiddelde voor astma: 0,84 (SD: 0,11); gemiddelde bronchiolitis: 0,62 (SD: 0,12); licht schedeltrauma: 0,63 (SD: 0,10). Samenvattend laat hoofdstuk 6 zien dat individuele en samengestelde ASKM's de prestaties van aiossen over patiëntcontacten heen kunnen vangen en onderscheiden, wat de deur opent voor het gebruik ervan in andere contexten en door andere specialismen.

## Hoofdstuk 7

Hoofdstuk 7 onderzoekt het verband tussen samengestelde ASKM-scores voor individuele patiëntcontacten op de afdeling kinder-SEH, entrustment decisions die supervisors nemen voor die patiëntcontacten (zoals blijkt uit de mate van toegewezen supervisie) en de ernst\* en complexiteit van deze patiëntcontacten. Voor de in dit hoofdstuk gerapporteerde studie gebruikten we de 83 aiossen en 299 contacten met astma-, bronchiolitis- en licht-schedeltraumapatiënten die in hoofdstuk 6 werden gepresenteerd. Om het verband tussen de samengestelde ASKM-scores en de andere relevante variabelen te meten, werden mixed models toegepast om de binnen aiossen geneste observaties te verklaren. Er bestond een positieve correlatie tussen entrustment decisions en de samengestelde ASKM-scores voor astma (bèta-coëfficiënt van 0,03;  $p = 0,0004$ ), maar er werd geen significante correlatie gevonden tussen samengestelde ASKM-scores en entrustment decisions voor bronchiolitis en licht schedeltrauma. Hoewel statistisch significant, heeft dit positieve lineaire verband voor astma waarschijnlijk geen onderwijskundige noch klinische relevantie, omdat de samengestelde ASKM-scores bij elke toename van 1 punt (op 7) in entrustment-score enkel met 0,03 (op 1) toenamen. Er wordt verondersteld dat er een positief verband bestaat tussen entrustment decisions en de samengestelde ASKM-scores. Aangezien ASKM's een meer objectieve maatstaf voor prestaties vormen, zouden we ons naar aanleiding van de bevindingen van deze studie moeten afvragen of entrustment decisions (al dan niet deels) gebaseerd moeten worden op uit ASKM's verkregen prestatiegegevens. De samengestelde ASKM-scores waren significant hoger naarmate de ernst toenam en significant

lager naarmate deze afnam voor zowel astma ( $p=0,0009$ ) als bronchiolitis ( $p=0,01$ ). De samengestelde ASKM-scores voor licht schedeltrauma bleven echter vrijwel identiek bij wisselende maten van ernst ( $p=0,94$ ). Er werd geen verband gevonden tussen samengestelde ASKM-scores en patiëntcomplexiteit.

