MASTER THESIS

Discrepancies in neuroradiology

How to explain and minimize discrepancies based on the experience and expertise of radiologists?

C.M.H. Degger

Student	C.M.H. Degger (s2149885)
Examination Committee	Dr. Anique Bellos-Grob
	Prof. Dr. Sabine Siesling
	Onno Vijlbrief
Study	Master Health Sciences
Institute	University of Twente
Faculty	Faculty of Science & Technology (TNW)
Document number	HS-20180018-2
Date	4-7-2019

Preface

Foremost, I would like to thank my supervisors dr. Anique Bellos-Grob, Prof. dr. Sabine Siesling and Onno Vijlbrief for their support during my research, their motivation, great knowledge and flexibility while working with me as a student. Also, I would like to express my sincere thanks to ZiekenhuisGroep Twente (ZGT) and Maatschap Radiologie Oost-Nederland (MRON), who enabled the data in order to carry out this study. Furthermore, I want to thank my family for their untiring and unconditional support during my master Health Sciences.

Summary

Purpose: Diagnostic errors often cause patient harm, due to delayed, incorrect or missed diagnoses. It is likely that radiologists' contribution is substantial, because diagnoses are often based on medical imaging. *Discrepancy* is a proper term to use, whenever talking about supposed errors or observer variation in radiology. The aim of this study is to analyse all neuroradiological (head and neck) discrepancies in order to study the current seriousness of discrepancies. The second aim of this study is to analyse neuroradiological discrepancies made by supervised or unsupervised radiology residents, to identify patterns for improvement and educational opportunities to reduce future discrepancies.

Methods: This retrospective quantitative descriptive study analysed 214 neuroradiological discrepancies out of 61,246 neuroradiological imaging studies and 88 neuroradiological discrepancies made by radiology residents out of 17,385 neuroradiological imaging studies assessed and reported by radiology residents. The Dutch ZiekenhuisGroep Twente gathered data for this study, between 1 April 2017 and 5 March 2019. This study was divided in two parts; part one focussed on all neuroradiological discrepancies and part two focussed on neuroradiological discrepancies made by radiology residents. Part 1: Subgroup analysis was performed, based on the variable *discrepancy made by*. Part 2: Subgroup analysis was performed, based on the variable *supervision.* Both parts included variables like *setting* and *imaging modality* and used the *Chi-square for independence test*, the *Fisher's exact test*, the *Fisher-Freeman-Halton exact test* and the *Independent-Samples Median Test.*

Results: Part 1: The overall discrepancy rate of all neuroradiological discrepancies is 0.35% (214). Most discrepancies are made by radiology residents (88; 41.1%). The number of discrepancies in CT-scans is significantly lower and the number of discrepancies in MRI-scans is significantly higher, for neuroradiologists and radiologists with a neuro subspecialisation. Neuroradiologist cause significantly more discrepancies in the morning and less in the evening/ night. Radiologists with a neuro subspecialisation cause significantly less discrepancies in the evening/ night. Radiology residents cause significantly less discrepancies in MRI-scans. Also, radiology residents cause significantly more discrepancies in the evening/ night. Part 2: The overall discrepancy rate of radiology residents is 0.5% (88). Discrepancies occur often without supervision (57; 64,8%) but also with supervision (31; 35.2%). Discrepancy rates are high in MRI-scans (13; 0.73%), the emergency (58; 0.54%) and outpatient setting (20; 0.64%). When causing discrepancies, unsupervised radiology residents are significantly further in their study (four years and two months) than supervised radiology residents (one year and 11 months).

Conclusion: Radiology residents cause most discrepancies. Further research must confirm whether the supervision and competence levels of the Dutch *radiology training plan (CORONA)* offer improvement and educational opportunities, for radiological departments and national guidelines.

Contents

List of acronyms

CORONA	Curriculum Opleiding Radiologie en Nucleaire Geneeskunde
MRON	Maatschap Radiologie Oost-Nederland
PACS	Picture Archiving and Communication System
SPSS	Statistical Package for Social Sciences
ZGT	ZiekenhuisGroep Twente

1 Introduction

Diagnostic errors often cause patient harm, due to delayed, incorrect or missed diagnoses. It is likely that radiologists' contribution is substantial, because diagnoses are often based on medical imaging [1, 2]. Errors might increase morbidity, mortality and length of hospital stay [3-5]. One billion radiological studies are performed worldwide annually with a 4% error rate, that would result in forty million radiological errors annually [1, 6, 7]. Zooming in on neuroradiological (head and neck) errors made by radiology residents shows overall error rates between 2.0 to 2.7% and significant error rates between 1.7% and 8.6% [8-11]. The overall error rate of radiology residents in an emergency setting was 3.3% [12]. Most significant errors were made by second-year radiology residents [10]. This study focusses on neuroradiological (head and neck) discrepancies, which will be referred to as *neuroradiology* as it is a common term in literature [13-16].

There is a difference between diagnostic error and observer variation. True 'errors' occur whenever the consensus of a radiologists' peers is substantially different from the discrepancy and there is no disagreement possible about the correct interpretation [1, 6, 7]. Variance is not a real 'error', it occurs whenever the correct interpretation is subject to a legitimate difference of opinion [7, 17]. Besides, decision-making in radiology is often not based on clear clinical presentations and imaging does not always show the diagnostic appearances of particular diagnoses. There is not one sensitivity and specificity degree for MRI-scans, CT-scans and X-rays, because it depends on the disease [18, 19]. So, it is impossible to say how often diseases are not shown. True errors, variance and unclear clinical presentations make it impossible to have clear distinctions between normal and abnormal imaging [7, 17, 20]. Therefore, the interpretation of radiologists is very important and with a high risk of variance. *Discrepancy* is a proper term to use, whenever talking about supposed errors or observer variation in radiology [6].

Dutch radiology residents follow a training program of five years, to become a general radiologist. Thereafter, they can become a fellow radiologist, in which they specialise themselves in two years in for example neuroradiology. During the five years training program, radiology residents are supervised based on competence levels of the Dutch *radiology training plan (CORONA)*. Radiology residents are not allowed to assess and report imaging independently (competence level 1) or unsupervised (competence level 2). Thereafter, radiology residents can assess imaging independently and receive limited supervision of which the initiative lies with radiology residents (competence level 3), radiology residents work independently but can ask for help (competence level 4) or radiology residents supervise others (competence level 5). These competence levels only apply in the Netherlands and radiologists determine the competence level of a radiology resident [21]. Rules regarding supervision of radiology residents are determined independently and individually by each country.

The average number of imaging that a radiologist must interpret per minute on an uninterrupted eight hours working day, has increased between 1999 and 2010 with almost a sevenfold per radiologist for CT-scans and a fourfold for MRI-scans [7, 22]. This trend might negatively influence the number of *false positive* or *false negative* discrepancies, since radiologists might take less breaks and experience increased stress and fatigue [7, 22, 23].

Radiological discrepancies can be due to multiple causes, like inadequate, incorrect or incomplete clinical information, different opinions of radiologists, cognitive biases, system failure, failure of attention, under reading, faulty reasoning etcetera [1, 4, 7, 17, 20, 24]. In order to prevent neuroradiological discrepancies, layers of defence are installed by radiology departments to offer protection against risks (appendix 1). However, holes in this defence expose risks and eventually cause a discrepancy that influences the quality of care negatively [25, 26]. Studies showed that between 60% and 80% of all radiology discrepancies are perceptual discrepancies, which are due to scanning, decision or recognition errors [1, 20, 27, 28]. "To err is human", but we should prevent them as much as possible [29]. Therefore, the Dutch healthcare inspectorate advised in 2016 to implement a discrepancy registration, to get insight in the current performances of radiology departments and to improve diagnostic care [30]. A discrepancy registration could help to determine minimum radiological quality standards and improvement opportunities for diagnostic care [3, 6].

American studies pointed out that neuroradiologists do not perform significantly better compared to unspecialised radiologists [5, 31]. Other American studies pointed out that radiologists do not perform more accurate than radiology residents, if they assess imaging outside their specialisation (like neuroradiology) [32]. Besides, radiology residents cause the most discrepancies in their second-year [10]. A study from Israel found that radiology residents cause more discrepancies in the emergency setting [13, 33].

Neuroradiological discrepancies have been studied in America but those studies are scarce in Europe and not yet performed in the Netherlands. It is scientifically relevant to study this topic as well in Europe, specifically in the Netherlands, because it enables us to compare performances of European countries with American countries. This would allow us to learn from each other and try to prevent discrepancies from happening again [13, 25].

The aim of this study is to analyse all neuroradiological discrepancies in order to study the current seriousness of discrepancies. The second aim of this study is to analyse neuroradiological discrepancies made by supervised or unsupervised radiology residents, to identify patterns for improvement and educational opportunities to reduce future discrepancies.

2 Methods

This chapter will discuss the study design, the dataset and the data analysis that were used during this study.

2.1 Study design

This retrospective quantitative descriptive study analysed 214 neuroradiological discrepancies made by radiologists and radiology residents in *ZiekenhuisGroep Twente* (ZGT). ZGT is a regional hospital located in the cities Hengelo and Almelo in the East of the Netherlands and facilitates prevention, diagnosis, treatment and nursing to stimulate approximately 250,000 patient's health annually. Radiology residents are trained in ZGT.

2.2 The dataset

ZGT implemented a discrepancy registration system in April 2017 and discrepancies in this study were discovered between April 2017 and March 2019. Included neuroradiological discrepancies were based on imaging from 1 January 2012 till 5 March 2019. The starting point of 2012 was chosen, because radiologists work since then based on their (neuro-) specialisation. The radiology department works with the *JiveX Picture Archiving and Communication System (PACS)* of *Alphatron Medical*, to assess and report imaging and to detect and registrate discrepancies.

Data has been made anonymously, therefore patients or reporting radiologist could not be identified and informed consent was not necessary. This study consisted of two parts. Part one provided an overview of information, based on all neuroradiological discrepancies made by radiologists and radiology residents. The variable of interest, *discrepancy made by*, was coded based on the scheduling and number of neuroradiological imaging assessed and reported by a specific radiologist between April 2017 and March 2019. The official Dutch classification of registered neuroradiologist and subspecialist were not used in this study, because training qualifications were different during the education period of older radiologists and it would underestimate their qualities nowadays. Recoding based on the output of radiologists is therefore more representative. Neuroradiologists assessed and reported ≥10% of all 16,246 neuroradiological imaging, radiologists with a neuro subspecialisation ≥2% and radiologists with another specialisation <2%. The fourth category included radiology residents, who registered discrepancies made by other radiology residents. Part two of this study focussed on the discrepancies made by radiology residents, to investigate whether this subgroup showed certain patterns that offer educational opportunities. The focus of interest, supervision, was coded as supervised or unsupervised, which was noted per imaging. Supervised radiology residents are corrected whenever their interpretation is incorrect, according to their supervisor during that shift. Discrepancies among supervised radiology residents indicated that both the radiology resident and the supervisor interpreted the imaging

'incorrect'. Unsupervised radiology residents are responsible to ask for help if they doubt their own judgement capabilities on certain imaging [21].

