

# **UNIVERSITY OF TWENTE.**

**Technische Geneeskunde** Multidisciplinaire Opdracht

## Transcutane bilirubinemetingen bij pasgeborenen

Kunnen er meer bloedafnames bij inividuele pasgeborenen voorkomen worden? 24 juni 2019

N. Doorn, S.G.J. Lip, N.A. Timmermans en L. Vanwinsen

**Begeleiders:** Medisch begeleider: Dr. H.L.M. van Straaten Technologisch begeleider UT: Dr. Ir. N. Bosschaart Technologisch begeleider instelling: Ir. L. Dam-Vervloet Procesbegeleider: drs. E.M. Walter Tutor: L. Zaremba

## Lijst met symbolen en afkortingen

$\lambda$	golflengte
$\mu_a$	absorptiecoëfficiënt
$\mu'_s$	gereduceerde verstrooiingscoëfficiënt
ANOVA	analysis of variance
С	fractie
G6PD	glucose-6-fosfaatdehyrogenase deficiëntie
GB	geconjugeerd bilirubine
IL20%	Intralipid <sup>®</sup> 20%
JM	jaundice meter
LG	lichtgele inkt
М	magenta inkt
MST	Medisch Spectrum Twente
NICU	neonatale intensive care unit
$R_a^2$	adjusted R-squared
RMSE	root mean square error
ROC	receiver operating characteristic
ТсВ	trancutane bilirubine
TSB	totaal serum bilirubine concentratie
Tukey's HSD	Tukey's honestly significant difference
UGB	ongeconjugeerd bilirubine
UGT	uridine 5'-difosfo-glucuronosyltransferase
UMCG	Universitair Medisch Centrum Groningen

## Voorwoord

Beste Lezer,

Voor u ligt het verslag van de Multidisciplinaire Opdracht van Technische Geneeskunde die wij in opdracht van de vakgroep *Biomedical Photonic Imaging Group* van de Universiteit Twente hebben uitgevoerd, in samenwerking met het Isala te Zwolle. De opdracht gaat over transcutane bilirubinemetingen bij pasgeborenen. Wij hebben de afgelopen tien weken met veel plezier aan deze opdracht gewerkt.

Er zijn een aantal personen die we graag zouden willen bedanken voor hun bijdrage aan de opdracht en het proces. We willen graag Nienke Bosschaart bedanken, omdat ze altijd met ons meedacht en bereikbaar was voor vragen en adviezen. Door haar hulp en kritische blik hebben we ons onderzoek naar een hoger niveau kunnen tillen. Daarnaast willen we graag Lida Dam-Vervloet bedanken voor het overdragen van haar passie voor de opdracht en de energie die ze in de opdracht heeft gestoken. Ze toonde altijd interesse in ons en de voortgang. Bovendien heeft ze ons ontzettend geholpen bij alles wat we nodig hadden voor het uitvoeren van onze experimenten. We willen ook graag Irma van Straaten bedanken, voor het delen van haar waardevolle kennis op het gebied van neonatologie, voor de enthousiaste rondleiding op de NICU, en voor alle inzet en hulp bij het verkrijgen van onderzoeksdata. Daarnaast willen we graag Loes Zaremba bedanken voor alle waardevolle persoonlijke en groepsbegeleiding en het vertrouwen dat ze in ons had. De gesprekken met haar waren altijd een welkome afwisseling tussen het harde werken door. Verder willen we graag Elfi Hofmeijer bedanken voor het feit dat ze altijd open stond en bereikbaar was voor vragen en hulp. Ze heeft ons erg geholpen bij de vormgeving en uitvoering van onze opdracht. Ten slotte willen we graag alle anderen bedanken die ons hebben geholpen bij het uitvoeren van ons onderzoek: Foky-Anna de Boer, Bart Titulaer, Martijn Kusters en Wilma Petersen.

Wij wensen u veel leesplezier toe.

Nina Doorn, Stefan Lip, Nienke Timmermans en Leen Vanwinsen

## Samenvatting

**Aanleiding:** Hyperbilirubinemie komt bij 60-80% van de pasgeborenen voor en kan irreversibele hersenschade (kernicterus) veroorzaken bij een zeer hoge totaal serum bilirubine concentratie (TSB). Kernicterus kan met behulp van fototherapie voorkomen worden. Om te bepalen of fototherapie gestart moet worden, dienen regelmatig invasieve TSB-bepalingen plaats te vinden. Een alternatief hiervoor zijn transcutane bilirubine (TcB) metingen met de Dräger transcutane bilirubinemeters (JM-103 en JM-105). Echter, doordat deze metingen onnauwkeurig zijn, worden de meters slechts gebruikt als screeningsmethode. Hofmeijer heeft een MATLAB model ontworpen die de TSB voorspelt aan de hand van de TcB-waarden, om daarmee meer invasieve TSB-bepalingen te kunnen vervangen door TcB-metingen. Voordat dit model in de kliniek kan worden toegepast, moeten de volgende vragen beantwoord worden:

- Wat is de inter-device reproduceerbaarheid van verschillende Dräger transcutane bilirubinemeters van het type JM-103 en JM-105?
- Welke verbeteringen kunnen er gedaan worden aan het model van Hofmeijer om het klinisch te kunnen toepassen bij individuele patiënten?
- Hoe groot is de klinische waarde van het model inclusief eventuele verbeteringen?

**Methode:** Om de reproduceerbaarheid van de bilirubinemeters te bepalen is er met dertien meters uit vier verschillende ziekenhuizen gemeten op zes verschillende fantomen die de optische eigenschappen van de neonatale huid nabootsen. Om het model te verbeteren is de gebruikte dataset uitgebreid, is patiëntoverlap tussen test- en trainset voorkomen en is meetdata van slechts één meter gebruikt. De klinische waarde van het verbeterde model is bepaald door te berekenen hoeveel procent van de huidige TSB-bepalingen veilig vervangen kan worden door TcB-metingen.

**Resultaten:** Het maximale verschil tussen de gemeten TcB-waarden van verschillende bilirubinemeters is 61.29  $\mu$ mol/L bij het fantoom met  $\mu_a$ (450 nm) = 3,28 mm<sup>-1</sup>. Wanneer de verschillen die veroorzaakt kunnen worden door het maken van verschillende fantomen en het bevestigen van verschillende Tegaderm<sup>TM</sup>-plakkers daarvan worden afgetrokken, is het maximale verschil in TcBwaarden tussen de dertien meters 56,44  $\mu$ mol/L. Het MATLAB model dat patiëntoverlap voorkomt en gebruik maakt van data van slechts één bilirubinemeter is het meest klinisch relevant en heeft een RMSE van 18,99  $\mu$ mol/L. Met dit model kan, bij gebruik van één meter, 19,64% van de huidig afgenomen TSB-bepalingen veilig vervangen worden door TcB-metingen.

**Conclusie:** Uit de resultaten kan worden geconcludeerd dat de transcutane bilirubinemeters niet reproduceerbaar zijn. Het MATLAB model is verbeterd, maar kan momenteel voor slechts één meter in de kliniek toegepast worden. Om het model algemeen toe te kunnen passen, is verder onderzoek nodig.

## Inhoudsopgave

Lij	jst me	et syml	oolen en afkortingen	i
Vc	orwo	ord		ii
Sa	men	vatting		iii
1	Intro	oductie		1
	1.1	Aanlei	ding	1
	1.2	Bilirub	ineproductie en -excretie	2
	1.3	Hyper	pilirubinemie	2
	1.4	Transc	utane bilirubinometrie	3
	1.5	Weten	schappelijk kader	5
	1.6	Doel v	an het onderzoek	6
2	Fant	toomst	udie	7
	2.1	Inleidi	ng	7
	2.2	Algem	ene methode fantoomstudie	7
		2.2.1	Samenstelling fantomen	7
		2.2.2	Meetopstelling	8
	2.3	Repro	duceerbaarheid fantomen	9
		2.3.1	Inleiding	9
		2.3.2	Methode	9
		2.3.3	Resultaten	10
		2.3.4	Conclusie	10
		2.3.5	Discussie	10
	2.4	Repro	duceerbaarheid Tegaderm <sup>TM</sup> $\ldots$	11
		2.4.1	Inleiding	11
		2.4.2	Methode	11
		2.4.3	Resultaten	11
		2.4.4	Conclusie	12
		2.4.5	Discussie	12
	2.5	Repro	duceerbaarheid transcutane bilirubinemeters	13
		2.5.1	Inleiding	13
		2.5.2	Methode	13
		2.5.3	Resultaten	14
		2.5.4	Conclusie	19
		2.5.5	Discussie	19
3	Mod	lel		24
	3.1	Inleidi	ng	24
	3.2	Metho	de	24
		3.2.1	Vergroten dataset	24
		3.2.2	Voorkomen van patiëntoverlap	25
		3.2.3	Gebruik van één bilirubinemeter	25

		3.2.4	Kwantificatie klinische waarde	25
	3.3	Result	aten	26
		3.3.1	Vergroten dataset	27
		3.3.2	Voorkomen van patiëntoverlap	27
		3.3.3	Gebruik van één bilirubinemeter	27
		334	Kwantificatie klinische waarde	28
	34	Concli		28
	3.5	Discus	2010 · · · · · · · · · · · · · · · · · ·	28
	0.0	351	Vergraten dataset	28
		252		20
		0.5.2		29
		3.5.3		29
		3.5.4		29
4	Con	clusie		30
5	Disc	ussie	en aanbevelingen	31
5	Disc	cussie	en aanbevelingen	31
5 Re	Disc eferei	cussie nties	en aanbevelingen	31 32
5 Re 6	Disc eferei Bijla	cussie nties agen	en aanbevelingen	31 32 35
5 Re 6	Disc eferer Bijla 6.1	cussie nties agen Bepali	en aanbevelingen ng absorptiecoëfficiënt inkt stockoplossing	<b>31</b> <b>32</b> <b>35</b> 35
5 Re 6	Disc eferer Bijla 6.1 6.2	<b>ties</b> agen Bepali Trendl	en aanbevelingen ng absorptiecoëfficiënt inkt stockoplossing	<ul> <li>31</li> <li>32</li> <li>35</li> <li>35</li> <li>36</li> </ul>
5 Re 6	Disc eferen Bijla 6.1 6.2 6.3	<b>ties</b> agen Bepali Trendl SPSS	en aanbevelingen ng absorptiecoëfficiënt inkt stockoplossing	<ul> <li>31</li> <li>32</li> <li>35</li> <li>36</li> <li>36</li> </ul>
5 Re 6	Disc eferer Bijla 6.1 6.2 6.3 6.4	cussie nties agen Bepali Trendl SPSS MATL/	en aanbevelingen ng absorptiecoëfficiënt inkt stockoplossing	<ul> <li>31</li> <li>32</li> <li>35</li> <li>36</li> <li>36</li> <li>39</li> </ul>
5 Re 6	Disc eferen Bijla 6.1 6.2 6.3 6.4	cussie nties agen Bepali Trendl SPSS MATL/ 6.4.1	en aanbevelingen ng absorptiecoëfficiënt inkt stockoplossing	<ul> <li>31</li> <li>32</li> <li>35</li> <li>36</li> <li>36</li> <li>39</li> <li>39</li> <li>39</li> </ul>
5 Re 6	Disc eferen Bijla 6.1 6.2 6.3 6.4	cussie nties agen Bepali Trendl SPSS MATL 6.4.1 6.4.2	en aanbevelingen ng absorptiecoëfficiënt inkt stockoplossing	<ul> <li>31</li> <li>32</li> <li>35</li> <li>36</li> <li>36</li> <li>39</li> <li>39</li> <li>44</li> </ul>
5 Re 6	Disc eferen Bijla 6.1 6.2 6.3 6.4	cussie nties agen Bepali Trendl SPSS MATL/ 6.4.1 6.4.2 6.4.3	en aanbevelingen ng absorptiecoëfficiënt inkt stockoplossing	<ul> <li>31</li> <li>32</li> <li>35</li> <li>36</li> <li>36</li> <li>39</li> <li>44</li> <li>47</li> </ul>
5 Re 6	Disc eferen 6.1 6.2 6.3 6.4	cussie of nties agen Bepali Trendl SPSS MATL/ 6.4.1 6.4.2 6.4.3 6.4.4	en aanbevelingen ng absorptiecoëfficiënt inkt stockoplossing	<ul> <li>31</li> <li>32</li> <li>35</li> <li>36</li> <li>36</li> <li>39</li> <li>44</li> <li>47</li> <li>48</li> </ul>
5 Re 6	Disc eferen Bijla 6.1 6.2 6.3 6.4	cussie nties agen Bepali Trendl SPSS MATL 6.4.1 6.4.2 6.4.3 6.4.4 BOC	en aanbevelingen ng absorptiecoëfficiënt inkt stockoplossing	<ul> <li>31</li> <li>32</li> <li>35</li> <li>36</li> <li>36</li> <li>39</li> <li>44</li> <li>47</li> <li>48</li> <li>50</li> </ul>

## 1 Introductie

In de introductie zal de aanleiding voor dit onderzoek besproken worden en zal achtergrondinformatie worden gegeven over bilirubineproductie en -excretie, hyperbilirubinemie en transcutane bilirubinometrie. Ten slotte zal het wetenschappelijk kader en het doel van het onderzoek worden beschreven.

## 1.1 Aanleiding

Geelzucht komt bij 60-80% van alle pasgeborenen voor en kenmerkt zich door een gele verkleuring van de huid en sclera [1]. Het wordt veroorzaakt door een verhoogde bilirubineconcentratie in het bloed (hyperbilirubinemie), waardoor bilirubine in de weefsels ophoopt. Geelzucht treedt op bij een totaal serum bilirubine concentratie (TSB) boven de 85  $\mu$ mol/L [1]. Bij een zeer hoge TSB is er vooral bij prematuren en zieke pasgeborenen een risico op irreversibele hersenschade, ook wel kernicterus genoemd [2].

In Nederland maakt ongeveer 1% van de à terme pasgeborenen een klinisch relevante hyperbilirubinemie door, waarvoor behandeling noodzakelijk is om kernicterus te voorkomen [3]. Deze behandeling bestaat uit fototherapie of een wisseltransfusie. Bij fototherapie wordt licht gebruikt met een golflengte van 400-520 nm [4]. Dit licht zorgt ervoor dat bilirubine wordt omgezet in oplosbare stoffen die vervolgens kunnen worden uitgescheiden. Bij een wisseltransfusie wordt het bloed van de pasgeborene vervangen door donorbloed [4].

De TSB-waarde bepaalt of fototherapie of een wisseltransfusie geïndiceerd is [5]. De therapiegrens is afhankelijk van de zwangerschapsduur, de leeftijd van het kind in uren en de aanwezigheid van risicofactoren. Bij prematuren heeft ook het geboortegewicht invloed op de therapiegrens [6]. Omdat het inschatten van de mate van hyperbilirubinemie aan de hand van de geelheid van de huid onbetrouwbaar is, dienen regelmatig invasieve bloedafnames plaats te vinden om de TSB te bepalen. Deze bloedafnames kennen een aantal belangrijke nadelen: het is tijdrovend, pijnlijk en stressvol voor de pasgeborene, leidt tot bloedverlies en verhoogt de kans op osteomyelitis en infecties rond de plek van afname [7].

Transcutane bilirubinometrie is een mogelijk niet-invasief alternatief voor de bloedafnames. Echter, doordat transcutane bilirubine (TcB) metingen onnauwkeurig zijn, worden deze momenteel slechts gebruikt als screeningsmethode [5]. Wanneer TcB-waarden 50  $\mu$ mol/L of minder onder de fotothe-rapiegrens liggen, dient alsnog de TSB regelmatig bepaald te worden. De onnauwkeurigheid van de TcB-metingen zorgt ervoor dat er nog relatief weinig bloedafnames vervangen kunnen worden door metingen met transcutane bilirubinemeters. Uit onderzoek van Van den Esker-Jonker blijkt dat door het gebruik van de transcutane bilirubinemeter slechts 38,5% van de bloedafnames voorkomen kan worden [8]. Dit onderzoek is uitgevoerd in onder andere het Isala te Zwolle bij patiënten met een minimale zwangerschapsduur van 32 weken. In ander onderzoek van Ercan et al bij à terme pasgeborenen wordt een vergelijkbaar percentage gevonden van 40% [9].

## 1.2 Bilirubineproductie en -excretie

Bilirubine is een afbraakproduct van hemoglobine [10]. Ongeveer 80% van de natuurlijke hoeveelheid bilirubine in het lichaam komt van de afbraak van oude erytrocyten en 20% van de afbraak van heembevattende eiwitten. De bilirubineproductie en -excretie zijn schematisch weergegeven in figuur 1.1. Bilirubine ontstaat door een tweedelige degradatiereactie die vooral plaatsvindt in het reticulo-endotheliale systeem. Het hemoglobinemolecuul wordt hier eerst afgebroken door heem oxygenase, waarbij ijzer, koolstofmonoxide en biliverdine ontstaan. Vervolgens wordt biliverdine gereduceerd tot bilirubine met behulp van biliverdine reductase. Het verkregen bilirubine wordt ongeconjugeerd bilirubine (UGB) genoemd en is hydrofoob. UGB heeft een hoge affiniteit voor plasma albumine en zal daar reversibel en covalent aan binden. Gebonden aan albumine komt UGB via de bloedbaan bij de lever terecht, zonder dat het wordt uitgescheiden door de nieren. In de lever splitst albumine af en gaat bilirubine de hepatocyten binnen. In het endoplasmatisch reticulum van de hepatocyten wordt UGB door UDP-glucuronosyltransferase (UGT) gekoppeld aan glucuronzuur, waardoor geconjugeerd bilirubine (GB) ontstaat. GB is hydrofiel en kan daarom geëxporteerd worden naar de galblaas [10].



Figuur 1.1: Schematisch overzicht van de bilirubineproductie en -excretie.