## Hoofdstuk 8

Toen het eerste bewijs van invoering van ASKM's op de afdeling kinder-SEH er eenmaal lag, trachtte hoofdstuk 8 na te gaan hoe ASKM-gegevens zouden kunnen worden gebruikt door mensen van buiten de afdeling kinder-SEH verantwoordelijk voor het nemen van beoordelingsbesluiten. Om dit doel te bereiken onderzocht de studie in dit hoofdstuk hoe individuele leden van de Klinische Competentiecommissie (KCC) de ASKM-data interpreteren, gebruiken en prioriteren wanneer deze in hun gebruikelijke beoordelingsprocedure worden opgenomen. In deze constructivistische-gefundeerde-theoriestudie namen we een doelgerichte en theoretische steekproef van 19 KCC-leden van de vervolgopleiding Kindergeneeskunde van het CCHMC. Participanten ontvingen een beoordelingsportfolio van een aios voorzien van beoordelingsscores en narratief commentaar over vijf klinische stages, alsook ASKM-gegevens over een van deze stages. Vervolgens werd hen gevraagd om deze te beoordelen en daarbij tot een besluit te komen over de mate waarin zij de aios in staat achtten om de zorg voor patiënten met veelvoorkomende acute problemen op zich te nemen (een entrustable professional activity voor Algemene Kindergeneeskunde). De dataverzameling besloeg twee fasen: 1) het observeren van en hardop denken door participanten tijdens het beoordelen van het portfolio van de aios; en 2) semigestructureerde interviews om participanten verder te bevragen over hun beoordelingen. De analyse ging van zorgvuldig lezen over op het coderen en ontwikkelen van thema's, wat uitmondde in een model dat in kaart bracht hoe de verschillende thema's zich onderling tot elkaar verhielden. We onderscheidden vijf dimensies voor de manier waarop participanten naar ASKM's keken en deze gebruikten: 1) Het zich kunnen oriënteren op de ASKM's: variërend van 'verwarrend' tot 'dat wijst zich vanzelf'; 2) De geneigdheid om de ASKM's daadwerkelijk te gebruiken: variërend van 'met terughoudendheid' tot 'met enthousiasme'; 3) De interpretatie van ASKM's: variërend van 'behoeft plaatsing in de juiste context' tot 'is duidelijk'; 4) ASKM's voor beoordelingsbesluiten: variërend van 'is niet lastig' tot 'is lastig'; en 5) Verwachtingen voor aiossen: variërend van 'het gebruik van ASKM's is mogelijk oneerlijk' tot 'is eerlijk'. Het samenspel van deze dimensies leidde tot drie profielen die weergaven hoe participanten gebruik maakten van de ASKM-gegevens, te weten: enthousiaste toepassing, gewillige toepassing en terughoudende toepassing. De eerste twee profielen kwamen het vaakst voor: de meeste participanten maakten in zekere mate gebruik van de ASKM's, wat spreekt vóór de opname van ASKM's ten behoeve van de beoordeling van aiossen door de KCC.

---

## Hoofdstuk 9

Naarmate ASKM's verder worden ontwikkeld, wordt het belangrijk dat we een optimale manier vinden om deze uit het elektronisch patiëntendossier (EPD) te halen, alsook om automatisch vast te stellen aan welke aiossen zij moeten worden toegeschreven in settings waar meerdere aiossen gemoeid zijn met de zorg voor een bepaalde patiënt, zoals de kliniek. Hoofdstuk 9 ging op zoek naar een methode waarmee de zorg voor individuele patiënten met behulp van "voetafdrukken" in het EPD aan individuele aiossen kon worden toegeschreven. In augustus 2017 en januari 2018 registreerden vijf dienstdoende specialisten van het University of Cincinnati Medical Center dagelijks de activiteiten van aiossen die zorgdroegen voor patiënten op een ziekenhuisafdeling Interne Geneeskunde. Deze gegevens beschouwden we als ultiem ijkpunt voor het identificeren van aiossen. Om te kunnen vaststellen in welke mate de aios bij de zorg betrokken was, keken we naar de volgende variabelen: vervolgopleidingsjaar, voortgangsnotitie van auteur, ontslagsamenvatting van auteur, door arts aangevraagde onderzoeken en activiteitenlogs uit het patiëntendossier. We zetten deze variabelen om in categorale kenmerken (bijv. voortgangsnotitie auteur: ja/nee), selecteerden informatieve kenmerken en rangschikten deze met behulp van een beslisboomalgoritme. Er werden in totaal 1.511 activiteiten geregistreerd, waarvan er 116 werden toegeschreven aan aiossen. Uit de resultaten bleek dat alle variabelen, behalve de 'ontslagsamenvatting auteur', ten minste enige mate van belangrijkheid vertoonden in de modellen. Het beste model bereikte 78,95% sensitiviteit, 97,61% specificiteit en een gebied onder de ROC-curve van circa 91%. Kortom, deze studie voorspelde met succes, door gebruik te maken van gegevens uit het EPD, de mate waarin aiossen binnen een ziekenhuisteam voor patiënten zorgdroegen. Daarbij kwam een uitstekend model naar voren. Hiermee wordt de basis gelegd voor het toeschrijven van patiënten aan aiossen zodat voor elke aios in kaart kan worden gebracht met welke patiëntdiagnoses deze effectief te maken heeft gehad. Deze kennis kan worden ingezet ten bate van voortdurende kwaliteitsverbeteringsdoeleinden in de medische vervolgopleiding en uiteindelijk helpen om ASKM's aan aiossen toe te schrijven.

