

**Aus dem Department für Frauengesundheit Tübingen -
Universitätsfrauenklinik**

**Pränatale Messung des fetalen Magenumfangs als
Indikator für eine Ösophagusatresie**

**Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Medizin**

**der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen**

vorgelegt von

Borgmeier, Felix

2018

Felix Borgmeier

Dissertation

Frauenklinik Tuebingen

Matrikelnr: 3502738

Praenataldiagnostik

Dekan: Professor Dr. I. B. Autenrieth

1. Berichterstatter: Professor Dr. M. Hoopmann

2. Berichterstatter: Privatdozent Dr. C Gille

Tag der Disputation: 04.07.2018

Felix Borgmeier

Dissertation

Frauenklinik Tuebingen

Matrikelnr: 3502738

Praenataldiagnostik

Meinen Eltern gewidmet.

Danksagung

An erster Stelle möchte ich Herrn Prof. Dr. med. Markus Hoopmann und Herrn Prof. Dr. med. Karl Oliver Kagan herzlich danken, die mir diese hochinteressante Arbeit ermöglicht und überlassen haben. Die zahlreichen Treffen und anregende Diskussionen während der Bearbeitungsphase der vorliegenden Arbeit waren stets hilfreich, wegweisend und motivierend. Für die kompetente, geduldige und professionelle Hilfe bei der Einleitung und der Diskussion der Studienergebnisse danke ich Herrn Prof. Dr. med. Markus Hoopmann sehr. Herrn Prof. Dr. med. Karl Oliver Kagan gilt dabei besonderer Dank für die tatkräftige Unterstützung bei der Statistik und der Erhebung der Studiendaten. In diesem Zusammenhang möchte ich ebenfalls dem ganzen Team der Pränataldiagnostik der Universitätsfrauenklinik Tübingen danken, welches mich fortwährend unterstützt hat.

Weiterhin gilt mein Dank auch Herrn Prof. Dr. med. Diethelm Wallwiener für die Möglichkeit, diese Arbeit an der Universitätsfrauenklinik Tübingen durchführen zu können und für das mir entgegengebrachte Vertrauen im Umgang mit sensiblen Datenerhebungen.

Auch bei Herrn Patrick Zanker möchte ich mich recht herzlich bedanken, der mir bei der Erhebung der Studiendaten bedeutsam geholfen hat, indem er die Datenerhebung gewissenhaft und unabhängig von mir wiederholt hat.

Meinen größten Dank möchte ich allerdings meiner Familie aussprechen. Sie zeigte nicht nur Interesse an meiner Arbeit, sondern stand mir jederzeit verlässlich mit Rat und Tat zur Seite und unterstützte mich unermüdlich in der Realisierung meines Vorhabens. Sie bestärkte mich auch in schwierigen Zeiten, den eingeschlagenen Pfad bis zum erfolgreichen Abschluss zu verfolgen, sodass mir die mit der Erstellung dieser Arbeit verbundenen Mühen spürbar erleichtert wurden. Ganz besonderer Dank gilt meinem Vater, Herrn Dr. Bernd Borgmeier,

mit dem ich die Anwendung und den Gebrauch statistischer Verfahren stets ausgiebig diskutieren konnte und der mich fortwährend geduldig und verständnisvoll unterstützt hat.

Inhaltsverzeichnis

1	Einleitung.....	1
1.1	Definition	1
1.2	Epidemiologie.....	1
1.3	Klassifikation	2
1.4	Ätiologie.....	4
1.4.1	Embryologie	4
1.4.2	Pathogenese	6
1.5	Klinik.....	7
1.5.1	Postnatale Klinik der Ösophagusatresie.....	7
1.5.2	Häufig assoziierte Fehlbildungen	8
1.5.3	Komplikationen.....	9
1.6	Diagnostik.....	9
1.7	Operative Maßnahmen.....	10
1.8	Prognose	10
1.9	Pränatale Diagnostik	12
1.10	Problematik und Fokus der Studie.....	15
2	Methodik	16
2.1	Studienkollektiv.....	16
2.2	Datenerhebung.....	17
2.3	Statistische Auswertung	20
3	Ergebnisse.....	22
3.1	Deskriptive Statistik	22

3.1.1	Messzeitpunkt gemessen in Schwangerschaftswochen	22
3.1.2	Testgröße Magenumfang	25
3.1.3	Rohdaten der Testgröße Magenumfang in Abhängigkeit vom Abdomenumfang28	
3.1.4	Gewichtsverteilung der Studienteilnehmerinnen.....	31
3.1.5	Auffälligkeiten in der Testgruppe.....	32
3.2	Konkordanzanalyse der gemessenen Daten.....	34
3.2.1	Reproduzierbarkeit der Kontrollgruppe: Untersucher F.....	34
3.2.2	Reproduzierbarkeit der Kontrollgruppe Untersucher P	36
3.2.3	Reproduzierbarkeit der Kontrollgruppe.....	38
3.3	Regressionsanalyse zur Identifikation einer Risikogruppe	41
3.3.1	Bivariate lineare Regressionsanalyse der Kontrollgruppe.....	41
3.3.2	Stichprobenvergleich mit standardisierten Residuen (z-Werte).....	46
3.3.3	Identifikation einer Risikogruppe anhand des fetalen Magenumfangs.....	50
3.3.4	Bewertung der Regressionsgeraden für die Kontrollgruppe.....	56
4	Diskussion.....	58
4.1	Der Magenumfang in Anhängigkeit vom Abdomenumfang als Indikator für eine Ösophagusatresie	58
4.2	Problematisch hohe „Falsch-Positiv Rate“.....	60
4.3	Suche nach einem „Pouch-sign“ zur Verbesserung der pränatalen Detektionsrate	62
4.4	Das Kernproblem: Ösophagusatresien mit tracheoösophagealer Fistel 63	
4.5	Die Polyhydramnie als wichtiges Zusatzkriterium.....	65
4.6	Diagnostische Alternativen zur Verbesserung der pränatalen Detektionsrate	67

4.7	Nutzen einer pränatalen Diagnose	69
4.8	Kritische Betrachtung dieser Arbeit	71
4.8.1	Keine Berücksichtigung der Dynamik der Magenfüllung	71
4.8.2	Der Magenumfang als ein weniger fehleranfälliger Parameter als die Magenfläche	72
4.8.3	Die Qualität der Schätzung ist entscheidend für die Aussagekraft der Analyse.	73
4.8.4	Ausreißer in der Subgruppe „TÖF“ des Testkollektivs	75
4.8.5	Testkollektiv spiegelt nicht die Verteilung der Grundgesamtheit wieder	77
4.8.6	Reproduzierbare Ergebnisse durch objektivierte Messung	77
4.9	Fazit für die Praxis.....	79
5	Referenzen.....	80
6	Anhang.....	87

Abbildungsverzeichnis

Abbildung 1 Klassifikation nach Vogt und nach Gross:.....	3
Abbildung 2 (Fausett and Klingensmith, 2012): Theorien zum Trennungsmechanismus zwischen Trachea und Ösophagus:.....	5
Abbildung 3 Biometrieebene für die Messung des fetalen Magenumfangs (Hoopmann et al., 2015).....	18
Abbildung 4 Boxplot: Verteilung Messzeitpunkte:	23
Abbildung 5 Histogramm: Verteilung Messzeitpunkte:.....	24
Abbildung 6 Boxplot des Magenumfangs in den Stichproben der Studie:.....	26
Abbildung 7 Histogramm des Magenumfangs in der Kontrollgruppe und Testgruppe:.....	28
Abbildung 8 Punktwolke Magenumfang zu Abdomenumfang:.....	30
Abbildung 9 Boxplot zum Gewicht vor der Schwangerschaft:.....	31
Abbildung 10 Histogramm zu Gewicht vor der Schwangerschaft:.....	32
Abbildung 11 Punktwolke der Kontrollgruppe für Untersucher F:.....	34
Abbildung 12 Bland-Altman-Plot der Kontrollgruppe Untersucher F:	35
Abbildung 13 Punktwolke der Kontrollgruppe Untersucher P:.....	37
Abbildung 14 Bland-Altman-Plot der Kontrollgruppe Untersucher P:.....	38
Abbildung 15 Punktwolke Kontrollgruppe Untersucher F vs. Untersucher P:..	40
Abbildung 16 Bland-Altman-Plot der Kontrollgruppe zum Vergleich von Untersucher F zu Untersucher P:	41
Abbildung 17 Bivariate lineare Regression: Magen- gegen Abdomenumfang der Kontrollgruppe:	43
Abbildung 18 Boxplot Residuen der Kontrollgruppe:.....	44
Abbildung 19 Histogramm Residuen der Kontrollgruppe:	45
Abbildung 20 Q-Q Plot Residuen der Kontrollgruppe:.....	45
Abbildung 21 Boxplot der Z-Werte:	48
Abbildung 22 Vergleich zwischen der Kontrollgruppe und der Testgruppe anhand der Regression der Kontrollgruppe:	52
Abbildung 23 Vergleich der Z-Werte in Abhängigkeit vom Abdomenumfang:..	54

Abkürzungen

AU	→	Abdomenumfang
ANOVA	→	eng. Analysis of Variance
bez.	→	bezüglich
ca.	→	circa
cm	→	Zentimeter
COR, r	→	Korrelationskoeffizient „r“
Diff.	→	Differenz
EA	→	eng., esophageal atresia
F1	→	1. Messung von Untersucher F
F2	→	2. Messung von Untersucher F
Fl	→	Femurlänge
kg	→	Kilogramm
KI	→	Konfidenzintervall
KU	→	Kopfumfang
Mean (eng.)	→	Mittelwert
ml	→	Milliliter
mm	→	Millimeter
MU	→	Magenumfang
MW	→	Mittelwert
neg.	→	negativ
ÖA	→	Ösophagusatresie
P1	→	1. Messung von Untersucher P
P2	→	2. Messung von Untersucher P
pos.	→	positiv
R ²	→	Bestimmtheitsmaß
SEE	→	eng., Standard Error of Estimate
SSW	→	Schwangerschaftswoche(n)
STABW	→	Standardabweichung
Stat. Kenngrößen	→	Statistische Kenngrößen
TEF	→	eng., tracheoesophageal fistula

TÖF → Tracheoösophageale Fistel

z.B. → zum Beispiel

1 Einleitung

1.1 Definition

Die Ösophagusatresie (ÖA) beschreibt eine angeborene Kontinuitätsunterbrechung des Ösophagus (AWMF online, 2012). Dieser kann entweder blind, in einem proximalen oder distalen Stumpf enden. Zusätzlich kann eine proximale/distale tracheoösophageale Fistel (TÖF) vorliegen. Abhängig vom Abstand der beiden Ösophagusenden unterscheidet man „kurzstreckige“ von „langstreckigen“ ÖA (AWMF online, 2012). Es werden bestimmte Parameter herangezogen (Zentimeterabstand, Wirbelkörperanzahl, chirurgische Machbarkeit), um zwischen „kurz- und langstreckigen“ ÖA/TÖF zu differenzieren (AWMF online, 2012).

1.2 Epidemiologie

Mit einer Inzidenz zwischen 1/2500 und 1/4000 Lebendgeburten stellt die ÖA/TÖF eine seltene Anomalie dar, die jedoch (Pinheiro et al., 2012, de Jong et al., 2010b, Ioannides et al., 2010) bezogen auf alle Ösophagusfehlbildungen am häufigsten festzustellen ist. In den meisten Fällen tritt die ÖA/TÖF sporadisch auf (Ioannides et al., 2010). Zwillingsstudien zeigen, dass das relative Risiko für Zwillinge 2,56-mal höher ist als für Einzelkinder (Pinheiro et al., 2012, Houben and Curry, 2008). Die Wiederholungswahrscheinlichkeit für weitere Kinder liegt bei 1% (Spitz, 2007).

Eine EUROCAT Working Group unternahm hierzu eine umfangreiche, europaweite, epidemiologische Studie. Diese Studie erfasste 1222 Fälle mit ÖA aus einer Population von 5 019 804 Geburten. Daraus ergibt sich eine Prävalenz innerhalb Europas von 2,43 pro 10 000 Geburten (95% KI 2,3 bis 2,57). Die Prävalenz ist seit den letzten 20 Jahren konstant. Jedoch existieren starke, interregionale Schwankungen von 1,27 bis 4,55 Fälle pro 10 000 Geburten. Von diesen 1222 Fällen mit ÖA/TÖF hatten 44,7% eine isolierte ÖA/TÖF, 31,6% zeigten multiple Malformationen und 23,7% konnten Syndromen zugeordnet

werden. Dementsprechend treten mehr als 40% der ÖA isoliert auf, allerdings mit einem signifikanten Anteil an begleitenden Fehlbildungen (Pedersen et al., 2012). Laut statistischem Bundesamt beträgt die Geburtenzahl 2015 in Deutschland 738 000 (Statistisches Bundesamt) Lebendgeburten. Unter der Prävalenzannahme von 2,43 pro 10 000 Geburten in Deutschland müssten statistisch 179 Neugeborene mit einer ÖA 2015 geboren worden sein.

1.3 Klassifikation

Die Klassifikation der ÖA erfolgt nach der Lokalisation der Atresie und nach dem Kriterium einer vorhandenen TÖF. Im 17. Jahrhundert wurde die ÖA, mit assoziierter TÖF, zum ersten Mal von Thomas Gibson erfasst. In der Klinik werden heutzutage zwei verschiedene Klassifikationen zur Beschreibung der ÖA verwendet: zum einen die Klassifikation nach Vogt, von 1929 und zum anderen die Klassifikation nach Gross (beides in Abbildung 1), von 1953 (Pinheiro et al., 2012). In Deutschland ist die Klassifikation nach Vogt am geläufigsten (AWMF online, 2012).

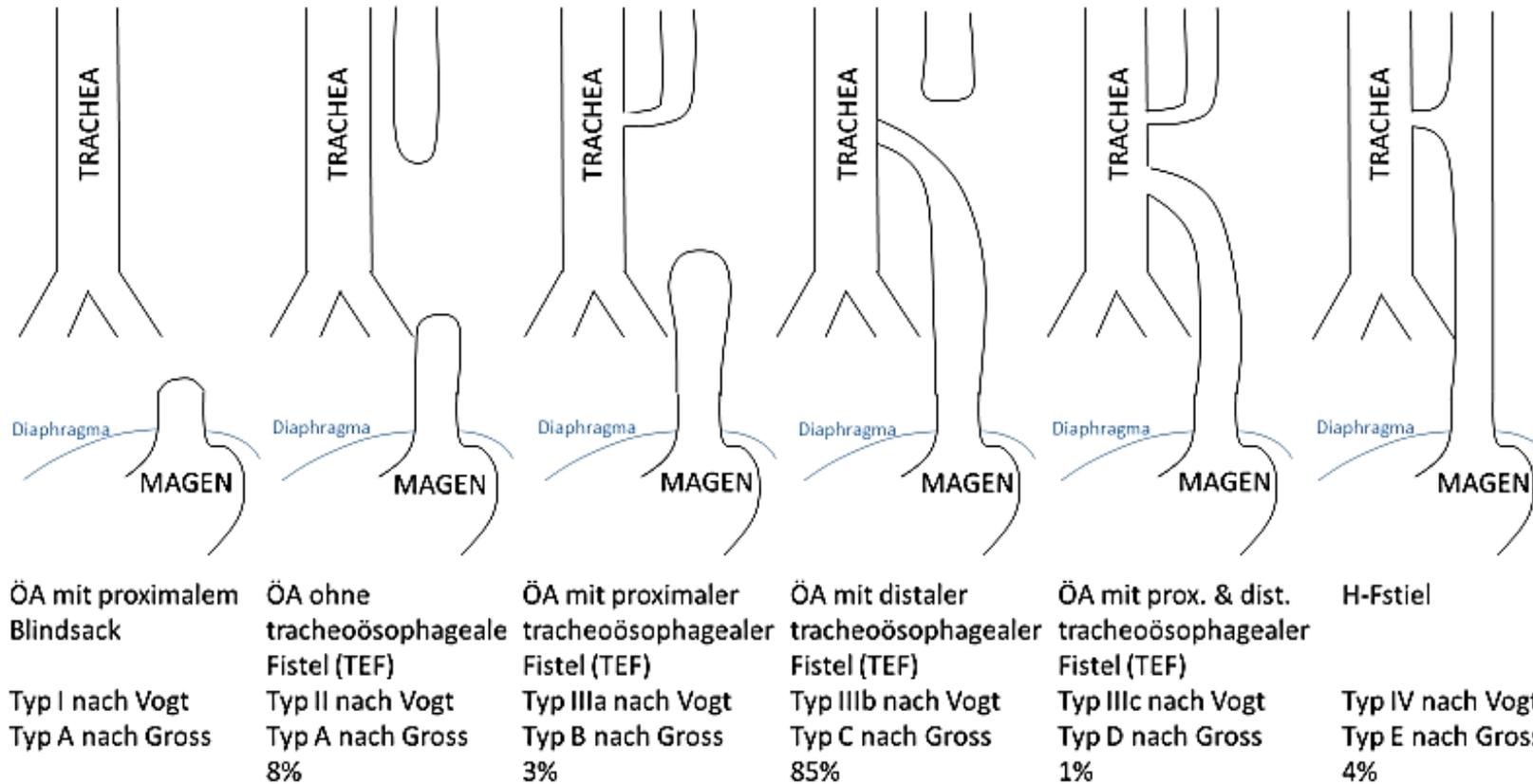


Abbildung 1 Klassifikation nach Vogt und nach Gross: Die Abbildung zeigt die verschiedenen Ösophagusatresietypen und deren Häufigkeiten (Houben and Curry, 2008, Pinheiro et al., 2012).

Mit 85% ist die proximale ÖA mit distaler TÖF, d.h. Typ IIIb nach Vogt und Typ C nach Gross, die häufigste Form der ÖA. Eine isolierte ÖA ohne TÖF (Vogt II, Gross A) stellt mit 8% die zweithäufigste Form dar, gefolgt von einer „H-Fistel“ (Vogt IV, Gross E) mit 4%. Die proximale TÖF mit distaler Atresie (Vogt IIIa, Gross B) ist mit 3% vertreten. Die proximale und distale TÖF (Vogt IIIc, Gross D) liegt bei unter 1% und gilt damit als Ausnahme. Die Bestimmung der anatomischen Variante der ÖA ist bedeutsam, weil davon das weitere medizinische Vorgehen sowie das operative Eingreifen abhängt (Pinheiro et al., 2012, Houben and Curry, 2008).

Zusammenfassend ist die ÖA mit proximaler Atresie und distaler TÖF mit Abstand die häufigste Form dieser Fehlbildung. In Folge dieser Erkenntnis müssten von insgesamt 179 Neugeborenen mit ÖA im Jahr 2015 theoretisch 152 Neugeborene eine TÖF aufweisen.