Via de galwegen komt GB in het duodenum [11]. In het terminale ileum en proximale colon wordt GB door darmbacteriën omgezet in het kleurloze urobilinogeen, waarbij glucuronzuur wordt verwijderd. Het grootste deel van het urobilinogeen wordt door andere darmbacteriën snel geoxideerd om zo stercobiline te vormen, dat de bruine kleur aan faeces geeft. De rest wordt in het colon geabsorbeerd. Een deel hiervan wordt in de lever opnieuw in het gal geëxcreteerd (enterohepatische circulatie) en het overige deel zal door de nieren worden geconverteerd naar uribiline, dat de gele kleur aan urine geeft. [11]

## 1.3 Hyperbilirubinemie

Bij pasgeborenen is de TSB vaak verhoogd [12]. Deze verhoging is meestal fysiologisch en heeft geen gevolgen voor het kind. De oorzaak voor deze verhoogde TSB is dat pasgeborenen een overproductie van bilirubine hebben en daarnaast een verlaagde hepatische opname, enzymatische conjugatie en galexcretie van bilirubine. Tijdens de zwangerschap neemt de lever van de moeder een groot deel van de bilirubineconjugatie op zich. Na geboorte heeft de lever van het kind tijd nodig om de volledige conjugatie over te nemen [13]. Fysiologische geelzucht wordt zichtbaar tussen 24-72 uur na geboorte en verdwijnt weer na 10-14 dagen. De TSB blijft hierbij meestal onder de 250  $\mu$ mol/L. [12]

Pathologische geelzucht kenmerkt zich door het optreden van geelzucht binnen 24 uur na de geboorte, een stijging in de TSB van minimaal 86  $\mu$ mol/L per dag of een TSB boven de 290 µmol/L [14]. Net als bij fysiologische hyperbilirubinemie is er sprake van overproductie van bilirubine of een verlaagde conjugatiesnelheid [15]. De overproductie van bilirubine kan worden veroorzaakt door een verhoogde hemolyse. Oorzaken hiervan zijn bloedgroepincompatibiliteit van moeder en kind, genetische afwijkingen of verworven hemolytische aandoeningen. Bloedgroepincompatibiliteit zorgt ervoor dat ervtrocyten worden afgebroken door antistoffen van de moeder. Een genetische afwijking die hemolyse veroorzaakt is bijvoorbeeld een deficiëntie van het G6PD-eiwit. Dit eiwit beschermt erytrocyten normaal gesproken tegen vervroegde afbraak. De conjugatiesnelheid kan verlaagd zijn doordat de lever van het kind de volledige conjugatie over moet nemen van de lever van de moeder. Normaal gesproken leidt dit tot een fysiologische hyperbilirubinemie, maar prematuren hebben vaak meer tijd nodig om de conjugatiesnelheid op peil te krijgen waardoor een ernstige hyperbilirubinemie kan ontstaan. Ook genetische afwijkingen kunnen tot een verlaagde of zelfs afwezige conjugatie leiden [16]. Bij het Gilbert Syndroom is de activiteit van UGT verlaagd en bij het Criggler-Najjar syndroom is de UGT concentratie in de lever minder dan 10% van normaal [17].

De gevolgen van een te hoge bilirubineconcentratie zijn neurologisch van aard [4]. UGB is een neurotoxine en wanneer het niet gebonden is aan albumine kan het de bloed-hersenbarrière passeren en zo de hersenen binnendringen. Doordat bij prematuren albumine een lagere affiniteit heeft voor UGB wordt dit proces bij hen versterkt. UGB kan focale necrose van neuronen en gliacellen veroorzaken. Deze necrose veroorzaakt schade, ook wel bilirubine-encephalopathie genoemd. In een vroeg stadium is de schade reversibel, maar wanneer er niet ingegrepen wordt, ontstaat een chronische bilirubine-encephalopathie, oftewel kernicterus [18]. De gebieden in de hersenen die het vaakst aangetast worden bij kernicterus zijn de basale ganglia en de nuclei voor oculomotorische en auditorische functies. Deze functies zullen dan ook als eerst uitvallen. [4]

Bij toenemende TSB kleurt eerst het hoofd geel, vervolgens de heup en tenslotte de extremiteiten [19]. Dit wordt de cephalocaudale spreiding van geelzucht genoemd. Het mechanisme van die spreiding is nog niet bekend, maar er zijn wel een aantal hypothesen. Een aannemelijke hypothese heeft te maken met de toename van de albumine-bilirubine affiniteit in de tijd, doordat albuminebilirubine complexen een conformatie-verandering ondergaan. Aangezien bilirubine, nadat het in het reticulo-endotheliale systeem is gevormd, via de circulatie eerst in de proximale delen van het lichaam terechtkomt, is de albumine-bilirubine affiniteit daar lager dan in de distale delen van het lichaam. In de proximale delen van het lichaam is er daarom meer bilirubine beschikbaar dat kan neerslaan in de huid. [19]

### 1.4 Transcutane bilirubinometrie

Transcutane bilirubinometrie is een methode die met optische spectroscopie de concentratie bilirubine in het subcutane weefsel meet [20]. De TcB-meter zendt licht uit met verschillende golflengten. Dit licht wordt in het weefsel deels geabsorbeerd en deels gereflecteerd. De mate van absorptie en reflectie wordt bepaald door de golflengte van het licht en de concentratie van diverse chromoforen in de huid, waaronder bilirubine, hemoglobine en melanine. Bilirubine absorbeert vooral licht met een golflengte van 450 nm (zie figuur 1.2) [21].



Figuur 1.2: Absorptiespectra van bilirubine, hemoglobine (Hb), gecarbamyleerd hemoglobine (cHb), geoxygeneerd hemoglobine (HbO) en methemoglobine (MetHb). Aangepast van [21]

Een veelgebruikte TcB-meter, de Dräger JM-103, zendt licht uit met golflengtes van 450 nm (blauw) en 550 nm (groen) [20]. Door het verschil in optische dichtheid van het gereflecteerde licht bij 450 en 550 nm te meten, kan gecorrigeerd worden voor de absorptie door hemoglobine. Daarnaast maakt de Dräger JM-103 bij de TcB-meting onderscheid tussen het gereflecteerde licht dat een korte en lange weg heeft afgelegd [22]. Licht dat wordt gereflecteerd vanuit de epidermis, de dermis en het oppervlakkige subcutane weefsel legt een korte weg af en wordt opgevangen in de kern van de tip van meter (zie figuur 1.3). Licht dat vanuit het dieper gelegen subcutane weefsel wordt gereflecteerd legt een lange weg af, waardoor het aan de rand wordt opgevangen. De twee spectra die in de fiber opgevangen worden, kunnen van elkaar afgetrokken worden. Na een correctie kan hieruit de absorptie door het dieper gelegen subcutane weefsel bepaald worden, waardoor de invloed van melanine in en onrijpheid van het oppervlakkig gelegen deel van de huid verminderd wordt [23]. De nieuwere versie van de Dräger JM-103, de Dräger JM-105, heeft hetzelfde werkingsmechanisme [24].

Doordat het subcutane weefsel voor minder dan 1% uit bloedvaten bestaat, draagt vooral het extravasculaire bilirubine bij aan de TcB-waarde [7]. Daarom kan de TcB-waarde niet direct gerelateerd worden aan de TSB, die bepaald wordt door het intravasculaire bilirubine. De transcutane bilirubinemeters maken gebruik van een correlatiecoëfficiënt, om de optische dichtheid van het gereflecteerde licht om te zetten in een schatting van de TSB [23]. In de handleiding van de Dräger JM-105 staat vermeld dat de bilirubinemeter de TSB schat met een onnauwkeurigheid van  $\pm$  25,5  $\mu$ mol/L bij een zwangerschapsduur boven de 35 weken en  $\pm$  27,4  $\mu$ mol/L bij een zwangerschapsduur van 24 tot 34 weken. [25].



Figuur 1.3: Werkingsmechanisme van de Dräger JM-103 en JM-105. Aangepast van [22].

## 1.5 Wetenschappelijk kader

Dit onderzoek is onderdeel van een groter onderzoek naar het beantwoorden van de volgende hoofdvraag:

## Hoe kan bij individuele pasgeborenen het percentage totaal serum bilirubinebepalingen dat vervangen wordt door transcutane bilirubinemetingen vergroot worden?

In 2018 heeft M. D. van Erk de eerste stap gezet in het beantwoorden van deze vraag [26]. Van Erk heeft onderzocht wat de invloed van de botdiepte en verschillende optische eigenschappen van de huid op transcutane bilirubinemetingen is. De botdiepte bleek van invloed te zijn op de TcB-waarde bij dieptes kleiner dan 3 mm. Ook een klinisch realistische variatie in de verstrooiing van de huid heeft invloed op de TcB-metingen en leidde tot een maximaal verschil van 72  $\mu$ mol/L in gemeten TcB-waarde. Verder is gebleken dat het verband tussen de absorptiecoëfficiënt bij 450 nm en de TcB-waarde logaritmisch is.

Daarnaast heeft Van Erk de reproduceerbaarheid van vier verschillende bilirubinemeters van hetzelfde type (Dräger JM-105) onderzocht [26]. Bij dit onderzoek is de botdiepte zo ingesteld dat deze geen invloed heeft op de metingen. Uit het onderzoek bleek dat de TcB-meetwaarden van alle vier de meters significant verschilden en dat maar drie van de vier meters vergelijkbare resultaten gaven. Bijna 40% van de TcB-metingen van de meter met serienummer B3601027 lag meer dan 25,5  $\mu$ mol/L onder de TcB-metingen van de andere drie meters. Het is daarom gewenst om de reproduceerbaarheid van meer verschillende transcutane bilirubinemeters door middel van een fantoomstudie systematisch te onderzoeken.

Voortbouwend op het onderzoek van Van Erk, heeft E.I.S. Hofmeijer begin 2019 onderzoek gedaan naar het beter voorspellen van de TSB aan de hand van TcB-metingen [27]. Een belangrijke deelvraag hierbij was of de cephalocaudale spreiding voor deze voorspelling gebruikt kon worden. In het onderzoek gebruikt Hofmeijer TSB-bepalingen en bijbehorende TcB-metingen op vijf plekken (voorhoofd, sternum, heup, tibia en enkel) van 44 premature pasgeborenen om een MATLAB *machine learning* model te trainen. Het resultaat is een model dat de TSB kan voorspellen aan de hand van informatie over het geboortegewicht, de zwangerschapsduur, de leeftijd en TcB-waarden op het voorhoofd, het sternum en de heup. Hierbij bleek dat de cephalocaudale spreiding niet significant bijdraagt aan de nauwkeurigheid van het model en dat een TcB-meting op het voorhoofd voldoende is om tot een nauwkeurige voorspelling van de TSB te komen. Verder kan het model van Hofmeijer als aanvulling hierop aan de hand van bestaande protocollen adviseren of er fototherapie gegeven moet worden.

Het model van Hofmeijer is zowel getraind als getest met data van 44 patiënten. Bij elke patiënt zijn meerdere TSB-bepalingen met bijbehorende TcB-waarden (dataparen) gemeten. Het model deelt de dataparen in een train- of testset in, waarbij het geen onderscheid maakt tussen verschillende patiënten. Daardoor kan een patiënt met verschillende dataparen in zowel de train- als testset voor-komen. Hierdoor is het model mogelijk getraind om bepaalde patiënten te herkennen en aan de hand van gelijke patiënten te testen dan waarmee het model getraind is of om maar één datapaar per patiënt te gebruiken. Daarnaast zijn in eerste instantie voor het trainen van dit model TcB-metingen van twee verschillende bilirubinemeters gebruikt. Nadat uit het onderzoek van Van Erk bleek dat de resultaten van deze meters significant kunnen verschillen, zijn de TcB-metingen met één meter voortgezet. Er zou onderzocht kunnen worden of het gebruik van TcB-metingen van slechts één meter het model kan verbeteren. Ten slotte is het gewenst om de klinische waarde van het model te bepalen.

## 1.6 Doel van het onderzoek

Het doel van dit onderzoek is om te bepalen of en hoe het model van Hofmeijer in de kliniek kan worden toegepast. Om dit doel te bereiken zullen de volgende deelvragen beantwoord moeten worden:

- Wat is de inter-device reproduceerbaarheid van verschillende Dräger JM-103 en JM-105 transcutane bilirubinemeters?
- Welke verbeteringen kunnen er gedaan worden aan het model van Hofmeijer om het klinisch te kunnen toepassen bij individuele patiënten?
- Hoe groot is de klinische waarde van het model inclusief eventuele verbeteringen?

In hoofdstuk 2 wordt het onderzoek naar de reproduceerbaarheid van de bilirubinemeters beschreven. Vervolgens wordt er in hoofdstuk 3 ingegaan op het model van Hofmeijer en de klinische waarde ervan. Ten slotte zullen deze twee delen geïntegreerd worden in een algemene conclusie en discussie, om het in een groter perspectief te plaatsen.

## 2 Fantoomstudie

## 2.1 Inleiding

Het doel van deze fantoomstudie is het bepalen van de inter-device reproduceerbaarheid van verschillende Dräger JM-103 en JM-105 transcutane bilirubinemeters. Om de meters op een nauwkeurige en reproduceerbare manier te vergelijken, zijn vloeibare fantomen gemaakt die de optische eigenschappen van de neonatale huid bij verschillende bilirubineconcentraties nabootsen. De meetopstelling en fantomen zijn gebaseerd op de fantoomstudie van Van Erk [26].

In dit hoofdstuk zal eerst de algemene methode van de fantoomstudie worden beschreven, waarbij de samenstelling van de fantomen en de meetopstelling worden toegelicht. Om de betrouwbaarheid van de methode te bepalen, is de reproduceerbaarheid van het maken van de fantomen en het uitvoeren van de metingen onderzocht. De resultaten daarvan zullen worden beschreven. Vervolgens zal toegelicht worden hoe de verschillende transcutane bilirubinemeters met elkaar vergeleken zijn en wat de resultaten daarvan zijn. Deze fantoomstudie is uitgevoerd in samenwerking met het Medisch Spectrum Twente (MST) te Enschede, het Isala te Zwolle, het Spaarne Gasthuis te Haarlem en het Universitair Medisch Centrum Groningen (UMCG).

## 2.2 Algemene methode fantoomstudie

#### 2.2.1 Samenstelling fantomen

De fantomen hebben een volume van 50 mL en bestaan uit een oplossing van lichtgele inkt (Talens Ecoline 201), magenta inkt (Talens Ecoline 337), IL20% (Intralipid<sup>®</sup>20%, Fresenius Kabi, Bad Homburg, Duitsland) en Milli-Q. Het fantoom bootst de absorptie- en gereduceerde verstrooiingscoëfficiënten van de neonatale huid bij golflengten van 450 en 550 nm na. De totale absorptiecoëfficiënten ( $\mu_{a,totaal}$ ) van de neonatale huid bij 450 en 550 nm zijn bepaald aan de hand van de absorptiecoëfficiënten van bilirubine en hemoglobine, die respectievelijk verkregen zijn uit de spectra van Veenstra et. al. en Prahl [28] [29]. Hierbij is de hemoglobineconcentratie in de huid geschat op 2,13 g/L en de fractie van hemoglobine dat geoxygeneerd is op 0,85 [26]. Hetzelfde standaardfantoom is gedefinieerd als bij Van Erk, hiervan staan de optische eigenschappen vermeld in tabel 2.1 [26].

$\lambda$ (nm)	$\mu_a$ (mm $^{-1}$ )	$\mu_s^\prime$ (mm $^{-1}$ )
450	1,55	2,00
550	0,35	1,63

Tabel 2.1: Optische eigenschappen van het standaardfantoom.

De lichtgele en magenta inkt zijn tien keer verdund om stockoplossingen te maken. De absorptiecoëfficiënten van de stockoplossingen bij 450 en 550 nm zijn bepaald met behulp van een spectrofotometer (zie bijlage 6.1). De fracties lichtgele en magenta inkt stockoplossing die nodig zijn om de gewenste  $\mu_{a,totaal}$  te bereiken zijn berekend aan de hand van de volgende formules:

$$\mu_{a,totaal}(\lambda) = \mu_{a,S_{LG}}(\lambda) * C_{LG} + \mu_{a,S_M}(\lambda) * C_M$$
(2.1)

 $\begin{array}{l} \mu_{a,totaal}(\lambda) = \text{De gewenste totale absorptiecoëfficiënt bij golflengte } \lambda. \\ \mu_{a,S_{LG}}(\lambda) = \text{De absorptiecoëfficiënt van de lichtgele inkt stockoplossing bij golflengte } \lambda, zie bijlage. \\ \mu_{a,S_{M}}(\lambda) = \text{De absorptiecoëfficiënt van de magenta inkt stockoplossing bij golflengte } \lambda, zie bijlage. \\ C_{LG} = \text{De benodigde fractie lichtgele inkt stockoplossing in het fantoom.} \\ C_{M} = \text{De benodigde fractie magenta inkt stockoplossing in het fantoom.} \end{array}$ 

 $\lambda = 550$  invullen in vergelijking 2.1 en herschrijven geeft:

$$C_M = \frac{\mu_{a,totaal}(550) - \mu_{a,S_{LG}}(550) * C_{LG}}{\mu_{a,S_M}(550)}$$
(2.2)

Vergelijking 2.2 en  $\lambda = 450$  invullen in vergelijking 2.1 geeft na herschrijven:

$$C_{LG} = \frac{\mu_{a,totaal}(450) - [\mu_{a,S_M}(450) * \mu_{a,totaal}(550)] / \mu_{a,S_M}(550)}{\mu_{a,S_LG}(450) - [\mu_{a,S_M}(450) * \mu_{a,S_LG}(550)] / \mu_{a,S_M}(550)}$$
(2.3)

Ten slotte kan  $C_{LG}$  ingevuld worden in formule 2.2 om  $C_M$  te berekenen. De berekende fracties lichtgele en magenta inkt zijn omgerekend naar volumes. Deze zijn vervolgens met behulp van een pipetboy in 50 mL buizen gepipetteerd en aangevuld met de juiste hoeveelheid Milli-Q.

De gereduceerde verstrooiingscoëfficiënt ( $\mu'_s$ ) van elk fantoom is 2,00 mm<sup>-1</sup> bij 450 nm en 1,63 mm<sup>-1</sup> bij 550 nm en is gebaseerd op de gemiddelde waarde van een neonatale huid, zoals gemeten door Bosschaart et. al. [30]. De fractie IL20% die nodig is om die verstrooiingscoëfficiënten te verkrijgen, is berekend aan de hand van formules bepaald door Michels et. al. [31]. Om te voorkomen dat een deel van de IL20% bij het pipetteren in de pipet achterblijft door de hoge viscositeit, is het IL20% verdund tot een 25% IL20% stockoplossing. Vervolgens is het juiste volume 25% IL20% met behulp van de pipetboy in een andere 50 mL buis gepipetteerd dan de inktoplossing.