This study included the independent variables; *discrepancy author, setting, imaging modality, discrepancy discovered by, time of assessing/ reporting imaging, supervised by, discrepancy classification, clinical relevance, training progress radiology residents (years; months)* and *discrepancy discovery duration (days).* The variable *discrepancy author* described whether discrepancies were discovered by neuroradiologists, radiologists with a neuro subspecialisation, radiologists with another specialisation or radiology residents. The coding was based on the variable *discrepancy made by*. The variable *setting* describes whether patients were referred to the radiology department from an *emergency setting, outpatient setting, inpatient setting* or by a *general practitioner*. The variable *imaging modality* was recoded based on the main neuroradiological modalities in ZGT; *CT-scan, MRI-scan* and *X-ray.* If the results encourage to analyse in more detail, it is possible to create new variables based on for example the focus of a CT-scan. The variable *discrepancy discovered by* showed how a discrepancy was found; *a second assessment of a radiologist, repeating the research, new advanced research* or by the *assessment of a non-radiologist.* The variable *time of assessing/ reporting imaging* showed whether a discrepancy was made in the *morning (6:00 – 12:00 h), afternoon (12:00 – 18:00 h)* or *evening/ night (18:00 – 6:00 h)*. The variable *supervised by* used the same three categories as the variable *discrepancy made by*, only the fourth category included for this variable all discrepancies without supervision. The variable *discrepancy classification* categorized causes of discrepancies in basically three classification types; cognition/ interpretation discrepancies, perceptual/ observation discrepancies and system related discrepancies [1, 4, 13]. Cognition/ interpretation discrepancies relate to different types of cognitive bias, like framing and satisfaction of search bias [25, 34]. Perception/ observation discrepancies relate to the interaction between the imaging and the radiologist and are often due to scanning, decision or recognition error [4, 20, 27, 35]. System discrepancies relate to technical or equipment failure, mental or physical fatigue, workplace distractions, procedures and policies [4, 25, 27]. The variable *clinical relevance* was based on the worldwide used RADPEER® scoring system and scored by radiologists of ZGT, prior to this study. The categories within the RADPEER® scoring system are based on the accuracy of a radiologists' interpretation, with one concurrence score and three different disagreements scores [17, 36-39]. Those four categories were recoded into two categories, *no/ low clinical relevance* and *significant/ severe clinical relevance*, in order to create proper category sizes. No/low clinical relevance implied no clinical relevance, or clinical relevance but without a necessary treatment adjustment. Significant/ severe clinical relevance implied a necessary treatment adjustment or a necessary treatment adjustment and a different definitive outcome for patients.

2.3 Data analysis

Part 1 of this study provided an overview of all neuroradiological discrepancies, including subgroup analysis for the dependent variable *discrepancy made by*. This was done in order to study the current situation of neuroradiological discrepancies. Part 2 studied the neuroradiological discrepancies made by radiology residents, including subgroup analysis for the dependent variable *supervision*. This was done to investigate the radiology resident discrepancies and the role of supervision more detailed. Besides, the overall discrepancy ratio and the ratio between the number of discrepancies and the variables *discrepancy made by* and *imaging modality* will be shown in part 1, to represent the current seriousness of discrepancies. Part 2 included the overall discrepancy ratio and the ratio between the number of discrepancies made by radiology residents and the variables *imaging modality* and *setting*. Part 1 did not include the variables *supervision, supervised by* and *training progress radiology residents (years; months).*

The *Chi-square test for independence* was used to perform subgroup analysis on the categorical variables. The *Fisher's exact test* (2x2 table) and the *Fisher-Freeman-Halton exact test* (tables larger than 2x2) were used, if the assumption *minimum expected cell frequency* was violated. The *phi coefficient* (2x2 table) and *Cramer's V* (tables larger than 2x2) were used to test the effect size of significant results. The *Independent-Samples Median Test* was used to perform subgroup analysis for continuous variables. Two sided *p* values were used and a $p < 0.05$ indicated statistically significant results [40]. Recoding and analysis have been done with Excel 2016 and *Statistical Package for Social Sciences (SPSS)* version 25. Improvement recommendations were made, if significant results urged to intervene.

3 Results

This chapter describes the results for part 1 and part 2 of this study.

3.1 Results part 1

A total of 61,246 neuroradiological imaging studies were performed and 214 neuroradiological discrepancies were found, between April 2017 and March 2019. This results in a 0.35% (214) overall discrepancy rate and an 0.15% (93) significant/ severe discrepancy rate. Of all 61,246 neuroradiological imaging, it was unknown for 137 imaging studies who assessed and reported it. Of 61,109 neuroradiological imaging, neuroradiologists assessed 32.6% (19,946), radiologists with a neuro subspecialisation assessed 32.7% (20,054), radiologists with another specialisation assessed 6.1% (3,724) and radiology residents assessed 28.4% (17,385). The discrepancy rate is 0.21% (42), 0.35% (70), 0.38% (14), and 0.51% (88) respectively. Overall, CT-scans are mostly interpreted by radiology residents (54.2%; 10.463). MRI-scans (45%; 8,593) and X-rays (33.9%;) are mostly interpreted by radiologists with a neuro subspecialisation. All radiologists and radiology residents together show a discrepancy rate of 0.48% (93/19,295) for CT-scans, 0.48% (92/19,081) for MRI-scans and 0.13% (29/22,870) for X-rays.

Table 1 shows a general descriptive analysis of the 214 neuroradiological discrepancies, divided in the categories of the dependent variable *discrepancy made by*. Most discrepancies are made by radiology residents (41.1%) and neuroradiologist find and register the most discrepancies (57%). Neuroradiological discrepancies are often of significant/severe (43.4%) clinical relevance, but still the most discrepancies are of low or no clinical relevance (56.5%). There were only three discrepancies of severe clinical relevance, which were all made by radiology residents. Perception/ observation discrepancies are most common (77.6%) and most discrepancies occur in an outpatient setting (51.4%), in the afternoon (56.1%), in MRI-scans (43%) and CT-scans (43.5%). Discrepancies are often discovered by a second assessment of a radiologist (36.4%) or by repeating the research (31.3%). Discrepancies are most often found after twenty days.

No assumptions were violated in the *Chi-square for independence test* for the variable *imaging modality*. The number of discrepancies in CT-scans made by neuroradiologists and radiologists with a neuro subspecialisation is significantly lower and the number of discrepancies in MRI-scans is significantly higher, compared to radiology residents ($p < 0.001$). The number of discrepancies made by radiology residents in CT-scans is significantly higher and the number of discrepancies in MRI-scans is significantly lower, compared to neuroradiologists and radiologists with a neuro subspecialisation ($p < 0.001$). The variables *discrepancy author*, *setting*, *discrepancy discovered by* and *time of assessing/ reporting imaging* violated the assumption of *minimum expected cell frequency*. The *Fisher-Freeman-Halton exact test* showed that discrepancies made by radiology residents were significantly less often

found by neuroradiologists (p = 0.035). Also, discrepancies made by radiologists with a neuro subspecialisation were significantly less often found by radiology residents and discrepancies made by

Table 1
Descriptive statistics radiologists and radiology residents (n = 214)

Radiologists and radiology residents			**Dependent categorical variable *discrepancy made by***							
Independent categorical variables	Total sample N	(%)	Neuroradiologist		Radiologists with a neuro subspecialisation		Radiologist with another specialisation		Radiology residents	
			42	(19,6%)	70	(32,7%)	14	(6,5%)	88	(41,1%)
Discrepancy author										
▪ Neuroradiologist	122	(57.0%)	24	(57.1%)	46	(65.7%)	10	(71.4%)	**42** **c**	(47.7%)
▪ Radiologists with a neuro subspecialisation	49	(22.9%)	10	(23.8%)	15	(21.4%)	3	(21.4%)	21	(23.9%)
▪ Radiologists with another specialisation	10	(4.7%)	3	(7.1%)	4	(5.7%)	1	(7.1%)	2	(2.3%)
▪ Radiology resident	33	(15.4%)	5	(11.9%)	**5** **c**	(7.1%)	0	(0.0%)	**23** **c**	(26.1%)
Clinical relevance										
▪ None/ low	121	(56.5%)	25	(59.5%)	37	(52.9%)	10	(71.4%)	49	(55.7%)
▪ Significant/ severe	93	(43.4%)	17	(40.5%)	33	(47.1%)	4	(28.6%)	39	(44.3%)
Discrepancy classification										
▪ Cognition/interpretation	48	(22.4%)	7	(16.7%)	19	(27.1%)	5	(35.7%)	17	(19.3%)
▪ Perception/observation	166	(77.6%)	35	(83.3%)	51	(72.9%)	9	(64.3%)	71	(80.7%)
Setting										
▪ Emergency department	67	(31.3%)	1	(2.4%)	5	(7.1%)	3	(21.4%)	58	(65.9%)
▪ Outpatient setting	110	(51.4%)	31	(73.8%)	52	(74.3%)	7	(50.0%)	20	(22.7%)
▪ General practitioner	20	(9.3%)	4	(9.5%)	6	(8.6%)	3	(21.4%)	7	(8.0%)
▪ Inpatient setting	17	(7.9%)	6	(14.3%)	7	(10.0%)	1	(7.1%)	3	(3.4%)
Imaging modality										
▪ CT-scan	93	(43.5%)	**9** **a**	(21.4%)	**19** **a**	(27.1%)	5	(35.7%)	**60** **a**	(68.2%)
▪ MRI-scan	92	(43.0%)	**27** **a**	(64.3%)	**45** **a**	(64.3%)	7	(50.0%)	**13** **a**	(14.8%)
▪ X-ray	29	(13.6%)	6	(14.3%)	6	(8.6%)	2	(14.3%)	15	(17.0%)
Discrepancy discovered by										
▪ Second assessment radiologist	78	(36.4%)	13	(31.0%)	27	(38.6%)	4	(28.6%)	34	(38.6%)
▪ New advanced research	60	(28.0%)	15	(35.7%)	16	(22.9%)	4	(28.6%)	25	(28.4%)
▪ Repeating research	67	(31.3%)	13	(31.0%)	23	(32.9%)	6	(42.9%)	25	(28.4%)
▪ Assessment of non-radiologist	9	(4.2%)	1	(2.4%)	4	(5.7%)	0	(0.0%)	4	(4.5%)
Time of assessing/ reporting imaging										
▪ Morning	71	(33.2%)	**20** **b**	(51.3%)	27	(39.7%)	6	(46.2%)	**18** **b**	(20.5%)
▪ Afternoon	120	(56.1%)	19	(48.7%)	41	(60.3%)	5	(38.5%)	55	(62.5%)
▪ Evening/ night	17	(7.9%)	**0** **b**	(0.0%)	**0** **b**	(0.0%)	2	(15.4%)	**15** **b**	(17.0%)
Missing	*6*	*(2.8%)*								