1.4 Ätiologie

Bis heute sind die genauen embryologischen Vorgänge der Ösophageal- und Trachealentwicklung weitestgehend ungeklärt, sodass die Ätiologie der ÖA, mit und ohne TÖF, als unvollständig anzusehen ist (de Jong et al., 2010b, Ioannides and Copp, 2009, Pinheiro et al., 2012, Kluth and Fiegel, 2003).

1.4.1 Embryologie

Ösophagus und Trachea sind endodermalen Ursprungs. In der vierten Schwangerschaftswoche (SSW) (Carnegie Stadien 13-16 (Ioannides et al., 2010)) lässt sich der Vorderdarm (Teil des primitiven Darmrohrs) in einen ventralen, respiratorischen und einen dorsalen, ösophagealen Teil unterteilen. Die Entwicklung der Lungenbläschen geschieht zeitlich vor der Trachealentwicklung. Der Trennungsmechanismus zwischen Trachea und Ösophagus ist nicht vollständig erforscht (de Jong et al., 2010b, Pinheiro et al., 2012, Ioannides and Copp, 2009, Kluth and Fiegel, 2003). Theorien über den Trennungsmechanismus sind in Abbildung 2 dargestellt. Die Septumhypothese

erscheint, basierend auf der aktuellen Studienlage, als die wahrscheinlichste Theorie, bei der es durch komplexe Faltungsprozesse bzw. Invaginationen zur Trennung von Ösophagus und Trachea kommt (Fausett and Klingensmith, 2012, Ioannides and Copp, 2009, Ioannides et al., 2010, Kluth and Fiegel, 2003, Spitz, 2007).

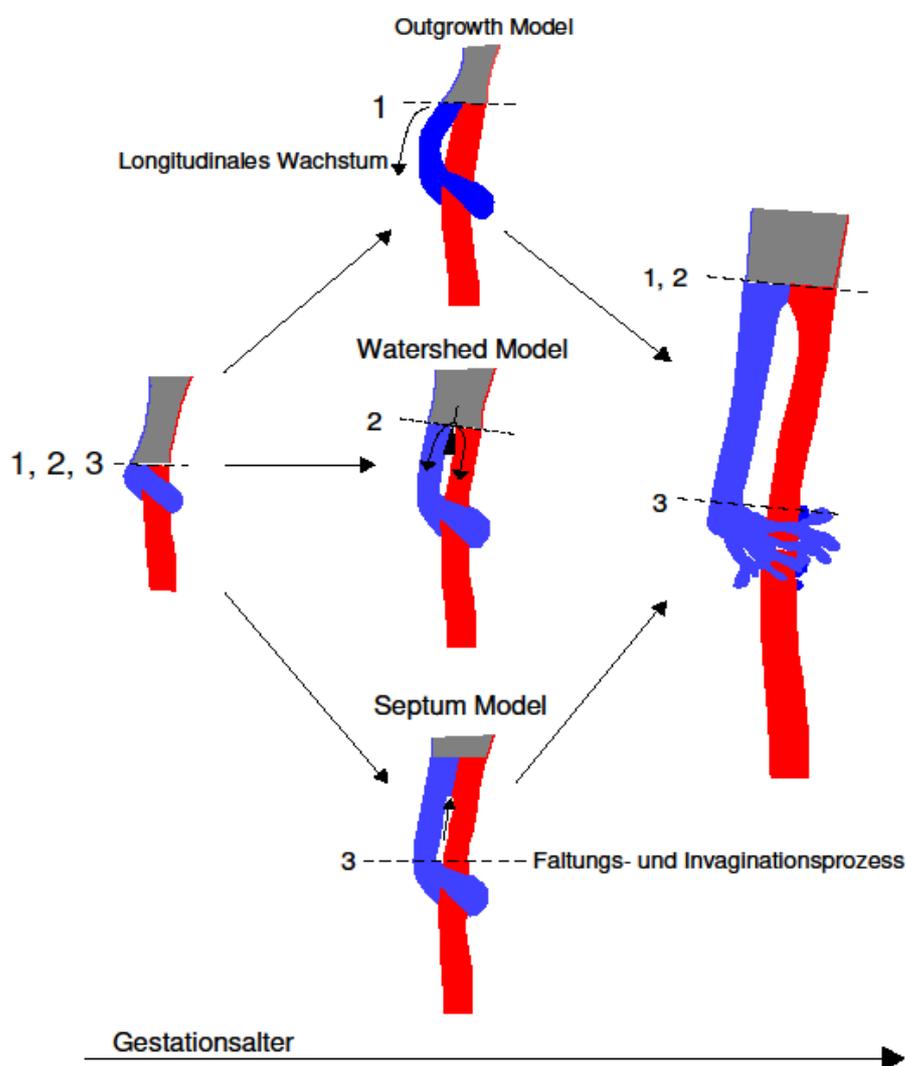


Abbildung 2 (Modifizierte Grafik aus Fausett and Klingensmith, 2012): Theorien zum Trennungsmechanismus zwischen Trachea und Ösophagus:

Die „**Outgrowth**“ – **Theorie (1)** vermutet eine schnelle Zellteilung mit longitudinalem Wachstum am proximalen Ende der Trachea. Dies würde dem Wachstumsmuster der Schilddrüse, des

Thymus oder des distalen bronchopulmonalen Systems entsprechen (Ioannides and Copp, 2009, Sasaki et al., 2001, Ioannides et al., 2010).

*Die **Watershed – Theorie (2)** geht von einer Keilformierung (Wedge) aus Mesenchymzellen, auf Höhe der Lungenbläschen, aus. Der stark proliferierende Vorderdarm wird über diesen „Mesenchymkeil“ geteilt (Fausett and Klingensmith, 2012, Ioannides et al., 2010).*

*Die **Septumhypothese (3)** basiert auf einem zur Medianlinie hin approximierenden Faltungsprozess, bzw. auf Invaginationen auf Höhe der Lungenbläschen. Rostral entstehen bilateral paarig angelegte Laryngealfalten und dorsal entwickelt sich eine singuläre, ventrale Tracheoösophagealfalte. Durch das Approximieren dieser Falten kommt es zur Trennung. Eine Pharyngoösophagealfalte bildet die Grenze zwischen Pharynx und Ösophagus (Fausett and Klingensmith, 2012, Ioannides and Copp, 2009, Ioannides et al., 2010, Kluth and Fiegel, 2003, Spitz, 2007).*

Interessant für den Trennungsprozess ist, wie die Natur molekularbiologisch zwischen Ösophagus und Trachea in deren Entwicklung differenziert. So spielen drei wichtige Akteure in der biochemischen Markierung von Ösophagus und Trachea eine große Rolle: Nkx 2.1, Sox2 und Shh. Für das respiratorische System existiert der tracheale Marker Nkx 2.1, dessen Expression durch Tbx4 (T-Box Transkriptionsfaktor) und dessen Produkt Fgf10 (fibroblast growth factor 10) reguliert wird (Sakiyama et al., 2003, Ioannides et al., 2010). Als Marker für den Ösophagus wird Sox2 exprimiert. Hierdurch entsteht eine klare ventro-dorsale Grenze. Falls Nullmutationen in einem dieser beiden Marker vorliegen, so wird der jeweils andere Marker hochexprimiert und es kann nicht mehr zwischen Ösophagus und Trachea differenziert werden (Ioannides et al., 2010, Ioannides and Copp, 2009). Generell in der Organogenese ist das Sonic Hedgehog (Shh) Glykoprotein von zentraler Bedeutung, so auch bei der Tracheal- und Ösophagealentwicklung. Während der Trennung kommt es zu einem Shh-shift. Hierbei wechselt die hohe Expression von Shh vom trachealen Teil in den ösophagealen Teil. Die zeitlichen und räumlichen Expressionsmuster sind dabei entscheidend (Ioannides et al., 2003, Ioannides et al., 2010, Ioannides and Copp, 2009). Bei großen Fehlbildungssyndromen wie z.B. dem VACTERL-Syndrom ist dieses Glykoprotein im Fokus der Forschung (Felix et al., 2007).

1.4.2 Pathogenese

Auf Grund der nicht vollständig geklärten Embryogenese bleibt auch die Pathogenese zunächst nicht vollständig erforscht. Zur realitätsnahen Simulation

und Erforschung der ÖA wird das Anthracyclin Adriamycin (Doxorubicin, DOX) Ratten- oder Hühnerembryos verabreicht. In 41,2% der Fälle entsteht eine ÖA, die epidemiologisch und anatomisch der ÖA im Menschen ähnelt (Fausett and Klingensmith, 2012, Goto et al., 2001, Qi et al., 1996).

Es besteht allgemeiner Konsens darüber, dass die Pathogenese multifaktoriell bedingt ist. Auf Grund der oben beschriebenen, vermuteten Embryogenese und molekularbiologischen Erkenntnisse kann eine fehlerhafte Septumbildung als die wahrscheinlichste Ursache vermutet werden (Fausett and Klingensmith, 2012, Ioannides et al., 2010). Fehlerhafte Expression oder Regulation der oben beschriebenen Marker können zu einer fehlerhaften Ösophagusanlage führen. Am Beispiel der Shh-Dynamik scheint der Shh-shift umgekehrt zu sein (Goyal et al., 2006, Ioannides and Copp, 2009, Ioannides et al., 2010).

Die Fistelbildung könnte auf Grund einer defekten Interaktion zwischen Epithelial- und Mesenchymalzellen entstehen (Crisera et al., 2000). Durch eine Fistel wird eine Kontinuität des Gastrointestinaltraktes sichergestellt (Ioannides and Copp, 2009).

1.5 Klinik

1.5.1 Postnatale Klinik der Ösophagusatresie

Typischerweise fällt postpartal eine ÖA durch das frustrane Sondieren des Neugeborenen auf. Hierbei ist ein federnder Widerstand in der Speiseröhre zu bemerken (AWMF online, 2012). Diese Sonde sollte unter konstantem Sog im ösophagealen Blindsack belassen werden, um eine Aspiration zu verhindern (Spitz, 2007). Des Weiteren können Dysphagie (repetitives Verschlucken), Dyspnoe und schaumiger Speichel vor Mund und Nase Hinweise auf eine ÖA sein. Falls der Verdacht bei einem Neugeborenen besteht, sollten jegliche Fütterungsversuche unterlassen werden (AWMF online, 2012, Houben and Curry, 2008, Spitz, 2007).

1.5.2 Häufig assoziierte Fehlbildungen

Begleitende Fehlbildungen sind bei einer ÖA keine Seltenheit und treten in über 50% der Fälle auf. Interessanterweise ist das Vorkommen von Fehlbildungen unterschiedlich häufig in den Ösophagusatresietypen verteilt. So kommt es in ÖA ohne Fistel in ca. 65% zu weiteren Fehlbildungen, wogegen diese nur in 10% der Fälle bei H-Fisteln auftreten (Holland and Fitzgerald, 2010, Pinheiro et al., 2012, Spitz, 2007).

Assoziierte Anomalien kommen oftmals in Zusammenhang mit Syndromen oder Chromosomenaberrationen vor. Allerdings können Fehlbildungen auch unabhängig von anderen Syndromen auftreten (Pedersen et al., 2012).

Falls ein Neugeborenes eine ÖA/TÖF und zwei weitere kongenitale Anomalien im Vertebral-, Anorektal-, Kardial-, Renal- oder Extremitätenbereich aufweist, kann dies dem VACTERL-Syndrom zugeordnet werden. Eine Fehlregulation/Expression des Shh-Gens wird in diesem Zusammenhang vermutet (Pinheiro et al., 2012). Bis zu 10% der Neugeborenen mit ÖA zeigen eine Assoziation zum VACTERL – Syndrom (Felix et al., 2007).

Bei zusätzlichem Auftreten von Kolobom, Herzfehler, Choanalatresie, Wachstumsrestriktion, genitaler Hypoplasie oder Ohrdeformitäten zu einer ÖA können diese unter dem Symptomkomplex des CHARGE-Syndroms zusammengefasst werden (Holland and Fitzgerald, 2010). 15-20% der CHARGE-Syndrome treten zusammen mit einer ÖA auf (Lee et al., 2008). Generell ist das CHARGE-Syndrom bei ÖA eher selten (1%) (Pedersen et al., 2012).

ÖA sind in ca. 6 – 10% der Fälle mit Chromosomenaberrationen, insbesondere in Form von Trisomien, assoziiert. Dies schließt alle lebensfähigen Trisomien ein (Felix et al., 2007).

Kongenitale kardiovaskuläre Malformationen sind die häufigsten assoziierten Anomalien, die begleitend zu einer ÖA vorkommen (Pinheiro et al., 2012). Dies ist wenig überraschend, da ein kongenitaler Herzfehler zum Symptomkomplex von allen oben genannten Syndromen gehört.

1.5.3 Komplikationen

Durch die Assoziation mit einem Polyhydramnion stellt die Frühgeburt ein klinisch relevantes Risiko für Komplikationen dar (Houben and Curry, 2008, Langer et al., 2001, Pedersen et al., 2012, Pinheiro et al., 2012, Spitz, 2007). In 38,5% der Schwangerschaften mit Polyhydramnion und ÖA kommt es zu einer Frühgeburt (Entbindung vor der 37+0. SSW) (Pedersen et al., 2012, Abele et al., 2012).

Bei Vorliegen einer distalen Fistel besteht ein Kurzschluss zwischen Trachea und Magen. So kann durch fetale Atemaktivität Luft in den Magen gelangen. Es besteht das Risiko einer massiven Überblähung des Magens mit einer daraus resultierender Rupturgefahr (AWMF online, 2012).

Weitere Komplikationen können in Zusammenhang mit assoziierten Syndromen (VACTERL, CHARGE), Fehlbildungen oder Chromosomenstörungen (Trisomie 13, 18 und 21) auftreten (Solomon, 2011, Wong et al., 2015, Roberts et al., 2016, Wyllie et al., 1994). Kardiale Komplikationen kommen bei Neugeborenen mit ÖA am häufigsten vor (Pinheiro et al., 2012). Solche Fehlbildungen können lebensbedrohlich sein.

1.6 Diagnostik

Häufig wird eine ÖA erst postpartal vermutet, basierend auf frustranen Fütterungs- oder auffälligen Sondierungsversuchen (AWMF online, 2012). Ein geäußelter Verdacht auf eine ÖA ergibt die Indikation für eine Röntgen-Thorax-Aufnahme, auf der zusätzlich der Oberbauch abgebildet ist. Oftmals ist das Sondenknäuel im proximalen Blindsack erkennbar (Houben and Curry, 2008). Bei luftgefülltem Magen muss eine Fistel zwischen Magendarmtrakt und Trachea bestehen (Typ IIIb, IIIc und IV nach Vogt). Bei Typen ohne Fistel (meist Typ II nach Vogt) präsentiert sich der Magen folglich luftleer. Durch die zusätzliche Darstellung des oberen Abdomens wird eine Distanzmessung zwischen den beiden Ösophagusenden möglich, die der Operationsplanung dient. Bei unklaren Befunden besteht die Option der Kontrastmittelgabe oder einer Bronchoskopie

und in letzter Konsequenz die intraoperative Sondierung (AWMF online, 2012). Auf Grund der Assoziation mit anderen Fehlbildungen muss im Anschluss an die Diagnosestellung hinsichtlich weiterer Anomalien untersucht werden.

1.7 Operative Maßnahmen

Die erste erfolgreiche operative Maßnahme erfolgte 1941 durch Cameron Haight (Pinheiro et al., 2012). Ein operativer Eingriff ist kurz nach der Geburt entscheidend für das Überleben des Neugeborenen. Die Operation sollte in einem spezialisierten Krankenhaus mit angeschlossener neonatologischer Intensivstation durchgeführt werden. Hierfür kann es nötig sein, Mutter und Kind nach der Geburt zu trennen, falls die Geburt in einer Einrichtung erfolgte, die diese Behandlungsmöglichkeiten nicht anbietet (Garabedian et al., 2015, Pinheiro et al., 2012). Die Dringlichkeit anschließender Behandlungen ist vom Vorliegen weiterer Fehlbildungen abhängig.

Operativ kann sowohl offen chirurgisch (Thorakotomie) als auch thorakoskopisch vorgegangen werden (AWMF online, 2012). Nach Möglichkeit wird dieser Eingriff elektiv durchgeführt. Bei vorliegender Fistelbildung wird das Ziel verfolgt, sie zu verschließen und eine Kontinuität des Gastrointestinaltraktes herzustellen. Falls der Abstand zwischen den beiden Ösophagusenden zu groß für eine spannungsfreie Anastomose ist, wird ein Magenhochzug durchgeführt. Eine durch die TÖF bedingte Rupturgefahr stellt die Hauptindikation zu einer Notfalloperation dar. Die Rupturgefahr besteht auch bei der intraoperativen Beatmung, wenn eine TÖF vorliegt (AWMF online, 2012, Houben and Curry, 2008).

1.8 Prognose

Bis 1944 betrug die Mortalitätsrate bei Kindern mit Oesophagusatresie 100%. Durch verbesserte anästhesiologische, operative und intensivmedizinische

Verfahren ist die Mortalität deutlich gesunken. Hierbei sind es die kongenital begleitenden Anomalien, die letztendlich zum Tode führen (Fausett and Klingensmith, 2012). Zur Einschätzung der Prognose entwickelte Spitz 1994 eine Risikogruppenklassifikation, welche auf einer vorherigen Arbeit von Waterston aufbaut (Spitz, 2007) und im Folgenden nach Spitz in drei Gruppen unterteilt wird:

- I. Geburtsgewicht über 1500g ohne schwerwiegenden Herzfehler
- II. Geburtsgewicht unter 1500g oder schwerwiegender Herzfehler
- III. Geburtsgewicht unter 1500g und zusätzlicher schwerwiegender Herzfehler

Die Bezeichnung „schwerwiegender Herzfehler“ definiert eine akut lebensbedrohliche Situation, die sofortiges medizinisches/chirurgisches Handeln verlangt oder palliativ begleitet werden muss (Spitz, 2007). Nach der vorliegenden Klassifikation wurden Überlebensraten für Gruppe I von 98%, für Gruppe II von 82% und für Gruppe III von 50% angeführt. Allerdings sind diese Daten nur orientierend zu verstehen, denn gerade die hohen Risikogruppen zeigen deutlich bessere Ergebnisse auf Grund von Erfolgen in der Neonatologie, Kinderanästhesie und Kinderchirurgie (AWMF online, 2012).

Insgesamt liegt die durchschnittliche Überlebensrate bei 86,9% in der ersten postnatalen Woche. Bei einer isolierten ÖA/TÖF mit einem Gestationsalter über der 38. SSW steigt die Überlebensrate in der ersten Woche auf 99,2% (Pedersen et al., 2012).

Die prognostische Auswirkung auf die Mortalität der pränatalen Detektion einer ÖA wird in der Literatur kontrovers diskutiert (Brantberg et al., 2007). Durch die pränatale Diagnostik wird ein größeres Zeitfenster geschaffen, sodass eine adäquate Beratung stattfinden und das Patientenmanagement anpasst werden kann.