De 50 mL buizen zijn telkens maximaal één dag voor de metingen gevuld met inktoplossing en 25% IL20%. Vervolgens zijn de gevulde buizen in een koeltasje en afgeschermd van licht naar de ziekenhuizen vervoerd. Voordat elke meting plaatsvond, is de inktoplossing bij de IL20% oplossing gevoegd. De metingen op één fantoom zijn telkens binnen dertig minuten uitgevoerd, omdat de tijd mogelijk invloed heeft op de optische eigenschappen van het fantoom. Uit eerder onderzoek bleek namelijk dat de TcB-meetwaarde met ongeveer 2,5  $\mu$ mol/L per uur toeneemt [26].

#### 2.2.2 Meetopstelling

De opstelling die voor het meten aan de fantomen is gebruikt, is gebaseerd op de fantoomstudie van Van Erk [26]. De meetopstelling is bedoeld om de meters op een reproduceerbare manier met elkaar te vergelijken. De opstelling is voornamelijk gemaakt van *Thorlabs*-elementen en bestaat uit een gefixeerd en een verstelbaar deel (zie figuur 2.1). Het gefixeerde deel bestaat uit vier pilaren en een aluminium plaat, waarin zich een gat met een diameter van 8,8 mm bevindt. Rond het gat bevindt zich een inkeping met een diameter van 14 mm, zodat er constante druk uitgeoefend kan worden op de tip van de Dräger JM-103 en JM-105. De onderkant van de aluminium plaat is zwart geverfd om het gereflecteerde licht te absorberen. Op het verstelbare deel is een bakje geplaatst met afmetingen 74 \* 104 \* 45 mm. Op de bodem van dit bakje is een botfantoom van schuimrubber geplaatst.



Figuur 2.1: De opstelling voor de fantoommeting. Aangepast van [26].

Bij het opstellen van elk nieuw fantoom is in het bakje het vloeibare fantoom gegoten. Om luchtbellen te verwijderen is met een gummetje met daaromheen een latex handschoen over de bodem geschraapt. Vervolgens is het verstelbare deel omhoog gebracht totdat de aluminiumplaat het botfantoom raakte, zodat de laatste luchtbellen verdwenen. In dit onderzoek is de invloed van de botdiepte niet meegenomen. Uit het onderzoek van Van Erk is gebleken dat het botfantoom vanaf een diepte van 3 mm geen invloed meer heeft op de metingen [26]. Om er zeker van te zijn dat het botfantoom zich niet in het meetvolume van de transcutane bilirubinemeter bevindt, is het botfantoom voor iedere meting op een afstand van minimaal 4,5 mm vanaf de aluminiumplaat ingesteld.

De bilirubinemeters zijn beschermd tegen vloeistofschade door een Tegaderm<sup>™</sup>-plakker op de tip van de bilirubinemeters te bevestigen. Deze plakkers werden voor het bevestigen uitgerekt en na het bevestigen werden luchtbellen met behulp van een tissue uit de Tegaderm<sup>™</sup>-plakker gestreken. Deze methode bleek, uit onderzoek van Van Erk, het meest reproduceerbaar [26]. Voordat de metingen met de transcutane bilirubinemeters zijn uitgevoerd, is gecontroleerd of de *checker values* binnen het bereik vielen dat op het *docking station* genoteerd staat. De metingen zijn handmatig genoteerd in Excel (Office 365). Voordat het volgende fantoom in het bakje werd gegoten, is het bakje en de opstelling schoon- en drooggemaakt met water en tissues.

## 2.3 Reproduceerbaarheid fantomen

### 2.3.1 Inleiding

De reproduceerbaarheid van het maken van de fantomen is onderzocht om te kunnen bepalen in welke mate de transcutane bilirubinemeters van verschillende ziekenhuizen onderling vergeleken kunnen worden. Voor de metingen zijn namelijk per ziekenhuis zes nieuwe fantomen gemaakt.

### 2.3.2 Methode

Er zijn tien standaardfantomen gemaakt op de manier zoals beschreven in sectie 2.2.1. Op elk van deze tien fantomen zijn 21 metingen uitgevoerd met de Dräger JM-105 (serienummer B3601050) in het MST te Enschede. De metingen zijn vervolgens geanalyseerd met behulp van Origin 2019.

#### 2.3.3 Resultaten

Doordat uit twee buizen IL20% gelekt was, zijn de metingen van slechts acht fantomen meegenomen in de analyse. De gemiddelde waarden en 95%-betrouwbaarheidsintervallen (95%-BI's) van de metingen bij deze acht fantomen zijn weergegeven in figuur 2.2.



Figuur 2.2: Gemiddelde TcB-meetwaarden met 95%-BI's van de Dräger JM-105 (serienummer B3601050) op acht standaardfantomen.

In de figuur is te zien dat het 95%-BI van de metingen bij fantoom 2 groter is dan bij de overige fantomen. Daarnaast is bij fantoom 8 een lagere waarde gemeten dan bij de andere zeven fantomen.

Het verschil tussen de minimale en maximale gemiddelde meetwaarde is 3,76  $\mu$ mol/L. De standaarddeviatie van de gemiddelde meetwaarden van de acht fantomen is 1,25  $\mu$ mol/L.

#### 2.3.4 Conclusie

Uit de resultaten is te concluderen dat een verschil van maximaal 3,76  $\mu$ mol/L tussen verschillende transcutane bilirubinemeters veroorzaakt kan worden door verschillen door het maken van verschillende fantomen.

#### 2.3.5 Discussie

Bij fantoom 2 lagen de eerste negen gemeten waarden rond de 106  $\mu$ mol/L en de laatste twaalf waarden rond de 110  $\mu$ mol/L. Hierdoor is het 95%-BI groter dan bij de overige fantomen. De lagere waarden aan het begin zouden verklaard kunnen worden door de aanwezigheid van een onopgemerkte luchtbel in het meetvolume die na meting 9 verdween. Om te voorkomen dat de meting afwijkt door de aanwezigheid van een luchtbel of stofdeeltje is er besloten om telkens na zes metingen het bakje te bewegen, zodat een eventueel aanwezige luchtbel of stofdeeltje verdwijnt. De lagere meetwaarde bij fantoom 8 zou voor een deel verklaard kunnen worden doordat deze metingen door een andere persoon zijn uitgevoerd dan de metingen op de eerste zeven fantomen.

Daarom zijn de metingen voor het vergelijken van de verschillende transcutane bilirubinemeters telkens door dezelfde persoon uitgevoerd.

De resultaten zouden betrouwbaarder zijn als er op meer fantomen zou zijn gemeten. Wegens gebrek aan tijd en materiaal is besloten om tien fantomen te maken, waarvan er twee geen bruikbare resultaten opleverden. Een andere beperking van dit onderzoek is dat de acht fantomen op hetzelfde moment zijn gemaakt, terwijl de fantomen voor verschillende ziekenhuizen binnen een tijdspanne van twee weken zijn gemaakt. Om te controleren of de absorptiecoëfficiënten van de inkt stockoplossingen in die tijd veranderd zijn, zijn de absorptiecoëfficiënten één en drie weken na de eerste meting opnieuw bepaald (zoals in bijlage 6.1 beschreven staat). Hieruit bleek dat er geen eenduidig verband was tussen de absorptiecoëfficiënt en de tijd.

In het onderzoek van Van Erk is ook de reproduceerbaarheid van het maken van fantomen onderzocht [26]. Daarbij is geen waarde gegeven voor het maximale verschil tussen fantomen, maar uit de figuren is af te lezen dat dit verschil ongeveer 12  $\mu$ mol/L bedraagt. Het verschil dat in dit onderzoek verkregen is, namelijk 3,76  $\mu$ mol/L, is aanzienlijk lager. Dit kan verklaard worden doordat in het onderzoek van Van Erk verschillende stockoplossingen zijn gebruikt voor het maken van fantomen, terwijl in dit onderzoek met één stockoplossing is gewerkt. Hieruit kan geconcludeerd worden dat het gebruik van één stockoplossing voor alle fantomen de meest nauwkeurige resultaten geeft.

## 2.4 Reproduceerbaarheid Tegaderm<sup>™</sup>

#### 2.4.1 Inleiding

De invloed van het bevestigen van verschillende Tegaderm<sup>TM</sup>-plakkers op de transcutane bilirubinemetingen is onderzocht. Dit is van belang om de meters te kunnen vergelijken, omdat op elke meter een andere Tegaderm<sup>TM</sup>-plakker is bevestigd.

### 2.4.2 Methode

Op de Dräger JM-105 (serienummer B3601050) in het MST te Enschede is vijf keer opnieuw een Tegaderm<sup>TM</sup>-plakker bevestigd, zoals beschreven in sectie 2.2.2, waarna telkens op één fantoom 21 metingen zijn uitgevoerd. Het bevestigen van de plakkers is gedurende de hele fantoomstudie door dezelfde twee personen uitgevoerd. De metingen zijn geanalyseerd met behulp van Origin 2019.

#### 2.4.3 Resultaten

De gemiddelde waarden en 95%-BI's bij de vijf verschillende Tegaderm<sup>TM</sup>-plakkers zijn weergegeven in figuur 2.3. In de figuur is te zien dat Tegaderm<sup>TM</sup>-plakker 4 een groter 95%-BI heeft en een lagere gemiddelde TcB-meetwaarde dan de andere vier plakkers. Het verschil tussen de minimale en maximale gemiddelde waarde is 1,09  $\mu$ mol/L. De standaarddeviatie van de gemiddelde waarden is 0,38  $\mu$ mol/L.





#### 2.4.4 Conclusie

Uit de resultaten is te concluderen dat een verschil van maximaal 1,09 µmol/L tussen verschillende transcutane bilirubinemeters veroorzaakt kan worden door het bevestigen van verschillende Tegaderm<sup>TM</sup>-plakkers.

#### 2.4.5 Discussie

In de meetwaarden bij Tegaderm<sup>TM</sup>-plakker 4 was te zien dat één set van drie waarden duidelijk lager is dan de andere waarden. Dit kan wellicht veroorzaakt zijn door een onopgemerkte luchtbel in het meetvolume, want na het bakje bewogen te hebben gaf de meter weer waarden die beter overeenkomen met de meetwaarden bij de rest van de Tegaderm<sup>TM</sup>-plakkers.

De resultaten zouden betrouwbaarder zijn als er vaker een Tegaderm<sup>TM</sup>-plakker bevestigd zou zijn, maar dit was niet mogelijk door de beperkte beschikbaarheid van Tegaderm<sup>TM</sup>-plakkers. Echter, de resultaten van de verschillende metingen liggen relatief dicht bij elkaar, waardoor deze metingen voldoende werden geacht om de reproduceerbaarheid van het bevestigen van de plakkers te bepalen.

In de fantoomstudie van Van Erk is de reproduceerbaarheid van het bevestigen van zes verschillende Tegaderm<sup>TM</sup>-plakkers onderzocht. Het verschil tussen de minimale en maximale gemiddelde TcB-meting was daarbij 1,4 µmol/L. Dit is vergelijkbaar met de reproduceerbaarheid die in dit onderzoek bepaald is.

## 2.5 Reproduceerbaarheid transcutane bilirubinemeters

### 2.5.1 Inleiding

De reproduceerbaarheid van transcutane bilirubinemeters is onderzocht door dertien meters van Dräger van het type JM-103 en JM-105 te vergelijken op fantomen die de optische eigenschappen van de neonatale huid nabootsen.

### 2.5.2 Methode

De transcutane bilirubinemeters uit het MST, Spaarne Gasthuis, Isala en UMCG die getest zijn, staan in tabel 2.2 weergegeven.

Ziekenhuis	Туре	Serienummer	Meetdatum
Medisch Spectrum Twente	JM-105	B3601005	23-05-2019
Spaarne Gasthuis	JM-105	B3601050	24-05-2019
	JM-103	3201007	24-05-2019
	JM-103	3201377	24-05-2019
Isala	JM-105	B3601081	27-05-2019
	JM-105	B3601086	27-05-2019
	JM-105	B3601107	27-05-2019
	JM-105	B3601137	27-05-2019
	JM-103	3201594	27-05-2019
Universitair Medisch Centrum Groningen	JM-105	B3601390	05-06-2019
	JM-105	B3601402	05-06-2019
	JM-103	3202711	05-06-2019
	JM-103	3203135	05-06-2019

Tabel 2.2: Serienummers, herkomst en meetdatums van de transcutane bilirubinemeters.

Deze meters zijn getest op zes fantomen die oplopende bilirubineconcentraties in de huid nabootsen. In tabel 2.3 staan de absorptiecoëfficiënten van de fantomen bij 450 en 550 nm weergegeven. Fantoom 3 komt overeen met het standaardfantoom. De fantomen zijn gemaakt zoals beschreven in sectie 2.2.1. In tabel 2.3 is ook de verwachte TcB-meting per fantoom weergegeven. Deze verwachte TcB-meting is een voorspelling aan de hand van eerdere metingen in de fantoomstudie door Van Erk met meter B3601086 [26].

Fantoom	$\mu_a$ bij 450 nm $(mm^{-1})$	$\mu_a$ bij 550 nm $(mm^{-1})$	Verwachte TcB-meting (µmol/L)
1	0,54	0,34	0
2	0,97	0,34	82
3	1,55	0,35	136
4	2,13	0,35	171
5	2,70	0,36	197
6	3,28	0,36	218

Tabel 2.3: Eigenschappen van de zes fantomen.

Per fantoom zijn met elke bilirubinemeter 21 metingen uitgevoerd door dezelfde persoon. Na elke zes metingen is het bakje met daarin het vloeibare fantoom voorzichtig heen en weer bewogen om een eventueel aanwezige luchtbel of een stofdeeltje uit het meetvolume te verwijderen. Om te bepalen of de optische eigenschappen van het fantoom in de tijd veranderd waren, zijn in het Isala, nadat met alle meters op een bepaald fantoom is gemeten, opnieuw negen metingen uitgevoerd met de eerste meter (met serienummer B3601086). In het Isala was de tijd tussen de metingen namelijk het grootst, omdat daar het grootste aantal meters getest werd. De metingen zijn vervolgens geanalyseerd met behulp van Origin 2019 en IBM SPSS Statistics 25. Voor de analyse is van de 21 metingen steeds het gemiddelde genomen. Als metingen duidelijk door een meetfout afwijken, zijn ze niet meegenomen bij het berekenen van het gemiddelde.

Er is een F-toets uitgevoerd over alle fantomen en alle meters om te achterhalen of er een significant verschil tussen TcB-meetwaarden wordt veroorzaakt door de verschillende bilirubinemeters. Er is voor een F-toets gekozen, omdat in de analyse meerdere groepen vergeleken moeten worden, namelijk de gemiddelde meetwaarden van de dertien verschillende meters. Hierbij is uitgegaan van een normale verdeling van alle gemiddelde meetwaarden. Bij deze toets was de nulhypothese: alle meters zijn gelijk. De alternatieve hypothese hierbij was: tenminste twee meters verschillen significant van elkaar. Wanneer de nulhypothese is verworpen, is er per fantoom een ANOVA met Tukey's HSD toets uitgevoerd om te achterhalen welke meters van elkaar verschillen. Hierbij is per fantoom iedere meter met alle andere meters individueel vergeleken, met telkens als nulhypothese: de twee vergeleken meters zijn gelijk. De alternatieve hypothese daarbij was: de twee meters verschillen significant van elkaar. Vervolgens is onderzocht of alle meters bij minstens één fantoom significant van elkaar verschillen. Bij de statistische analyses is een significantieniveau van 5% gehanteerd.

#### 2.5.3 Resultaten

In figuur 2.4 zijn de gemiddelde TcB-meetwaarden van alle meters uitgezet tegen de absorptiecoëfficiënt van de fantomen bij 450 nm. De eerste zes metingen op fantoom 3 die gedaan zijn met bilirubinemeter B3601005 uit het MST zijn niet meegenomen bij het berekenen van het gemiddelde van deze meter. Er is voor elke meter een logaritmische trendlijn toegevoegd. In tabel 6.1 zijn de formules van die trendlijnen weergegeven. In die tabel is te zien dat de  $R_a^2$  voor alle trendlijnen groter is dan 0,99. De  $R_a^2$  is een maat voor hoe goed de trendlijn overeenkomt met de meetpunten. Daarnaast zijn in tabel 2.5 de standaarddeviaties van elke meter per fantoom weergegeven. In die tabel is te zien dat de standaarddeviatie van de metingen varieert tussen 0 en 3,93  $\mu$ mol/L. Vooral de meters met serienummer 3201594, B3601081 en 3201377 hebben een hoge standaarddeviatie. In figuur 2.4 is te zien dat het verschil tussen de TcB-meetwaarden van de verschillende meters steeds groter wordt bij toenemende absorptieoëficiënt. Bij fantoom 1 ( $\mu_a$ (450 nm) = 0,54 mm<sup>-1</sup>) is het maximale verschil tussen twee metingen van verschillende transcutane bilirubinemeters 3,48  $\mu$ mol/L. Bij fantoom 6 ( $\mu_a$ (450 nm) = 3,28 mm<sup>-1</sup>) is dit maximale verschil 61,29  $\mu$ mol/L.



Figuur 2.4: TcB-meetwaarden van de dertien transcutane bilirubinemeters bij de zes fantomen, uitgezet tegen de absorptiecoëfficiënt van de fantomen bij 450 nm. Standaarddeviaties van de meters per fantoom zijn te vinden in tabel 2.5.

Serienummer	а	b	с	$R_a^2$
B3601005	56,75237	98,6388	-0,01659	0,99744
B3601050	69,83599	95,83432	0,06038	0,9988
3201007	81,1806	82,17756	0,17102	0,99742
3201377	88,0042	75,99881	0,2134	0,99798
B3601081	82,10492	118,64882	0,03194	0,99808
B3601086	74,97443	113,52724	0,02639	0,99872
B3601107	69,31562	112,68147	0,003	0,99833
B3601137	57,51974	119,24012	-0,07159	0,99719
3201594	84,32623	75,48647	0,21593	0,99582
B3601390	34,28629	103,50314	-0,16375	0,99116
B3601402	22,45601	116,82302	-0,27204	0,99499
3202711	36,8554	107,30502	-0,15712	0,99395
3203135	27,23581	120,94114	-0,24566	0,99475

**Tabel 2.4:** Logaritmische trendlijnen van de TcB-metingen, waarbij  $TcB = a + b * ln(\mu_a(450) + c)$ in  $\mu$ mol/L.  $R_a^2$  = adjusted R-squared.