Independent continues variables	Total sample Median (SD)		Neuroradiologist	Radiologists with a neuro subspecialisation	Radiologist with another specialisation	Radiology residents
Discrepancy discovery duration (days)	20.0	(554.22)	**84.5** **d** (674.22)	**34.5** **d** (537.82)	**699.5** **d** (798.10)	**5.50** **d** (370.23)
▪ Range	2493		2493	2471	2098	1597

***** System related discrepancies were excluded, due to small category sizes (n=4)
a Chi-square for independence; χ^2 (6, n = 214) = 51.918, Cramer's V = 0.348, $p < 0.001$
b Fisher-Freeman-Halton exact test; (6, n = 208) = 31.028, Cramer's V= 0.265, Exact Sig. (2-sided) $p < 0.001$
c Fisher-Freeman-Halton exact test; (9, n = 214) = 16.860, Exact Sig. (2-sided) p = 0.035
d Independent-Samples Median Test; test statistic (3, n = 214) = 21.048, $p < 0.001$

radiology residents were significantly more often found by radiology residents ($p = 0.035$). The number of discrepancies made by neuroradiologists is significantly higher in the morning and lower in the evening/ night ($p < 0.001$). The number of discrepancies made by radiologists with a neuro subspecialisation is significantly lower in the evening/ night ($p < 0.001$). The number of discrepancies made by radiology residents is significantly lower in the morning and significantly higher in the evening/ night ($p < 0.001$). No significant results were found on the other categorical variables. The *Fisher-Freeman-Halton exact test* could not be executed for the variable *setting*, due to insufficient memory.

The *Independent-Samples Median Test* for the variable *discrepancy discovery duration (days)* shows that the median of *radiology residents* (5.5 days) is significantly lower than the median of *neuroradiologists* (84.5 days), *radiologists with a neuro subspecialisation* (34.5 days) and *radiologists with another specialisation* (699.5 days) ($p < 0.001$).

3.2 Results part 2

Radiology residents assessed and reported, supervised or unsupervised, 28.4% (17,385) of all neuroradiological imaging studies between April 2017 and March 2019. This resulted in 88 discrepancies made by radiology residents and an overall discrepancy rate of 0.5%. The significant/ severe discrepancies show a discrepancy rate of 0.22% (39/17,385). Of all imaging assessed and reported by radiology residents, 60.2% were CT-scans, 10.3% were MRI-scans and 29.5% were X-rays. The discrepancy rate for the modalities is 0.57%, 0.73% and 0.29% respectively. The radiology residents assessed and reported imaging in four different settings, 62% (10,778) in an emergency setting, 18% (3,135) in an outpatient setting, 14.8% (2,573) in a general practitioner setting and 7.2% (899) in an inpatient setting. The discrepancy rate for the settings is 0.54%, 0.64%, 0.27% and 0.33% respectively.

Table 2 shows general descriptive analysis of the neuroradiological discrepancies made by radiology residents in total and divided over the categories of the dependent variable *supervision*. Most discrepancies occur among radiology residents without supervision (64.8%). From the 31 discrepancies made under supervision, radiology residents were supervised 17 times by a radiologist with a neuro subspecialisation. Most radiology residents were three years and five months in their training to become a radiologist, when they made a discrepancy. Neuroradiologist find and register most of the discrepancies made by radiology residents (47.7%). The clinical relevance of discrepancies was often significant/ severe (44,3%). The median of the discrepancy discovery duration is 5.5 days and 80.7% of the discrepancies made by radiology residents are perception/ observation discrepancies. Most discrepancies occur in the emergency setting (65.9%), in the afternoon (62.5%) and in CT-scans (68.2%). Discrepancies occur often in CT-scans, so a new variable *type of CT-scan* was created to investigate this finding in more detail. Of all 60 CT-scan discrepancies made by radiology residents, 35 discrepancies were made in brain CT-scans. Discrepancies of radiology residents are most often

Table 2

Descriptive statistics radiology residents (n = 88)

Radiology residents			**Dependent categorical variable *supervision***			
Independent categorical variables	**Total sample N**	**(%)**	Supervision		No supervision	
			31	(35.2%)	57	(64.8%)
Supervised by	-				-	
▪ Neuroradiologist			9	(10.2%)		
▪ Radiologists with a neuro subspecialisation			17	(19.3%)		
▪ Radiologists with another specialisation			5	(5.7%)		
▪ No supervision			57	(64.8%)		
Discrepancy author						
▪ Neuroradiologist	42	(47.7%)	13	(41.9%)	29	(50.9%)
▪ Radiologists with a neuro subspecialisation	20	(22.7%)	9	(29.0%)	11	(19.3%)
▪ Radiologists with another specialisation	3	(3.4%)	1	(3.2%)	2	(3.5%)
▪ Radiology resident	23	(26.1%)	8	(25.8%)	15	(26.3%)
Setting						
▪ Emergency department	58	(65.9%)	**15** **b**	(48.4%)	**43** **b**	(75.4%)
▪ Outpatient setting	20	(22.7%)	**14** **b**	(45.2%)	**6** **b**	(10.5%)
▪ General practitioner	7	(8.0%)	2	(6.5%)	5	(8.8%)
▪ Inpatient setting	3	(3.4%)	0	(0.0%)	3	(5.3%)
Imaging modality						
▪ CT-scan	60	(68.2%)	18	(58.1%)	42	(73.7%)
▪ MRI-scan	13	(14.8%)	**11** **a**	(35.5%)	**2** **a**	(3.5%)
▪ X-ray	15	(17.0%)	2	(6.5%)	13	(22.8%)
Type of CT-scan						
▪ CT brain	35	(39.8%)	10	(32.3%)	25	(43.9%)
▪ CT vertebral column	12	(13.6%)	**1** **c**	(3.2%)	**11** **c**	(19.3%)
▪ CT angiography	9	(10.2%)	4	(12.9%)	5	(8.8%)
▪ CT throat, nose, ears, head, neck	4	(4.5%)	3	(9.7%)	1	(1.8%)
▪ Other; MRI-scan, X-ray	28	(31.8%)	13	(41.9%)	15	(26.3%)
Discrepancy discovered by						
▪ Second assessment radiologist	34	(38.6%)	11	(35.5%)	23	(40.4%)
▪ New advanced research	25	(28.4%)	9	(29.0%)	16	(28.1%)
▪ Repeating research	25	(28.4%)	10	(32.3%)	15	(26.3%)
▪ Assessment of non-radiologist	4	(4.5%)	1	(3.2%)	3	(5.3%)
Time of assessing/ reporting imaging						
▪ Morning	18	(20.5%)	6	(19.4%)	12	(21.1%)
▪ Afternoon	55	(62.5%)	23	(74.2%)	32	(56.1%)
▪ Evening/ night	14	(15.9%)	2	(6.5%)	13	(22.8%)
*Discrepancy classification **						
▪ Cognition/interpretation	17	(19.3%)	9	(29.0%)	8	(14.0%)
▪ Perception/observation	71	(80.7%)	22	(71.0%)	49	(86.0%)
Clinical relevance						
▪ None/ low	49	(55.7%)	15	(48.4%)	34	(59.6%)
▪ Significant/ severe	39	(44.3%)	16	(51.6%)	23	(40.4%)
Independent continuous variables	Total sample Median	(SD)	Supervision		No supervision	
Discrepancy discovery duration (days)	5.50	(370.23)	7.0	(520.04)	3.0	(216.65)
▪ Range	1597		1597		968	
Training progress radiology residents (years; months)	3;5	(1;7)	**1;11** **d**	(1;5)	**4;2** **d**	(1;4)
▪ Range	5;7		5;0		5;4	

***** System related discrepancies were excluded, due to small category sizes (n=4)

a Chi-square for independence; χ^2 (2, n = 88) = 17.767, Cramer's V = 0.449, $p < 0.001$

b Fisher-Freeman-Halton exact test; (3, n = 88) = 13.381, Cramer's V= 0.407 Exact Sig. (2-sided) $p < 0.002$

c Fisher-Freeman-Halton exact test; (4, n = 88) = 9.085, Exact Sig. (2-sided) $p < 0.049$

d Independent-Samples Median Test; test statistic (1, n = 88) = 21.963, $p < 0.001$

discovered by a second assessment of a radiologist (38.6%). More descriptive statistics about supervised and non-supervised radiology residents are shown in table 2. No assumptions were violated during the *Chi-square for independence test* for the variables *imaging modality*, *time of assessing/ reporting imaging*, *discrepancy classification* and *clinical relevance*. Supervised radiology residents cause significantly more discrepancies in MRI-scans, than unsupervised radiology ($p < 0.001$). The variables *discrepancy author*, *setting*, *type of CT-scan* and *discrepancy discovered by* violated the assumption of *minimum expected cell frequency*. The *Fisher-Freeman-Halton exact test* showed for the variable *setting* that unsupervised radiology residents cause significantly more discrepancies in an emergency setting, than supervised radiology residents ($p = 0.002$). Besides, supervised radiology residents cause significantly more discrepancies in an outpatient setting, than unsupervised radiology residents ($p = 0.002$). It is also found that supervised radiology residents cause significantly less discrepancies in a CT-vertebral column, than unsupervised radiology residents ($p = 0.049$). No significant results were found on the other categorical variables.

The *Independent-Samples Median Test* showed that unsupervised radiology residents are significantly further in their study when they make discrepancies (four years and two months) compared to supervised radiology residents (one year and 11 months) ($p < 0.001$). No statistically significant result was found for the variable discrepancy discovery duration (days).

4 Discussion

This study analysed 214 neuroradiological discrepancies (0.35% discrepancy rate) and 88 neuroradiological discrepancies made by radiology residents (0.5% discrepancy rate). All discrepancies were discovered between 1 April 2017 and 5 March 2019. The discrepancy ratio is much lower for neuroradiologists (0.21%) than for radiologists with a neuro subspecialisation (0.35%) or radiologists with another specialisation (0.38%). This is not in line with literature, as it was not expected to see more accurate performances of specialised radiologists compared to unspecialised radiologists [5, 31]. There is no evidence-based explanation for this. However, it might be due to the high-quality fellow training in the Netherlands in which radiologists specialise themselves in a certain radiology area like neuroradiology. This finding shows that workings based on specialisation results in less discrepancies.