Bezüglich der langfristigen Prognose können verschiedene Komplikationen auftreten. Im Vordergrund stehen dabei Schluckstörungen, Atemwegsinfektionen (bei Typen mit Fistel) und ösophageale Neoplasien (AWMF online, 2012).

1.9 Pränatale Diagnostik

Die präzise sonographische, antenatale Diagnostik der ÖA erweist sich bis heute als eine Herausforderung für Mensch und Maschine. Von einer pränatalen Diagnose verspricht man sich einen Zeitgewinn für eine verbesserte und ausführlichere Beratung und Aufklärung der Eltern sowie eine verbesserte medizinische Planung und Behandlung (Brantberg et al., 2007, Garabedian et al., 2015).

Im Vordergrund der Pränataldiagnostik steht die Ultraschalluntersuchung. Im Rahmen der von gesetzlichen Krankenkassen angebotenen Vorsorgeuntersuchungen können Befundkonstellationen auftreten, die den Verdacht auf eine ÖA ergeben. So erweist sich die Kombination aus fehlender Magenblase und Polyhydramnion in mehrfach wiederholten pränatalen Ultraschalluntersuchungen als hoch suspekt für das Vorliegen einer ÖA. Diese Befundkonstellation findet sich meistens bei einer ÖA ohne Fistel. Fälle mit Fistel können oftmals durch dieses Vorgehen nicht identifiziert werden. Der positive-prädiktive Wert für diese Kombination aus Polyhydramnion und nicht-darstellbarer Magenblase liegt zwischen 25% und 100% (Houben and Curry, 2008, Kunisaki et al., 2014, Bradshaw et al., 2016). Solche Schwankungen können zumeist mit niedrigen Fallzahlen und unterschiedlichen Häufigkeiten der Fälle ohne Fistel erklärt werden (Kunisaki et al., 2014). Die Sensitivität und Spezifität eines fehlenden bzw. kleinen Magens mit oder ohne Polyhydramnion auf eine ÖA variiert stark zwischen Krankenhäusern. In Level I Perinatalzentren kann mit einer Sensitivität von 57%, bei einer Spezifität von 99%, gerechnet werden (Bradshaw et al., 2016). Die pränatale Detektionsrate der ÖA liegt im europaweiten Durchschnitt dennoch nur bei 36,5%, mit deutlichen interregionalen Schwankungen (Bereich: <10 – 50%) (Pedersen et al., 2012,

Houben and Curry, 2008). Hierbei muss beachtet werden, dass für die Diagnosestellung alle diagnostischen Verfahren zur Verfügung standen (Pedersen et al., 2012). Bei geäußertem Verdacht auf eine ÖA können weitere diagnostische Maßnahmen in Erwägung gezogen werden. Hierzu gehören z.B. das fetale MRT und eine Fruchtwasseruntersuchung (Garabedian et al., 2014, Langer et al., 2001). Allerdings ist eine sichere Diagnosestellung oft nur postnatal möglich.

Die Gründe für die niedrige Detektionsrate sind vielschichtig. Zunächst erweist sich die pränatale Darstellbarkeit des Ösophagus als äußerst schwierig (Blaas et al., 2013, Malinger et al., 2004). In seltenen Fällen lässt sich der Ösophagus in der 9. SSW als ein heller Streifen hinter dem Herzen darstellen. Zu diesem Zeitpunkt ist der Magen gelegentlich als echoarme Struktur aufzufinden. Im Rahmen des ersten Screenings zwischen der 11. und 14. SSW ist der Ösophagus manchmal als eine echoreiche Doppelkontur zu erkennen, die im darstellbaren Magen endet (Blaas et al., 2013).

Die Einschätzung der fetalen Magengröße in der abdominalen Biometrieebene erfolgt zudem oft subjektiv, obwohl Studien zur objektiven Magenentwicklung (Wachstumskurven) existieren (Goldstein et al., 1987, Nagata et al., 1990, Sase et al., 2002). Qualitativ wird dabei meist zwischen „normal“, „klein“ und „nicht darstellbar“ unterschieden (Bradshaw et al., 2016, Houben and Curry, 2008, Stringer et al., 1995, Kunisaki et al., 2014). Ein nicht-darstellbarer Magen im pränatalen Ultraschall ist ein sehr seltenes Phänomen (0,07%) (Brumfield et al., 1998). Von diesen 0,07% sind 63-65% der Fälle ein Hinweis auf eine Anomalie (McKenna et al., 1995, Brumfield et al., 1998). Ein nicht darstellbarer Magen ist bei den ÖA meistens mit einer Anatomie ohne Fistel assoziiert (Sase et al., 2002). Fälle mit Fistel stellen eine Schwierigkeit in der pränatalen Diagnostik dar (Garabedian et al., 2015). Sobald es zu einer Magenfüllung kommt, erweist sich die subjektive Einschätzung als problematisch, da keine objektiven Grenzwerte existieren. Das macht die Beurteilungsqualität untersucherabhängig und damit

schwer reproduzier- und vergleichbar. Als Resultat ergibt sich eine hohe Anzahl falsch-negativer und falsch-positiver Befunde. Zudem ist ein kleiner/fehlender Magen ein indirekter Hinweis auf eine ÖA, der eine Vielzahl an Differentialdiagnosen zulässt. Hierzu gehören Anomalien des respiratorischen, gastrointestinalen Systems, des zentralen Nervensystems (ZNS) und des Karyotyps sowie neuromuskuläre Syndrome (Houben and Curry, 2008).

Die Messung der Fruchtwassermenge gehört zu den Standardmessungen im pränatalen Ultraschall. Dabei zählt ein Polyhydramnion zu einem anerkannten Risikofaktor für eine ÖA. Allerdings stellt ein Polyhydramnion, ähnlich wie die Magengröße, lediglich ein indirektes Hinweiszeichen dar. Unter physiologischen Bedingungen erhöht sich die Fruchtwassermenge um 10 ml/Tag bis zur 34. SSW und sinkt sodann wieder ab. Die Fruchtwassermenge wird anhand des „Amniotic fluid index (AFI)“ beschrieben, wobei die Normalwerte vom Gestationsalter abhängig sind. Es werden vertikale Fruchtwasserdepots in vier Quadranten des mütterlichen Bauchs in Zentimeter gemessen und als Summe ausgedrückt (Houben and Curry, 2008). Ein Polyhydramnion ist definiert bei einem AFI über der 97,5. Perzentile, bei einem AFI über 24 cm oder bei einem Fruchtwasserdepot von mindestens 8 cm in einem Quadranten (Reddy et al., 2014). Dass es vor der 24. SSW zur Ausprägung eines Polyhydramnions kommt, gilt eher als untypisch (Kunisaki et al., 2014, Houben and Curry, 2008).

Die Prävalenz eines Polyhydramnions beläuft sich auf 0,4 bis 1,5%. Oftmals liegt ein marginaler Befund vor (AFI 25,0 – 29,9 cm), der mit Makrosomie assoziiert werden kann (Stoll et al., 1991, Houben and Curry, 2008). Bei AFI – Werten über 30 cm spricht man von einem ausgeprägten Polyhydramnion. Die Ätiologie ist nicht vollends bekannt. Obwohl die Ätiologie nicht vollends bekannt ist, kann eine positive Korrelation zwischen einer erfolgreichen Ursachensuche und dem Ausmaß des Polyhydramnions konstatiert werden (Dashe et al., 2002, Hill et al., 1987). Im Allgemeinen kennzeichnet ein Polyhydramnion ein unspezifisches Symptom, welches eine Vielzahl an Differentialdiagnosen zulässt, meistens allerdings pränatal idiopathischer Natur ist (40%) (Abele et al., 2012). Von den

zunächst pränatal idiopathischen Fällen können bei 9,3% postpartal Anomalien festgestellt werden (Abele et al., 2012). Bei einem Drittel der Schwangerschaften mit Polyhydramnion lassen sich bereits pränatal Anomalien ermitteln. Hierzu gehören gastrointestinale Obstruktionen (32,6%) sowie ZNS und kardiale Anomalien (Dashe et al., 2002, Abele et al., 2012). Zu den Fehlbildungen, die im pränatalen Ultraschall häufig nicht auffallen, gehören besonders ÖA mit TÖF, kardiale Septumdefekte und LKG-Spalten (Dashe et al., 2002). Dennoch ist unter ÖA ein Polyhydramnion eine häufige pränatale Begleiterscheinung und tritt in bis zu 73% der Fälle auf (Kunisaki et al., 2014).

1.10 Problematik und Fokus der Studie

Epidemiologisch ist die ÖA eine sehr seltene kongenitale Anomalie. Anatomisch lässt sich diese Fehlbildung in Atresien mit und ohne Fistel unterteilen, wobei die Ausprägungen mit Fistel dominieren (ca. 85%) (AWMF online, 2012). Der Verdacht auf eine ÖA im pränatalen Ultraschall basiert maßgeblich auf der indirekten, unspezifischen und zum Teil subjektiven Befundkonstellation aus nicht darstellbarer oder kleiner Magenblase und/oder Polyhydramnion. Ein nicht-darstellbarer Magen führt meist zum Verdacht auf eine ÖA ohne Fistel, die allerdings die Minderheit der ÖA darstellt (Bradshaw et al., 2016, Stringer et al., 1995). Die Sensitivität und Spezifität dieser Befundkonstellation sinkt sobald es zu einer Magenfüllung kommt, z.B. durch eine Fistel (Garabedian et al., 2015). Hierbei ist die subjektive Beurteilung problematisch, die zur niedrigen pränatalen Detektionsrate (Pedersen et al., 2012) und den hohen falsch-negativen und falsch-positiven Befunden beiträgt (Bradshaw et al., 2016, Garabedian et al., 2015). So wird in der vorliegenden Arbeit der pränataldiagnostische Stellenwert der objektiven Magengröße für die Diagnose einer ÖA unter Berücksichtigung der obig erwähnten Problematik untersucht.

2 Methodik

2.1 Studienkollektiv

Durchführungsort dieser Studie war die Pränatalambulanz der Universitätsfrauenklinik Tübingen. Das Patientenkollektiv besteht aus zwei unabhängigen Stichproben: eine Kontrollgruppe und eine Testgruppe. Die Kontrollgruppe dient zur Erhebung von Normwerten, mit denen die Testgruppe verglichen werden kann. Die Datenerhebung wurde retrospektiv durchgeführt, sodass den Ultraschalluntersuchern zum Zeitpunkt der Untersuchung die Verwendung ihrer dokumentierten Bilder für diese Studie nicht bekannt war. Daher ermöglicht diese Vorgehensweise die Simulation einer routinemäßigen klinischen Praxis.

Die Kontrollgruppe umfasst ein Patientinnenkollektiv von 374 schwangeren Frauen, die einen routinemäßigen, pränatalen Ultraschalltermin zwischen den Monaten Januar und Oktober des Jahres 2013 in der Pränataldiagnostik der Universitätsfrauenklinik Tübingen wahrnahmen und sich zwischen der 19. und 40. SSW befanden. Es wurde eine Liste mit allen Patientinnen, die in diesem Zeitraum untersucht wurden, erstellt. Konsekutiv wurden anhand dieser Liste die Patientinnen retrograd untersucht und diejenigen in die Studie involviert, die sich über die Einschlusskriterien qualifizierten. Als Einschlusskriterium für die Kontrollgruppe galt eine unauffällige Einlingsgravidität mit der Geburt eines gesunden Neugeborenen und eine standardmäßige sonographische Abbildung des fetalen Abdomens. Es wurde damit sichergestellt, dass die biometrisch festgehaltenen Daten der Schwangerschaft innerhalb des Referenzbereiches liegen. Zudem durften keine Komplikationen im Schwangerschaftsverlauf vom Untersucher vermerkt worden sein.

Die Testgruppe setzt sich aus 59 Schwangerschaften zusammen, die in der Pränataldiagnostik der Universitätsfrauenklinik Tübingen zwischen den Jahren 2001 und 2013 betreut wurden. Als Einschlusskriterium galt eine postpartal

diagnostizierte ÖA, unabhängig davon, ob diese isoliert oder im Rahmen anderer Syndrome (z.B. VACTERL- oder CHARGE-Syndrom) vorkam. Es wurden ausschließlich die sonographischen Bilder berücksichtigt, die in der ersten pränatalen Untersuchung an der Universitätsfrauenklinik Tübingen erhoben wurden, sodass ausnahmslos sonographische Bilder ausgewertet wurden, die aus einem Level I Perinatalzentrum stammten.

2.2 Datenerhebung

Der Referenzstandard für diese Studie wurde mithilfe der Kontrollgruppe zentrumspezifisch festgelegt. Es wurden die Magenfläche und der Magenumfang von 374 unterschiedlichen Feten sonographisch gemessen, die alle eine unauffällige Schwangerschaft durchlaufen haben und gesund geboren wurden. Die Messungen wurden bei Frauen in unterschiedlichen Schwangerschaftswochen (zwischen der 19. und 40. SSW) durchgeführt. Aus diesen Messungen ergaben sich statistisch verwertbare Normwerte für eine „normale“ fetale Magenfläche und einen normalen Magenumfang. Diese Normwerte dienten schließlich als Vergleich zu den gemessenen Werten von Magenumfang und Magenfläche von Feten mit ÖA mit und ohne Fistelanlage.

Grundlage der retrospektiven Datenerhebung war ein digitales, medizinisches Datenverarbeitungsprogramm namens ViewPoint von GE Healthcare. Dieses IT-Programm wird an der Frauenklinik der Universität Tübingen verwendet, um die Schwangerschaftsdaten von Routineuntersuchungen zu dokumentieren. Mithilfe dieses Programmes wurde zu Beginn dieser Promotionsarbeit eine Excel Tabelle erstellt, die retrospektiv die Untersuchungstermine von schwangeren Frauen in der chronologisch korrekt terminierten Reihenfolge auflistete. Anhand dieser Liste wurden retrograd, konsekutiv die Schwangerschaftsverläufe der Patientinnen in den Blick genommen und diejenigen, die die oben beschriebenen Einschlusskriterien erfüllten, in die Kontrollgruppe dieser Studie aufgenommen. Insgesamt wurden 1781 Patientinnen gesichtet und davon 374 in die Kontrollgruppe eingeschlossen.

Weiterhin wurden die routinemäßig festgehaltenen Ultraschallbilder der Abdomenumfangsmessung auf die Abbildung einer korrekten Biometrieebene begutachtet, anhand der erforderlichen anatomischen Landmarken. Als Kriterien für eine solche Darstellung werden der Querschnitt der Wirbelsäule, ein parallel angeschnittenes Rippenpaar, ein echoarmer gefüllter Magen und die Vena umbilicalis im Übergang zum Ductus venosus und zur Vena portae angeführt.

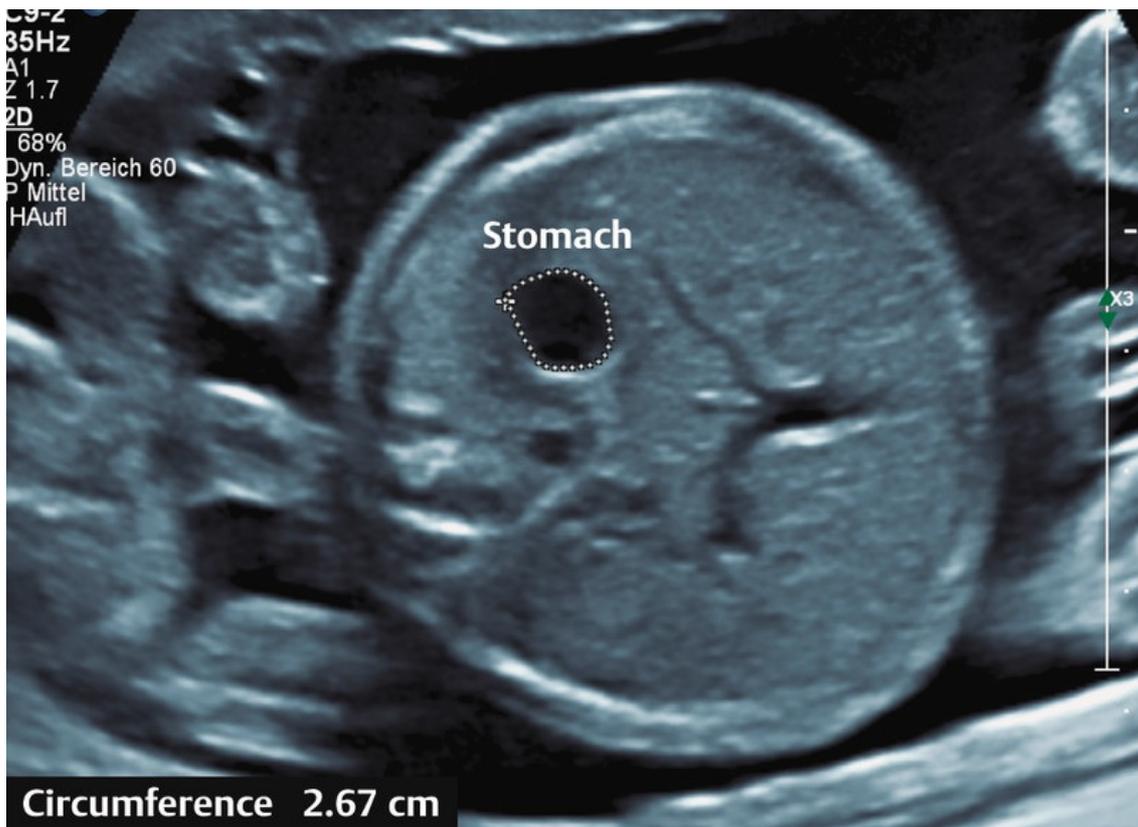


Abbildung 3 Biometrieebene für die Messung des fetalen Magenumfangs (Hoopmann et al., 2015)

Es wurde stets jenes Ultraschallbild ausgewählt, welches sich in der Biometrieebene zur Abdomenumfangsmessung befand und den oben genannten Kriterien für eine korrekte Darstellung entsprach. Im Rahmen dieser Studie wird jeder Fet nur zu einem Zeitpunkt in der Schwangerschaft vermessen. Der Magenumfang (MU, Millimeterangabe) wurde mit der Tracer-Funktion bestimmt und das IT-Programm ViewPoint errechnete daraus die Magenfläche (Quadratmillimeter). Zur Erfassung der Normwerte wurde jedes sonographische

Bild von zwei Doktoranden vermessen, die beide jedes Bild zweimal zu unterschiedlichen Zeitpunkten vermessen haben. Dieses Verfahren führten beide Doktoranden unabhängig voneinander aus und ohne die Kenntnis der Messwerte des anderen. Neben der Vermessung des Magenumfangs und der Magenfläche wurden folgende weitere Parameter für die Kontroll- sowie für die Testgruppe festgehalten: Gestationsalter (Schwangerschaftswochen, SSW), Abdomenumfang (AU in Millimetern), Gravida und Para, Biparietaldurchmesser (BPD in Millimetern), occipito-frontal Durchmesser (OFD in Millimetern), Kopfumfang (KU in Millimetern), anterior-posterior Abdominaldurchmesser (APAD in Millimetern), transversaler abdominaler Durchmesser (TAD in Millimetern), Femurlänge (FI in Millimetern) und das Gewicht der Schwangeren (in Kilogramm).