Serienummer	F1	F2	F3	F4	F5	F6
B3601086 (JM-105)	0	0,77	0,56	1,31	0,85	1,61
B3601137 (JM-105)	0	0,51	0,60	2,20	0,68	1,34
B3601107 (JM-105)	0	0	0,56	0,85	1,43	1,91
3201594 (JM-103)	0	0,72	1,29	1,58	2,93	2,80
B3601081 (JM-105)	0,51	0,57	1,08	3,92	2,71	2,14
B3601005 (JM-105)	0	0,61	1,02	0,94	1,56	0,78
3201377 (JM-103)	0,98	3,04	1,07	2,14	1,04	1,26
3201007 (JM-103)	0	0,60	1,45	1,31	1,62	1,39
B3601050 (JM-105)	0	0,40	0	0,48	0,36	0,51
B3601390 (JM-105)	0	0,40	1,12	1,11	0,74	0,74
B3601402 (JM-105)	0	0,22	0,48	0,74	0,80	1,11
3202711 (JM-103)	0	0,58	0,70	0,85	1,37	1,24
3203135 (JM-103)	0	0,50	0,59	1,76	0,83	0,51

**Tabel 2.5:** Standaarddeviaties van alle dertien meters per fantoom in  $\mu$ mol/L. F = fantoom.

In figuur 2.5 zijn de verschillende meters per ziekenhuis weergegeven. Hierin is te zien hoe meters, gemeten op exact dezelfde fantomen, zich onderling verhouden. In het Isala is het maximale verschil tussen twee metingen van verschillende bilirubinemeters 55,62  $\mu$ mol/L bij fantoom 6. In het Spaarne Gasthuis is dit 8,43  $\mu$ mol/L en in het Universitair Medisch Centrum Groningen is dit 17,86  $\mu$ mol/L.



Figuur 2.5: TcB-meetwaarden van de transcutane bilirubinemeters bij de zes fantomen, uitgezet tegen de absorptiecoëficiënt van de fantomen bij 450 nm, ingedeeld per ziekenhuis.

In figuur 2.6 zijn de meters per type weergegeven. Het maximale verschil tussen twee metingen van verschillende Dräger JM-103 meters is 20,86  $\mu$ mol/L bij fantoom 2 ( $\mu_a$ (450 nm) = 0,97 mm<sup>-1</sup>). Bij de Dräger JM-105 meters is dit 61,29  $\mu$ mol/L bij fantoom 6. In de figuur is te zien dat het verloop van de grafieken verschilt tussen beide typen. De meetwaarden bij het type JM-105 liggen verder uit elkaar bij hogere absorptiecoëfficiënten, zoals ook te zien was in figuur 2.4. Bij het type JM-103 lijken de meetwaarden bij hogere absorptiecoëfficiënten juist dichter bij elkaar te liggen.



Figuur 2.6: TcB-meetwaarden van de transcutane bilirubinemeters bij de zes fantomen, uitgezet tegen de absorptiecoëficiënt van de fantomen bij 450 nm, weergegeven per type.

In figuur 2.7 zijn gemiddelde TcB-meetwaarden van alle meters uitgezet tegen de verwachte TcBmeetwaarden van meter B3601086, waarbij lineaire trendlijnen zijn toegevoegd. De formules van die trendlijnen zijn in bijlage 6.2 weergegeven. Hierin is te zien dat de meetwaarden van meter B3601086 iets onder de verwachte TcB-meetwaarden van die meter liggen. Behalve de meetwaarden van meter B3601081, liggen alle meetwaarden onder de verwachte TcB-meetwaarden van meter B3601086.



Figuur 2.7: TcB-meetwaarden van de transcutane bilirubinemeters bij de zes fantomen, uitgezet tegen de verwachte TcB-meting.

Uit de F-toets is gebleken dat bij elk fantoom tenminste één van de meters significant verschilt van de rest. Vervolgens is uit de Tukey's HSD toetsen gebleken dat bij elk fantoom sommige meters niet significant van elkaar verschillen, maar dat alle meters bij minstens drie fantomen wel significant van elkaar verschillen. De SPSS output van de F-toets en Tukey's HSD toetsen is terug te vinden in bijlage 6.3.

In figuur 2.8 is per fantoom te zien wat het verschil in gemiddelde TcB-meetwaarde was tussen de eerste en laatste metingen met meter B3601086. Dit geeft weer of de optische eigenschappen van elk fantoom veranderd zijn gedurende de metingen met alle meters. Er zat bij de fantomen 2 t/m 6 een tijdsduur van 11 minuten tussen deze eerste en laatste meting. Bij het eerste fantoom zat er 17 minuten tussen, omdat er gewacht moest worden totdat de vijfde meter beschikbaar was. In deze tijd is met alle andere meters gemeten op de fantomen. In de figuur is te zien dat de TcB-waarde bij de fantomen 3 tot en met 6 maximaal 2,89  $\mu$ mol/L is toegenomen, terwijl de TcB-waarde bij fantoom 2 juist is afgenomen.



**Figuur 2.8:** Verschil in  $\mu$ mol/L tussen de eerste gemiddelde TcB-meting en de laatste gemiddelde TcB-meting met de JM-105 (B3601086) per fantoom.

#### 2.5.4 Conclusie

Uit de resultaten kan geconcludeerd worden dat metingen met de Dräger bilirubinemeters niet reproduceerbaar zijn. Alle dertien meters verschilden bij tenminste drie fantomen significant van elkaar. Het maximale verschil tussen de gemeten TcB-waarden is bij de Dräger JM-105 meters 61,29  $\mu$ mol/L bij  $\mu_a$ (450nm) = 3,28 mm<sup>-1</sup>. Bij de Dräger JM-103 meters is dat 20,86  $\mu$ mol/L bij  $\mu_a$ (450 nm) = 0,97 mm<sup>-1</sup>. Wanneer de verschillen die veroorzaakt kunnen worden door het maken van verschillende fantomen en het bevestigen van verschillende Tegaderm<sup>TM</sup>-plakkers daarvan worden afgetrokken, is het maximale verschil in TcB-waarden tussen de dertien meters 56,44  $\mu$ mol/L.

#### 2.5.5 Discussie

#### De metingen

In het UMCG was uit twee buizen ongeveer 1,5 mL 25% IL20% gelekt. Deze buizen zijn gebruikt voor het maken van fantoom 1 en 2. De verwachte TcB-meting bij fantoom 1 is hierbij niet veranderd, maar de verwachte TcB-meting bij fantoom 2 is van 82  $\mu$ mol/L toegenomen tot 85  $\mu$ mol/L. Tijdens het analyseren van de resultaten is onderzocht of het buiten beschouwing laten van fantoom 2 van invloed is op het verkregen verband tussen de TcB-meting en de absorptiecoëfficiënt bij 450 nm. Dit bleek niet zo te zijn, dus is besloten om het fantoom mee te nemen in de resultaten.

Met elke meter is op elk fantoom 21 keer gemeten. Opvallend hierbij was dat vooral de meters met serienummer 3201594, B3601081 en 3201377 een hoge standaarddeviatie hadden. Het vermoeden ontstond dat bilirubinemeters waarvan de tip meer beschadigd was, een hogere standaarddeviatie hadden. Omdat het aannemelijk is dat er een verband is tussen de leeftijd van de meters en de mate van beschadigingen, is de correlatie tussen leeftijd van de meters en standaarddeviatie onderzocht. Van slechts vijf meters was de installatiedatum in het ziekenhuis bekend. Deze vijf meters zijn in figuur 2.9 weergegeven met hun gemiddelde standaarddeviatie uitgezet tegen hun leeftijd in maanden.



Figuur 2.9: De gemiddelde standaarddeviatie van vijf bilirubinemeters, berekend over zes fantomen, uitgezet tegen het aantal maanden sinds installatie in het ziekenhuis (leeftijd).

Deze figuur is moeilijk te interpreteren, omdat er ten eerste verschillende typen meters in weergegeven zijn. Ten tweede is de leeftijd bepaald aan de hand van de installatiedatum van de meters in het ziekenhuis. Dat de meters op de installatiedatum zijn geïnstalleerd in het ziekenhuis wil niet zeggen dat de meters toen daadwerkelijk in gebruik zijn genomen. Ook kan het zijn dat de meters al eerder in een ander ziekenhuis of onderzoeksopzet gebruikt zijn. Verder zegt niet alleen de leeftijd iets over de mate van beschadiging, maar ook hoe intensief de meters gebruikt zijn. Er wordt verwacht dat de variatie in standaarddeviatie geen invloed heeft gehad op de resultaten van dit onderzoek, omdat het gemiddelde van de 21 metingen is meegenomen in de analyse van de resultaten.

Uit figuur 2.8 bleek dat de gemiddelde TcB-meting bij vier fantomen iets was toegenomen tijdens de metingen op die specifieke fantomen, terwijl de gemiddelde TcB-meting bij fantoom 2 juist was afgenomen. Omdat het toegenomen verschil maximaal 2,89  $\mu$ mol/L was en er bij fantoom 2 juist een afname was, wordt er aangenomen dat de optische eigenschappen van de fantomen nauwe-lijks veranderen in de tijd dat er gemeten wordt op één fantoom.

Bij het vergelijken van de gemeten TcB-waarden met de verwachte TcB (figuur 2.7) bleek dat de gemeten TcB van meter B3601086 niet precies overeenkomen met de verwachte waarden van deze meter, zoals berekend aan de hand van metingen door Van Erk [26]. Het verschil in meetwaarden zou verklaard kunnen worden door onnauwkeurigheid in het bepalen van de absorptiecoëfficiënt van de inkt stockoplossingen.

#### Reproduceerbaarheid van de meters

Verschillen in gemeten TcB-waarden tussen bilirubinemeters kunnen veroorzaakt zijn door verschillen tussen de afzonderlijke meters, zoals kleine verschillen in de bouw van de meters. Wanneer bijvoorbeeld de oppervlakte van de sensoren in de tip van de meters verschilt, leidt dit tot een verschil in de hoeveelheid licht die opgevangen wordt. Daarnaast kan de hoeveelheid licht die wordt uitgezonden door de bilirubinemeters, en de bijdrage van verschillende golflengten hierin, verschillen. Doordat de golflengte invloed heeft op de mate van absorptie en verstrooiing, kan de gemeten TcB-waarde verschillen. Ten slotte bestaat er ook het vermoeden dat de meters verschillen door verschillende gradaties van beschadigingen aan de tip van de meters.

De resultaten in verschillende ziekenhuizen zijn op verschillende momenten verkregen, waardoor er ook andere factoren zijn die invloed kunnen hebben gehad op de gemeten verschillen. Ten eerste is er voor ieder ziekenhuis een nieuwe reeks fantomen gemaakt en is er op elke meter een andere Tegaderm<sup>TM</sup>-plakker bevestigd. Om rekening te kunnen houden met het verschil dat is veroorzaakt door verschillen tussen fantomen en Tegaderm<sup>TM</sup>-plakkers, is de reproduceerbaarheid hiervan onderzocht. Deze onderzoeken worden in de secties 2.3 en 2.4 beschreven. De reproduceerbaarheid van het maken van de fantomen is bepaald aan de hand van acht fantomen met de beoogde optische eigenschappen van het standaardfantoom. Hierdoor is er alleen iets te concluderen over de reproduceerbaarheid van het standaardfantoom (fantoom 3). Echter, er wordt verwacht dat deze resultaten gelden voor de reproduceerbaarheid van alle zes de fantomen, omdat deze gemaakt zijn op eenzelfde manier met dezelfde materialen.

Ten tweede kunnen verschillen in omgevingsfactoren in de verschillende ziekenhuizen van invloed zijn geweest op de resultaten. In het MST waren er bijvoorbeeld veel stofdeeltjes in het fantoom aanwezig, die afkomstig waren van de tissue die daar gebruikt werd om de opstelling droog te maken. Daarom is ook het verschil tussen de transcutane bilirubinemeters per ziekenhuis geanalyseerd. De grote verschillen tussen de meters die te zien zijn in figuur 2.5a en de verschillen die tussen de meters uit het Spaarne Gasthuis en het UMCG gevonden zijn, ondersteunen de bewering dat de verschillen tussen alle meters niet toe te schrijven zijn aan het verschil in meetlocatie en -tijdstip.

Ten slotte kunnen de gemeten verschillen veroorzaakt zijn door verschillen tussen de twee typen meters die gebruikt zijn. Om deze reden zijn de verkregen resultaten ook per type meter geanalyseerd en weergegeven in figuur 2.6. Opvallend is dat het maximale verschil dat gemeten is, tussen twee meters van het type JM-105 is gemeten. De JM-105 is een nieuwere versie van de JM-103, maar beide typen zouden hetzelfde werkingsmechanisme moeten hebben [24]. Toch lijkt er een verschil te zijn tussen de twee typen. Aangezien er slechts met een beperkt aantal meters per type is gemeten, zou dit verschil ook door toeval veroorzaakt kunnen zijn.

#### Klinische betekenis van de resultaten

In de kliniek wordt de transcutane bilirubinemeter slechts gebruikt als screeningsmethode. Wanneer de gemeten TcB-waarde minder dan 50  $\mu$ mol/L onder de fototherapiegrens voor een individuele patiënt ligt, wordt de TSB bepaald door middel van een bloedafname [5]. Uit de resultaten van dit onderzoek is gebleken dat bij hogere absorptiecoëfficiënten de gemeten TcB-waarden van verschillende meters steeds verder uit elkaar liggen. Hieruit valt af te leiden dat de transcutane bilirubinemeters onderling minder overeenkomen naarmate de bilirubineconcentratie hoger is. Deze hogere bilirubineconcentraties zijn juist belangrijk om nauwkeurig te meten, omdat hogere concentraties schadelijk kunnen zijn voor pasgeborenen [4].

Uit het onderzoek is geconcludeerd dat het maximale verschil tussen verschillende bilirubinemeters, rekening houdend met de reproduceerbaarheid van de fantomen en het Tegaderm<sup>™</sup>, 56,44  $\mu$ mol/L is bij  $\mu_a$ (450 nm) = 3,28 mm<sup>-1</sup>. In de handleiding van de Dräger JM-105 staat vermeld dat de onnauwkeurigheid van de bilirubinemeter  $\pm$  25,5  $\mu$ mol/L is bij een zwangerschapsduur boven de 35 weken en  $\pm$  27,4  $\mu$ mol/L bij een zwangerschapsduur van 24 tot 34 weken. Dit is bepaald door de gemeten TcB-waarde op pasgeborenen te vergelijken met de TSB. Daarbij is niet aangegeven of de onnauwkeurigheid is bepaald met behulp van één of meerdere bilirubinemeters.

Wanneer Dräger de onnauwkeurigheid heeft bepaald met behulp van één bilirubinemeter, betekent dit dat de onnauwkeurigheid van één meter opgeteld moet worden bij het verschil dat gevonden is tussen de verschillende bilirubinemeters. In het uiterste geval zou een bepaalde bilirubinemeter de TSB met maximaal 56,44 + 27,4 = 83,84  $\mu$ mol/L kunnen onderschatten. Dit zou ervoor kunnen zorgen dat er bij een individuele patiënt te laat of zelfs niet gestart wordt met fototherapie. Daarom is in dit geval de gebruikte marge van 50  $\mu$ mol/L onder de fototherapiegrens niet meer veilig.

Als Dräger met behulp van meerdere bilirubinemeters de onnauwkeurigheid heeft bepaald, is hierbij ook het verschil tussen meters meegenomen. Dit betekent dat het verschil tussen de bilirubinemeters maximaal 51  $\mu$ mol/L (2\*25,5) zou moeten zijn. Uit de resultaten van dit onderzoek blijkt echter dat deze waarde hoger ligt. In dat geval zou een bepaalde bilirubinemeter uit dit onderzoek de TSB met maximaal 28,22  $\mu$ mol/L (56,44/2) onderschatten. De marge van 50  $\mu$ mol/L zou wel een veilige marge zijn, maar toch kan de onnauwkeurigheid van de meters ook in dat geval gevolgen hebben voor de individuele patiënt. In de kliniek kan het betekenen dat de ene transcutane bilirubinemeter wel een indicatie geeft tot het bepalen van een TSB, waar een andere meter deze indicatie niet geeft. Hierdoor zal bij gebruik van de ene meter vaker een TSB bepaald worden dan bij de andere meter.

#### Aanbevelingen voor vervolgonderzoek

Ondanks het feit dat er een duidelijke conclusie uit de resultaten getrokken kan worden, zijn er verschillende manieren waarop onderzoek naar de reproduceerbaarheid van transcutane bilirubinemeters uitgebreid zou kunnen worden. Ten eerste is in dit onderzoek slechts met dertien meters in vier ziekenhuizen getest. Het zou nuttig zijn om meer meters uit verschillende ziekenhuizen te testen, om zo tot een breder toepasbare conclusie over de reproduceerbaarheid van de bilirubinemeters te komen. Verder is dit onderzoek enkel uitgevoerd met de Dräger JM-103 en JM-105. Er zijn andere transcutane bilirubinemeters op de markt die getest moeten worden, zoals de BiliChek<sup>®</sup> van Philips.

Daarnaast loopt het bereik van TcB-waarden die in dit onderzoek gemeten is tot ongeveer 220  $\mu$ mol/L. De grens voor fototherapie ligt echter boven de 220  $\mu$ mol/L bij een zwangerschapsduur van minimaal 35 weken, een leeftijd van minimaal 48 uur en afwezigheid van een risicofactor [5]. Omdat het verschil tussen de meters mogelijk verder toeneemt bij hogere TcB-waarden is verder onderzoek nodig naar het verschil tussen bilirubinemeters in het volledige klinische bereik.

Een andere beperking van het uitgevoerde onderzoek is dat de verschillende meters in verschillende ziekenhuizen zijn getest. Hierdoor kunnen verschillen in omgevingsfactoren en verschillen in fantomen van invloed zijn geweest op de uitkomsten. Over de invloed van omgevingsfactoren op de metingen, maar ook op het fantoom is weinig bekend. Als hier meer onderzoek naar gedaan wordt, kan met deze factoren rekening worden gehouden tijdens de metingen en bij de analyse van de resultaten. Voor de verschillen in fantomen kan gecorrigeerd worden door te testen hoe reproduceerbaar het maken van deze fantomen is, zoals beschreven in sectie 2.2. Idealiter zou dit niet nodig zijn en zouden alle meters op dezelfde locatie op dezelfde fantomen getest worden. Bij dit onderzoek was dat niet mogelijk, omdat de bilirubinemeters niet meegenomen konden worden uit de ziekenhuizen in verband met de verzekering ervan.

Uit dit onderzoek is te concluderen dat er een verschil is tussen de transcutane bilirubinemeters. Het is echter niet bekend door welk mechanisme de meters verschillen. Een aantal mogelijkheden zijn eerder in de discussie genoemd. Er wordt aanbevolen om onderzoek te doen naar de oorzaak achter de gemeten verschillen en een eventuele manier om de verschillen tussen de meters te voorkomen.