Radiology residents are responsible for 41.1% of all neuroradiological discrepancies, this is interesting since they only assessed and reported 28.4% of the neuroradiological imaging. The imaging rate of radiology residents (28.4%) is lower compared to all radiologists together (71.6%), but the aggregate discrepancy rate of radiology residents is higher (0.51%) compared to all radiologists together (0.29%). It is possible that this result is partly due to some kind of registration bias, in which discrepancies made by radiology residents are earlier discovered and sooner registered, compared to discrepancies made by neuroradiologist, radiologists with a neuro subspecialisation or radiologists with another specialisation. This will be discussed in more detail later in this chapter. Other explanations might be found in the supervision area. Unsupervised radiology residents made 64.8% of the discrepancies. The other 35.2% of the neuroradiological discrepancies were made by supervised radiology residents. Detailed investigation showed that discrepancies in an emergency setting were significantly more often made by unsupervised radiology residents, than supervised radiology residents. Radiology residents who assess imaging in an emergency setting need at least competence level four, where they are expected to be able to work independently [21]. Significantly more discrepancies were made unsupervised which draws the hypothesis that radiology residents work unsupervised to quickly and that competence level four is inappropriate in an emergency setting. Literature already stated that radiology residents in emergency settings need better supervision [41]. Besides, this study is in line with literature and confirms that radiology residents cause most discrepancies in an emergency setting (65.9%) [13, 33]. These findings urge to further investigate the competence levels of the *radiology training plan (CORONA)*, to test the hypothesis.

In the outpatient setting was found, that supervised radiology residents make significantly more discrepancies than unsupervised radiology residents. Radiologists with a neuro subspecialisation are most involved in supervised discrepancies made by radiology residents. These findings support the hypothesis that supervision needs improvement, but further investigation should test whether these findings are also visible within other radiology specialisations, like abdominal discrepancies, to confirm

the hypothesis with more certainty. This might open educational and improvement opportunities for the supervision of radiology residents.

For the *imaging modality* was found that the discrepancy rate of the X-ray (0.13%) is much lower compared to the discrepancy rate of the CT-scan (0.48%) and MRI-scan (0.48%). This might indicate that X-rays are less sensitive to discrepancies. It was also found that neuroradiologists and radiologists with a neuro subspecialisation cause significantly more MRI-scan discrepancies, which can be explained because they assess most imaging in this modality. Therefore, it is also logical that the number of CT-scan discrepancies is lower for neuroradiologists and radiologists with a neuro subspecialisation, since most CT-scans are assessed by radiology residents (54.2%). Focussing only on radiology residents, shows that the majority of the neuroradiological discrepancies made by radiology residents were in CT-scans (68.2%). This result could be expected, since radiology residents also assessed and reported most imaging in this modality (60.2%). However, hindsight bias might affect the registration of CT-scan discrepancies. It is unknown whether CT-scan discrepancies were about realistic or unrealistic expectations of prospective findings. It can also be that the probability to find certain diagnosis in prospect was overestimated, which would cause hindsight bias and unrealistic expectations towards radiologists [25, 42]. So, some CT-scan discrepancies might be based on unrealistic expectations and should not be classified as discrepancies. With these comments in mind, further research into CT-scan discrepancies made by radiology residents is desirable, to decrease the number of discrepancies in the future. The significantly lower number of supervised CT-vertebral column discrepancies might indicate proper supervision and more supervision might decrease the number of discrepancies further. The significantly higher number of supervised MRI-scan discrepancies is again in line with the hypothesis that supervision needs improvement. Future studies need to test this hypothesis, to identify educational opportunities for radiological departments towards the supervision of radiology residents.

This study cannot confirm that second-year radiology residents cause the most discrepancies as it was found in the study of Huntley et al. [10]. This study found that third-year radiology residents caused the most discrepancies. Besides, the descriptive analysis of this study showed that unsupervised radiology residents were significantly further in their training, then supervised radiology residents. The Dutch *radiology training plan (CORONA)* might prevent most discrepancies made by radiology residents in their first and second year of training, because the supervision is stricter in the beginning of the training due to competence level one and two [21]. Further investigation is necessary to confirm this hypothesis.

In order to minimize patients' harm caused by discrepancies, it is important to know who is best in finding discrepancies and how this can be done as soon as possible. It is unknown why discrepancies made by radiology residents were significantly less often found by neuroradiologists and why discrepancies made by radiologists with a neuro subspecialisation were significantly less often found by

radiology residents. Discrepancies made by radiology residents were also significantly more often found by radiology residents. There is no evidence-based explanation for this result, but it might be due to the radiology residents who help each other and therefore find mistakes they have made. Discrepancies made by radiology residents are found significantly sooner than discrepancies made by neuroradiologists, radiologists with a neuro subspecialisation and radiologists with another specialisation. This is probably due to the fact that 64.8% of the discrepancies made by radiology residents were made without supervision. There might be control on the work of unsupervised radiology residents that results in this quick discovery of discrepancies. The high median of the discovery duration of discrepancies made by radiologists with another specialisation cannot be explained. However, detailed investigation showed that 10 out of 14 discrepancies made by radiologists with another specialisation had no/low clinical relevance. Further investigation must discover how discrepancies made by radiologists with another specialisation can be discovered sooner, in order to prevent patient harm.

This study showed that the number of discrepancies made by neuroradiologist and radiologists with a neuro subspecialisation is significantly lower in the evening/ night. This can be explained, since radiologists work less often evening/ night shifts than day shifts. It can also be explained that radiology residents cause significantly more discrepancies in the evening/ night than in the morning, since radiology residents work many night shifts.

The majority of all neuroradiological discrepancies (77,6%) and all neuroradiological discrepancies made by radiology residents (80,7%) of this study are perceptual/ observational discrepancies, as was expected based on previous studies [1, 20, 27, 28].

The aggregate discrepancy rate for all neuroradiological discrepancies (0.35%) was much lower compared to the discrepancy rates found in other studies worldwide (4%) [1, 6, 7]. This is also the case for the aggregate discrepancy rate for radiology residents (0.5%) and the aggregate discrepancy rate for significant/ severe discrepancies made by radiology residents (0.22%), which were respectively between 2.0 - 2.7% and 1.7 - 8.6% in literature [1, 6-11]. The aggregate discrepancy rate of radiology residents (0.51%) is a bit higher compared to the aggregate discrepancy rate of radiologists with another specialisation (0.38%). This is not in line with a previous study, as it was expected to see no different performances [32]. The overall discrepancy rate of radiology residents in an emergency setting is much lower (0,54%) than described in literature (3.3%) [12]. However, the discrepancy rate for radiology residents is higher in an outpatient than in an emergency setting (0.64%; 0.54% respectively). So, even though the absolute number of discrepancies is higher in the emergency setting, the relative number is higher in the outpatient setting. For discrepancies made by radiology residents, both the emergency and outpatient setting show interesting results that need to be further investigated. There is no evidence-based explanation for the lower discrepancy rates in this study, compared to previous (American)

studies. However, it is unknown when discrepancies made by radiology residents are simply part of their learning process or whether it are 'real' discrepancies that were not expected to occur in the training phase of a specific radiology resident. Besides, there might be a registration bias. Radiologists might registrate discrepancies made by radiology residents easier than discrepancies made by other, sometimes even more specialised, radiologists. So, a difference in discrepancy registration might explain the difference between the discrepancy rates of this study and those in literature. However, there is no evidence for this assumption available.

To successfully use the educational opportunities for radiology departments identified by this study, it is important to focus on changeable factors. Besides, hypotheses made in this study must be tested in future studies to identify proper targets of future interventions. So, the supervision of radiology residents should be further investigated as well as the competence levels described in CORONA. Future studies should include more radiological specialisations, like abdominal discrepancies, to establish evidence-based conclusions. It is also discovered that working with specialised neuroradiologists is useful. This is not an educational opportunity, but it does confirm that the current working patterns is an effective barrier to prevent or at least decrease the number of neuroradiological discrepancies. This study has shown that analysing discrepancies of a certain time period is useful. It allows hospitals to discover possible patterns and future study and educational opportunities.

4.1 Limitations of this study

This study is related to some limitations. First, a 100% registration of discrepancies will never happen, due to several reasons like a lack of focus on discrepancy registration, no time to register discrepancies (emergency setting), the medical culture of autonomy and personal responsibility, unintended forgetting to register discrepancies and some discrepancies might never be found because not every imaging will be assessed twice [17, 41]. The number of discrepancies is probably underestimated and relatively low in this study compared to literature. Some patterns might therefore never be discovered, and further investigation is necessary to draw evidence-based conclusions. Besides, it is unknown if the results are representative for all hospitals in the Netherlands. Second, there is variance in the opinion of radiologists. This makes it impossible to define a true error, because that implies that someone knows what the truth is. Third, some discrepancies could not be prevented due to the inability of imaging modalities to show minimalistic deviations. Fourth, the clinical relevance of discrepancies is based on radiologists' expectations but cannot be stated with certainty. For example, missed lung cancer might not influence the definitive outcome because whether the disease is discovered now or in three months does not affect the chance of a cure. These limitations must be considered while interpreting the results.

5 Conclusion and recommendations

This study indicates that improvement and educational opportunities lie within the competence levels of the Dutch CORONA radiology training plan and the supervision, especially in an emergency and outpatient setting and in MRI-scans. It is recommended to increase the awareness of supervisors and radiology residents about the high discrepancy rates here, to stimulate earlier feedback, of which the effectiveness needs to be monitored. Further investigation is necessary to confirm the hypotheses made in this study and to provide recommendations to decrease the number of discrepancies in an outpatient and emergency setting and in MRI-scans.

Finally, radiology departments should work with specialised radiologists, like neuroradiologist. This decreases the number of discrepancies. Also, a compliment system besides the discrepancy registration system might stimulate radiologists in their daily work, because this will highlight the outstanding interpretations, instead of only the doubtful interpretations. This study identified not only possible educational opportunities for radiology residents, but also for radiological departments and national guidelines. A feedback loop to analyse neuroradiological discrepancies annually should be implemented in hospitals, to discover new patterns, improvement possibilities and educational opportunities to improve diagnostic care.

Reference list

[1] Bruno, M.A., E.A. Walker, and H.H. Abujudeh, *Understanding and confronting our mistakes: the epidemiology of error in radiology and strategies for error reduction.* Radiographics, 2015. **35**(6): p. 1668-1676.