Die Vermessung der 59 pathologischen Bilder geschah ebenfalls zweimal, jedoch nur von einem Doktoranden. Es wurden jeweils die Bilder ausgewählt, die bei der ersten Untersuchung der Schwangeren an der Universitätsfrauenklinik Tübingen aufgenommen wurden. Die Kriterien zur Auswahl der Bilder und der Auswahl zur Bestimmung der Normwerte waren identisch. Aus auswertungstechnischen Gründen wurde bei einem nicht erkennbaren Magen der Magenumfang mit 0,1 mm festgehalten. Zusätzlich zu den oben genannten festgehaltenen Daten wurden bei den pathologischen Schwangerschaften noch folgende Werte erhoben: Polyhydramnion (größtes vertikales Fruchtwasserdepot über 8 cm), Durchführung einer Entlastungspunktion, Gestationsalter bei der ersten Entlastungspunktion, Karyotyp (falls dieser bestimmt wurde) und Vorliegen einer Fistel.

Zwei Doktoranden nahmen die Vermessung der fetalen Mägen auf den Ultraschallbildern des gesunden Patientenkollektivs vor. Beide waren während der Vermessungsphase im klinischen Abschnitt des Medizinstudiums an der Universität Tübingen eingeschrieben, sowie seit zwei Jahren angestellte wissenschaftliche Hilfskräfte in der Pränataldiagnostik an der

Universitätsfrauenklinik. Prof. Dr. Karl Oliver Kagan wies beide Doktoranden detailliert ein in die ordnungsgemäße Vorgehensweise der Vermessungsdurchführung. Zudem begleitete er die ersten zehn Vermessungen unter direkter Supervision und kontrollierte weitere Vermessungen stichprobenartig.

2.3 Statistische Auswertung

Für die statistische Auswertung werden die Rohdaten der zwei unabhängigen Stichproben anhand der deskriptiven Statistik miteinander verglichen und dabei standardmäßig folgende statistische Kenngrößen für die untersuchten Variablen angegeben: Mittelwert, Median, Modalwert, oberes und unteres 95% - Konfidenzintervall (Upper/lower 95%-KI), Standardabweichung (STABW), Standardfehler des Mittelwertes (Standarderror of the Mean) als Maß für die Präzision des Mittelwertes für die Grundgesamtheit, oberes und unteres Limit des Mittelwertes (oberes/unteres Limit = Mittelwert \pm 1,96*Standardfehler des Mittelwertes), Variationskoeffizient als Maß für die Streuung der Variablen, Schiefe und Excess Kurtosis. Es folgt die Prüfung der Reproduzierbarkeit der Daten anhand von Konkordanzanalysen. Hierfür werden die verschiedenen Messungen von Untersucher F und Untersucher P unter- und gegeneinander auf systematische Trends untersucht. Graphisch dient dazu ein Bland – Altman – Plot und mathematisch eine Regressionsanalyse. Zur Darstellung eines statistischen Zusammenhangs zwischen Magenumfang und Abdomenumfang wird eine Regressionsanalyse durchgeführt. Um auf eine zentrale Tendenz zu testen, wurden die Residuen der Kontroll- und Testgruppe zur Regressionsgeraden der Kontrollgruppe in z-Werte transformiert. Diese werden unter Annahme einer Normalverteilung anhand von T-Tests auf signifikante Unterschiede untersucht. Hierfür wird ein Signifikanzniveau von 0,05 angesetzt. Zur Quantifizierung der Fälle, die unter der 5. Perzentile liegen, wird die 5. Perzentile in die Regressionsanalyse integriert. Die Ergebnisse werden schließlich als Sensitivität und Spezifität angegeben. Um die Qualität der Regressionsgeraden zu bewerten, werden folgende statistische Größen

angegeben: Steigung der Regressionsgeraden, Signifikanztest der Steigung mittels t-Test, Schnittpunkt der y-Achse, Bestimmtheitsmaß, Signifikanztest des Bestimmtheitsmaßes anhand einer F-Statistik, Standardfehler der Schätzung (Standard Error of Estimate, SEE) und der Quotient aus SEE und Mittelwert der Residuen, ausgedrückt in Prozent.

3 Ergebnisse

3.1 Deskriptive Statistik

Die statistische Auswertung beinhaltet zwei unabhängige Stichproben. Das gesunde Patientenkollektiv bzw. die Kontrollgruppe umfasst 374 Patientinnen und die Testgruppe 59 Patientinnen. Die Testgruppe kann weiterhin aufgeteilt werden in 38 (64,4%) Fälle mit ÖA und TÖF und 21 (35,6%) Fälle mit ÖA ohne Fistel.

3.1.1 Messzeitpunkt gemessen in Schwangerschaftswochen

Das mediane Gestationsalter (Tabelle 1) der Kontrollgruppe, gemessen in Schwangerschaftswochen (SSW), liegt bei 22,4 SSW (Bereich 18,4 – 39,1 SSW). 64,7% der Patientinnen befanden sich bei der Ultraschalluntersuchung vor und 35,3% nach der 27. SSW (Abbildung 5). Wie in Abbildung 5 zu sehen ist, erfolgt eine Vielzahl der Messungen zwischen der 21. und der 23. SSW. Die Testgruppe insgesamt zeigt ein durchschnittliches Gestationsalter von 28,8 SSW mit einem Median von 28,7 SSW (Bereich: 19,4 – 37,9 SSW). Wie aus Abbildung 4 und Abbildung 5 zu entnehmen, ist die Messzeitpunktverteilung deutlich „gleichmäßiger“ bei der Test- als bei der Kontrollgruppe. Der Mittelwert des Gestationsalters ist bei der Testgruppe mit 28,8 SSW fast vier Wochen später als der der Kontrollgruppe (25 SSW, Tabelle 1).

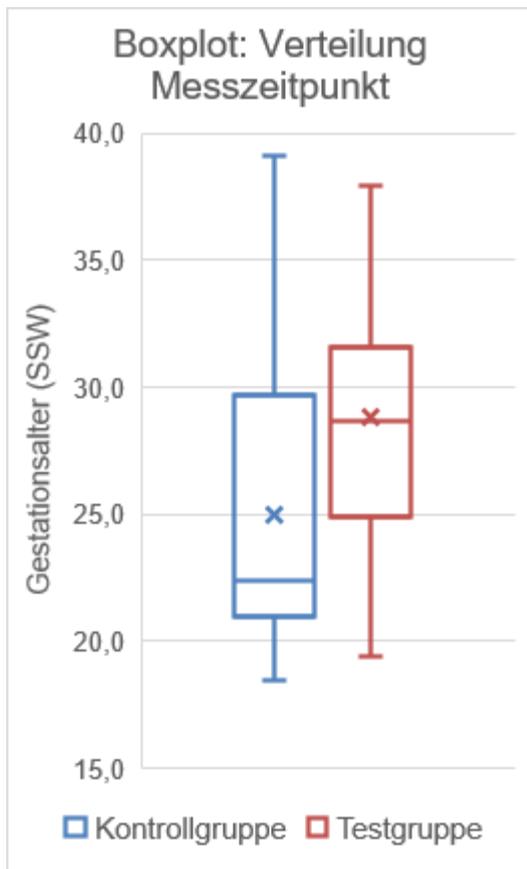


Abbildung 4 Boxplot: Verteilung Messzeitpunkte: Die Kontrollgruppe ist in blau und die Testgruppe in Rot dargestellt. Das Kreuz stellt den Mittelwert dar und der horizontale Balken im Kasten den Median. Die Box umfasst jeweils das 25%- bis 75%-Quantil. Zudem sind mit den Endpunkten das Minimum und das Maximum für beide Stichproben markiert.

Auffällig ist, dass Mittelwert und Median der Kontrollgruppe fast vier Wochen auseinanderliegen, weil sich in der Kontrollgruppe die Messungen im zweiten Trimenon stark häufen. Im Vergleich dazu sind die Messpunkte der Testgruppe deutlich „gleichmäßiger“ verteilt. Hier liegen Mittelwert und Median sehr nah beieinander.

Im Vergleich zeigt sich eine ähnliche Messzeitpunktverteilung zwischen TÖF und ÖA ohne Fistel (Tabelle 1).

Stat. Kenngrößen	Gestationsalter (SSW)			
	Kontrollgruppe	Testgruppe	TÖF	ÖA ohne Fistel
Mittelwert	25,0	28,8	28,5	29,5
Median	22,4	28,7	28,4	29,7
Minimum	18,4	19,4	19,4	20,6
Maximum	39,1	37,9	37,7	37,9
Standardabweichung (STABW)	5,1	4,8	4,5	5,4
Schiefe	0,9	0,1	0,2	-0,1
Excess-Kurtosis	-0,5	-0,6	-0,2	-0,9
Anzahl	374	59	38	21

Tabelle 1 Deskriptive Statistik: Gestationsalter Es werden statistische Kenngrößen der verschiedenen Gruppen gezeigt. ÖA = Ösophagusatresie, TÖF = Tracheoösophageale Fistel

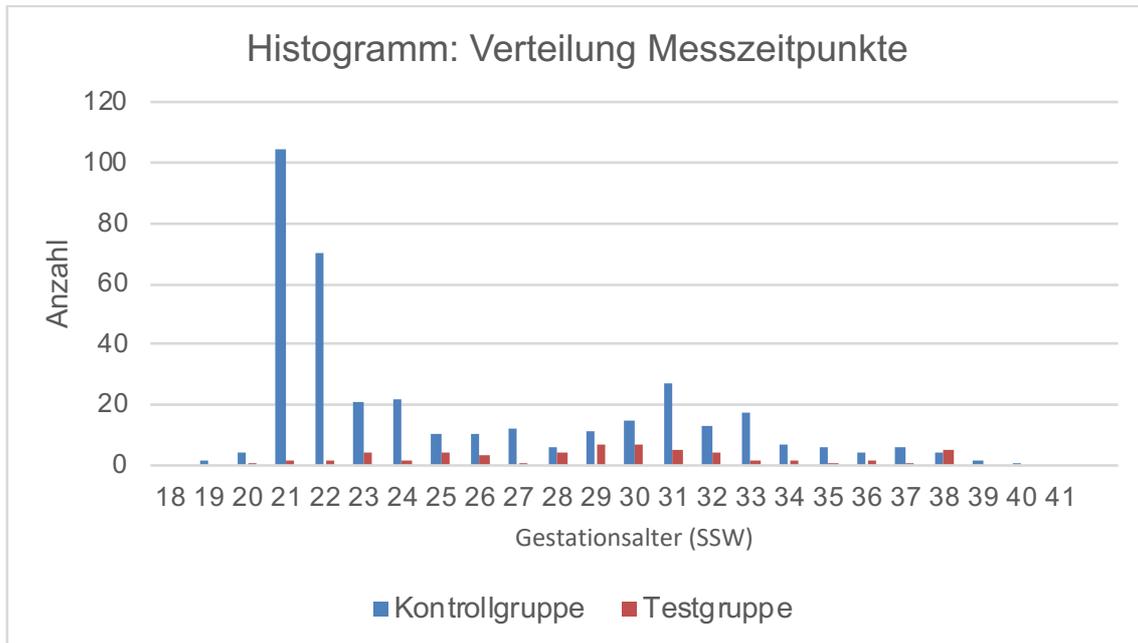


Abbildung 5 Histogramm: Verteilung Messzeitpunkte: Es wird die Verteilung der Messzeitpunkte auf der Zeitachse in Schwangerschaftswochen angezeigt. Die blauen Balken stellen die Kontrollgruppe dar, die roten Balken bilden die Testgruppe ab. Relativ viele Messungen der Kontrollgruppe wurden vor der 27. SSW durchgeführt.

3.1.2 Testgröße Magenumfang

Als Testgröße wurde der Magenumfang in Millimetern (mm) gemessen. Tabelle 2 zeigt eine Zusammenfassung der statistischen Kenngrößen für den Magenumfang in der Kontrollgruppe, der Testgruppe sowie die Testgruppe aufgeteilt in TÖF und ÖA ohne Fistel.

Magenumfang (MU in mm)				
Stat. Kenngrößen	Kontrollgruppe	Testgruppe	TÖF	ÖA ohne Fistel
COR(GA ~ MU)	0,6693	0,074	0,259	0,032
COR(AU ~ MU)	0,6857	0,177	0,264	0,205
Mittelwert	43,045	22,059	33,754	0,898
Median	40,225	18,300	32,650	0,100
Minimum	20,350	0,1	0,1	0,1
Maximum	99,650	140,700	140,700	16,850
STABW	13,738	24,178	22,683	3,655
Standardfehler des Mittelwertes	0,710	3,148	3,680	0,798
Oberes Limit	44,437	28,229	40,966	2,461
Unteres Limit	41,652	15,890	26,542	-0,666
Variationskoeffizient	0,319	1,096	0,672	0,246
Schiefe	1,161	2,128	2,857	4,583
Anzahl	374	59	38	21

Tabelle 2 Statistische Kenngrößen zu Magenumfang in den Studienkollektiven: Die ersten beiden Zeilen geben die Korrelation des Magenumfangs (MU) einmal zum Gestationsalter (GA) und einmal zum Abdomenumfang (AU) an. Mittelwert, Median, Minimum, Maximum und Standardabweichung (STABW) sind Excelberechnungen und entsprechen den üblichen Formeln für diese Lagemaße zur Beschreibung von diskreten Verteilungen (Fahrmeir et al., 2011a). Der Standardfehler des Mittelwertes berechnet sich als Mittelwert über die Quadratwurzel der Stichprobenlänge und gilt als (Standard-)Maß dafür, wie „nah“ der (unbekannte) Mittelwert der Grundgesamtheit (Population) am Mittelwert der vorliegenden Stichprobe liegt. Der Standardfehler des Mittelwertes liegt für die Kontrollgruppe mit 0,71 mm deutlich unter der Testgruppe. Daraus folgt beispielsweise für die Kontrollgruppe, dass mit einer Konfidenz von 95% der Mittelwert der Population zwischen dem oberen (44,437 mm) und unteren Limit liegt (41,652 mm). Der Variationskoeffizient ist ein standardisiertes Maß für die Streuung (STABW über Mittelwert). Hiernach streuen die Werte der Kontrollgruppe vergleichsweise deutlich weniger als die verschiedenen Ausprägungen der Testgruppe. Die positiven Werte zur Schiefe (Excelberechnung) zeigen, dass alle hier untersuchten Stichprobenverteilungen rechtsschief sind.

Die ersten beiden Zeilen stellen die Korrelation des Magenumfangs dar, einmal zum Gestationsalter und einmal zum Abdomenumfang. Der Magenumfang der Kontrollgruppe zeigt eine signifikante, positive Korrelation zum Gestationsalter ($r = 0,669$, $p < 0,0001$) sowie zum Abdomenumfang ($r = 0,686$, $p < 0,0001$). Es bestehen ebenfalls signifikante Korrelationen des Magenumfangs zum

Kopfumfang ($r = 0,659$, $p < 0,0001$) und zur Femurlänge ($r = 0,671$, $p < 0,001$) (Zahlen in Tabelle nicht enthalten). Die stärkste, signifikante Korrelation liegt demnach zwischen Magenumfang und Abdomenumfang vor.

Im Vergleich dazu besteht eine deutlich niedrigere Korrelation der gesamten Testgruppe, als auch aufgeteilt in TÖF und ÖA ohne Fistel, zum Gestationsalter und Abdomenumfang. Die üblichen Verlaufsparemeter in der Schwangerschaft, wie Kopfumfang und Femurlänge zeigen Korrelationswerte zum Abdomenumfang und Gestationsalter von deutlich über 0,9.

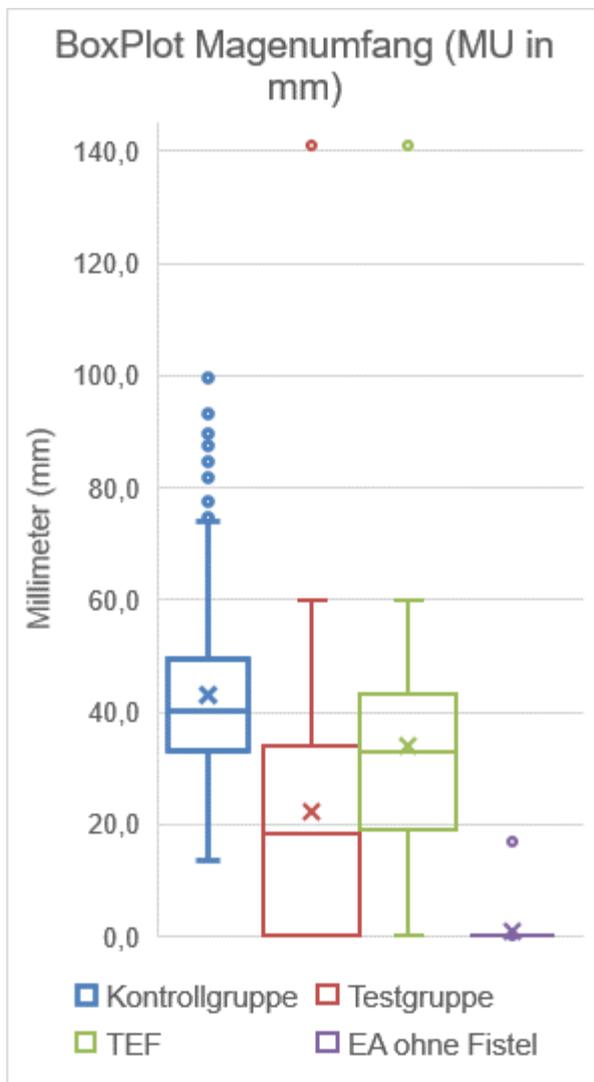


Abbildung 6 *Boxplot des Magenumfangs in den Stichproben der Studie: Es werden die Kontrollgruppe und die Testgruppe dargestellt. Zusätzlich wird die Testgruppe aufgeteilt in TEF und EA ohne Fistel.*

Der Mittelwert und Median der Kontrollgruppe (43,045 mm, 40,225 mm) sind ca. doppelt so hoch wie in der Testgruppe (Tabelle 2, Abbildung 6). Bei der aufgeteilten Testgruppe zeigt sich, dass die Gruppe ÖA ohne Fistel (Mittelwert = 0,898 mm, Median = 0,1 mm) sich deutlich von der Kontrollgruppe unterscheidet (Abbildung 6), da in dieser Untergruppe kein Magen im Ultraschall feststellbar ist und als Erkennungsgröße der Wert von 0,01 mm jeweils eingetragen wurde. Die Gruppe der TÖF liegt im Mittelwert (33,754 mm) und Median (32,65 mm) näher an der Kontrollgruppe (Tabelle 2), jedoch immer noch 8 mm (Median) bzw ca. 10 mm (Mittelwert) unter der Kontrollgruppe (Tabelle 2, Abbildung 6). Die Datenverteilung der Kontrollgruppe und der TÖF zeigt eine Überlappung, da der Mittelwert der TÖF weniger als eine Standardabweichung vom Mittelwert der Kontrollgruppe entfernt ist. Die Standardabweichung (STABW) sowie insbesondere der Variationskoeffizient als Maße der Streuung verdeutlichen, dass die Magenumfangsmessungen in der Testgruppe, besonders bei den TÖF, stärker streuen als in der Kontrollgruppe (Tabelle 2, Abbildung 6).