## Hoofdstuk 10

Hoofdstuk 10 beschrijft opnieuw de herhaalde vraag naar een betere waarborging van kwalitatief hoogwaardige zorg en stelt dat het medisch onderwijs een belangrijke bijdrage kan en moet leveren aan het bereiken van dit doel. Wij zijn van mening dat we ons hiervoor ten volle en voortdurend moeten richten op de patiënt, iets dat voor onderzoek van medisch onderwijs of de medische-onderwijspraktijk van oudsher nooit een grote focus is geweest. Patiëntgerichte toetsmethoden zoals ASKM's kunnen helpen om dit doel te bereiken. De kernboodschap van ons onderzoek is dat ASKM's die met behulp van consensusgroepsmethoden zijn ontwikkeld grotendeels de handelingen van aiossen weerspiegelen en dat de verrichte handelingen sterk van aios tot aios en van patiëntcontact tot patiëntcontact verschillen. Bij de toekomstige ontwikkeling van ASKM's moet worden overwogen om andere belanghebbenden in het proces te betrekken en dit uit te breiden naar andere settings en specialismen. Ons onderzoek

reikt al enkele mogelijke toekomstige settings op het gebied van Kindergeneeskunde aan, waaronder ziekenhuisgeneeskunde en de poli Algemene Kindergeneeskunde, maar het is van belang dat er ook onderzoek wordt verricht binnen andere medische specialismen, alsmede chirurgische en ziekenhuisspecialismen. Om dit werk te voltooien zouden belanghebbenden die we tot nog toe niet in ons werk hebben opgenomen, zoals verplegers, patiënten en hun families, betrokken moeten worden. Onze bevindingen met betrekking tot het verband tussen de ASKM's en entrustment decisions over aiossen liepen uiteen. Voor alle aandoeningen (astma, licht schedeltrauma en alle andere aandoeningen samen), behalve bronchiolitis, gold dat de samengestelde ASKM-scores significant hoger waren wanneer aiossen hun verantwoordelijkheden zonder supervisie mochten uitvoeren of wanneer zij belast waren met de supervisie van anderen dan wanneer zij onder een grotere mate van supervisie stonden. Dit levert validiteitsbewijs voor de samengestelde ASKM-scores, hoewel het absolute verschil gering was. Voorts zagen we dat, wanneer de supervisieniveau's afzonderlijk werden beschouwd in plaats van deze in twee categorieën op te delen (1: mag op zijn minst taken zonder supervisie uitvoeren, en 2: heeft ongeacht het niveau nog supervisie nodig), er alleen voor astma een significant positief lineair verband bestond tussen de samengestelde ASKM-scores en entrustment decisions. Waarschijnlijk heeft deze observatie voor astma echter geen onderwijskundige noch klinische relevantie. Naar onze mening vormt deze bevinding eerder een te overwegen bedreiging voor het gebruik van entrustment als beoordelingskader dan voor het gebruik van ASKM's ten behoeve van de beoordeling van aiossen. Anders gezegd, we stellen dat ASKM's kunnen helpen de nauwkeurigheid van entrustment decisions te verbeteren. Ondanks de behoefte aan verder inzicht en validiteitsbewijs, vinden wij dat patiëntgerichte toetsmethoden zoals ASKM's en entrustment moeten worden omarmd aangezien zij, door hun aandacht voor de patiënt, kwalitatief hoogwaardige zorg beloven te borgen. Het gebruik van deze toetsmethoden is echter niet zonder risico. De overgang op nieuwe benaderingen van toetsing kan met enige weerstand gepaard gaan. Dit gezegd te hebben, beloven recente discussies waarin ervoor gepleit wordt de psychometrie achter te laten en in plaats daarvan het subjectieve in prestatiebeoordelingen te omarmen, dat verandering wél mogelijk is. Dankzij de herhaalde vraag naar een betere waarborging van kwalitatief hoogwaardige zorg alsmede een betere afstemming tussen de missies van het medisch onderwijs en de zorg, staat het medisch onderwijs mogelijk aan de vooravond van een keerpunt. Vernieuwingen zoals ASKM's kunnen de gemeenschap helpen om een duidelijke stap te zetten richting het einddoel om de patiënt in onze beoordelingsmethoden centraal te stellen.