[2] Wachter, R.M., *Why diagnostic errors don't get any respect—and what can be done about them.* Health Affairs, 2010. **29**(9): p. 1605-1610.

[3] Pinto, A. and L. Brunese, *Spectrum of diagnostic errors in radiology.* World journal of radiology, 2010. **2**(10): p. 377.

[4] Lee, C.S., et al., *Cognitive and system factors contributing to diagnostic errors in radiology.* American Journal of Roentgenology, 2013. **201**(3): p. 611-617.

[5] Jordan, M.J., J.B. Lightfoote, and J.E. Jordan, *Quality outcomes of reinterpretation of brain CT imaging studies by subspecialty experts in neuroradiology.* Journal of the National Medical Association, 2006. **98**(8): p. 1326.

[6] Brady, A.P., *Error and discrepancy in radiology: inevitable or avoidable?* Insights into imaging, 2017. **8**(1): p. 171-182.

[7] Waite, S., et al., *Interpretive error in radiology.* American Journal of Roentgenology, 2017. **208**(4): p. 739-749.

[8] Verdoorn, J.T., et al., *Increasing neuroradiology exam volumes on-call do not result in increased major discrepancies in primary reads performed by residents.* The open neuroimaging journal, 2014. **8**: p. 11.

[9] Weinberg, B.D., et al., *Radiology resident preliminary reporting in an independent call environment: multiyear assessment of volume, timeliness, and accuracy.* Journal of the American College of Radiology, 2015. **12**(1): p. 95-100.

[10] Huntley, J.H., et al., *Opportunities for targeted education: critical neuroradiologic findings missed or misinterpreted by residents and fellows.* American Journal of Roentgenology, 2015. **205**(6): p. 1155-1159.

[11] Stevens, K.J., et al., *Discordance rates between preliminary and final radiology reports on cross-sectional imaging studies at a level 1 trauma center.* Academic radiology, 2008. **15**(10): p. 1217-1226.

[12] Miyakoshi, A., et al., *Accuracy of preliminary interpretation of neurologic CT examinations by on-call radiology residents and assessment of patient outcomes at a level I trauma center.* Journal of the American College of Radiology, 2009. **6**(12): p. 864-870.

[13] Caranci, F., et al., *Errors in neuroradiology.* La radiologia medica, 2015. **120**(9): p. 795-801.

[14] Babiarz, L. and D. Yousem, *Quality control in neuroradiology: discrepancies in image interpretation among academic neuroradiologists.* American Journal of Neuroradiology, 2012. **33**(1): p. 37-42.

[15] Ginat, D.T., et al., *Identification of neuroradiology MRI protocol errors via a quality-driven categorization approach.* Journal of the American College of Radiology, 2016. **13**(5): p. 545-548.

[16] Johnson, B.A., *Avoiding diagnostic pitfalls in neuroimaging.* Applied Radiology-The Journal of Practical Medical Imaging and Management, 2016. **45**(3): p. 24-29.

[17] FitzGerald, R., *Radiological error: analysis, standard setting, targeted instruction and teamworking.* European radiology, 2005. **15**(8): p. 1760-1767.

[18] Schwarz, S.T., et al., *The 'Swallow Tail'Appearance of the Healthy Nigrosome–A New Accurate Test of Parkinson's Disease: A Case-Control and Retrospective Cross-Sectional MRI Study at 3T.* PloS one, 2014. **9**(4): p. e93814.

[19] Hsu, S.-H., et al., *Investigation of a method for generating synthetic CT models from MRI scans of the head and neck for radiation therapy.* Physics in Medicine & Biology, 2013. **58**(23): p. 8419.

[20] Brady, A., et al., *Discrepancy and error in radiology: concepts, causes and consequences.* The Ulster medical journal, 2012. **81**(1): p. 3.

[21] J.P.J. van Schaik, R.J. Bennink, *CORONA Opleidingsplan Radiologie.* https://www.radiologen.nl/system/files/bestanden/documenten/corona_opleidingsplan_radiologie.pdf , 2015 (accessed 27 June 2019).
[22] McDonald, R.J., et al., *The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload.* Academic radiology, 2015. **22**(9): p. 1191-1198.
[23] Krupinski, E.A., et al., *Long radiology workdays reduce detection and accommodation accuracy.* Journal of the American College of Radiology, 2010. **7**(9): p. 698-704.
[24] Donald, J.J. and S.A. Barnard, *Common patterns in 558 diagnostic radiology errors.* Journal of medical imaging and radiation oncology, 2012. **56**(2): p. 173-178.
[25] Itri, J.N., et al., *Fundamentals of Diagnostic Error in Imaging.* RadioGraphics, 2018. **38**(6): p. 1845-1865.
[26] Larson, D.B., et al., *Key concepts of patient safety in radiology.* Radiographics, 2015. **35**(6): p. 1677-1693.
[27] Pow, R.E., C. Mello-Thoms, and P. Brennan, *Evaluation of the effect of double reporting on test accuracy in screening and diagnostic imaging studies: a review of the evidence.* Journal of medical imaging and radiation oncology, 2016. **60**(3): p. 306-314.
[28] Degnan, A.J., et al., *Perceptual and Interpretive Error in Diagnostic Radiology—Causes and Potential Solutions.* Academic radiology, 2018.
[29] Kohn, L.T., J. Corrigan, and M.S. Donaldson, *To err is human: building a safer health system*. Vol. 6. 2000: National academy press Washington, DC.
[30] G. Guit, B.J. van Kelckhoven, W. de Monyé, I. Chesaru, Radiologen moeten missers vastleggen, in Medisch Contact. https://www.medischcontact.nl/nieuws/laatste-nieuws/artikel/radiologen-moeten-missers-vastleggen.htm, 2016 (accessed 25 April 2019).
[31] Jordan, Y.J., et al., *Quality outcomes of reinterpretation of brain CT studies by subspecialty experts in stroke imaging.* American Journal of Roentgenology, 2012. **199**(6): p. 1365-1370.
[32] Branstetter IV, B.F., et al., *Preliminary reports in the emergency department: is a subspecialist radiologist more accurate than a radiology resident?* Academic radiology, 2007. **14**(2): p. 201-206.
[33] Yaniv, G., et al., *Common sites and etiologies of residents' misinterpretation of head CT scans in the emergency department of a level I trauma center.* Isr Med Assoc J, 2013. **15**(5): p. 221.
[34] Itri, J.N. and S.H. Patel, *Heuristics and cognitive error in medical imaging.* American Journal of Roentgenology, 2018. **210**(5): p. 1097-1105.
[35] Kundel, H.L., C.F. Nodine, and D. Carmody, *Visual scanning, pattern recognition and decision-making in pulmonary nodule detection.* Investigative radiology, 1978. **13**(3): p. 175-181.
[36] Cooper, V.F., et al., *Radiology resident interpretations of on-call imaging studies: the incidence of major discrepancies.* Academic radiology, 2008. **15**(9): p. 1198-1204.
[37] Borgstede, J.P., et al., *RADPEER quality assurance program: a multifacility study of interpretive disagreement rates.* Journal of the American College of Radiology, 2004. **1**(1): p. 59-65.
[38] Ruma, J., et al., *Cross-sectional examination interpretation discrepancies between on-call diagnostic radiology residents and subspecialty faculty radiologists: analysis by imaging modality and subspecialty.* Journal of the American College of Radiology, 2011. **8**(6): p. 409-414.
[39] Hanna, T.N., et al., *Effect of shift, schedule, and volume on interpretive accuracy: a retrospective analysis of 2.9 million radiologic examinations.* Radiology, 2017. **287**(1): p. 205-212.
[40] Pallant, J., *SPSS Survival Manual. A step by step guide to data analysis using IBM SPSS.* 2016.
[41] Guly, H., *Diagnostic errors in an accident and emergency department.* Emergency Medicine Journal, 2001. **18**(4): p. 263-269.
[42] Gunderman, R.B., *Biases in radiologic reasoning.* American Journal of Roentgenology, 2009. **192**(3): p. 561-564.

Appendix 1

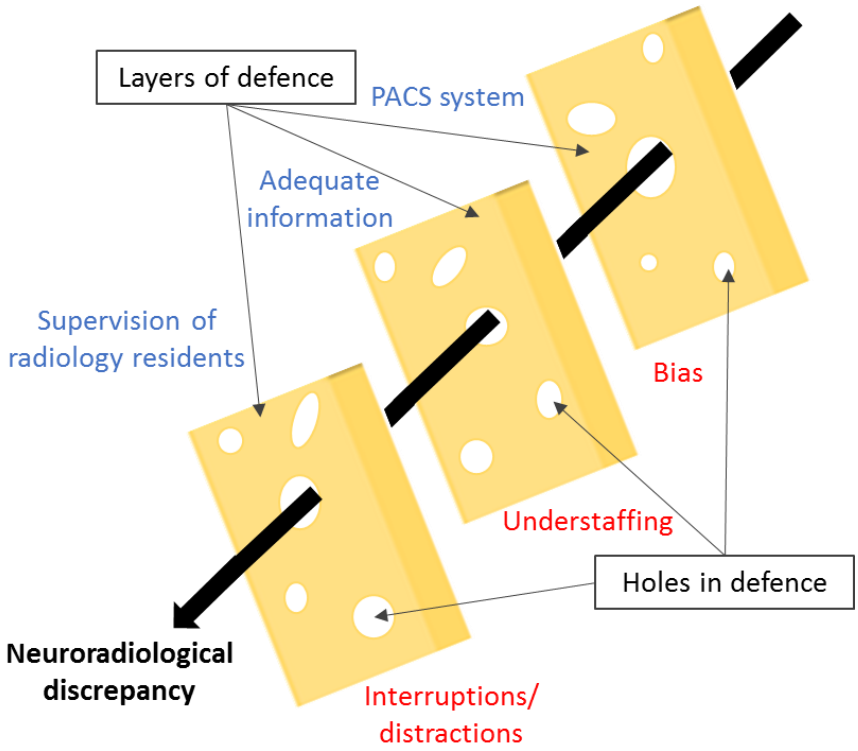

Figure 1: Swiss cheese model neuroradiological discrepancies. Adjusted from Itri et al. [25] and Larson et al. [26]

Discrepancies in neuroradiology made by radiology residents: educational opportunities

C.M.H. Degger (c.m.h.degger@student.utwente.nl)
A.T.M. Bellos-Grob (a.t.m.bellos-grob@utwente.nl)
S. Siesling (s.siesling@iknl.nl)
O. Vijlbrief (o.vijlbrief@zgt.nl)

Abstract

Purpose: Diagnostic errors often cause patient harm, due to delayed, incorrect or missed diagnoses. Radiologists' contribution is probably substantial, because diagnoses are often based on medical imaging. *Discrepancy* is a proper term to use, for supposed errors or observer variation in radiology. This study aims to analyse neuroradiological (head and neck) discrepancies made by supervised or unsupervised radiology residents, to identify patterns for improvement and educational opportunities to reduce future discrepancies.