Bereits Mittelwert und Median lassen in allen Gruppen eine Rechtsschiefe vermuten, die durch die jeweils errechnete Schiefe bestätigt wird (Abbildung 7). Auffällig ist das hohe Maximum in der Gruppe der TÖF, welches mehr als das Vierfache über seinem Mittelwert liegt und mehr als 40 mm oberhalb des Maximums der Kontrollgruppe (Tabelle 2, Abbildung 6). Diese Auffälligkeit wird im Punkt 4.8.4 diskutiert.

Zusammenfassend zeigt sich eine signifikante Korrelation des Magenumfangs mit dem Abdomenumfang in der Kontrollgruppe. Durchschnittlich ist der Magenumfang der Testgruppe kleiner als in der Kontrollgruppe, auch bei den TÖF. Auffällig ist ebenfalls die Rechtsschiefe des Magenumfangs, die möglicherweise auch die Auswirkung des Magenwachstums über die Schwangerschaft in der Kontrollgruppe reflektiert.

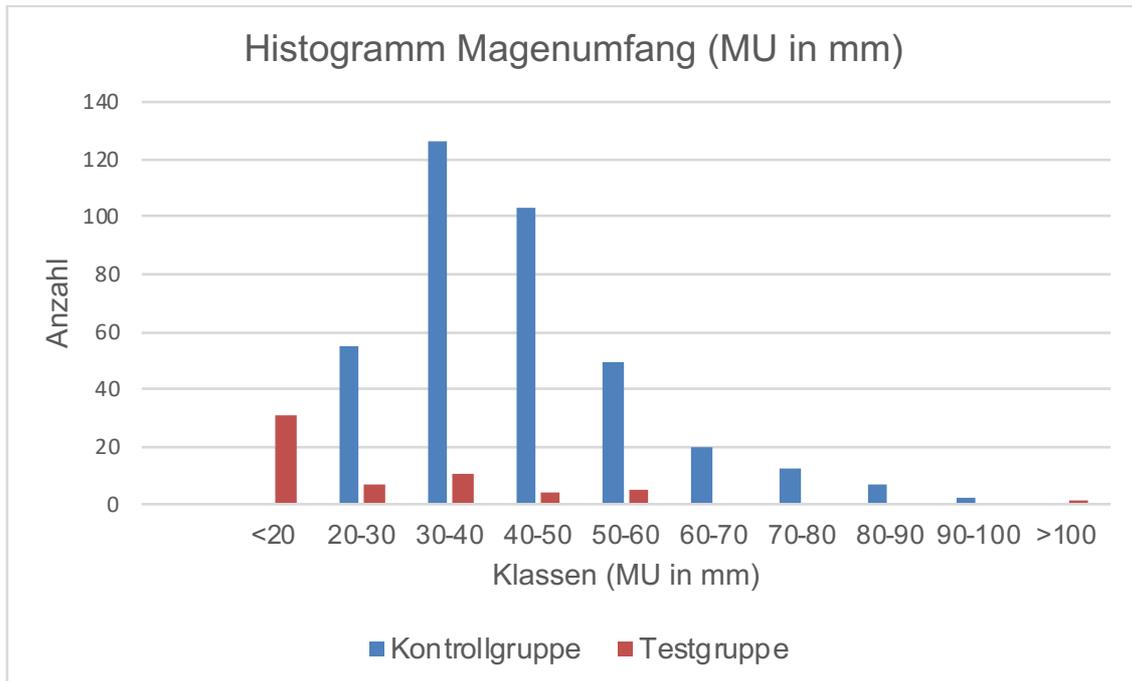


Abbildung 7 Histogramm des Magenumfangs in der Kontrollgruppe und Testgruppe: Die blauen Balken kennzeichnen die Kontrollgruppe, während die roten Balken die Testgruppe darstellen. Am häufigsten wurden in der Kontrollgruppe Magenumfänge zwischen 30 und 50 mm gemessen. Die Mehrheit des Testkollektivs zeigt Magenumfangsmessungen von kleiner als 20 mm.

3.1.3 Rohdaten der Testgröße Magenumfang in Abhängigkeit vom Abdomenumfang

Die gemessenen Rohdaten werden in Abbildung 8 dargestellt: der Magenumfang auf der Y-Achse und der Abdomenumfang auf der X-Achse. Die blauen Punkte geben die Kontrollgruppe wieder, die roten Dreiecke die der TÖF aus der Testgruppe und die schwarzen Kreuze die der ÖA ohne Fistel. Die oben beschriebene Konzentration der Messungen vor der 27. SSW in der Kontrollgruppe ist hier erneut in der blauen Punktwolke sichtbar. Die Kontrollgruppe weist eine auffällige Konzentration der Messungen bei niedrigem Abdomenumfang auf, ähnlich der Verteilung der Messzeitpunkte (Abbildung 5). Weiterhin zeigt sich tendenziell ein trichterförmiges Streuungsbild, sodass augenscheinlich die Streuung der Daten in dieser Stichprobe bei größer werdendem Abdomenumfang leicht zunimmt. Diese Tatsache und die o.g.

Rechtsschiefe der Kontrollgruppe würden zum erwarteten Wachstumseffekt passen.

Die Verteilung der TÖF weist in der Punktwolke (rote Dreiecke mit roter Trendlinie) zu dem korrespondierenden Abdomenumfang trendmäßig einen kleineren Magenumfang auf als die Kontrollgruppe (blaue Punkte mit blauer Trendlinie). Der um circa 10 mm kleinere Mittelwert der TÖF im Vergleich zur Kontrollgruppe in der univariaten Betrachtung (Tabelle 2, Mittelwert, Median) wird also hier durch die bivariate Analyse von Magenumfang in Abhängigkeit vom Abdomenumfang bestätigt (siehe Trendlinien in Abbildung 8). Weiterhin wird in diesen Stichproben an den Steigungen der Trendlinien optisch sowie mathematisch deutlich, dass die Magengröße in der Kontrollgruppe anders auf den Abdomenumfang reagiert als die der Subgruppe der TÖF. Wenn der Abdomenumfang um 10 mm steigt, nimmt der Magenumfang trendmäßig in der Kontrollgruppe um 1,785 mm und in der Subgruppe der TÖF nur um 1,1118 mm zu, also um 37,7% weniger. Im Ergebnis liegen also die Mittelwerte der Trends und die Steigungen der Trendlinien für die Subgruppe der TÖF unter denen der Kontrollgruppe. Dieses Ergebnis ergibt sich trotz des Ausreißers im Kollektiv der TÖF, der mit 140 mm deutlich oberhalb aller anderen Stichprobenmessungen liegt.

In der Gruppe der ÖA ohne Fistel ist in 20 von 21 Fällen kein Magen im Ultraschall darstellbar (Erkennungswert 0,01 mm). Lediglich in einem Fall wurde ein Messwert für den Magenumfang von 16,85 mm (zum Abdomenumfang von 270 mm) eingetragen.

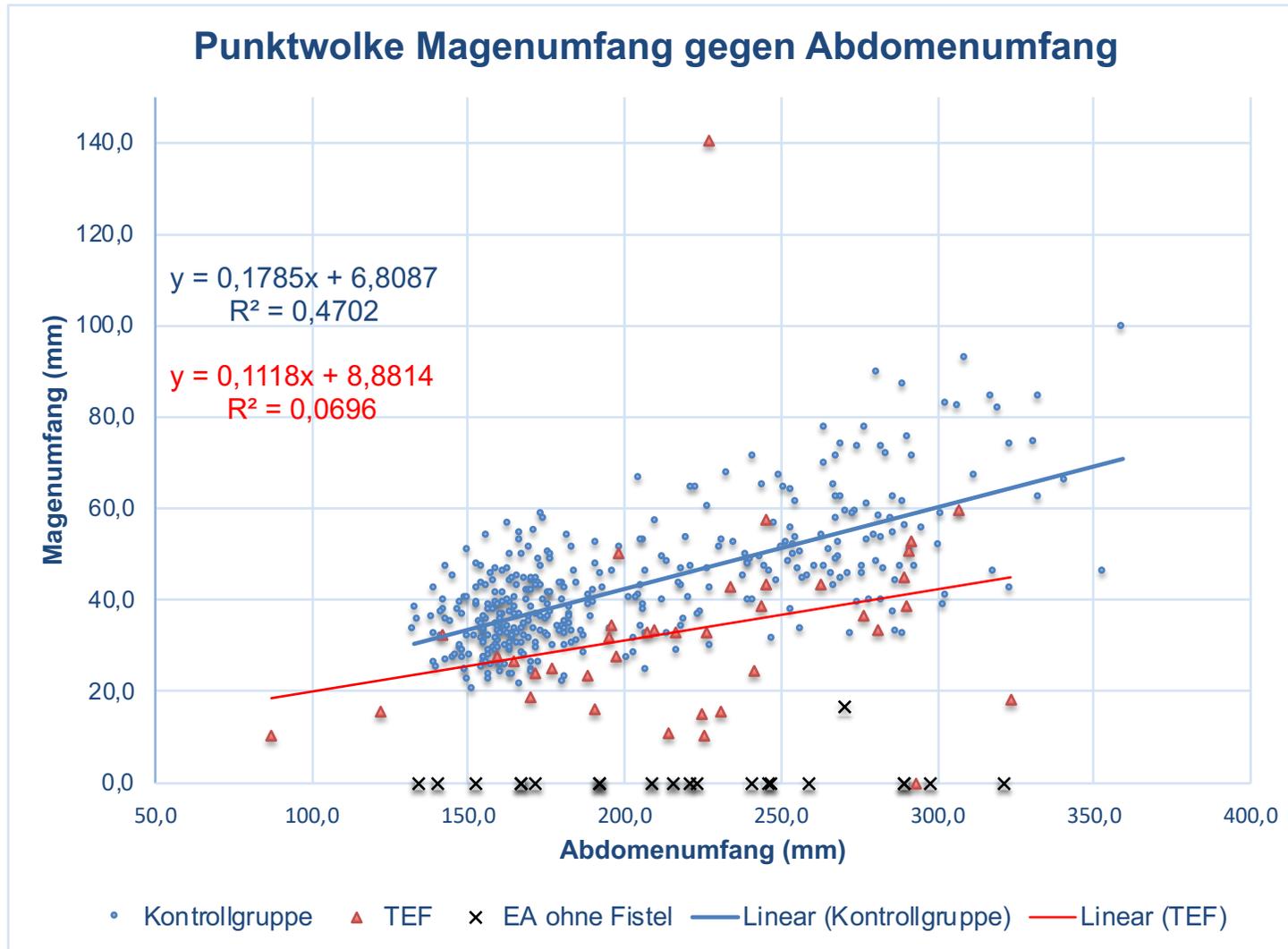


Abbildung 8 Punktwolke Magenumfang zu Abdomenumfang: Der Magenumfang auf der y-Achse wird mit dem dazugehörigen Abdomenumfang auf der x-Achse in Millimeter (mm) dargestellt. Die Kontrollgruppe wird als blaue Punktwolke, die Gruppe der TEF als rote Dreiecke und die Gruppe der EA ohne Fistel als schwarze Kreuze angegeben. Zusätzlich werden Trendlinien zur Verteilung der Kontroll- und TEF-Gruppe angegeben. Die Formel zur jeweiligen Trendlinie ist in der Farbe der Trendlinie angegeben (blau = Kontrollgruppe, rot = TEF-Gruppe).

3.1.4 Gewichtsverteilung der Studienteilnehmerinnen

Gewicht (kg) vor Schwangerschaft

Stat. Kenngrößen	Kontrollgruppe	Testgruppe
Mittelwert	66,1	65,5
Median	62	64
Minimum	38	35
Maximum	163	106
STABW	15,5	12,5
Anzahl	367	59

Tabelle 3 Statistische Kenngrößen zum Gewicht vor der Schwangerschaft: Zwischen den beiden unabhängigen Stichproben ist kaum ein nennenswerter Unterschied zu erkennen. Lediglich das Maximalgewicht ist in der Kontrollgruppe deutlich größer.

Die Gewichtsangaben der Frauen beziehen sich auf die letzte Messung vor der Schwangerschaft. In der Kontrollgruppe konnte von 98,1% der Schwangeren das Gewicht erfasst werden, wohingegen in der Testgruppe das Gewicht bei 100%

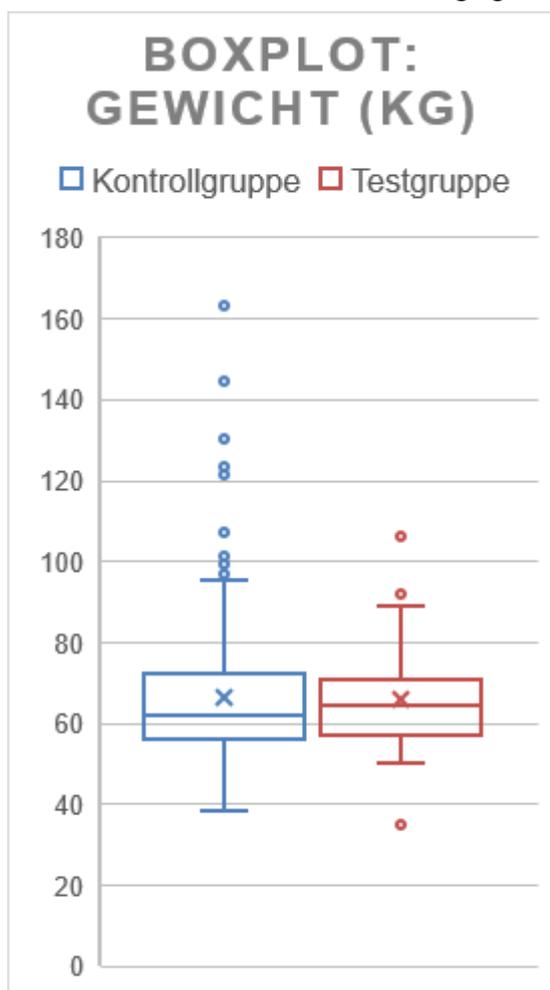


Abbildung 9 Boxplot zum Gewicht vor der Schwangerschaft: Die Gewichtsverteilung erscheint ähnlich. Allerdings gibt es in der Kontrollgruppe mehr Ausreißer nach oben.

der Probanden bekannt war. Die statistischen Kenngrößen ergeben ähnliche Werte zwischen den beiden Gruppen (Tabelle 3). Auffällig ist lediglich das deutlich höhere Maximalgewicht in der Kontrollgruppe. Der Boxplot (Abbildung 9) sowie das Histogramm (Abbildung 10) zeigen eine ähnliche Verteilung des Gewichts.

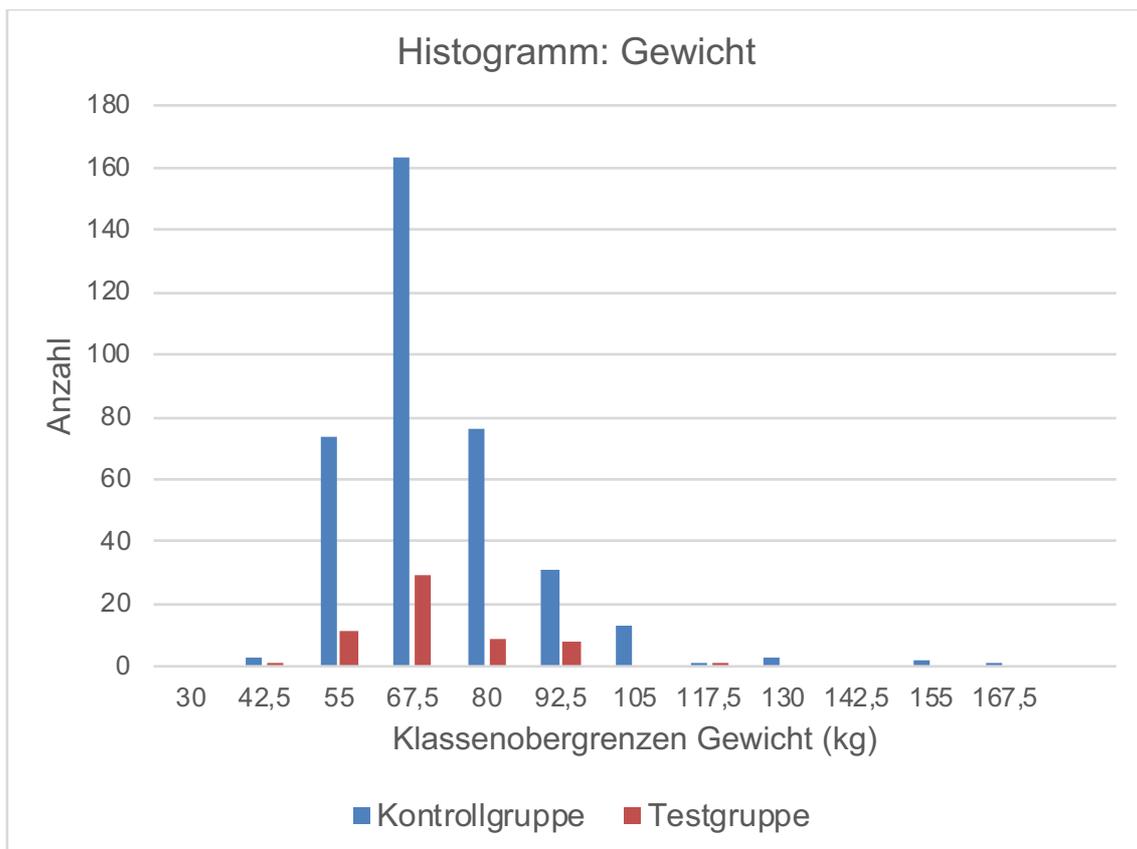


Abbildung 10 Histogramm zu Gewicht vor der Schwangerschaft: Die x-Achse gibt die Klassenobergrenzen für das Gewicht an und die y-Achse die dazugehörige Anzahl. Die blauen Balken stellen die Kontrollgruppe dar, die roten Balken zeigen die Testgruppe. Keine Patientin wiegt unter 30 kg und 3 Patientinnen der Kontrollgruppe weisen ein Gewicht zwischen 30 und 42,5 kg auf. Außerdem wird der Unterschied in der Kollektivgröße deutlich (374 Fälle in der Kontrollgruppe vs. 59 in der Testgruppe). Allerdings beinhalten beide Gruppen die Mehrzahl der Patientinnen in der Klasse bis 67,5 kg (Klassenbreite: 55 – 67,5 kg).