Ten slotte ontstond bij de experimenten het vermoeden dat er een verband was tussen de mate van beschadiging van de meters en de standaarddeviatie. In figuur 2.9 is de standaarddeviatie van een klein aantal meters uitgezet tegen het aantal maanden sinds installatie. Om een verband te kunnen bevestigen of uit te sluiten, zal er gebruik gemaakt moeten worden van data van een groter aantal meters. Zoals eerder genoemd is het aantal maanden sinds installatie geen goede maat voor beschadiging. Het is daarom aan te bevelen op een andere manier onderzoek te doen naar de correlatie tussen beschadiging en standaarddeviatie. Dit kan bijvoorbeeld door de gebruiksfrequentie van de meters te achterhalen, of door de beschadigingen te kwantificeren door te bepalen welk deel van het oppervlak beschadigd is.

## 3 Model

## 3.1 Inleiding

Zoals in sectie 1.5 beschreven, heeft Hofmeijer een model ontwikkeld dat de TSB schat aan de hand van gemeten TcB-waarden op het voorhoofd, het sternum en de heup. Het doel van dit model is om invasieve TSB-bepalingen te kunnen vervangen door metingen met de transcutane bilirubinemeter. Om het model dichter bij toepassing in de kliniek te brengen, zijn aan het model van Hofmeijer een aantal aanpassingen gedaan en is vervolgens de klinische waarde beoordeeld.

Ten eerste is het *machine learning* model van Hofmeijer getraind met een grotere dataset, om de betrouwbaarheid van het model te vergroten. Vervolgens is voorkomen dat dezelfde patiënt in zowel de traindata als de testdata voorkomt. Wanneer het model in de kliniek wordt toegepast, zal er al een getraind model beschikbaar moeten zijn. Dat gegevens van patiënten zowel in de train- als de testdata voorkomen, is daarom klinisch niet realistisch. Als aanvulling op de genoemde aanpassingen is onderzocht of het gebruik van meetdata van verschillende transcutane bilirubinemeters van invloed is op het model.

Om te onderzoeken of het model in de kliniek van toegevoegde waarde kan zijn, is de klinische waarde van het model gekwantificeerd. Dit is gedaan door te bepalen welk percentage TSBbepalingen met behulp van het verbeterde model veilig vervangen kan worden door TcB-metingen.

## 3.2 Methode

Nadat een bepaalde aanpassing aan het model is gedaan, zijn er opnieuw tien iteraties van de train- en testfase uitgevoerd. Voor het trainen en testen van het model zijn ook de dataparen meegenomen waarbij de pasgeborenen fototherapie ondergingen. Uit de resultaten van Hofmeijer bleek namelijk dat het model kan corrigeren voor de aanwezigheid van fototherapie [27]. Per iteratie zijn de volgende uitkomstwaarden bepaald: RMSE's uit de train- en testfase, nauwkeurigheid, sensitiviteit en specificiteit. De RMSE geeft aan hoe nauwkeurig het model de TSB kan voorspellen aan de hand van patiëntgegevens en de bekende TcB-waarden. Een lagere RMSE betekent een meer nauwkeurige schatting. De nauwkeurigheid, sensitiviteit en specificiteit zijn bepaald door zowel met behulp van de voorspelde TSB als met de werkelijke TSB aan de hand van het protocol te beslissen of er gestart moet worden met fototherapie en die besluiten met elkaar te vergelijken. Met behulp van de sensitiviteit en specificiteit is een ROC-curve gemaakt. Uiteindelijk is de RMSE gebruikt om de klinische waarde te evalueren, omdat daarmee bepaald kan worden in welke mate de huidige marge van 50  $\mu$ mol/L onder de fototherapiegrens kan worden verlaagd. De sensitiviteit en specificiteit zijn minder van belang bij het evalueren van de klinische waarde, omdat het model op dit moment nog niet alle TSB-bepalingen kan vervangen.

### 3.2.1 Vergroten dataset

Hofmeijer heeft haar model getraind en getest met behulp van meetgegevens van 44 patiënten. In dit onderzoek is eerst nieuwe meetdata uit het Isala toegevoegd aan de dataset. Hierna is het model van Hofmeijer, getraind met 80% van de dataparen van de 44 oorsprongelijke patiënten, getest met de nieuw toegevoegde data. Het aangepaste MATLAB script is terug te vinden in bijlage 6.4.1. Het MATLAB model is vervolgens opnieuw getraind en getest met alle beschikbare data, waarbij net als bij het model van Hofmeijer 80% van de data is gebruikt om het model te trainen en 20% om het model te testen. De uitkomstwaarden hiervan zijn met elkaar vergeleken. Bij de rest van de aanpassingen aan het model is gebruik gemaakt van de vergrote dataset.

#### 3.2.2 Voorkomen van patiëntoverlap

In het bestaande model van Hofmeijer wordt 80% van de dataparen gebruikt om het model te trainen, en 20% om het model te testen. Omdat één enkele patiënt meedere dataparen kan hebben kunnen gegevens van bepaalde patiënten in beide sets terecht komen. Hierdoor ontstaat het risico dat het model patiëntgegevens herkent en gebruikt om de TSB te schatten. Om te voorkomen dat verschillende dataparen van één patiënt in zowel de train- als de testset terechtkomen, is in MATLAB de definitie van de train- en testset zodanig aangepast dat er getraind is met 80% van de patiënten in plaats van 80% van de dataparen. De overige 20% van de patiënten is gebruikt om het model te testen. Het aangepaste MATLAB script is terug te vinden in bijlage 6.4.2.Bij de rest van de aanpassingen aan het model is ook op deze manier patiëntoverlap voorkomen. Om de kans op bias door het gebruik van meerdere dataparen per patiënt nog verder te verkleinen, is het model vervolgens getraind en getest met slechts één willekeurig datapaar per patiënt. Het aanpaste MATLAB script daarvan is terug te vinden in bijlage 6.4.3. De uitkomstwaarden zijn vergeleken met de uitkomstwaarden die verkregen zijn na alleen het vergroten van de dataset.

#### 3.2.3 Gebruik van één bilirubinemeter

Het model is opnieuw getraind en getest met data die verkregen is door slechts één transcutane bilirubinemeter te gebruiken. Dit is gedaan door alleen data te gebruiken die is verkregen nadat één van de twee bilirubinemeters van de afdeling werd gehaald. De uitkomstwaarden zijn vergeleken met de uitkomstwaarden die verkregen waren nadat de dataset vergroot was en patiëntoverlap was voorkomen.

#### 3.2.4 Kwantificatie klinische waarde

Voor het kwantificeren van de klinische waarde van het model wordt de gemiddelde RMSE van de testfase van het, uit de resultaten verkregen, meest relevante model gebruikt. De gemiddelde RMSE wordt verdubbeld en gebruikt als nieuwe marge onder de fototherapiegrens (veiligheidsmarge), voordat er een TSB-bepaling gedaan dient te worden. Dit is zo gedaan, omdat de huidige veiligheidsmarge van 50  $\mu$ mol/L, die in ziekenhuizen gehanteerd wordt, is gekozen aan de hand van de door Dräger bepaalde afwijking van  $\pm$  25,5  $\mu$ mol/L. De nieuwe veiligheidsmarge is vergeleken met de huidige toegepaste marge van 50  $\mu$ mol/L. Vervolgens is met behulp van de dataset, die ook gebruikt is voor het trainen en testen van het model, bepaald welk percentage van de huidige TSB-bepalingen vervangen kan worden door TcB-metingen. Hiervoor is per patiënt aan de hand van de veiligheidsmarge een kritische TcB-waarde bepaald. Daarna zijn alle TcB-metingen vergeleken met die kritische TcB-waarde boven de kritische TcB-waarde ligt. Met deze aanname is bepaald hoeveel minder TSB-bepalingen er gedaan dienen te worden wanneer de marge van 50  $\mu$ mol/L omlaag kan worden gebracht. Het MATLAB script is terug te vinden in bijlage 6.4.4.

### 3.3 Resultaten

In de tabellen 3.1 en 3.2 zijn de RMSE's van de train- en testfase, nauwkeurigheid, sensitiviteit en specificiteit van de verschillende aangepaste modellen weergegeven. Per aanpassing zal worden besproken welke waarden uit die tabellen opvallend zijn. De ROC-curves zijn terug te vinden in bijlage 6.5. Ten slotte zijn in figuur 3.1 de RMSE's van de testsets van de verschillende modelaanpassingen per iteratie in een grafiek weergegeven.

	RMSE trainset (µmol/L)	RMSE testset (µmol/L)
Model Hofmeijer	17,90 (17,29 - 18,78)	20,35 (19,87 - 20,63)
Vergrote dataset	19,83 (19,19 - 20,38)	19,82 (18,42 - 21,24)
Geen patiëntoverlap	19,94 (19,19 - 20,27)	21,67 (18,00 - 23,43)
Eén datapaar	22,08 (19,17 - 25,63)	22,13 (12,64 - 35,96)
Eén meter	19,45 (18,63 - 20,23)	18,99 (16,57 - 21,40)

 

 Tabel 3.1: De gemiddelde RMSE's van zowel de test- als de trainset zijn weergegeven per modelaanpassing. Ook de minimale en maximale RMSE zijn toegevoegd.

	Nauwkeurigheid (%)	Sensitiviteit (%)	Specificiteit (%)
Model Hofmeijer	82,2 (81,7 - 83,0)	88,3 (86,4 - 90,9)	99,9 (99,3 - 100,0)
Vergrote dataset	85,1 (77,2 - 91,3)	89,4 (80,0 - 94,1)	97,0 (93,1 - 100,0)
Geen patiëntoverlap	75,7 (70,4 - 82,4)	87,7 (77,1 - 100,0)	92,2 (86,4 - 96,5)
Eén datapaar	78,6 (64,3 - 85,7)	86,3 (33,3 - 100,0)	98,8 (87,5 - 100,0)
Eén meter	82,5 (76,9 - 88,5)	91,6 (78,3 - 100,0)	93,2 (88,2 - 100,0)

 

 Tabel 3.2: De gemiddelde nauwkeurigheid, sensitiviteit en specificiteit zijn weergegeven per modelaanpassing. Ook de minimale en maximale waarden zijn toegevoegd.



Figuur 3.1: RMSE's van de testsets van de verschillende modelaanpassingen per iteratie, met de onnauwkeurigheid van de Dräger JM-105 als referentie [25].

#### 3.3.1 Vergroten dataset

De dataset is aangevuld tot 70 patiënten. Alle pasgeborenen geïncludeerd in deze studie hadden een gemiddeld geboortegewicht (gemiddelde  $\pm$  standaard afwijking) van 1492.8 ( $\pm$ 469,1) gram met een bereik van 675 - 3280 gram. Het geboortegewicht van één pasgeborene was niet bekend. De gemiddelde zwangerschapsduur was 30.6 ( $\pm$ 1.8) weken met een bereik van 27.7- 35.7 weken. De TSB was gemiddeld 177.6 ( $\pm$ 53.6)  $\mu$ mol/L met een bereik van 7- 277  $\mu$ mol/L. Van de patiënten heeft 71.4% fototherapie ondergaan. Karakteristieken van alle pasgeborenen zijn samengevat in tabel 3.3.

Populatiegrootte	n = 70	
Geslacht	Man(41) Vrouw (29)	
Etniciteit	Kaukasisch (51) Afrikaans (1) Aziatisch (1)	
	Latijns-Amerikaans (1) Anders (2) Onbekend (14)	
TcB*	n = 932	
TSB	n = 677	
TcB-TSB-paren	n = 462	

\*Alle TcB-metingen op verschillende locaties op één moment worden gezien als één TcB-meting

**Tabel 3.3:** Patiënt karakteristieken en populatie groottes.

In de tabellen 3.1 en 3.2 staan de uitkomstwaarden die verkregen zijn bij de verificatie van het model van Hofmeijer met 80% van de dataparen van de 44 oorspronkelijke patiënten als trainset en de 26 toegevoegde patiënten als testset. Daarin staan ook de uitkomstwaarden die verkregen zijn nadat het model is getraind met 80% en getest met 20% van alle dataparen van de 70 patiënten. In tabel 3.1 is te zien dat de RMSE van de trainfase toeneemt en die van de testfase afneemt. Wel is het bereik van de RMSE van de testfase toegenomen. De RMSE van de train- en testfase zijn na het vergroten van de dataset bijna exact gelijk. In tabel 3.2 is te zien dat de nauwkeurigheid, sensitiviteit en specificiteit ongeveer gelijk blijven.

#### 3.3.2 Voorkomen van patiëntoverlap

De uitkomstwaarden die verkregen zijn nadat het model is getraind met data van 80% van de patiënten en getest met data van 20% van de patiënten is terug te vinden in de tabellen 3.1 en 3.2. Hierin is te zien dat de RMSE's iets zijn toegenomen ten opzichte van de RMSE's die verkregen zijn na het vergroten van de dataset, voornamelijk de RMSE van de testfase. De nauwkeurigheid, specificiteit en sensitiviteit zijn afgenomen.

Ook de uitkomstwaarden die verkregen zijn nadat het model is getraind met één datapaar per patiënt zijn weergegeven in tabellen 3.1 en 3.2. Hierin is te zien dat de RMSE's nog meer zijn toegenomen. Daarnaast hebben alle uitkomstwaarden een groot bereik.

#### 3.3.3 Gebruik van één bilirubinemeter

In tabellen 3.1 en 3.2 is te zien dat de RMSE's die verkregen zijn nadat het model is getraind en getest met metingen door één meter zijn afgenomen ten opzichte van de RMSE's die verkregen zijn na het vergroten van de dataset en het voorkomen van patiëntoverlap. De nauwkeurigheid, sensitiviteit en specificiteit zijn toegenomen.

#### 3.3.4 Kwantificatie klinische waarde

De veiligheidsmarge is bepaald aan de hand van de gemiddelde RMSE van de testfase van het model waarbij één bilirubinemeter is gebruikt en patiëntoverlap is voorkomen. Door deze RMSE van 18,99  $\mu$ mol/L te verdubbelen, is er een nieuwe veiligheidsmarge van 37,98  $\mu$ mol/L onder de fototherapiegrens vastgesteld. Met deze nieuwe veiligheidsmarge bleek dat 19,64% van de TSB-bepalingen die op dit moment gedaan worden, met behulp van het model vervangen kunnen worden door TcB-metingen.

## 3.4 Conclusie

Het vergroten van de dataset en het gebruik van de data van één meter zorgen voor een meer betrouwbare en nauwkeurige schatting van de TSB. Het voorkomen van overlap in patiënten zorgt daarnaast voor een klinisch toepasbaar model. Het model dat gebruik maakt van de vergrote dataset, de data van één meter en patiëntoverlap voorkomt, heeft een RMSE in de testfase van 18,99  $\mu$ mol/L. Hierdoor kan de veiligheidsmarge van 50  $\mu$ mol/L naar 37,98  $\mu$ mol/L verlaagd worden. Met dit verbeterde model kan 19,64% van de huidige TSB-bepalingen voorkomen worden.

## 3.5 Discussie

In het Isala zijn voor het onderzoek steeds drie TcB-metingen op het voorhoofd, het sternum en de heup gedaan. In het model van Hofmeijer wordt van deze drie metingen steeds de mediaan genomen. Opvallend aan de data was dat drie opvolgende metingen soms tot wel 60  $\mu$ mol/L van elkaar verschilden. De mediaan kan dus erg afwijken van de werkelijke TcB-waarde. Het vermoeden is dat het hanteren van de meter erg belangrijk is voor de gemeten TcB-waarden.

### 3.5.1 Vergroten dataset

Dat de RMSE van de trainfase toeneemt na het vergroten van de dataset kan verklaard worden doordat het bij meer gegevens moeilijker is om een model goed bij alle gegevens te laten passen. De afgenomen gemiddelde RMSE van de testfase na het vergroten van de dataset wijst erop dat het model hierna een betere TSB-schatting kan maken. Het bereik van de RMSE van de testfase is echter na het vergroten van de dataset behoorlijk toegenomen, waardoor de maximale RMSE hoger is dan de gemiddelde RMSE van de testfase vóór het vergroten van de dataset. Dat dit bereik na het vergroten van de dataset groter is geworden, kan verklaard worden doordat het model van Hofmeijer elke iteratie is getest met dezelfde dataset, terwijl het aangevulde model elke iteratie is getest met een willekeurige 20% van de data. Het feit dat de RMSE van de train- en testfase na het vergroten van de dataset even groot zijn, betekent dat het model voor zowel de train- als de testset even goed werkt en dus voor verschillende datasets waarschijnlijk als even nauwkeurig kan worden beschouwd.

De nieuwe data is gebruikt om het model van Hofmeijer te testen, omdat hiermee het model ook voor grotere patiëntgroepen geverifieerd kan worden. Nadat vervolgens het model met 80% van de vergrote dataset is getraind, is het getest met 20% van de vergrote dataset. De testset voor het model op basis van de vergrote dataset is dus kleiner dan de testset voor het model van Hofmeijer. Daarom zijn de uitkomstwaarden van de twee modellen minder goed te vergelijken.

#### 3.5.2 Voorkomen van patiëntoverlap

Het is niet onverwacht dat de RMSE's zonder patiëntoverlap hoger zijn dan de RMSE's bij het model met patiëntoverlap. Bij overlap kan het model rekening hebben gehouden met gegevens van bepaalde patiënten die zowel in de train- als de testset voorkomen, waardoor het een betere voorspelling heeft kunnen doen. Dit betekent niet dat het model met overlap van patiënten betrouwbaarder is. Wanneer het model klinisch toegepast kan worden, is de trainfase namelijk afgerond. De patiënten waarop het model dan toegepast kan worden, zullen nooit hetzelfde zijn als de patiënten waarmee het model getraind is.

Dat de nauwkeurigheid, sensitiviteit en specificiteit zijn afgenomen ten opzichte van het model met slechts de vergrote dataset, gaat gepaard met een toegenomen RMSE van de testfase. Deze RMSE duidt namelijk een groter verschil tussen geschatte TSB en daadwerkelijke TSB aan, waardoor dus ook een minder goede voorspelling voor fototherapie gegeven wordt. Het verschil in RMSE tussen de train- en testfase kan verklaard worden doordat de train- en testset minder verge-lijkbaar zijn nadat overlap in patiënten voorkomen is.