Methods: Out of 17,385 neuroradiological imaging studies assessed and reported by radiology residents, 88 neuroradiological discrepancies were found and analysed. The Dutch ZiekenhuisGroep Twente gathered data for this retrospective quantitative descriptive study, between April 2017 and March 2019. Radiology residents can follow the five-year training program in this hospital. Subgroup analysis was performed, divided by supervision or not, including variables like *setting* and *imaging modality*. The *Chi-square for independence, Fisher-Freeman-Halton exact test, Independent-Samples Median Test* and *logistic regression analysis* were used.

Results: Discrepancies occur without supervision (64,8%) but also with supervision (35.2%). Discrepancy rates are high in MRI-scans (13; 0.73%), the emergency (58; 0.54%) and outpatient setting (20; 0.64%). If radiology residents are one year further in their training, it becomes 0.5 times less likely that they are supervised when causing discrepancies. The 0.5% (88/17,385) aggregate discrepancy rate for radiology residents, is lower compared to literature (2.0–2.7%).

Conclusions: Future research must confirm whether the supervision and competence levels within the Dutch *radiology training plan (CORONA)* offer improvement and educational opportunities, for radiological departments and national guidelines.

Keywords

Radiology residents • neuroradiological discrepancies • educational opportunities

1. Introduction

Diagnostic errors often cause patient harm, due to delayed, incorrect or missed diagnoses. It is likely that radiologists' contribution is substantial, because diagnoses are often based on medical imaging [1, 2]. There is a difference between diagnostic error and observer variation. True 'errors' occur whenever the interpretation of radiologists' peers differs substantially from the original report and the correct interpretation is indisputable [1, 3, 4]. Variance is no real 'error', it occurs whenever the correct interpretation is subject to a legitimate difference of opinion [4, 5]. *Discrepancy* is a proper term to use, whenever talking about supposed errors or observer variation in radiology [3].

Dutch radiology residents follow a five years training program, to become a general radiologist. Thereafter, they can become a fellow radiologist to specialise themselves in two years in for example

neuroradiology [6]. One billion radiological studies are performed worldwide annually with a 4% error rate, resulting in forty million radiological errors [1, 3, 4]. Zooming in on neuroradiological errors made by radiology residents shows aggregate error rates from 2.0 to 2.7% and significant error rates from 1.7% to 8.6% [7-10]. Radiology residents cause more discrepancies in an emergency setting, with an aggregate discrepancy rate of 3.3% [11-13]. Most significant discrepancies were made by second-year radiology residents [9]. Rules regarding the supervision of radiology residents are independently and individually determined by each country. The Netherlands use competence levels, stated in the Dutch *radiology training plan (CORONA).* Radiologists determine the competence level of a radiology resident. The competence levels start with radiology residents are not allowed to assess imaging independently (competence level 1) or not without supervision (competence level 2). Thereafter, radiology residents can assess imaging independently and receive limited supervision of which the initiative lies with radiology residents (competence level 3), radiology residents work independently but can ask for help (competence level 4) or radiology residents supervise others (competence level 5) [6]. *Neuroradiology* is a common term in literature and will therefore be used to refer to the specialisation neuro- (head and neck) radiology [12, 14-16].

"To err is human", but we should prevent them as much as possible [17]. The Dutch healthcare inspectorate advised in 2016 to implement a discrepancy registration system [18]. This will provide insight in the current performances of radiology departments and enables them to determine minimum quality standards, to improve diagnostic care [3, 18, 19]. Registered discrepancies should be analysed and discussed to reach quality improvement. The aim of this study is to analyse neuroradiological discrepancies made by supervised or unsupervised radiology residents, to identify patterns for improvement and educational opportunities to reduce the number of future discrepancies.

2. Material and methods

2.1 Study design

This retrospective quantitative descriptive study analysed 88 neuroradiological discrepancies made by radiology residents in *ZiekenhuisGroep Twente* (ZGT). ZGT is a regional hospital located in the cities Hengelo and Almelo in the East of the Netherlands and facilitates prevention, diagnosis, treatment and nursing to stimulate approximately 250,000 patient's health annually. Radiology residents are trained in ZGT.

2.2 The dataset

ZGT implemented a discrepancy registration system in April 2017 and discrepancies in this study were discovered between 1 April 2017 and 5 March 2019. Included neuroradiological discrepancies were based on imaging from 1 January 2012 till 5 March 2019. The starting point of 2012 was chosen, because radiologists work since then based on their (neuro-)specialisation to increase expertise. The radiology department works with the *JiveX Picture Archiving and Communication System* of *Alphatron Medical,* to assess and report imaging and to detect and registrate discrepancies.

Data was made anonymously, therefore patients or reporting radiologist could not be identified and informed consent was not necessary. The focus of interest, supervision, was coded as supervised or unsupervised, which was noted per imaging. Supervised radiology residents are corrected whenever their interpretation is incorrect, according to their supervisor during that shift. Discrepancies among supervised radiology residents

indicated that both the radiology resident and the supervisor interpreted the imaging 'incorrect'. Unsupervised radiology residents are responsible to ask for help if they doubt their own judgement capabilities on certain imaging. [6] The following independent categorical variables *discrepancy author, supervised by, setting, imaging modality, discrepancy discovered by, time of assessing/ reporting imaging, discrepancy classification* and *clinical relevance* and the independent continuous variables *training progress radiology residents (years; months)* and *discrepancy discovery duration (days)* were included in this study.

The variable *discrepancy author* describes whether discrepancies were discovered by neuroradiologists, radiologists with a neuro subspecialisation, radiologists with another specialisation or radiology residents. The recoding was based on the scheduling and number of neuroradiological imaging assessed and reported by a specific radiologist, between April 2017 and March 2019. The official Dutch classification of registered neuroradiologist and subspecialist were not used in this study, because training qualifications were different during the education period of older radiologists and it would underestimate their qualities nowadays. Recoding based on the output of radiologists is therefore more representative. Neuroradiologists assessed and reported ≥10% of all 61,246 neuroradiological imaging, radiologists with a neuro subspecialisation ≥2% and radiologists with another specialisation <2%. The fourth category included radiology residents, who registered discrepancies made by other radiology residents. The variable *supervised by* used the categories neuroradiologist, radiologists with a neuro subspecialisation and radiologists with another specialisation as well, only the fourth category included for this variable all unsupervised discrepancies. The variable *setting* describes whether patients were referred to the radiology department from an *emergency setting, outpatient setting, inpatient setting* or by a *general practitioner.* The variable *imaging modality* was recoded based on the main neuroradiological modalities in ZGT; *CT-scan, MRI-scan* and *X-ray.* If the results encourage to analyse in more detail, it is possible to create new variables based on for example the focus of a CT-scan. The variable *discrepancy discovered by* showed how a discrepancy was found; *a second assessment of a radiologist, repeating the research, new advanced research* or by the *assessment of a non-radiologist.* The variable *time of assessing/ reporting imaging* showed whether a discrepancy was made in the *morning (6:00–12:00 h), afternoon (12:00–18:00 h)* or *evening/ night (18:00–6:00 h).* The discrepancy classification was determined by radiologists prior to this study; *perception/observation, cognition/ interpretation* or *system related* discrepancies. The *clinical relevance* was based on the worldwide used RADPEER® scoring system and scored by radiologists of ZGT prior to this study. The four categories of the RADPEER® scoring system, based on the accuracy of a radiologists' interpretation, [5, 20-23] were recoded into two categories to create proper category sizes; *no/ low clinical relevance* and *significant/ severe clinical relevance*. No/low clinical relevance implied no clinical relevance, or clinical relevance but without a necessary treatment adjustment. Significant/ severe clinical relevance implied a necessary treatment adjustment or a necessary treatment adjustment and a different definitive outcome for patients.

2.3 Data analysis

The neuroradiological discrepancies made by radiology residents were studied, to investigate the radiology resident discrepancies and the role of supervision more detailed. The ratio between the number of discrepancies and the total number of imaging on the three modalities and the four settings will also be shown, to represent the current seriousness of discrepancies.

This study performed subgroup analysis, based on the dependent variable *supervision*. The *Chi-square test for independence* was used to perform subgroup analysis for categorical variables. The *Fisher's exact test* (2x2

table) and the *Fisher-Freeman-Halton exact test* (tables larger than 2x2) were used, if the assumption *minimum expected cell frequency* was violated. The *phi coefficient* (2x2 table) and *Cramer's V* (tables larger than 2x2) were used to test the effect size of significant results. The *Independent-Samples Median Test* was used to perform subgroup analysis for continuous variables. Thereafter, *univariate logistic regression analyses* were performed for the dependent variable *supervision*, including the *Cox and Snell R Square* and the *Nagelkerke R Square*. The reference categories for the categorical variables were determined based on the category with the most cases. The variables *setting*, *imaging modality*, *time of assessing/ reporting imaging* and *training progress radiology residents (years; months)* were included in the logistic regressions, because it is expected that these variables influence the supervision. Two sided *p* values were used and $p < 0.05$ indicated statistically significant results [24]. Recoding and analysis was done with Excel 2016 and *Statistical Package for Social Sciences (SPSS)* version 25. Improvement recommendations were made, if significant results urged to intervene.

3. Results

3.1 Results descriptive statistics

In total there are 61,246 neuroradiological imaging studies performed between April 2017 and March 2019. Radiology residents assessed 28.4% (17,385) of those images, either supervised or unsupervised. This resulted in 88 discrepancies made by radiology residents and a 0.5% aggregate discrepancy rate. The significant/ severe discrepancies show a discrepancy rate of 0.22% (39). Of all imaging assessed and reported by radiology residents, 60.2% were CT-scans, 10.3% were MRI-scans and 29.5% were X-rays. The discrepancy rate for the modalities is 0.57%, 0.73% and 0.29% respectively. The radiology residents assessed and reported imaging in four different settings, 62% (10,778) in an emergency setting, 18% (3,135) in an outpatient setting, 14.8% (2,573) in a general practitioner setting and 7.2% (899) in an inpatient setting. The discrepancy rate for the settings is 0.54%, 0.64%, 0.27% and 0.33% respectively. Table 1 shows the descriptive analysis of the neuroradiological discrepancies made by radiology residents and divided over the categories of the dependent variable *supervision*. This analysis included *Chi-square for independence tests*, *Fisher-Freeman-Halton exact tests* and an *Independent-Samples Median Test*.