3.1.5 Auffälligkeiten in der Testgruppe

Bezüglich des Testkollektivs wurde bei 20 (33,9%) von 59 Feten ein auffälliger Karyotyp registriert. In 18 (90%) Fällen davon liegt eine Trisomie vor und bei

zweien (10%) wurde der auffällige Karyotyp mit „sonstige“ beschrieben. Von den 20 auffälligen Karyotypen befinden sich 9 (23,7%) in der Gruppe der TÖF und 11 (52,4%) bei ÖA ohne Fistel. Eine Trisomie 18 liegt bei 15 (75%) von den 20 auffälligen Karyotypen vor. In der Gruppe der TÖF sind 7 (18,4%) Feten mit einer Trisomie 18 und in der Gruppe der ÖA ohne Fistel 8 (38,1%). Jeweils ein Fall von Trisomie 13 (ÖA ohne Fistel) und einer von Trisomie 21 kommen vor.

In der Testgruppe liegt bei 26 (44,1%) von 59 Fällen ein Polyhydramnion vor. Aufgeteilt finden sich 14 (36,8%) davon in der Gruppe der TÖF und 12 (57,1%) in der Gruppe ÖA ohne Fistel. Lediglich ein Fetus mit einem Polyhydramnion befindet sich vor der 24. SSW. In 8 Fällen trat eine Chromosomenstörung zusammen mit einem Polyhydramnion auf. Davon hatten 7 eine Trisomie 18 und einer eine Trisomie 21. Alle diese Fälle wurden nach der 24. SSW registriert. In der Subgruppe der TÖF zeigten zwei Feten ein Polyhydramnion und eine Chromosomenstörung (Trisomie 18 und 21) und bei den Feten mit ÖA ohne Fistel kam es in sechs Fällen zu einer solchen Konstellation (Alle Trisomie 18). Im Umkehrschluss kam es in der Mehrzahl der Fälle (60%, bzw. 12 von 20) mit Chromosomenstörungen nicht zur Ausbildung eines Polyhydramnions. Von denen hatten 75% (9 von 12) ihren Untersuchungszeitpunkt vor dem dritten Trimenon.

3.2 Konkordanzanalyse der gemessenen Daten

3.2.1 Reproduzierbarkeit der Kontrollgruppe: Untersucher F

Untersucher F hat die Magenumfänge der Feten dieser Stichprobe zweimal gemessen. Beide Messungen sind in einer Punktwolke (Abbildung 11) und in einem Bland-Altman-Plot (Abbildung 12) zusammengefasst, um zu untersuchen, inwiefern beide Messungen übereinstimmen. Auf Grund der relativ kleinen Abstände zur Winkelhalbierenden besteht augenscheinlich eine hohe Übereinstimmung beider Messungen.

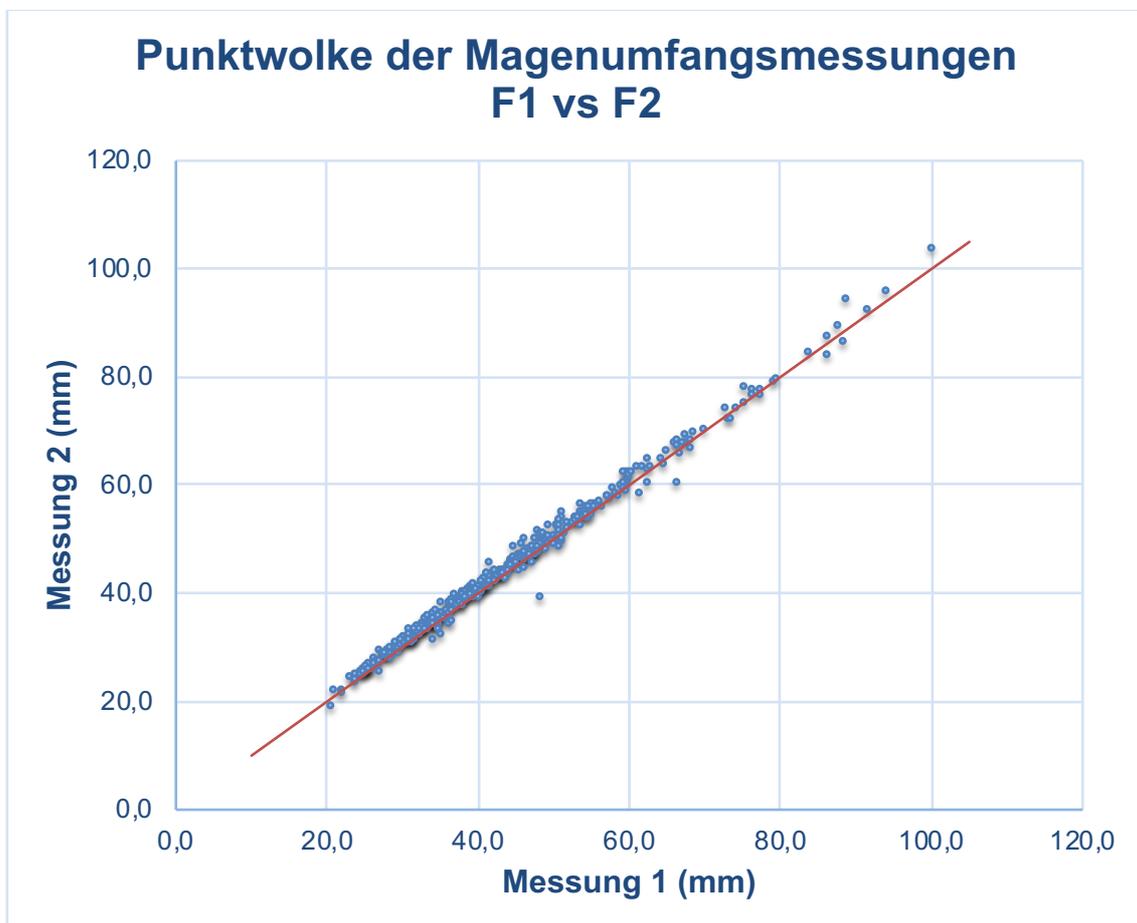


Abbildung 11 Punktwolke der Kontrollgruppe für Untersucher F: Messung 1 und 2 von Untersucher F werden als blaue Punktwolke im xy-Diagramm veranschaulicht. Die rote Gerade stellt eine Winkelhalbierende dar (45°-Linie). Identische Messungen befinden sich exakt auf der Geraden. Aus der vorliegenden Abbildung ergibt sich somit die Schlussfolgerung, dass offensichtlich beide Messungen eng beieinander liegen.

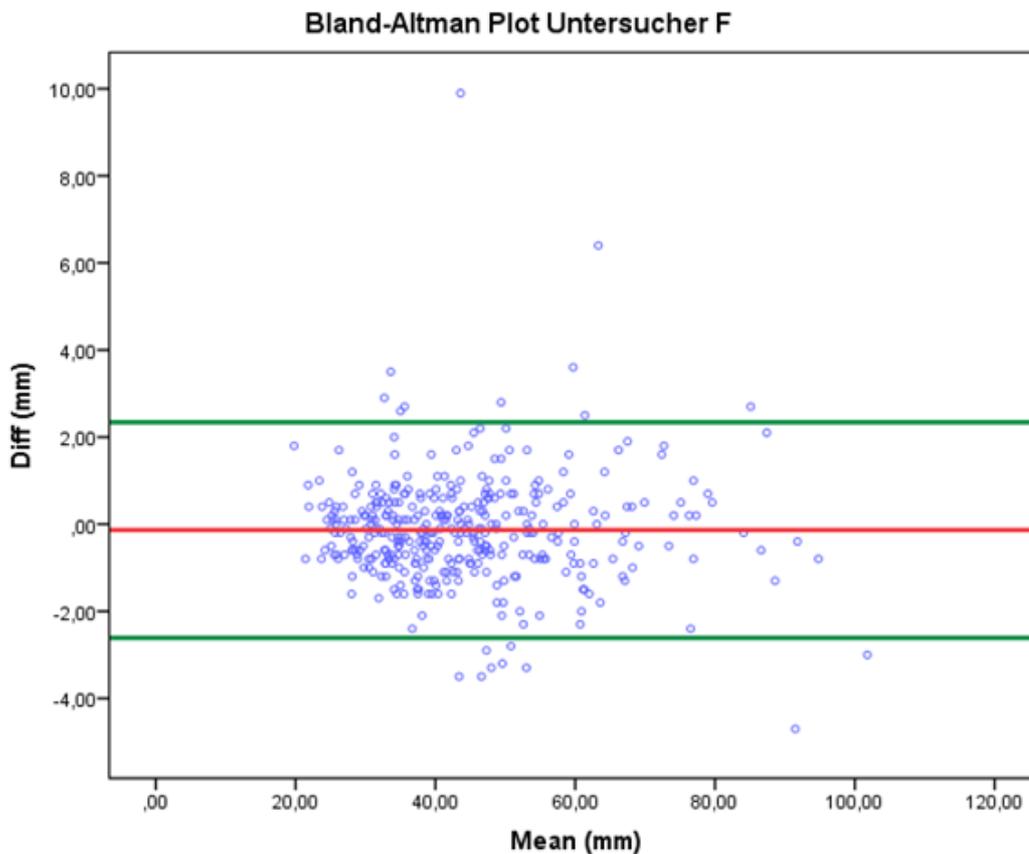


Abbildung 12 Bland-Altman-Plot der Kontrollgruppe Untersucher F: Auf der y-Achse wird die Differenz der beiden Messungen, die pro Fall durchgeführt wurden, ersichtlich (in Millimeter (mm) angegeben). Die x-Achse zeigt den Mittelwert aus den beiden Messungen (Mean) von Untersucher F. Der Mittelwert der Differenzen liegt nahezu bei 0 (rote Gerade). Unter der Annahme der Normalverteilung der Differenzen schwanken 95% (Punkte zwischen den beiden grünen Geraden) der Messungen um weniger als 4,95 mm. Das Diagramm zeigt keinerlei Trendmuster in der Punktwolke, die auf systematische Abweichungen hinweisen würden.

Im Bland-Altman-Diagramm (Abbildung 12) liegt der Mittelwert der Differenzen bei minus 0,133 mm (Standardabweichung = 1,26 mm) und ist als rote Horizontale eingezeichnet. Bei normalverteilten Differenzen ist davon auszugehen, dass die Differenzen von Untersucher F in 95% der Fälle um weniger als 4,95 mm (95% Übereinstimmungsbereich zwischen grünen Horizontalen) schwanken. Weiterhin sind im Diagramm sind keine Trendmuster erkennbar, das durch eine Regressionsanalyse zwischen den beiden Variablen bestätigt wird – Differenz der Messungen als abhängige Variable (y) und „Mean“

als unabhängige Variable (x). Die Regressionsanalyse zeigt, dass keine signifikante lineare Beziehung zwischen den beiden Variablen besteht (R -Quadrat = 0,001, p -Wert = 0,602) und keine signifikante Steigung der Regressionsgeraden vorliegt (Regressionskoeffizient (b) = -0,002, p -Wert = 0,602).

3.2.2 Reproduzierbarkeit der Kontrollgruppe Untersucher P

Ähnlich wie bei Untersucher F orientiert sich die Punktwolke der beiden Messungen von Untersucher P an der Winkelhalbierenden (Abbildung 13). Der Mittelwert der Differenzen ist minus 0,365 mm (Standardabweichung = 1,4 mm) und wird als rote Horizontale in Abbildung 14 dargestellt. Unter Annahme von normalverteilten Differenzen ist anzunehmen, dass die Differenzen in 95% (Punkte zwischen den grünen Horizontalen) der Fälle um weniger als 5,49 mm schwanken. Trendmuster als Zeichen einer systematischen Abweichung sind in Abbildung 14 nicht erkennbar. Die Regressionsanalyse bestätigt, dass keine signifikante Korrelation zwischen der Differenz und dem Mittelwert besteht (R -Quadrat < 0,001, p -Wert = 0,730) und auch kein signifikanter Regressionskoeffizient vorliegt (Regressionskoeffizient (b) = -0,002, p -Wert = 0,730).

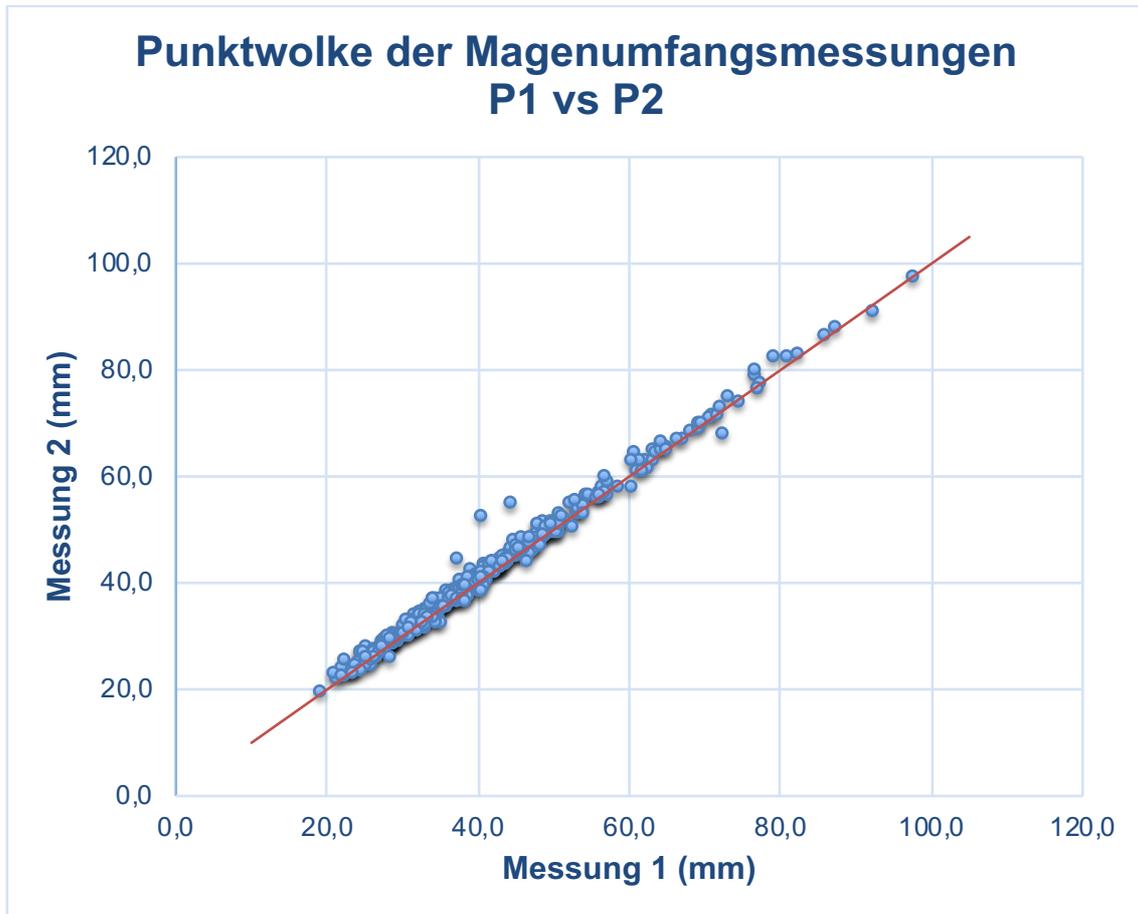


Abbildung 13 Punktwolke der Kontrollgruppe **Untersucher P**: Messung 1 und 2 von **Untersucher P** werden als blaue Punktwolke im xy-Diagramm dargestellt. Die rote Gerade stellt eine Winkelhalbierende dar (45°-Linie). Falls die Messungen identisch sind, liegen diese exakt darauf. Offensichtlich liegen beide Messungen eng beieinander

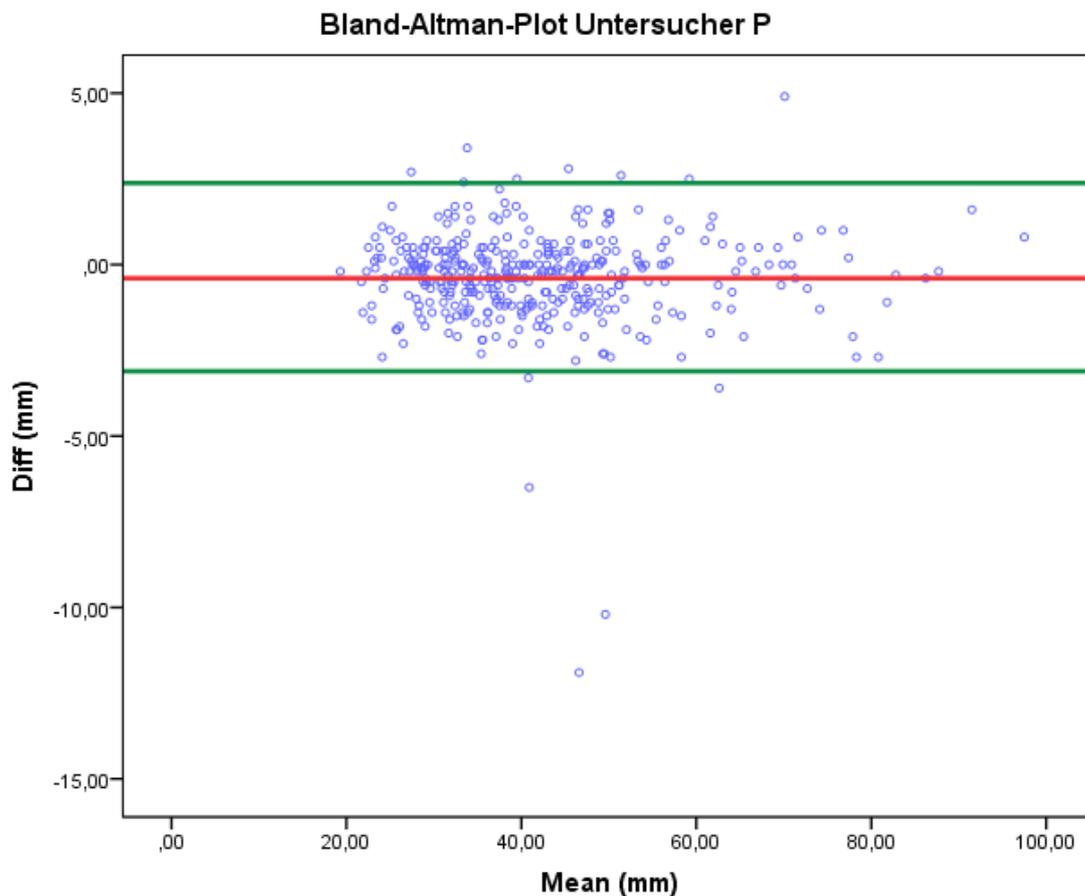


Abbildung 14 Bland-Altman-Plot der Kontrollgruppe Untersucher P: Auf der y-Achse ist die Differenz der beiden Messungen, die pro Fall durchgeführt wurden, in Millimeter (mm) angegeben. Die x-Achse gibt den Mittelwert aus den beiden Messungen (Mean) von Untersucher P an. Der Mittelwert der Differenzen liegt nahezu bei 0 (rote Gerade). Unter der Annahme von normalverteilten Differenzen schwanken 95% (Punkte zwischen den beiden grünen Geraden) der Messungen um weniger als 5,49 mm. Das Diagramm zeigt keinerlei Trendmuster in der Punktwolke, die auf systematische Abweichungen hinweisen würden.