Wanneer er maar één datapaar per patiënt wordt gebruikt, wordt het model met minder data getraind en getest. Dit kan dan ook verklaren waarom dit model een hogere RMSE heeft voor beide fases en dus minder nauwkeurig de TSB kan voorspellen. Het feit dat het bereik van alle uitkomstwaarden zo groot is, is ook te wijten aan de kleine hoeveelheid data. Voor dit onderdeel is het dus nuttig om extra patiënten toe te voegen.

#### 3.5.3 Gebruik van één bilirubinemeter

Het feit dat de uitkomstwaarden van het model waarbij de data van één meter wordt gebruikt zijn verbeterd ten opzichte van het model zonder patiëntoverlap, kan verklaard worden doordat de bilirubinemeters onderling verschillen.

#### 3.5.4 Kwantificatie klinische waarde

De dataset die gebruikt is om het percentage van de huidige TSB-bepalingen dat vervangen kan worden door TcB-metingen te bepalen, is in een onderzoekssetting op de NICU in het Isala verkregen. Hier is voor gekozen, omdat patiënten die op de NICU liggen een centrale lijn hebben waaruit direct de TSB bepaald kan worden. Bij deze patiënten hoeft dus geen extra bloedafname gedaan te worden om de TSB te bepalen. Het is voor patiënten op de NICU daarom ook niet relevant om gebruik te maken van transcutane bilirubinemeters in combinatie met het model. Het gebruik hiervan is wel gewenst op de *Medium* of *High Care* neonatale afdelingen. Bij patiënten op deze afdelingen moet namelijk wel bloed afgenomen worden om de TSB te bepalen. Bij het verkrijgen van de onderzoeksdata van de NICU is van tevoren vastgesteld hoe vaak de TcB-waarde en de TSB bij elke patiënt bepaald moesten worden. Deze dataset geeft dus geen goede weergave van hoe vaak TSB-bepalingen op een *Medium* of *High Care* afdeling zouden plaatsvinden. Daarnaast bestaat de dataset uit slechts 70 patiënten. Voor een representatiever en betrouwbaarder beeld moet de klinische waarde geëvalueerd worden met behulp van een grote dataset met informatie over het aantal bloedafnames op een *Medium* of *High Care* afdeling, en bijbehorende TcB-waarden.

## 4 Conclusie

Het doel van dit onderzoek was om te bepalen of en hoe het model van Hofmeijer in de kliniek kan worden toegepast. Om die vraag te kunnen beantwoorden, is de inter-device reproduceerbaarheid van verschillende Dräger transcutane bilirubinemeters van het type JM-103 en JM-105 onderzocht, zijn enkele verbeteringen aan het model van Hofmeijer gedaan en is de klinische waarde van het model gekwantificeerd.

Uit de fantoomstudie is gebleken dat de verschillende Dräger transcutane bilirubinemeters van het type JM-103 en JM-105 niet reproduceerbaar zijn. Het maximale verschil in gemeten TcB-waarde is 61,29  $\mu$ mol/L bij  $\mu_a$ (450 nm) = 3,28 mm<sup>-1</sup>. Het verschil dat hierbij veroorzaakt zou kunnen worden door het maken van verschillende fantomen en het bevestigen van verschillende Tegaderm<sup>TM</sup>-plakkers is ongeveer 4,85  $\mu$ mol/L. Als dat verschil wordt afgetrokken van het gemeten verschil, wordt het maximale verschil in TcB-waarde tussen de meters 56,44  $\mu$ mol/L bij  $\mu_a$ (450 nm) = 3,28 mm<sup>-1</sup>.

Er is gebleken dat het model van Hofmeijer verbeterd kan worden door de dataset die gebruikt is voor het trainen van het model uit te breiden, door patiëntoverlap te voorkomen tussen de trainen testset en door meetdata van slechts één transcutane bilirubinemeter te gebruiken. De RMSE van het verbeterde model is 18,99  $\mu$ mol/L. De klinische waarde van het model inclusief verbeteringen is gekwantificeerd door te bepalen welk percentage van de huidige TSB-bepalingen met behulp van dit model veilig vervangen kan worden door TcB-metingen. Daaruit bleek dat 19,64% van de huidige TSB-bepalingen vervangen kan worden.

## 5 Discussie en aanbevelingen

Het uiteindelijke doel van dit onderzoek is om de transcutane bilirubinemeters samen met het model in de kliniek toe te passen, om zo meer invasieve TSB-bepalingen veilig te vervangen door TcBmetingen. Echter, het model moet de TSB dan nauwkeurig kunnen voorspellen aan de hand van TcB-metingen van verschillende bilirubinemeters uit verschillende ziekenhuizen. Uit de fantoomstudie is gebleken dat de bilirubinemeters onderling verschillen. Het uiteindelijke model is gevormd met TcB-metingen van maar één bilirubinemeter. Hieruit is te concluderen dat het huidige model slechts bruikbaar is voor de bilirubinemeter waarmee het model getraind is en bilirubinemeters die overeenkomen met deze meter. Om de transcutane bilirubinemeters en het model algemeen te kunnen toepassen, moet het model op iedere locatie aangepast kunnen worden aan de daar gebruikte bilirubinemeter. Het zou daarom ideaal zijn als er van elke bilirubinemeter gemakkelijk een ijkwaarde berekend kan worden, die vervolgens ingevoerd kan worden in het model om daarmee het model bruikbaar te maken voor die specifieke bilirubinemeter. Het wordt aanbevolen om hier onderzoek naar te doen, zodat het model in elk ziekenhuis en met elke meter gebruikt kan worden om een betere voorspelling te doen van de TSB.

Het huidige model kan alleen gebruikt worden in MATLAB, waarbij het invoeren van de data tijdsintensief is. Voor toepassing in de kliniek zou het daarom gewenst zijn als het model verwerkt kan worden in een eenvoudige applicatie. In deze applicatie kunnen dan de TcB-waarden, geboortegewicht, zwangerschapsduur en leeftijd worden ingevoerd, waarna de applicatie de geschatte TSB geeft. Daarbij zou ook het protocol voor fototherapie meegenomen kunnen worden, waardoor de applicatie kan aangeven of er fototherapie gegeven moet worden of niet. Daarvoor moet eerst onderzoek gedaan worden naar hoe het protocol wordt toegepast in de kliniek, bijvoorbeeld wanneer er van dit protocol afgeweken wordt. Het is echter de vraag of het klinisch wenselijk is dat de applicatie aangeeft of er fototherapie gegeven moet worden, of dat artsen en verpleegkundigen dit zelf willen bepalen aan de hand van een voorspelde TSB. Om dat te onderzoeken, zouden interviews gedaan kunnen worden met neonatologen en verpleegkundigen.

Bij het invoeren van de data die gebruikt is voor het model is opgevallen dat drie opvolgende TcBmetingen soms vrijwel gelijk waren maar soms ook tot wel 60  $\mu$ mol/L van elkaar verschilden. Omdat de metingen verkregen zijn door verschillende zorgverleners, doet dit vermoeden dat het gebruik van de meters per zorgverlener verschilt. Als de meter onder een hoek wordt geplaatst, kan dat al leiden tot een afwijkende TcB-meting [25]. In de praktijk zou het verschil in gebruik gevolgen kunnen hebben voor de individuele patiënt. In het UMCG wordt bij elke patiënt drie keer gemeten en aan de hand van de hoogste gemeten waarde wordt besloten of er een TSB-bepaling moet plaatsvinden. Als er per toeval drie keer verkeerd wordt gemeten, kan een TcB-waarde onderschat worden, waardoor de TSB-bepaling te laat gedaan wordt en het starten van fototherapie uitgesteld wordt. Daarom is het ook belangrijk dat zorgverleners beter opgeleid worden in het gebruik van de meters, om in de toekomst meer TSB-bepalingen te kunnen vervangen door TcB-metingen.

## Referenties

- [1] Ansong-Assoku B, Ankola PA. Neonatal Jaundice. StatPearls Publishing; 2018.
- [2] Maisels MJ, Watchko JF, Bhutani VK, Stevenson DK. An approach to the management of hyperbilirubinemia in the preterm infant less than 35 weeks of gestation. Journal Of Perinatology. 2012 jun;32:660.
- [3] van den Akker ESA, de Beer JJA, Dijk PH, Fetter WPF, Flikweert S. Richtlijn preventie, diagnostiek en behandeling van hyperbilirubinemie bij de pasgeborene, geboren na een zwangerschapsduur van meer dan 35 weken. Nederlandse Vereniging voor Kindergeneeskunde; 2008.
- [4] Ali R, Ahmed S, Qadir M, Ahmad K. Icterus Neonatorum in Near-Term and Term Infants: An overview. Sultan Qaboos University medical journal. 2012 may;12(2):153–160.
- [5] Dijk PH, de Vries TW, de Beer JJAH. Richtlijn Preventie, diagnostiek en behandeling van hyperbilirubinemie bij de pasgeborene, geboren na een zwangerschapsduur van meer dan 35 weken'. Nederlans Tijdschrift voor Geneeskunde. 2009;153.
- [6] Bilicurves prematuren < 35 wkn; Available from: http://babyzietgeel.nl/kinderarts/ hulpmiddelen/diagnostiek/bilicurves{\_}prematuren.php.
- [7] Bosschaart N, Kok J, Newsum A, Ouweneel D, Mentink R, van Leeuwen T, et al. Limitations and Opportunities of Transcutaneous Bilirubin Measurements. Pediatrics. 2012 mar;129(4):689–694. doi:10.1542/peds.2011-2586.
- [8] van den Esker-Jonker B, den Boer L, Pepping RMC, Bekhof J. Transcutaneous Bilirubinometry in Jaundiced Neonates: A Randomized Controlled Trial. Pediatrics. 2016 dec;138(6):e20162414. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27940715. doi:10.1542/peds.2016-2414.
- [9] Ercan, Özgün G. The accuracy of transcutaneous bilirubinometer measurements to identify the hyperbilirubinemia in outpatient newborn population. Clinical Biochemistry. 2018 may;55:69–74. doi:10.1016/J.CLINBIOCHEM.2018.03.018.
- [10] Boron WF, Boulpaep EL. Uptake, processing, and secretion of compounds by hepatocytes. In: Medical Physiology. 2nd ed. Saunders, Elsevier; 2012. p. 986 – 993.
- [11] Bhagavan NV, Ha CE, Bhagavan NV, Ha CE. Metabolism of Iron and Heme. In: Essentials of Medical Biochemistry. Academic Press; 2015. p. 511–529. doi:10.1016/B978-0-12-416687-5.00027-0.
- [12] Ullah S, Rahman K, Hedayati M. Hyperbilirubinemia in Neonates: Types, Causes, Clinical Examinations, Preventive Measures and Treatments: A Narrative Review Article. Iranian journal of public health. 2016 may;45(5):558–568.
- [13] Polin RA, Rowitch DH, Fox WW, Abman SH, Benitz WE. Liver and Bilirubin Metabolism. In: Fetal and Neonatal Physiology. 5th ed. Elsevier; 2017. p. 909–952.
- [14] L Porter M, L Dennis B. Hyperbilirubinemia in the term newborn. American family physician. 2002 mar;65:599–606.

- [15] Cashore WJ. Neonatal Bilirubin Metabolism. In: Fetal and Neonatal Physiology. 5th ed. Elsevier; 2017. p. 929–933. doi:10.1016/B978-0-323-35214-7.00096-2.
- [16] Roche SP, Kobos R. Jaundice in the Adult Patient. American Family Physician. 2004;69(2):299–304.
- [17] Chaubal AN, Patel R, Choksi D, Shah K, Ingle M, Sawant P. Management of pregnancy in Crigler Najjar syndrome type 2. World journal of hepatology. 2016 apr;8(11):530–532. doi:10.4254/wjh.v8.i11.530.
- [18] Fakhri M, Farhadi R, Mousavinasab N, Hosseinimehr SJ, Yousefi SS, Davoodi A, et al. Preventive effect of purgative manna on neonatal jaundice: A double blind randomized controlled clinical trial. Journal of Ethnopharmacology. 2019 may;236:240–249. doi:10.1016/J.JEP.2019.03.009.
- [19] Knudsen A. The cephalocaudal progression of jaundice in newborns in relation to the transfer of bilirubin from plasma to skin. Early Human Development. apr;(1):23–28. doi:10.1016/0378-3782(90)90022-B.
- [20] Engle WD, Jackson GL, Engle NG. Transcutaneous bilirubinometry. Seminars in Perinatology. 2014 nov;38(7):438–451. doi:10.1053/J.SEMPERI.2014.08.007.
- [21] Mcewen M, Reynolds K. Noninvasive detection of bilirubin using pulsatile absorption. vol. 29; 2006.
- [22] El-Beshbishi SN, Shattuck KE, Mohammad AA, Petersen JR. Hyperbilirubinemia and Transcutaneous Bilirubinometry. Clinical Chemistry. 2009 jul;55(7):1280 LP – 1287. doi:10.1373/clinchem.2008.121889.
- [23] Dräger Medical System. Gebruiksaanwijzing JM-103; 2008.
- [24] Draeger Medical Systems, Inc. Jaundice Meter JM-103 and Jaundice Meter JM-105 Recalled Due to Misinterpretation of Display Messages for Out of Range Values; 2018. Available from: https://www.fda.gov/medical-devices/medical-device-recalls/ draeger-medical-systems-inc-jaundice-meter-jm-103-and-jaundice-meter-jm-105-recalled-due.
- [25] Dräger Medical System. Gebruiksaanwijzing JM-105; 2013.
- [26] van Erk M. Influence of bone depth and skin optical properties on transcutaneous bilirubinometer readings. Enschede: Biomedical Photonic Imaging Group, University of Twente; 2018.
- [27] Hofmeijer EIS. Internship report: bilirubin study. Zwolle: University of Twente, Isala Klinieken Zwolle; 2019.
- [28] Veenstra C, Petersen W, Vellekoop IM, Steenbergen W, Bosschaart N. Spatially confined quantification of bilirubin concentrations by spectroscopic visible-light optical coherence tomography. Biomed Opt Express. 2018 Aug;9(8):3581–3589.
- [29] Prahl S. Hemoglobin spectrum as compiled by S. Prahl; 1999. Available from: https://omlc. org/spectra/hemoglobin/summary.html.
- [30] Bosschaart N, Mentink R, van Leeuwen T, Aalders M, Kok J. Optical properties of neonatal skin measured in vivo as a function of age and skin pigmentation. Journal of biomedical optics. 2011;16(9). doi:10.1117/1.3622629.

[31] Michels R, Foschum F, Kienle A. Optical properties of fat emulsions. Opt Express. 2008 Apr;16(8):5907–5925.

## 6 Bijlagen

### 6.1 Bepaling absorptiecoëfficiënt inkt stockoplossing

Om de absorptiecoëfficiënten ( $\mu_a$ ) bij 450 en 550 nm van de lichtgele en magenta inkt stockoplossingen te bepalen, is een spectrofotometer (Shimadzu's UV-2600) gebruikt. De volgende stappen zijn drie keer uitgevoerd om de nauwkeurigheid te vegroten. De stockoplossing is met Milli-Q 100 keer verdund tot een 1%-stockoplossing, zodat de absorptiewaarden in het meetbereik van de spectrofotometer liggen. Vervolgens is deze 1%-stockoplossing weer met Milli-Q verdund tot 0,83%- stockoplossing en 0,67%-stockoplossing. Van elk van deze verdunningen is drie keer de transmissie bepaald met behulp van de spectrofotometer. Dit staat schematisch weergegeven in figuur 6.1.





Van de drie gemeten transmissies per verdunning bij 450 en 550 nm is het gemiddelde berekend. Met behulp van de volgende vergelijking is  $\mu_a$  uit de gemiddelde transmissie berekend:

$$T = I/I_0 = e^{-\mu_a d}$$
(6.1)

T = transmissie I = doorgelaten lichtintensiteit  $I_0 = \text{uitgezonden lichtintensiteit}$   $\mu_a = \text{absorptiecoëfficiënt}$ d = padlengte

De breedte van de cuvetten is 10 mm, dus d=10 mm is ingevuld in vergelijking 6.1. Vervolgens is voor de verschillende concentraties de absorptiecoëfficiënt berekend. De absorptiecoëfficiënten

zijn uitgezet tegen de concentraties en van dat verband is de richtingscoëfficiënt bepaald. Hierbij is meegenomen dat de absorptiecoëfficiënt nul is als de concentratie ook nul is. De richtingscoëfficiënt is gelijk aan de absorptiecoëfficiënt van de stockoplossing.

Serienummer	а	b	$R_a^2$
B3601005	-5,65762	0,8094	0,99276
B3601050	-4,06315	0,84133	0,99632
3201007	-1,77021	0,80603	0,99728
3201377	3,03475	0,78243	0,998
B3601081	-3,06544	1,01518	0,99497
B3601086	-5,38185	0,96666	0,99541
B3601107	-5,76793	0,94017	0,99436
B3601137	-7,36207	0,93595	0,99089
3201594	-0,58002	0,77944	0,99614
B3601390	-8,16389	0,75803	0,98205
B3601402	-9,3675	0,79509	0,9815
3202711	-8,14094	0,78973	0,98471
3203135	-9,57373	0,83744	0,98222

## 6.2 Trendlijnen grafiek 2.7

**Tabel 6.1:** Lineaire trendlijnen van de TcB-metingen, waarbij TcB = a + b \* verwachte TcB-meting in  $\mu$ mol/L.  $R_a^2$  = adjusted R-squared.

## 6.3 SPSS Outputs

#### Tests of Between-Subjects Effects

Dependent Variable:	Gemeten_TcB				
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	322344.930 <sup>a</sup>	17	18961.466	347.016	.000
Intercept	925098.242	1	925098.242	16930.342	.000
Fantoom_nummer	312524.604	5	62504.921	1143.911	.000
Meter_Nummer	9820.326	12	818.361	14.977	.000
Error	3278.486	60	54.641		
Total	1250721.659	78			
Corrected Total	325623.417	77			

a. R Squared = .990 (Adjusted R Squared = .987)

Figuur 6.2: Uitkomsttabel van de F-toets over alle meters en alle fantomen.

			Subse	et for alpha =	0.05
	Meter_Nummer	Ν	1	2	3
Tukey HSD <sup>a</sup>	B3601086 (JM105)	21	.0000		
	B3601137 (JM105)	21	.0000		
	B3601107 (JM105)	21	.0000		
	3201594 (JM103)	21	.0000		
	B3601005 (JM105)	21	.0000		
	3201007 (JM103)	21	.0000		
	B3601050 (JM105)	21	.0000		
	B3601390 (JM105)	21	.0000		
	B3601402 (JM105)	21	.0000		
	3202711 (JM103)	21	.0000		
	3203135 (JM103)	21	.0000		
	B3601081 (JM105)	21		2.5714	
	3201377 (JM103)	21			3.4762
	Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 21.000.