Discrepancies often occurred without supervision (64.8%). Supervised radiology residents cause significantly more discrepancies in MRI-scans, than unsupervised radiology ($p < 0.001$). It is also found that supervised radiology residents cause significantly less discrepancies in a CT-vertebral column, than unsupervised radiology residents ($p = 0.049$). For the variable *setting* was found that unsupervised radiology residents cause significantly more discrepancies in an emergency setting, than supervised radiology residents ($p = 0.002$). Besides, supervised radiology residents cause significantly more discrepancies in an outpatient setting, than unsupervised radiology residents ($p = 0.002$). When causing discrepancies, unsupervised radiology residents are significantly further in their study (four years and two months) than supervised radiology residents (one year and 11 months) ($p < 0.001$). No significant results were found on the other variables (table 1).

3.2 Results univariate logistic regression (L.R.)

The outcomes of three univariate logistic regressions are presented in table 2. All logistic regression models were statistically significant ($p < 0.01$) and were therefore able to distinguish between discrepancies made by supervised and unsupervised radiology residents. Logistic regression three, including the variable *training progress radiology*

Table 1
Descriptive statistics radiology residents (n = 88)

Radiology residents			**Dependent categorical variable *supervision***			
Independent categorical variables	**Total sample N**	**(%)**	Supervision		No supervision	
			31	(35.2%)	57	(64.8%)
Supervised by	-				-	
▪ Neuroradiologist			9	(10.2%)		
▪ Radiologists with a neuro subspecialisation			17	(19.3%)		
▪ Radiologists with another specialisation			5	(5.7%)		
▪ No supervision			57	(64.8%)		
Discrepancy author						
▪ Neuroradiologist	42	(47.7%)	13	(41.9%)	29	(50.9%)
▪ Radiologists with a neuro subspecialisation	20	(22.7%)	9	(29.0%)	11	(19.3%)
▪ Radiologists with another specialisation	3	(3.4%)	1	(3.2%)	2	(3.5%)
▪ Radiology resident	23	(26.1%)	8	(25.8%)	15	(26.3%)
Setting						
▪ Emergency department	58	(65.9%)	**15** **b**	(48.4%)	**43** **b**	(75.4%)
▪ Outpatient setting	20	(22.7%)	**14** **b**	(45.2%)	**6** **b**	(10.5%)
▪ General practitioner	7	(8.0%)	2	(6.5%)	5	(8.8%)
▪ Inpatient setting	3	(3.4%)	0	(0.0%)	3	(5.3%)
Imaging modality						
▪ CT-scan	60	(68.2%)	18	(58.1%)	42	(73.7%)
▪ MRI-scan	13	(14.8%)	**11** **a**	(35.5%)	**2** **a**	(3.5%)
▪ X-ray	15	(17.0%)	2	(6.5%)	13	(22.8%)
Type of CT-scan						
▪ CT brain	35	(39.8%)	10	(32.3%)	25	(43.9%)
▪ CT vertebral column	12	(13.6%)	**1** **c**	(3.2%)	**11** **c**	(19.3%)
▪ CT angiography	9	(10.2%)	4	(12.9%)	5	(8.8%)
▪ CT throat, nose, ears, head, neck	4	(4.5%)	3	(9.7%)	1	(1.8%)
▪ Other; MRI-scan, X-ray	28	(31.8%)	13	(41.9%)	15	(26.3%)
Discrepancy discovered by						
▪ Second assessment radiologist	34	(38.6%)	11	(35.5%)	23	(40.4%)
▪ New advanced research	25	(28.4%)	9	(29.0%)	16	(28.1%)
▪ Repeating research	25	(28.4%)	10	(32.3%)	15	(26.3%)
▪ Assessment of non-radiologist	4	(4.5%)	1	(3.2%)	3	(5.3%)
Time of assessing/ reporting imaging						
▪ Morning	18	(20.5%)	6	(19.4%)	12	(21.1%)
▪ Afternoon	55	(62.5%)	23	(74.2%)	32	(56.1%)
▪ Evening/ night	14	(15.9%)	2	(6.5%)	13	(22.8%)
*Discrepancy classification **						
▪ Cognition/interpretation	17	(19.3%)	9	(29.0%)	8	(14.0%)
▪ Perception/observation	71	(80.7%)	22	(71.0%)	49	(86.0%)
Clinical relevance						
▪ None/ low	49	(55.7%)	15	(48.4%)	34	(59.6%)
▪ Significant/ severe	39	(44.3%)	16	(51.6%)	23	(40.4%)
Independent continuous variables	Total sample Median	(SD)	Supervision		No supervision	
Discrepancy discovery duration (days)	5.50	(370.23)	7.0	(520.04)	3.0	(216.65)
▪ Range	1597		1597		968	
Training progress radiology residents (years; months)	3;5	(1;7)	**1;11** **d**	(1;5)	**4;2** **d**	(1;4)
▪ Range	5;7		5;0		5;4	

***** System related discrepancies were excluded, due to small category sizes (n=4)

a Chi-square for independence; χ^2 (2, n = 88) = 17.767, Cramer's V = 0.449, $p < 0.001$

b Fisher-Freeman-Halton exact test; (3, n = 88) = 13.381, Cramer's V= 0.407 Exact Sig. (2-sided) $p = 0.002$

c Fisher-Freeman-Halton exact test; (4, n = 88) = 9.085, Exact Sig. (2-sided) $p = 0.049$

d Independent-Samples Median Test; test statistic (1, n = 88) = 21.963, $p < 0.001$

Table 2
Univariate logistic regression radiology residents' supervision (n = 88)

Radiology residents	**Dependent categorical variable *supervision***			
			OR 95% C.I.	
Independent variables	OR	Sig.	Lower	Upper
Logistic Regression 1				
Setting				
▪ Emergency department	Ref.			
▪ Outpatient setting	**6.7**	0.001	2.177	20.552
▪ General practitioner	1.1	0.878	0.201	6.546
▪ Inpatient setting	-	-	-	
Logistic Regression 2				
Imaging modality				
▪ CT-scan	Ref.			
▪ MRI-scan	**12.8**	0.002	2.579	63.859
▪ X-ray	0.4	0.206	0.073	1.756
Logistic Regression 3				
Training progress radiology residents (years; months)	**0.5**	0.001	0.320	0.665

residents, explains the most variance in supervision (*Cox & Snell R Square* = 22.8%; *Nagelkerke R Square* = 31.4%). L.R. 1 explains between 18.5% (Cox & Snell R Square) and 25.4% (Nagelkerke R Square) of the variance in supervision. L.R. 2 explains between 15.7% (Cox & Snell R Square) and 21.7% (Nagelkerke R Square) of the variance in supervision. The category MRI-scan within the variable *imaging modality* (L.R. 2), was the strongest predictor of *supervision*, with a significant odds ratio of 12.8 (p = 0.002). The category outpatient setting within the variable *setting* (L.R. 1), showed a statistically significant odds ratio of 6.7 (p = 0.001). The category *inpatient setting* did not have supervised discrepancies and could therefore not be included in the logistic regression. The variable *training progress radiology residents* (L.R. 3), showed a statistically significant odds ratio of 0.5 (p < 0.001). No other significant results were found.

4. Discussion

This study analysed 88 (0.5% discrepancy rate) neuroradiological discrepancies made by radiology residents, discovered between April 2017 and March 2019. Unsupervised radiology residents made 64.8% of the discrepancies. The other 35.2% of the neuroradiological discrepancies were made by supervised radiology residents. Detailed investigation showed that discrepancies in an emergency setting were significantly more often made by unsupervised radiology residents, than supervised radiology residents. Radiology residents who assess imaging in an emergency setting need at least competence level four, where they are expected to be able to work independently [6]. Significantly more discrepancies were made unsupervised, which draws the hypothesis that radiology residents work unsupervised to quickly and that competence level four is inappropriate in an emergency setting. Literature already stated that radiology residents in emergency settings need better supervision [25]. Besides, this study confirms in line with literature that radiology residents cause most discrepancies in an emergency setting (65.9%) [12, 13]. These findings urge to investigate the competence levels described in CORONA, to test the hypothesis.

In the outpatient setting was found, that supervised radiology residents make significantly more discrepancies than unsupervised radiology residents. Besides, if radiology residents assess imaging in an outpatient

setting, it is 6.7 times more likely that supervised radiology residents cause discrepancies, compared to supervised radiology residents who assess imaging in an emergency setting. It was impossible to control for other factors that might affect the supervision, because there were not enough cases in this study. Radiologists with a neuro subspecialisation are most involved in supervised discrepancies made by radiology residents. These findings draw the hypothesis that the supervision needs improvement. Further investigation should test whether these findings are also visible within other radiology specialisations to confirm the hypothesis with more certainty. This might open educational opportunities for the supervision of radiology residents.

The majority of the neuroradiological discrepancies made by radiology residents were in CT-scans (68.2%). This result could be expected, since radiology residents also assessed and reported most imaging in this modality (60.2%). However, hindsight bias might affect the registration of CT-scan discrepancies. It is unknown whether CT-scan discrepancies were about realistic or unrealistic expectations of prospective findings. It can also be that the probability to find certain diagnosis in prospect was overestimated, which would cause hindsight bias and unrealistic expectations of radiologists [26, 27]. So, some CT-scan discrepancies might be based on unrealistic expectations and should not be classified as discrepancies. With these comments in mind, further research into CT-scan discrepancies made by radiology residents is desirable to decrease the number of discrepancies in the future. The significantly lower number of supervised CT-vertebral column discrepancies might indicate proper supervision and more supervision might decrease discrepancies further. The significantly higher number of supervised MRI-scan discrepancies is again in line with the hypothesis that supervision needs improvement. Future studies need to test this hypothesis, to identify educational opportunities for radiological departments. The discrepancy rate of X-rays is lower (0.29%) compared to CT-scans (0.57%) and MRI-scans (0.73%), this might indicate that X-rays are less sensitive to discrepancies. If radiology residents assess MRI-scans, it is 12.8 times more likely that radiology residents are supervised, compared to radiology residents who assess CT-scans. Again, it was impossible to control for other factors that might affect the supervision, because there were not enough cases in this study.

This study cannot confirm that second-year radiology residents cause the most discrepancies as it was found in the study of Huntley et al. [9]. This study found that third-year radiology residents caused the most discrepancies. Besides, if radiology residents are one year further in their training, it becomes 0.5 times less likely that they are supervised when causing discrepancies. However, it was impossible to control for other factors that might affect the supervision, because there were not enough cases in this study. The Dutch *radiology training plan (CORONA)* might prevent most discrepancies made by radiology residents in their first and second year of training, because the supervision is stricter in the beginning of the training due to competence level one and two [6]. Further investigation is necessary to confirm this hypothesis.