3.2.3 Reproduzierbarkeit der Kontrollgruppe

Im folgenden Teil der Arbeit werden die durchschnittlichen Werte der Magenumfangsmessung (Mittelwert von Messung 1 und 2) von Untersucher F mit denen von Untersucher P verglichen. Bereits die Streuung der Punktwolke (Abbildung 15, interpersonell) um die Winkelhalbierende erscheint größer als bei den einzelnen Untersuchern selbst (Abbildung 11 und Abbildung 13, intrapersonell). Da die meisten Punkte augenscheinlich unter der Winkelhalbierenden liegen, hat Untersucher F den Magenumfang

durchschnittlich größer gemessen als Untersucher P, nämlich um 2,3 mm (Mittelwert Untersucher F = 44,2 mm; Mittelwert Untersucher P = 41,9 mm) und F (x-Achse) liegt damit um diesen Betrag rechts der Winkelhalbierenden (Abbildung 16). Die Standardabweichung der Mittelwertdifferenzen zwischen Untersucher F und P beträgt 3,41 mm. Unter Annahme der Normalverteilung der Differenzen liegen demnach 95% der Differenzen in einem Übereinstimmungsbereich von 13,35 mm (Abbildung 16). In der Regressionsanalyse hierzu (Differenz vom Mittelwert von Untersucher P und F pro Fetus als abhängige Variable und Mittelwert aller Messungen von P und F pro Fetus als unabhängige Variable) wurde eine signifikante Korrelation (R-Quadrat = 0,068, P-Wert < 0,001) zwischen den Differenzen und des Mittelwertes festgestellt, mit einer eher schwachen, aber signifikanten Steigung (Regressionskoeffizient = 0,065, P-Wert < 0,001). Diese Auffälligkeit zeigt geringe, aber statistisch messbare unterschiedliche Messergebnisse zwischen den Untersuchern.

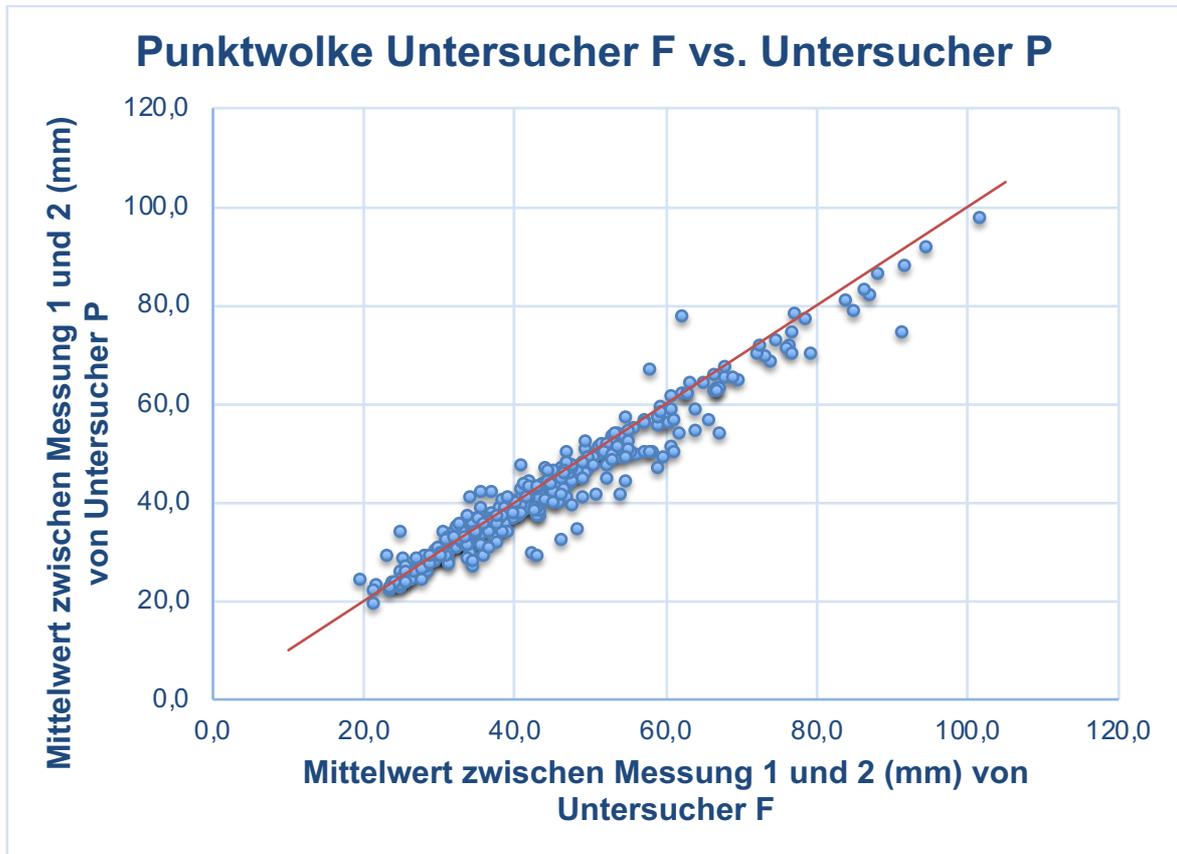


Abbildung 15 Punktwolke Kontrollgruppe Untersucher F vs. Untersucher P: Es werden die Mittelwerte der einzelnen Messungen von Untersucher F (x-Achse) mit denen von Untersucher P (y-Achse) dargestellt. Die rote Gerade stellt eine winkelhalbierende Gerade dar. Falls die Werte von Untersucher F mit denen von Untersucher P exakt übereinstimmen, würden die Punkte genau auf der Geraden liegen. Es scheint, als wären die Messungen eng um die winkelhalbierende Gerade gestreut, sodass ein geringer Unterschied vermutet werden kann.

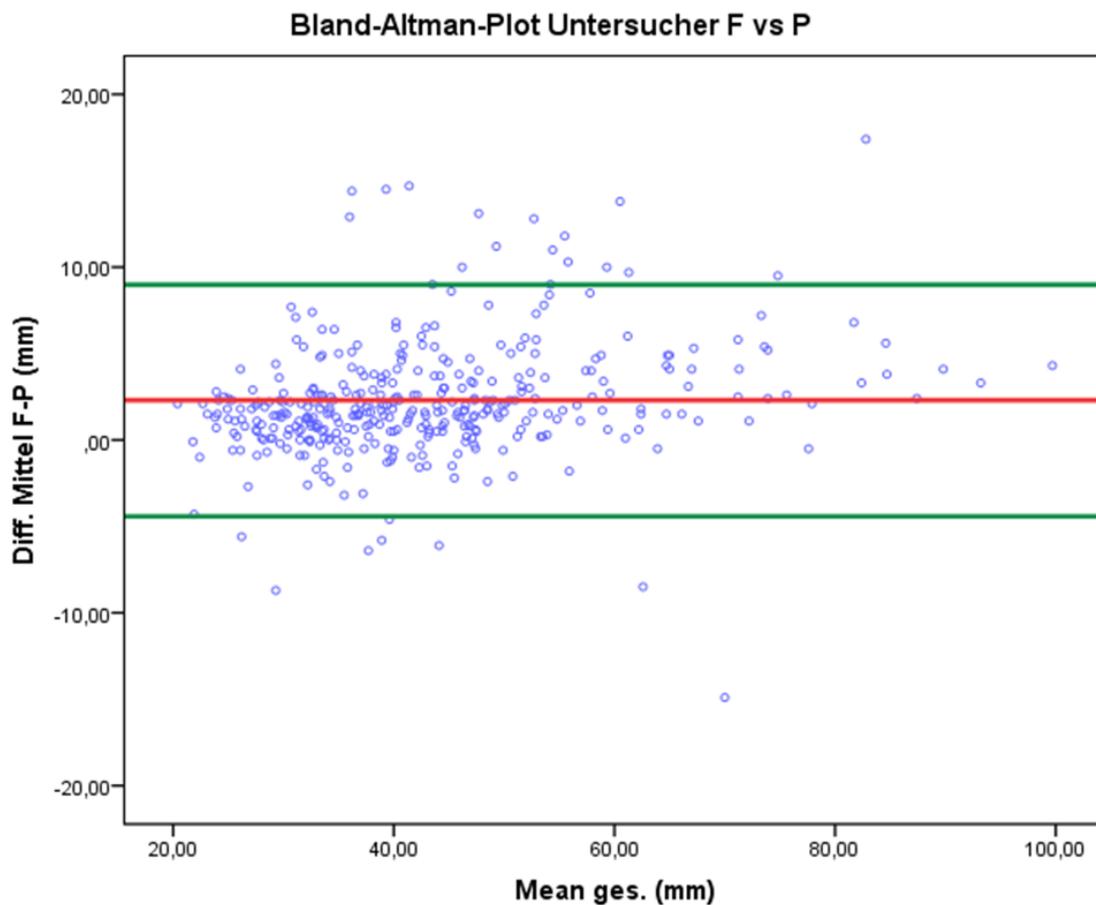


Abbildung 16 Bland-Altman-Plot der Kontrollgruppe zum Vergleich von Untersucher F zu Untersucher P: Auf der y-Achse ist die Differenz der Mittelwerte der einzelnen Fälle in Millimeter (mm) dargestellt. Die x-Achse gibt den Mittelwert aus den Messungen von Untersucher F und Untersucher P der einzelnen Fälle an. Der Mittelwert der Differenzen liegt nahezu bei 0 (rote Gerade). Unter Annahme der Normalverteilung schwanken 95% (Punkte zwischen den beiden grünen Geraden) der Messungen um 13,35 mm. Das Diagramm zeigt keine systematischen Trends in der Punktwolke, die auf systematische Abweichungen hinweisen würden.

3.3 Regressionsanalyse zur Identifikation einer Risikogruppe

3.3.1 Bivariate lineare Regressionsanalyse der Kontrollgruppe

Auf Grund der guten Korrelation zwischen der Zielgröße Magenumfang und dem Abdomenumfang ($r = 0,686$, $p < 0,0001$) wird deren Zusammenhang weiter in einer bivariaten linearen Regression untersucht. Die bivariate Regression hat das Ziel, „kleine“ Magenumfänge von Feten bezogen auf deren Abdomenumfang

objektiv zu identifizieren, um ein Kriterium für eine Risikogruppe zu definieren. Regressiert wird der Magenumfang (abhängige Variable) gegen den Abdomenumfang (unabhängige Variable) der Kontrollgruppe (Abbildung 17).

Die Regressionsgerade in Abbildung 17 hat die Formel $MU = 0,1785 \cdot AU + 6,8087$. Hiermit lassen sich zu gegebenen Abdomenumfängen zu erwartende Magenumfänge schätzen. Laut dieser Funktion nimmt der geschätzte Magenumfang um 0,1785 mm (Regressionskoeffizient „b“, $p < 0,0001$) zu, wenn der Abdomenumfang um 1 mm steigt. Das signifikante Bestimmtheitsmaß beträgt 0,4702 ($p < 0,0001$): Somit lässt sich die Veränderung des Magenumfangs zu 47,02% mit der Veränderung des Abdomenumfangs erklären.

Um Aussagen zu Perzentilen und Signifikanzniveaus der durchgeführten statistischen Regression treffen zu können, müssen neben anderen Annahmen (siehe Seite 73) auch die Residuen (beobachteter Magenumfang minus Schätzwert Magenumfang) normalverteilt sein (Urban and Mayerl, 2011d).

Die Tabelle 4 und die Abbildung 18, Abbildung 19 und Abbildung 20 veranschaulichen Untersuchungen der Residuen der Kontrollgruppe auf Normalverteilung. Die statistischen Kenngrößen Mittelwert (0,00 mm), Median (-1,175 mm) und Schiefe (0,376) sowie das Histogramm (Abbildung 19) lassen eine Rechtsschiefe erkennen. Der Boxplot (Abbildung 18) zeigt zusätzlich mehrere Ausreißer nach oben, die für die Rechtsschiefe mitverantwortlich sind. Abschließend wird die Normalverteilung anhand eines Q-Q Plots (Abbildung 20) getestet. Die Verteilung der Residuen folgt im Wesentlichen einer geraden Linie im Q-Q Plot, welche typisch für eine Normalverteilung ist. Dementsprechend wird eine Normalverteilung der Residuen angenommen. Mit der angenommenen Normalverteilungseigenschaft der Residuen lassen sich die grünen Perzentilgeraden (für ein Konfidenzintervall von 95%) in Abbildung 17 errechnen. Beide sind Parallelen zur Regressionsgeraden der Kontrollgruppe mit einem konstanten Abstand von 16,448 mm nach oben und unten.

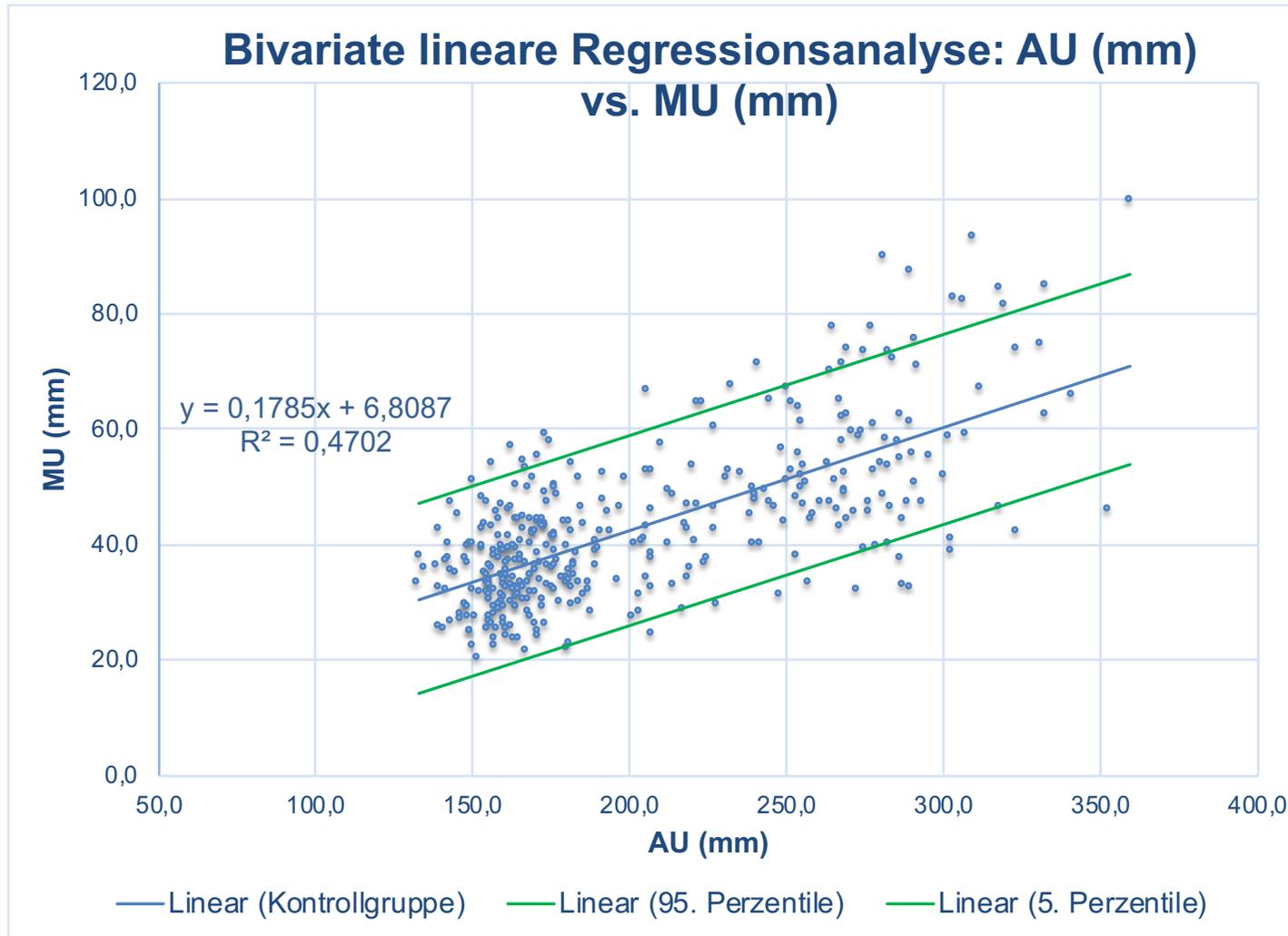


Abbildung 17 Bivariate lineare Regression: Magen- gegen Abdomenumfang der Kontrollgruppe: Die x-Achse gibt den Abdomenumfang (AU) in Millimetern (mm) an, die y-Achse den Magenumfang (MU) in Millimetern (mm). Die Rohdaten der Kontrollgruppe sind die blauen Punkte. Die Formel $MU = 0,1785 \cdot AU + 6,8087$ sowie das Bestimmtheitsmaß ($R^2 = 0,4702$) der Regression sind in der Abbildung dargestellt. Die 5. und 95. Perzentile (grüne Geraden) beziehen sich auf die Residuen (Normalverteilungsannahme) der Kontrollgruppe.

Residuen (mm)

Stat. Kenngrößen	Kontrollgruppe
Mittelwert	0,000
Median	-1,175
Minimum	-25,786
Maximum	32,824
Oberes 95% - KI	19,600
Unteres 95% - KI	-19,600
STABW	10,000
Standardfehler des Mittelwertes	0,517
Oberes Limit	1,013
Unteres Limit	-1,013
Schiefe	0,376
Standardfehler Schiefe	0,127
Grobregel Schiefe <1,96	2,969
Excess-Kurtosis	0,363
Standardfehler Excess Kurtosis	0,253
Grobregel Excess Kurtosis <1,96	1,432
Anzahl	374

Tabelle 4 Stat. Kenngrößen der Residuen der Kontrollgruppe: Mittelwert und Median haben nicht denselben Wert. Die Standardabweichung (STABW) beträgt 10 mm. Das Ausmaß der Schiefe und Excess Kurtosis werden anhand der Grobregeln untersucht. Die Rechtsschiefe liegt oberhalb des Grenzwertes von 1,96 (Urban and Mayerl, 2011b).