#### Figuur 6.3: Tukey's HSD toets fantoom 1.

							Subset for a	alpha = 0.05				
	Meter_Nummer	N	1	2	3	4	5	6	7	8	9	10
Tukey HSD <sup>a</sup>	B3601390 (JM105)	21	40.8095									
	B3601402 (JM105)	21		42.0476								
	3202711 (JM103)	21			43.6667							
	3203135 (JM103)	21				45.0476						
	B3601005 (JM105)	21		42.0476         Image: Constraint of the system         Image: Constand of the system								
	B3601137 (JM105)	21						58.8095				
	B3601050 (JM105)	21						58.8095				
	3201007 (JM103)	21						59.4762				
	3201594 (JM103)	21						59.7143				
	B3601107 (JM105)	21							63.0000			
	3201377 (JM103)	21								64.5238		
	B3601086 (JM105)	21									66.0000	
	B3601081 (JM105)	21										71.8571
	Sig.		1.000	1.000	1.000	1.000	1.000	.139	1.000	1.000	1.000	1.000
Means for gro	ups in homogeneous su	ubsets are	displayed.									

a. Uses Harmonic Mean Sample Size = 21.000.

Figuur 6.4: Tukey's HSD toets fantoom 2.

								Subset for	r alpha = 0.05					
	Meter_Nummer	N	1	2	3	4	5	6	7	8	9	10	11	12
M Tukey HSD <sup>a,b</sup> B 33 33 33 8 33 8 33 33 8 8 33 8 8 8 8	B3601390 (JM105)	21	91.9524											
	B3601402 (JM105)	21		95.3333										
	3202711 (JM103)	21			97.0952									
	3203135 (JM103)	21				99.9524								
	B3601005 (JM105)	21					102.3333							
	3201594 (JM103)	21						107.1905						
	B3601050 (JM105)	15						108.0000						
	3201007 (JM103)	21							109.0000					
	3201377 (JM103)	21								110.3810				
	B3601137 (JM105)	21									115.5714			
	B3601107 (JM105)	21										118.7143		
	B3601086 (JM105)	21											123.2857	
	B3601081 (JM105)	21												131.1905
	Sig.		1.000	1.000	1.000	1.000	1.000	.185	1.000	1.000	1.000	1.000	1.000	1.000

a. Uses Harmonic Mean Sample Size = 20.373.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.



	Meter_Nummer B3601390 (JM105) B3601402 (JM105) 3202711 (JM103) B3601005 (JM105) 3201594 (JM103) 3203135 (JM103) 3201007 (JM103) B3601050 (JM105) B3601137 (JM105)					Subset for a	alpha = 0.05			
	Meter_Nummer	Ν	1	2	3	4	5	6	7	8
Tukey HSD <sup>a</sup>	B3601390 (JM105)	21	126.1429							
	B3601402 (JM105)	21	127.6190	127.6190						
	3202711 (JM103)	21		129.1429						
	B3601005 (JM105)	21			134.9048					
	3201594 (JM103)	21			136.0952					
	3203135 (JM103)	21			136.2857					
	3201007 (JM103)	21				139.2857				
	3201377 (JM103)	21				141.0000	141.0000			
	B3601050 (JM105)	21					141.6667			
	B3601137 (JM105)	21						156.6667		
	B3601107 (JM105)	21						158.2857		
	B3601086 (JM105)	21							162.7143	
	B3601081 (JM105)	21								174.9048
	Sig.		.213	.174	.309	.068	.989	.111	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 21.000.

#### Figuur 6.6: Tukey's HSD toets fantoom 4.

						Subs	et for alpha =	0.05			
	Meter_Nummer	N	1	2	3	4	5	6	7	8	9
Tukey HSD <sup>a</sup>	B3601390 (JM105)	21	142.9524								
	3202711 (JM103)	21		150.2381							
	B3601402 (JM105)	21		150.3810							
	3201594 (JM103)	21			155.9048						
	B3601005 (JM105)	21			156.0476						
	3201377 (JM103)	21			156.9048	156.9048					
	3203135 (JM103)	21				158.0952					
	3201007 (JM103)	21				158.3333					
	B3601050 (JM105)	21					164.1429				
	B3601137 (JM105)	21						177.1905			
	B3601107 (JM105)	21							179.3333		
	B3601086 (JM105)	21								184.8571	
	B3601081 (JM105)	21									197.3810
	Sig.		1.000	1.000	.621	.104	1.000	1.000	1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 21.000.



							Subs	et for alpha =	0.05				
	Meter_Nummer	Ν	1	2	3	4	5	6	7	8	9	10	11
Tukey HSD <sup>a</sup>	B3601390 (JM105)	21	158.6190										
	3201594 (JM103)	21		164.2857									
	3202711 (JM103)	21			165.9524								
	B3601402 (JM105)	21				167.6667							
	3201377 (JM103)	21					171.0000						
	3201007 (JM103)	21					171.3333						
	B3601005 (JM105)	21					171.7143						
	3203135 (JM103)	21						176.4762					
	B3601050 (JM105)	21							179.4286				
	B3601137 (JM105)	21								200.0952			
	B3601107 (JM105)	21									201.9524		
	B3601086 (JM105)	21										208.0000	
	B3601081 (JM105)	21											219.9048
	Sig.		1.000	1.000	1.000	1.000	.939	1.000	1.000	1.000	1.000	1.000	1.000

a. Uses Harmonic Mean Sample Size = 21.000.

Figuur 6.8: Tukey's HSD toets fantoom 6.

### 6.4 MATLAB scripts

#### 6.4.1 Script voor het toevoegen van nieuwe testdata

Hier volgt het script van Hofmeijer, dat is aangepast voor het toevoegen van nieuwe testdata.

```
1 % This script trains and tests 10 models using the input data table.
2 % It therefore also determines the predicted TSB. According to the protocol
3 % it also determines if phototherapy is necessary from the predicted TSB
4 % and the true TSB. From this information it calculates the confusion
5 % matrix and in turn the accuracy, sensitivity and specificity can be
  % calculated from this matrix. The sensitivity and specificity are used to
6
  % create a ROC-curve with as a varying parameter and added bias.
7
  %
8
  %NPUT: - Table containing the following parameters: BirthWeight,
9
              GestationalAge, PostnatalAge, Forehead, Sternum, Hip,
  %
10
  %
              Phototherapy, TSB
11
            (- It is possible to use less locations if also BiliModel.m is
  %
12
            adjusted accordingly)
  %
13
            CAUTION: It is important that the table contains the parameters in
  %
14
            the same order.
15
  %
            NOTICE: If data is imported manually (copy/paste) into the
16
  %
            workspace, make sure it is imported as a Table called "Data" and
17
  %
            the columns have the correct names as described above.
18
  %
19
  %
  % OUPUT: - A Struct called Validate with all the information per iteration
20
  %
              form the trained model with RMSE to the calculated sensitivity,
21
  %
              etc.
22
            - Figure contain the ROC curve of the 10 iterations
  %
23
  %
24
25
  %
   \%- Risk factor has now been put to standard risk (Standard Risk factor = 0;). If this
26
       needs to be changed per subject an extra line has to be added before determining
27
  %
       the weight category and TSBcritical.
28
  %
29
  % Insert the old data as Data1 and the new added data as Data2 in te workspace.
30
31
  %% Create data struct with 10 random train sets from old data (Data1), create test set
32
       with
  % new added data (Data2), train model and predict TSB using the trained model and test set
33
```

```
for h = 1:10 % number of iterations
34
       random_data = Data1(randperm(size(Data1, 1)), :); %Determine random row numbers
35
       sized1 = size(random_data);
36
       lengthd = sized1(:,2);
37
       lengtha = sized1(:,1);
38
       sized2 = size(Data2);
39
       lengthc = sized2(:,2);
       lengthr = sized2(:,1);
       %Create train set with random 80% of the original data
       Train = random_data(1:(round(0.8*lengtha)),:);
43
       %Create test with new data
       Test = Data2(:,1:(lengthd -1));
45
       %Keep the True TSB of the test set for later validation
46
       TSBTrue = table2array(Data2(:,lengthd));
47
       %Create validate struct with:
48
       Validate(h).Trainset = Train; %train data
49
       Validate(h).Testset = Test; %testdata
50
       %Train model using the GPR algorithm
51
       [trainedModel, validationRMSE] = BiliModel(Train);
52
       Validate(h).TrainedModel = trainedModel; %trained model
53
       Validate(h).RMSETrain = validationRMSE; %final RMSE of trained model
54
       %Predict new TSB using the trained model
55
       TSBpred = trainedModel.predictFcn(Test);
56
57
       Validate(h).TSBpred = TSBpred; %predicted TSB values
       Validate(h).TSBTrue = TSBTrue; %true TSB values
58
   end
59
60
  %% Determine the RMSE of newly predicted TSB using the True TSB
61
62
   for t = 1:length(Validate)
63
      Pred = Validate(t).TSBpred;
64
      True = Validate(t).TSBTrue;
65
      RMSETest = sqrt(nanmean(((Pred-True).^2))); %calculate RMSE
66
      Validate(t).RMSETest = RMSETest; %save into validate struct
67
   end
68
69
  % Determine if, using the predicted TSB, phototherapy is necessary (yes/no)
70
  % In order to predict this the protocol for giving phototherapy from the
71
  % UMOG has been used. The critical TSB value depends on the Risk Factor and
72
  % the birth weight of the infant.
73
   for v = 1:length(Validate)
74
       RiskFactor = 0; % for the first tests only the standard Risk Factor has been used
75
       for i = 1:length(Validate(v).Testset.Birthweight)
76
            if BiskFactor == 0
77
           % Determine Birthweight category and TSBcritical for standard risk factor
78
           % infant
79
                if Validate(v).Testset.Birthweight(i) < 1000 %Weight category <1000 grams
80
                    WeightCategory(i) = 1;
81
                    if Validate(v).Testset.PostnatalAge(i) == 1 %postnatal age in days
82
                    TSBcriticalSR = 75;
83
                    else
84
                    TSBcriticalSR = 100;
85
                    end
86
                elseif Validate(v).Testset.Birthweight(i) >= 1000 && Validate(v).Testset.
87
                    Birthweight(i) < 1250
                    WeightCategory(i) = 2;
88
                    if Validate(v).Testset.PostnatalAge(i) == 1
89
                    TSBcriticalSR = 85;
90
                    elseif Validate(v).Testset.PostnatalAge(i) == 2
91
```

```
TSBcriticalSR = 135;
92
                     else
93
                          TSBcriticalSR = 150;
94
                     end
95
                 elseif Validate(v).Testset.Birthweight(i) >= 1250 && Validate(v).Testset.
96
                     Birthweight(i) < 1500
                     WeightCategory(i) = 3;
97
                     if Validate(v).Testset.PostnatalAge(i) == 1
98
                          TSBcriticalSR = 95;
99
                     elseif Validate(v).Testset.PostnatalAge(i) == 2
100
                          TSBcriticalSR = 165;
101
                     else
102
                          TSBcriticalSR = 190;
103
                     end
104
                 elseif Validate(v).Testset.Birthweight(i) >= 1500 && Validate(v).Testset.
105
                     Birthweight(i) < 2000
                     WeightCategory(i) = 4;
106
                     if Validate(v).Testset.PostnatalAge(i) == 1
107
                          TSBcriticalSR = 105;
108
109
                     elseif Validate(v).Testset.PostnatalAge(i) == 2
                          TSBcriticalSR = 190;
110
111
                     else
                          TSBcriticalSR = 220;
112
                     end
113
                 else
114
                     WeightCategory(i) = 5;
115
                     if Validate(v).Testset.PostnatalAge(i) == 1
116
                          TSBcriticalSR = 115;
117
                     elseif Validate(v).Testset.PostnatalAge(i) == 2
118
                          TSBcriticalSR = 210;
119
                     else
120
                          TSBcriticalSR = 240;
121
                     end
122
                 end
123
                 TSBcritical(i,1) = TSBcriticalSR; %Add all critical TSB values into one array
124
            else
125
            % Determine Birthweight category and TSBcritical for high risk factor
126
            % infant
127
                 if Validate(v).Testset.Birthweight(i) < 1000
128
                     WeightCategory(i) = 1;
129
                     if Validate(v).Testset.PostnatalAge(i) == 1
130
                         TSBcriticalHR = 75;
131
                     else
132
                          TSBcriticalHR = 100;
133
                     end
134
                 elseif Validate(v).Testset.Birthweight(i) >= 1000 && Validate(v).Testset.
135
                     Birthweight(i) < 1250
                     WeightCategory(i) = 2;
136
                     if Validate(v).Testset.PostnatalAge(i) == 1
137
                          TSBcriticalHR = 75;
138
                     elseif Validate(v).Testset.PostnatalAge(i) == 2
139
                         TSBcriticalHR = 100;
140
                     else
141
                          TSBcriticalHR = 100;
142
                     end
143
                 elseif Validate(v). Testset. Birthweight(i) >= 1250 && Validate(v). Testset.
144
                     Birthweight(i) < 1500
                     WeightCategory(i) = 3;
145
                     if Validate(v).Testset.PostnatalAge(i) == 1
146
```

```
TSBcriticalHR = 85;
147
                     elseif Validate(v).Testset.PostnatalAge(i) == 2
148
                          TSBcriticalHR = 135;
149
                     else
150
                         TSBcriticalHR = 150;
151
                     end
152
                 elseif Validate(v).Testset.Birthweight(i) >= 1500 && Validate(v).Testset.
153
                     Birthweight(i) < 2000
                     WeightCategory(i) = 4;
154
                     if Validate(v).Testset.PostnatalAge(i) == 1
155
                          TSBcriticalHR = 95;
156
                     elseif Validate(v).Testset.PostnatalAge(i) == 2
157
                          TSBcriticalHR = 165;
158
                     else
159
                          TSBcriticalHR = 190;
160
                     end
161
                 else
162
                     WeightCategory(i) = 5;
163
                     if Validate(v).Testset.PostnatalAge(i) == 1
164
165
                          TSBcriticalHR = 105;
                     elseif Validate(v).Testset.PostnatalAge(i) == 2
166
                         TSBcriticalHR = 190;
167
168
                     else
                         TSBcriticalHR = 220;
169
170
                     end
                 end
171
                 TSBcritical(i,1) = TSBcriticalHR;
172
            end
173
        end
174
        %Save the critical TSB values into the validate struct
175
        Validate(v).TSBcritical = TSBcritical;
176
177
   end
178
   %% Determine if phototherapy should be applied or not
179
   % Both the TSBpred and TSBTrue will be run throught this for-loop to
180
   % determine for both value arrays if phototherapy is necessary or not.
181
    for v = 1:length(Validate)
182
        i = 1;
183
        s = 1;
184
        for m = 0.4:40 %Used to add a value to TSBpred, to see if the model systematically
185
            % underestimates the TSBTrue
186
            TSBpred = Validate(v).TSBpred + m;
187
             for k = 1:length(Validate(v).TSBcritical)
188
                % If the infant already receives phototherapy, the critical value goes
189
                \% \mbox{ down by 100 umol/L to indicate that the treatment can be stopped }
190
                 if Validate(v).Testset.Phototherapy(k,1) == 1 && TSBpred(k,1) >= (Validate(v).
191
                     TSBcritical(k, 1) - 100)
                          PhototherapyPred(k,1) = 1; %Yes
192
                 else
                     %If NaN is present in the TSBpred, the corresponding
194
                     %phototherapyPred automatically gets NaN.
195
                     if isnan(TSBpred(k,1)) == 1
196
                          PhototherapyPred(k,1) = NaN;
197
                     elseif TSBpred(k) >= Validate(v).TSBcritical(k)
198
                         PhototherapyPred(k,1) = 1;
                                                          % Yes
199
                     else
200
                          PhototherapyPred(k,1) = 0;
                                                          % No
201
                     end
202
                 end
203
```

```
Validate (v). Phototherapy(s). Pred = PhototherapyPred; %save to validate struct
204
             end
205
             %Classify the TSBTrue values into phototherapy yes or no
206
             for k = 1:length(Validate(v).TSBcritical)
207
                  if Validate(v).Testset.Phototherapy(k,1) == 1 && Validate(v).TSBTrue(k,1) >= (
208
                       Validate (v). TSBcritical (k, 1) - 100
                           PhototherapyTrue(k, 1) = 1; %Yes
209
                  else
210
                       if isnan(Validate(v).TSBTrue(k,1)) == 1
211
                           PhototherapyTrue(k, 1) = NaN;
212
                       elseif Validate(v).TSBTrue(k) >= Validate(v).TSBcritical(k)
213
                           PhototherapyTrue(k, 1) = 1;
                                                             % Yes
214
                       else
215
                        PhototherapyTrue(k, 1) = 0;
                                                         % No
216
                      end
217
                  end
218
                  Validate(v). Phototherapy(s). True = PhototherapyTrue; % ave to validate struct
219
220
             end
221
         s = s + 1:
222
223
224
225
   % Compare the phototherapyPred (yes/no) with phototherapyTrue (yes/no) - and get
226
   % sensitivity, specificity and create ROC curve
227
        % Determine confusion matrix
228
         [C(j).conf,order] = confusionmat(PhototherapyTrue, PhototherapyPred, 'Order', [1,0]);
229
230
        % Determine sensitivity, specificity, error rate and accuracy
231
         Accuracy(j,1) = (C(j).conf(1,1)+C(j).conf(2,2))/length(PhototherapyTrue);
232
         \operatorname{ErrorRate}(j,1) = (C(j).\operatorname{conf}(2,1)+C(j).\operatorname{conf}(1,2))/\operatorname{length}(\operatorname{PhototherapyTrue});
233
         Sensitivity (j, 1) = C(j) \cdot conf(1, 1) / (C(j) \cdot conf(1, 1) + C(j) \cdot conf(1, 2));
234
         Specificity (j, 1) = C(j) \cdot conf(2, 2) / (C(j) \cdot conf(2, 1) + C(j) \cdot conf(2, 2));
235
         j = j + 1;
236
         end
237
        % Save all the information into the validate struct
238
         Validate(v). Accuracy = Accuracy;
239
         Validate(v). ErrorRate = ErrorRate;
240
         Validate(v).Sensitivity = Sensitivity;
241
         Validate(v).Specificity = Specificity;
242
         Validate(v).ConfMat = C;
243
    end
244
   %% Create ROC-curve using the sensitivity and 1-specificity
245
246
247
    figure()
    for v = 1:length(Validate)
248
         y = [0 (Validate(v).Sensitivity') 1];
249
         x = [0 ((1 - Validate(v) . Specificity)') 1];
250
        name = string(v);
251
         plot(x,y, 'LineStyle', '-','color', rand(1,3), 'marker', 's','MarkerSize', 5,...
252
              DisplayName', 'Iteration '+name )
253
         hold on
254
    end
255
256
    x = [0 \ 1];
257
    y = [0 \ 1];
    plot(x,y,
                 — r', 'DisplayName', 'Random guess')
258
    xlabel('1-Specificity')
259
260 ylabel('Sensitivity')
```

```
261 title ('ROC curve for model trained with phototherapy data (FSH)') %Changing name is possible
262 legend ('Location', 'southeast')
```

```
263 hold off
```