The discrepancy rates for radiology residents overall (0.5%) and for significant/ severe discrepancies (0.22%) are both much lower compared to literature, respectively between 2.0–2.7% and 1.7–8.6% [1, 3, 4, 7-10]. The aggregate discrepancy rate of radiology residents in an emergency setting is also much lower (0,54%) compared to literature (3.3%) [11]. More interesting, the outpatient setting shows a higher discrepancy rate than the emergency setting (0.64%; 0.54% respectively). This indicates that the relative number of discrepancies is higher in the outpatient setting. As mentioned before, both the emergency and outpatient setting need to be further investigated within radiology departments. There is no evidence-based explanation for the lower discrepancy rates in this study, compared to previous (American) studies. However, it is unknown when discrepancies made by

radiology residents are part of their learning process or whether it are 'real' discrepancies that were not expected to occur in the training phase of a specific radiology resident. So, a difference in discrepancy registration might explain the difference between the discrepancy rates of this study and those in literature. Besides, there might be a registration bias. Radiologists might registrate discrepancies made by radiology residents easier than discrepancies made by other, sometimes even more specialised, radiologists. Again, there is no evidence for this assumption available.

To successfully use the educational opportunities for radiology departments identified by this study, the focus must be on changeable factors. Besides, hypotheses made in this study must be tested in future studies to identify proper targets of future interventions. So, the supervision of radiology residents should be further investigated as well as the competence levels described in CORONA. Future studies should include more radiological specialisations, like abdominal discrepancies, to establish evidence-based conclusions. This study has shown that analysing discrepancies of a certain time period is useful, to discover patterns and future study and educational opportunities.

This study is related to some limitations. First, a 100% registration of discrepancies will never happen, due to several reasons like a lack of focus on discrepancy registration, no time to register discrepancies (emergency setting), the medical culture of autonomy and personal responsibility and some discrepancies might never be found [5, 25]. The number of discrepancies is probably underestimated and relatively low in this study compared to literature. Some patterns might therefore never be discovered, and further investigation is necessary to draw evidence-based conclusions. Besides, it is unknown if the results are representative for all Dutch hospitals. Second, there is variance in the opinion of radiologists. This makes it impossible to define a true error, because that implies that someone knows what the truth is. Third, some discrepancies could not be prevented due to the inability of modalities to show minimalistic deviations. Fourth, the clinical relevance of discrepancies is based on radiologists' expectations but cannot be stated with certainty. These limitations must be considered while interpreting the results.

5. Conclusion and recommendations

This study indicates that improvement and educational opportunities lie within the competence levels of the Dutch CORONA radiology training plan and the supervision, especially in an emergency and outpatient setting and in MRI-scans. It is recommended to increase the awareness of supervisors and radiology residents about the high discrepancy rates in an emergency and outpatient setting and in MRI-scans, to stimulate earlier feedback, of which the effectiveness needs to be monitored. Further investigation is necessary to confirm the hypotheses made in this study and to provide recommendations to decrease the number of discrepancies in an outpatient and emergency setting and in MRI-scans.

This study identified not only possible educational opportunities for radiology residents, but also for radiological departments and national guidelines. A feedback loop to analyse neuroradiological discrepancies annually should be implemented in hospitals, to improve diagnostic care by discovering new patterns, improvement possibilities and educational opportunities.

References

[1] M.A. Bruno, E.A. Walker, H.H. Abujudeh, Understanding and confronting our mistakes: the epidemiology of error in radiology and strategies for error reduction. Radiographics. 35(6) (2015) 1668-1676. https://doi.org/10.1148/rg.2015150023.

[2] R.M. Wachter, Why diagnostic errors don't get any respect—and what can be done about them. Health Affairs. 29(9) (2010) 1605-1610. https://doi.org/10.1377/hlthaff.2009.0513.

[3] A.P. Brady, Error and discrepancy in radiology: inevitable or avoidable? Insights into imaging. 8(1) (2017) 171-182. https://doi.org/10.1007/s13244-016-0534-1.

[4] S. Waite, J. Scott, B. Gale, T. Fuchs, S. Kolla, D. Reede, Interpretive error in radiology. American Journal of Roentgenology. 208(4) (2017) 739-749. https://doi.org/10.2214/AJR.16.16963.

[5] FitzGerald, R., Radiological error: analysis, standard setting, targeted instruction and teamworking. European radiology. 15(8) (2005) 1760-1767. https://doi.org/10.1007/s00330-005-2662-8.

[6] J.P.J. van Schaik, R.J. Bennink, *CORONA Opleidingsplan Radiologie*. https://www.radiologen.nl/system/files/bestanden/documenten/corona_opleidingsplan_radiologie.pdf , 2015 (accessed 27 June 2019).

[7] J.T. Verdoorn, C.H. Hunt, M.T. Luetmer, C.P. Wood, L.J. Eckel, K.M. Schwartz, F.E. Diehn, D.F. Kallmes, Increasing neuroradiology exam volumes on-call do not result in increased major discrepancies in primary reads performed by residents. The open neuroimaging journal. (2014) 11. https://doi.org/10.2174/1874440001408010011.

[8] B.D. Weinberg, M.D. Richter, J.G. Champine, M.C. Morriss, T. Browning, Radiology resident preliminary reporting in an independent call environment: multiyear assessment of volume, timeliness, and accuracy, Journal of the American College of Radiology. 12(1) (2015) 95-100. https://doi.org/10.1016/j.jacr.2014.08.005.

[9] J.H. Huntley, M. Carone, D.M. Yousem, L.S. Babiarz, Opportunities for targeted education: critical neuroradiologic findings missed or misinterpreted by residents and fellows. American Journal of Roentgenology. 205(6) (2015) 1155-1159. https://doi.org/10.2214/AJR.15.14905.

[10] K.J. Stevens, K.L. Griffiths, J. Rosenberg, S. Mahadevan, L.M. Zatz, A.N. Leung, Discordance rates between preliminary and final radiology reports on cross-sectional imaging studies at a level 1 trauma center. Academic radiology. 15(10) (2008) 1217-1226. https://doi.org/10.1016/j.acra.2008.03.017.

[11] A. Miyakoshi, Q.T. Nguyen, W.A. Cohen, L.B. Talner, Y. Anzai, Accuracy of preliminary interpretation of neurologic CT examinations by on-call radiology residents and assessment of patient outcomes at a level I trauma center. Journal of the American College of Radiology. 6(12) (2009) 864-870. https://doi.org/10.1016/j.jacr.2009.07.021.

[12] F. Caranci, E. Tedeschi, G. Leone, A. Reginelli, G. Gatta, A. Pinto, E. Squillaci, F. Briganti, L. Brunese, Errors in neuroradiology. La radiologia medica. 120(9) (2015) 795-801. https://doi.org/10.1007/s11547-015-0564-7.

[13] G. Yaniv, O. Mozes, G. Greenberg, M. Bakon, C. Hoffmann, *Common sites and etiologies of residents' misinterpretation of head CT scans in the emergency department of a level I trauma center*. Isr Med Assoc J. 15(5) (2013) 221.

[14] L.S. Babiarz, D.M. Yousem, Quality control in neuroradiology: discrepancies in image interpretation among academic neuroradiologists. American Journal of Neuroradiology. 33(1) (2012) 37-42. https://doi.org/10.3174/ajnr.A2704.

[15] D.T. Ginat, P. Uppuluri, G. Christoforidis, G. Katzman, S.K. Lee, Identification of neuroradiology MRI protocol errors via a quality-driven categorization approach. Journal of the American College of Radiology. 13(5) (2016) 545-548. https://doi.org/10.1016/j.jacr.2015.08.027.

[16] B.A. Johnson, Avoiding diagnostic pitfalls in neuroimaging. Applied Radiology-The Journal of Practical Medical Imaging and Management. 45(3) (2016) 24-29.

[17] L.T. Kohn, J. Corrigan, M.S. Donaldson, To err is human: building a safer health system. Sixth ed., National academy press Washington, DC, 2000.

[18] G. Guit, B.J. van Kelckhoven, W. de Monyé, I. Chesaru, Radiologen moeten missers vastleggen, in Medisch Contact. https://www.medischcontact.nl/nieuws/laatste-nieuws/artikel/radiologen-moeten-missers-vastleggen.htm, 2016 (accessed 25 April 2019).

[19] A. Pinto, L. Brunese, Spectrum of diagnostic errors in radiology. World journal of radiology. 2(10) (2010) 377. https://doi.org/10.4329/wjr.v2.i10.377.

[20] V.F. Cooper, L.A. Goodhartz, A.A. Nemcekk Jr, R.K. Ryu, Radiology resident interpretations of on-call imaging studies: the incidence of major discrepancies. Academic radiology. 15(9) (2008) 1198-1204. https://doi.org/10.1016/j.acra.2008.02.011.

[21] J.P. Borgstede, R.S. Lewis, M. Bhargavan, J.H. Sunshine, RADPEER quality assurance program: a multifacility study of interpretive disagreement rates. Journal of the American College of Radiology. 1(1) (2004) 59-65. https://doi.org/10.1016/S1546-1440(03)00002-4.

[22] J. Ruma, K.A. Klein, S. Chong, J. Wesolowski, E.A. Kazerooni, J.H. Ellis, J.D. Myles, Cross-sectional examination interpretation discrepancies between on-call diagnostic radiology residents and subspecialty faculty radiologists: analysis by imaging modality and subspecialty. Journal of the American College of Radiology. 8(6) (2011) 409-414. https://doi.org/10.1016/j.jacr.2011.01.012.

[23] T.N. Hanna, C. Lamoureux, E.A. Krupinski, S. Weber, J.O. Johnson, Effect of shift, schedule, and volume on interpretive accuracy: a retrospective analysis of 2.9 million radiologic examinations. Radiology. 287(1) (2017) 205-212. https://doi.org/10.1148/radiol.2017170555.

[24] J. Pallant, SPSS Survival Manual. A step by step guide to data analysis using IBM SPSS. Sixth ed., McGraw-Hill Education, United Kingdom, 2016.

[25] H.R. Guly, *Diagnostic errors in an accident and emergency department.* Emergency Medicine Journal. 18(4) (2001) 263-269.

[26] R.B. Gunderman, *Biases in radiologic reasoning.* American Journal of Roentgenology. 192(3) (2009) 561-564. https://doi.org/10.2214/AJR.08.1220.

[27] J.N. Itri, R.R. Tappouni, R.O. McEachern, A.J. Pesch, S.H. Patel, *Fundamentals of Diagnostic Error in Imaging.* RadioGraphics. 38(6) (2018) 1845-1865. https://doi.org/10.1148/rg.2018180021