#### 6.4.2 Script voor het voorkomen van patiëntoverlap

In het bovenstaande script, het script van Hofmeijer, moeten de regels 30 tot en met 177 vervangen worden door het volgende.

```
1 % Load data or insert data manually into workspace in the form described
  % above
2
  % load ('Data.mat') % If data is put manually into workspace, comment this line
3
  %Define an array in which all patient ID's occur only once.
5
   AllPatientID = table2array(Data(:,1));
6
   PatientID = unique(AllPatientID);
7
  % Create data struct with 10 random train and test sets from Data, train model
9
  % and predict TSB using the trained model and test set.
10
11
   for h = 1:10 %number of iterations
12
       random_patientID = PatientID(randperm(size(PatientID,1)),:); %randomize pati ntID 's
13
       sizeID = size(random_patientID);
14
       lengthID = sizeID(:,1);
15
       TrainID = random_patientID(1:round(0.8*lengthID)); % select 80% of the patients
16
       Agree = ismember(Data.ID, TrainID); %see whether pati nt ID occurs in train set
17
       Data.Agree = Agree; %save column with 1's for every row that's supposed to be in the
18
           train set
       sized = size(Data);
19
       lengthc = sized(:,2); %number of columns
20
       lengthr = sized(:,1); %number of rows
21
       n = 1:
22
23
       m = 1:
24
  %Create train set with 80% of the patients and test set with 20% of the patients
25
  %Make sure the train set does not include Patient ID
26
  %Make sure the test set does not include Patient ID and TSBtrue
27
28
       for p = 1:lengthr
29
            if Data.Agree(p) == 1
30
                 Train(n,:) = Data(p,2:lengthc-1);
31
                n = n + 1;
32
            else
33
                Test(m,:) = Data(p,2:lengthc-2);
34
                %Keep the True TSB of the test set for later validation
35
                TSBTrue(m,:) = table2array(Data(p,lengthc-1));
36
                m = m + 1;
37
            end
38
       end
39
       %Create validate struct with:
40
       Validate(h). Trainset = Train; %train data
41
       Validate(h).Testset = Test; %test data
42
       %Train model using the GPR algorithm
43
       [trainedModel, validationRMSE] = BiliModel(Train);
44
       Validate(h).TrainedModel = trainedModel; %trained model
45
       Validate(h).RMSETrain = validationRMSE; %final RMSE of trained model
46
       %Predict new TSB using the trained model and save data to validate
47
```

```
%struct
48
        TSBpred = trainedModel.predictFcn(Test);
49
        Validate(h).TSBpred = TSBpred; %predicted TSB values
50
        Validate(h).TSBTrue = TSBTrue; %true TSB values
51
52
       % clearing all data is needed because the lengths of the sets differ
53
       % every iteration
54
        clear Train Test TSBTrue
55
   end
56
57
       %% Determine the RMSE of newly predicted TSB using the True TSB
58
59
   for t = 1:length(Validate)
60
       Pred = Validate(t).TSBpred;
61
       True = Validate(t).TSBTrue;
62
       RMSETest = sqrt(nanmean(((Pred-True).^2))); %calculate RMSE
63
       Validate(t).RMSETest = RMSETest; %save into validate struct
64
   end
65
66
67
   % Determine if, using the predicted TSB, phototherapy is necessary (yes/no)
   % In order to predict this the protocol for giving phototherapy from the
68
   % UMOG has been used. The critical TSB value depends on the Risk Factor and
69
70 % the birth weight of the infant.
   for v = 1:length(Validate)
71
        RiskFactor = 0; %for the first tests only the standard Risk Factor has been used
72
        for i = 1:length(Validate(v).Testset.BirthWeight)
73
            if RiskFactor == 0
74
            % Determine Birthweight category and TSBcritical for standard risk factor
75
            % infant
76
                 if Validate(v).Testset.BirthWeight(i) < 1000 %Weight category <1000 grams
77
                     WeightCategory(i) = 1;
78
                     if Validate(v).Testset.PostnatalAge(i) == 1 %postnatal age in days
79
                     TSBcriticalSR = 75;
80
                     else
81
                     TSBcriticalSR = 100;
82
                    end
83
                elseif Validate(v).Testset.BirthWeight(i) >= 1000 && Validate(v).Testset.
84
                     BirthWeight(i) < 1250
                     WeightCategory(i) = 2;
85
                     if Validate(v).Testset.PostnatalAge(i) == 1
86
                     TSBcriticalSR = 85;
87
                     elseif Validate(v).Testset.PostnatalAge(i) == 2
88
                         TSBcriticalSR = 135;
89
                     else
90
                         TSBcriticalSR = 150;
91
                    end
92
                elseif Validate(v).Testset.BirthWeight(i) >= 1250 && Validate(v).Testset.
93
                     BirthWeight(i) < 1500
                     WeightCategory(i) = 3;
94
                     if Validate(v).Testset.PostnatalAge(i) == 1
95
                         TSBcriticalSR = 95;
96
                     elseif Validate(v).Testset.PostnatalAge(i) == 2
97
                         TSBcriticalSR = 165;
98
                     else
99
                         TSBcriticalSR = 190;
100
                    end
101
                elseif Validate(v). Testset. BirthWeight(i) \geq 1500 & Validate(v). Testset.
102
                     BirthWeight(i) < 2000
                     WeightCategory(i) = 4;
103
```

```
if Validate(v).Testset.PostnatalAge(i) == 1
104
                          TSBcriticalSR = 105;
105
                     elseif Validate(v).Testset.PostnatalAge(i) == 2
106
                          TSBcriticalSR = 190;
107
                     else
108
                          TSBcriticalSR = 220;
109
                     end
110
                 else
111
                     WeightCategory(i) = 5;
112
                     if Validate(v).Testset.PostnatalAge(i) == 1
113
                          TSBcriticalSR = 115;
114
                     elseif Validate(v).Testset.PostnatalAge(i) == 2
115
                          TSBcriticalSR = 210;
116
                     else
117
                          TSBcriticalSR = 240;
118
                     end
119
                 end
120
                 TSBcritical(i,1) = TSBcriticalSR; %Add all critical TSB values into one array
121
122
            else
            % Determine Birthweight category and TSBcritical for high risk factor
123
            % infant
124
                 if Validate(v).Testset.BirthWeight(i) < 1000
125
126
                     WeightCategory(i) = 1;
                     if Validate(v).Testset.PostnatalAge(i) == 1
127
128
                          TSBcriticalHR = 75;
                     else
129
                          TSBcriticalHR = 100;
130
                     end
131
                 elseif Validate(v).Testset.BirthWeight(i) >= 1000 && Validate(v).Testset.
132
                     BirthWeight(i) < 1250
                     WeightCategory(i) = 2;
133
                     if Validate(v).Testset.PostnatalAge(i) == 1
134
                          TSBcriticalHR = 75;
135
                     elseif Validate(v).Testset.PostnatalAge(i) == 2
136
                          TSBcriticalHR = 100;
137
                     else
138
                          TSBcriticalHR = 100;
139
                     end
140
                 elseif Validate(v).Testset.BirthWeight(i) >= 1250 && Validate(v).Testset.
141
                     BirthWeight(i) < 1500
                     WeightCategory(i) = 3;
142
                     if Validate(v).Testset.PostnatalAge(i) == 1
143
                         TSBcriticalHR = 85;
144
                     elseif Validate(v).Testset.PostnatalAge(i) == 2
145
                         TSBcriticalHR = 135;
146
                     else
147
                          TSBcriticalHR = 150;
148
                     end
                 elseif Validate(v).Testset.BirthWeight(i) >= 1500 && Validate(v).Testset.
150
                     BirthWeight(i) < 2000
                     WeightCategory(i) = 4;
151
                     if Validate(v).Testset.PostnatalAge(i) == 1
152
                          TSBcriticalHR = 95;
153
                     elseif Validate(v).Testset.PostnatalAge(i) == 2
154
                         TSBcriticalHR = 165;
155
                     else
156
                          TSBcriticalHR = 190;
157
                     end
158
                 else
159
```

```
WeightCategory(i) = 5;
160
                      if Validate(v).Testset.PostnatalAge(i) == 1
161
                          TSBcriticalHR = 105;
162
                      elseif Validate(v).Testset.PostnatalAge(i) == 2
163
                          TSBcriticalHR = 190;
164
                      else
165
                          TSBcriticalHR = 220;
166
                      end
167
                 end
168
                 TSBcritical(i,1) = TSBcriticalHR;
169
             end
170
        end
171
        %Save the critical TSB values into the validate struct
172
        Validate(v).TSBcritical = TSBcritical;
173
        clear TSBcritical
174
    end
175
```

#### 6.4.3 Script voor het trainen met één datapaar per patiënt

In het script van Hofmeijer moeten de regels 30 tot en met 59 vervangen worden door het volgende.

```
% Load data or insert data manually into workspace in the form described
1
2 % above
  % load data.mat;% If data is put manually into workspace, comment this line
3
5
   for h = 1:10 %number of iterations
6
       random_data = Data(randperm(size(Data, 1)), :); %Determine random row numbers
7
       [X, ia] = unique(random_data.ID);
8
       unique_data = random_data(ia,:); %Table that contains only one data-pair per patient
9
       sized = size(unique_data);
10
       lengthc = sized(:,2); %number of columns
11
       lengthr = sized(:,1); %number of rows
12
13
       %Create train set with 80% of the data. Make sure that the train set
14
       %does not contain the ID and Sepsis.
15
       Train = unique_data(1:(round(0.8*lengthr)),2:lengthc);
16
       %Create test set with 20% of the data. Make sure that
17
       %the test set does not contain the ID, TSB and Sepsis.
18
       Test = unique_data ((lengthr-round(0.2*lengthr)+1):lengthr, 2:(lengthc-1));
19
       %Keep the True TSB and Sepsis of the test set for later validation
20
       TSBTrue = table2array(unique_data((lengthr-round(0.2*lengthr)+1):lengthr,lengthc));
21
22
       %Create validate struct with:
23
       Validate(h).Trainset = Train; %train data
24
       Validate(h).Testset = Test; %testdata
25
       %Train model using the GPR algorithm
26
       [trainedModel, validationRMSE] = BiliModel(Train);
27
       Validate(h).TrainedModel = trainedModel; %trained model
28
       Validate(h).RMSETrain = validationRMSE; %final RMSE of trained model
29
       %Predict new TSB using the trained model
30
       TSBpred = trainedModel.predictFcn(Test);
31
       Validate(h).TSBpred = TSBpred; %predicted TSB values
32
       Validate(h).TSBTrue = TSBTrue; %true TSB values
33
   end
34
```

#### 6.4.4 Script voor het kwantificeren van de klinische waarde

```
1 % This script determines the percentage of TSB measurements that can be
  % replaced by TcB measurements, by using a new margin instead of the old margin of 50 umol
2
       /L.
3 % First, a TSB critical is determined for every measurement moment.
  % Then a margin under the TSB critical is chosen to get the TcB critical and
  % the TcB measerument at the forehead is compared to the TcB critical to determine
  % if a TSB measurement is necessary. Lastly, the number of heel punctures
  % that are necessary by using the new margin is compared to the number of
  % heel punctures that are necessary by using the old margin.
8
9
10
  % INPUT: - Table containing the following parameters: BirthWeight,
11
  %
              PostnatalAge, Sepsis, Voorhoofd. Save the table as Data.mat
  % OUTPUT: - Percentage of TSB measurements that can be replaced by TcB
12
  %
               measurments.
13
14
15
  % Determine for every measerument the TSB critical for phototherapy
16
   for i = 1:length(Data.BirthWeight)
17
       Riskfactor = 0; %Only the standard Risk factor has been used
18
           if Riskfactor == 0
19
           % Determine Birthweight category and TSBcritical if no clinical
20
           % sepsis is present
21
                if Data.BirthWeight(i) < 1000 %Weight category <1000 grams
22
                    WeightCategory(i) = 1;
23
                    if Data.PostnatalAge(i) == 1 %postnatal age in days
24
                    TSBcriticalSR = 75;
25
                    else
26
                    TSBcriticalSR = 100;
27
28
                    end
                elseif Data.BirthWeight(i) >= 1000 && Data.BirthWeight(i) < 1250
29
                    WeightCategory(i) = 2;
30
31
                    if Data.PostnatalAge(i) == 1
                    TSBcriticalSR = 85;
32
                    elseif Data.PostnatalAge(i) == 2
33
                        TSBcriticalSR = 135;
34
                    else
35
                        TSBcriticalSR = 150;
36
                    end
37
                elseif Data.BirthWeight(i) >= 1250 && Data.BirthWeight(i) < 1500
38
                    WeightCategory(i) = 3;
39
                    if Data.PostnatalAge(i) == 1
40
                        TSBcriticalSR = 95;
41
                    elseif Data.PostnatalAge(i) == 2
42
                        TSBcriticalSR = 165;
43
                    else
11
                        TSBcriticalSR = 190;
45
                    end
46
                elseif Data.BirthWeight(i) >= 1500 && Data.BirthWeight(i) < 2000
47
                    WeightCategory(i) = 4;
48
                    if Data.PostnatalAge(i) == 1
49
                        TSBcriticalSR = 105;
50
                    elseif Data.PostnatalAge(i) == 2
51
                        TSBcriticalSR = 190;
52
                    else
53
                        TSBcriticalSR = 220;
54
                    end
55
```

```
else
56
                     WeightCategory(i) = 5;
57
                      if Data.PostnatalAge(i) == 1
58
                          TSBcriticalSR = 115;
59
                      elseif Data.PostnatalAge(i) == 2
60
                          TSBcriticalSR = 210;
61
                      else
62
                          TSBcriticalSR = 240;
63
                     end
64
                 end
65
                 TSBcritical(i,1) = TSBcriticalSR; %Add all critical TSB values into one array
66
            else
67
            % Determine Birthweight category and TSBcritical if clinical sepsis
68
            % is present
69
                 if Data.BirthWeight(i) < 1000
70
                     WeightCategory(i) = 1;
71
                      if Data.PostnatalAge(i) == 1
72
                          TSBcriticalHR = 75;
73
                     else
74
                          TSBcriticalHR = 100;
75
76
                     end
                 elseif Data.BirthWeight(i) >= 1000 && Data.BirthWeight(i) < 1250
77
78
                     WeightCategory(i) = 2;
                      if Data.PostnatalAge(i) == 1
79
                          TSBcriticalHR = 75;
80
                      elseif Data.PostnatalAge(i) == 2
81
                          TSBcriticalHR = 100;
82
                     else
83
                          TSBcriticalHR = 100;
84
                     end
85
                 elseif Data.BirthWeight(i) >= 1250 && Data.BirthWeight(i) < 1500
86
                     WeightCategory(i) = 3;
87
                      if Data.PostnatalAge(i) == 1
88
                          TSBcriticalHR = 85;
89
                      elseif Data.PostnatalAge(i) == 2
90
                          TSBcriticalHR = 135;
91
                     else
92
                          TSBcriticalHR = 150;
93
                     end
94
                 elseif Data.BirthWeight(i) >= 1500 && Data.BirthWeight(i) < 2000
95
                     WeightCategory(i) = 4;
96
                      if Data.PostnatalAge(i) == 1
97
                          TSBcriticalHR = 95;
98
                      elseif Data.PostnatalAge(i) == 2
99
                          TSBcriticalHR = 165;
100
                     else
101
                          TSBcriticalHR = 190;
102
                     end
103
                 else
104
                     WeightCategory(i) = 5;
105
                      if Data.PostnatalAge(i) == 1
106
                          TSBcriticalHR = 105;
107
                      elseif Data.PostnatalAge(i) == 2
108
                          TSBcriticalHR = 190;
109
                     else
110
                          TSBcriticalHR = 220;
111
                     end
112
                 end
113
                 TSBcritical(i,1) = TSBcriticalHR;
114
```

```
end
115
116
   end
        Data.TSBcritical = TSBcritical;
117
118
    %% Determine for every measurement the TcB critical for TSB measurement
119
     Margin1 = 50; % The current margin
120
     Data. TcBcritical1 = Data. TSBcritical - Margin1;
121
     Margin2 = 37.98; % The new margin: 2 times the RMSE of the trained model
122
     Data. TcBcritical2 = Data. TSBcritical - Margin2;
123
124
    %% Determine if the TcB measurement at the forehead exceeds the TcB critical
125
     for j = 1 : length(Data.Voorhoofd)
126
         if Data.Voorhoofd(j) < Data.TcBcritical1(j)
127
             Data.Heelpuncture1(j) = 0;
128
             Data.Heelpuncture2(j) = 0;
129
         elseif Data.Voorhoofd(j) >= Data.TcBcritical1(j) && Data.Voorhoofd(j) < Data.</pre>
130
             TcBcritical2(j)
             Data.Heelpuncture1(j) = 1;
131
             Data.Heelpuncture2(j) = 0;
132
133
         else
             Data.Heelpuncture1(j) = 1;
134
             Data.Heelpuncture2(j) = 1;
135
136
         end
137
    end
138
    %% Determine the percentage TSB measurements that can be replaced by TcB measurements
139
   number_heelpunctures1 = sum(Data.Heelpuncture1);
140
   number_heelpunctures2 = sum(Data.Heelpuncture2);
141
   difference = number_heelpunctures1 - number_heelpunctures2;
142
   percentage = (difference)/number_heelpunctures1*100
143
```

#### 6.5 ROC-curves



Figuur 6.9: ROC-curve van het model van Hofmeijer getest met nieuwe patiëntdata.



Figuur 6.10: ROC-curve van het model getraind en getest met de vergrote dataset van 70 patiënten.



Figuur 6.11: ROC-curve van het model getraind en getest met gescheiden patiëntgegevens.



Figuur 6.12: ROC-curve van het model getraind en getest met één datapaar per patiënt.



Figuur 6.13: ROC-curve van het model getraind en getest met TcB-metingen door één bilirubinemeter.