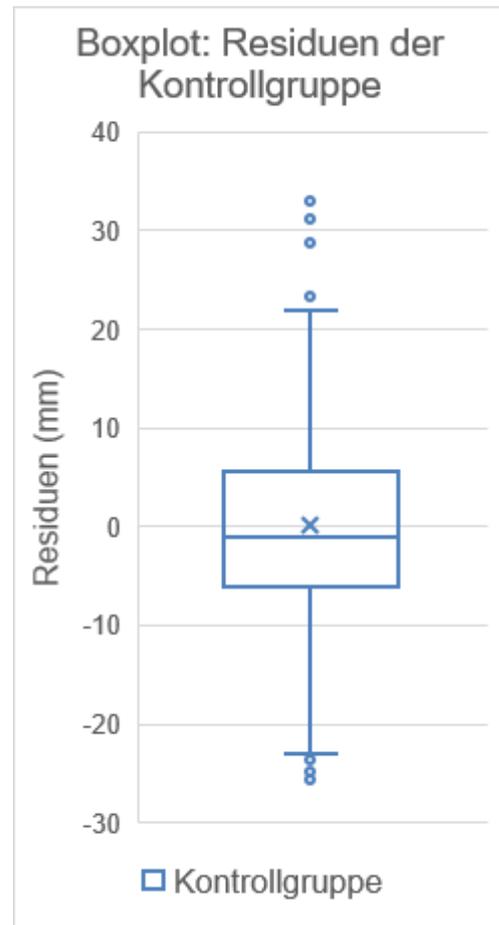


Abbildung 18 Boxplot Residuen der Kontrollgruppe: Das Kreuz stellt den Mittelwert der Residuen dar und der mittlere Balken den Median. Zu erkennen sind einige Extremwerte nach oben und drei nach unten.

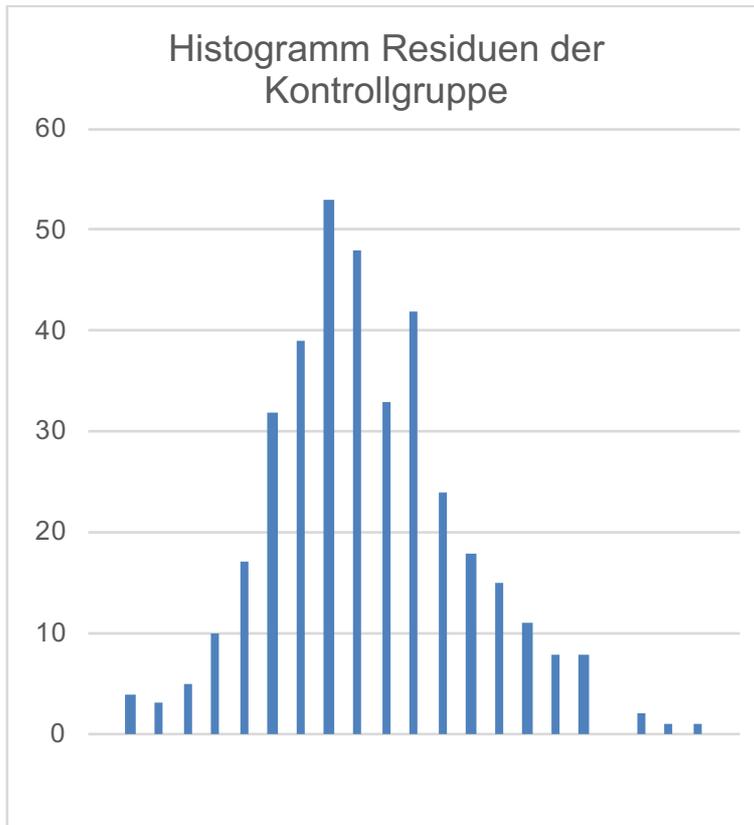
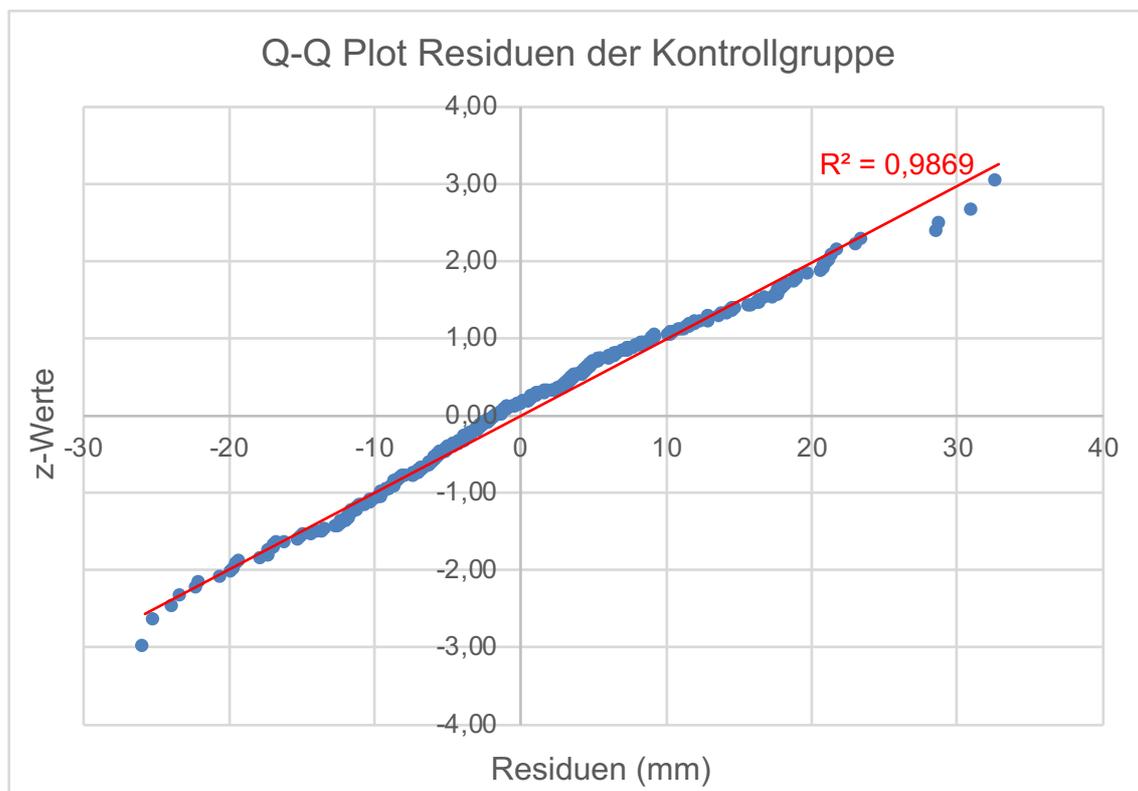


Abbildung 19 Histogramm Residuen der Kontrollgruppe: Es ist eine Rechtsschiefe in der Verteilung zu erkennen.

Abbildung 20 Q-Q Plot Residuen der Kontrollgruppe: Die Residuen auf der x-Achse werden gegen z-Werte der Normalverteilung (y-Achse) angezeigt (Fahrmeir et al., 2011b). Die Residuen liegen augenscheinlich nah an der Trendlinie mit einem Bestimmtheitsmaß von 0,99 und rechtfertigen somit die Normalverteilungsannahme.



3.3.2 Stichprobenvergleich mit standardisierten Residuen (z-Werte)

Für einen standardisierten Stichprobenvergleich mit den Residuen werden diese in z-Werte transformiert. Voraussetzung für eine sinnvolle Anwendung einer Z-Transformation sind normalverteilte Residuen der Kontrollgruppe. Der Vorteil einer Z-Transformation liegt in der Standardisierung der Messeinheit. Die Messungen der Kontrollgruppe und die der Testgruppe werden über die Messeinheit „Standardabweichungen der Kontrollgruppe“ miteinander vergleichbar gemacht. Dafür erfolgt eine Umskalierung der Messeinheit (von bisher Abstand in Millimeter) in Standardabweichungen der Kontrollgruppe. Im Ergebnis erhält man für jede Messung die Information, um wie viele Standardabweichungen der Kontrollgruppe die betreffende Messung vom Schätzwert (der Regressionsgeraden) der Kontrollgruppe abweicht.

Mithilfe von Formel 1 werden die Residuen der Kontroll- und Testgruppe, zur Regressionsgeraden der Kontrollgruppe, in z-Werte überführt.

$$z_i = \frac{(R_i - MW(R))}{\sigma(R)}$$

Formel 1 Z-Transformation: Für jedes Residuum (R_i) wird ein z-Wert (z_i) erstellt. Dafür wird die Differenz zwischen R_i und dem Mittelwert (MW) der Kontrollgruppen-Residuen (R) durch die Standardabweichung (σ) der Residuen der Kontrollgruppe geteilt. Der Mittelwert der Kontrollgruppen-Residuen ist per definitionem gleich null. Dementsprechend ist die Einheit dieser z-Werte die „Standardabweichungen der Kontrollgruppen-Residuen“.

Die statistischen Kenngrößen der z-Werte der Residuen sind in Tabelle 5 zusammengefasst und im Boxplot (Abbildung 21) graphisch dargestellt. Mittelwert und Median der Kontrollgruppe sind wie zu erwarten 0 und 1. Dadurch werden Unterschiede zwischen den Stichproben offensichtlich (Abbildung 21): Der Mittelwert der Testsubgruppe TÖF liegt um 1,275 Standardabweichungen unter der Kontrollgruppe, beim Median verhält es sich ähnlich. In der Subgruppe „ÖA ohne Fistel“ sind diese Abstände erwartungsgemäß noch größer (4,5

Standardabweichungen beim Mittelwert). Allerdings ist bei der TÖF Subgruppe auch die Standardabweichung mehr als doppelt so hoch wie die der Kontrollgruppe (2,217-fache), sodass die Überlappung beider Stichproben tendenziell zunimmt. Die Diskrepanz beim Mittelwert zusammen mit der hohen Standardabweichung der TÖF-Gruppe (2,217) lassen vermuten, dass hier ein tatsächlicher Unterschied zwischen den Stichproben besteht. Ein zweiseitiger T – Test für unabhängige Stichproben prüft diesen Unterschied auf seine Signifikanz (Tabelle 6). Hierfür muss eine Normalverteilung der Residuen der Kontrollgruppe angenommen werden. Demnach besteht zwischen allen in Tabelle 5 und Abbildung 21 dargestellten Stichproben ein signifikanter Unterschied auf einem 5%-Niveau (Kontrollgruppe vs. TÖF $p = 0,001$, Kontrollgruppe vs. ÖA ohne Fistel $p < 0,0001$ und TÖF vs ÖA ohne Fistel $p < 0,0001$).

Z-Score der Residuen

Stat. Kenngrößen	Kontrollgruppe	TEF	EA ohne Fistel
Mittelwert	0,00	-1,275	-4,537
Median	-0,118	-1,257	-4,523
Minimum	-2,579	-5,899	-6,401
Maximum	3,282	9,336	-3,061
Oberes 95% - KI	1,960	3,070	-2,640
Unteres 95% - KI	-1,960	-5,620	-6,435
STABW	1	2,217	0,968
Standardfehler des Mittelwertes	0,052	0,360	0,211
Oberes Limit	0,101	-0,570	-4,123
Unteres Limit	-0,101	-1,980	-4,951
Schiefe	0,376	2,722	-0,299
Standardfehler Schiefe	0,127	0,397	0,535
Grobregel Schiefe <1,96	2,969	6,851	-0,559
Excess-Kurtosis	0,363	14,356	-0,854
Standardfehler Excess Kurtosis	0,253	0,795	1,069
Grobregel Excess Kurtosis <1,96	1,432	18,064	-0,799
Anzahl	374	38	21

Tabelle 5 Statistische Kenngrößen der Z-Werte: Die Kontrollgruppe kann hier mit der aufgeteilten Testgruppe in TEF und EA ohne Fistel verglichen werden. Mittelwert und Median unterscheiden sich in allen Stichproben. Im Mittel liegen die Residuen der TEF Gruppe 1,3 und die der Gruppe ohne Fistel 4,5 Standardabweichungen unter der Kontrollgruppe. Auffällig ist das hohe Maximum bei TEF. Der dritte Teil, beginnend mit der Zeile Schiefe, gibt Anhalte auf die Normalverteilung der Z-Werte. Ausmaß von Schiefe und Excess Kurtosis werden mithilfe von Grobregeln untersucht. In der Kontrollgruppe liegt die Schiefe oberhalb des Grenzwertes (Urban and Mayerl, 2011b).

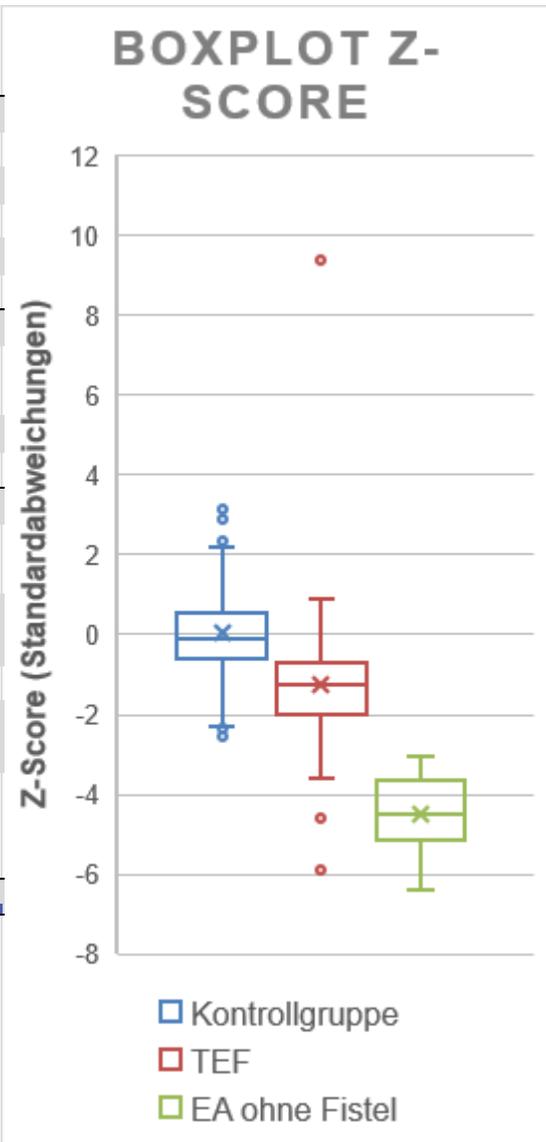


Abbildung 21 Boxplot der Z-Werte: Alle drei untersuchten Gruppen stellen sich nebeneinanderversetzt voneinander dar. Die Verteilungen von TEF und EA ohne Fistel liegen unterhalb der Kontrollgruppe. Auch TEF und EA ohne Fistel fallen erwartungsgemäß deutlich auseinander. In der Gruppe der TEF zeigen sich erneut das hohe Maximum und zwei Ausreißer nach unten.

Gruppenstatistiken

	Ösophagus	Anzahl	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
Z-Score	Kontrollgruppe	374	0	1	0,052
	TÖF	38	-1,275	2,217	0,360
	ÖA ohne Fistel	21	-4,537	0,968	0,211

T-Test bei unabhängigen Stichproben

Z-Score		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit		
		F	Signifikanz	T	df	Sig. (2-seitig)
Kontrollgruppe vs. TÖF	Varianzen sind gleich	7,749	0,006	6,465	410	0,000
	Varianzen sind nicht gleich			3,533	38,555	0,001
Kontrollgruppe vs. ÖA ohne Fistel	Varianzen sind gleich	0,007	0,935	20,318	393	0,000
	Varianzen sind nicht gleich			21,065	22,504	0,000
TÖF vs. ÖA ohne Fistel	Varianzen sind gleich	0,813	0,371	6,434	57	0,000
	Varianzen sind nicht gleich			7,874	54,786	0,000

Tabelle 6 T-Test bei unabhängigen Stichproben: Die erste Tabelle (Gruppenstatistiken) zeigt relevante statistischen Kenngrößen. Es werden die Mittelwerte der einzelnen Gruppen miteinander verglichen. Der zweite Teil veranschaulicht die relevanten Ergebnisse des T-Tests. Zunächst erfolgt ein F-Test zur Prüfung der Varianzgleichheit. Liegt dieser unter 0,05 wird die Annahme der Varianzgleichheit verworfen und es muss das T-Test-Ergebnis der zweiten Reihe „Varianzen sind nicht gleich“ betrachtet werden. Alle Gruppen unterscheiden sich auf einem 0,05-Niveau signifikant voneinander.

3.3.3 Identifikation einer Risikogruppe anhand des fetalen

Magenumfangs

Die Konstruktion der Risikogruppe basiert auf der Feststellung, dass für die Stichprobenverteilung der Testgruppe der Magenumfang in Abhängigkeit vom Abdomenumfang signifikant kleiner ist als in der Kontrollgruppe (Abbildung 8, S.31). Daraus folgt, dass je kleiner der Magenumfang abhängig vom Abdomenumfang gemessen wird, umso geringer ist die Wahrscheinlichkeit, zur Gruppe der „Gesunden“ zu gehören, wenn man als Gradmesser hierfür die Stichprobenverteilung der Kontrollgruppe heranzieht. Im Umkehrschluss lässt sich vermuten, dass bei (besonders) „kleinen“ Magen-/Abdomenumfangsmessungen die Schwangere zur Risikogruppe zählt.

Die Risikogruppe definiert sich also über die Grenze, wann bezogen auf den Abdomenumfang der betreffende Magenumfang (lineare Regression) auffällig klein ist. Als Grenze zur Risikogruppe wird hier die 5. Perzentile in der Regressionsanalyse der Kontrollgruppe (Abbildung 22) definiert. Abbildung 22 basiert auf der Darstellung von Abbildung 17. Die Rohdaten der Magenumfänge der Kontrollgruppe (blaue Punkte in Abbildung 17) werden hier jedoch ausgeblendet. Stellvertretend sind dafür die Regressionsgerade sowie die 5. und 95. Perzentilgerade der Kontrollgruppe abgebildet. Zusätzlich sind nun die Rohdaten der Testgruppe eingezeichnet (Rote Dreiecke = TÖF, schwarze Kreuze = ÖA ohne Fistel). Visuell erkennbar werden daher diejenigen Fälle der Testgruppe (TÖF und ÖA ohne Fistel), bei denen der gemessene Magenumfang lediglich eine Wahrscheinlichkeit von 5% oder weniger aufweist, noch zur Verteilung der Kontrollgruppe zu gehören (Schneider et al., 2010). In Abbildung 22 sind es all die Messungen, die unterhalb oder auf der 5. Perzentilgeraden liegen und mathematisch all die Fälle, die einen Abstand von mindestens -16,448 mm (grüne Parallelen) zur Regressionsgeraden der Kontrollgruppe haben. Die Fälle ÖA ohne Fistel sind eindeutig, da alle (100%) deutlich unterhalb der 5. Perzentile liegen.

Durch die Überlappung der Stichprobenverteilung der Kontrollgruppe mit der Gruppe der TÖF ist hier die Diskriminierung schwieriger. Rechnerisch und visuell sind exakt 13 Fälle mit TÖF unterhalb der 5. Perzentile nachzuweisen. Sie würden über das Screening als Risikokandidaten sichtbar. Im Umkehrschluss verbergen sich 25 Fälle der TÖF-Subgruppe in der Verteilung der Kontrollgruppe oberhalb der 5. Perzentilgeraden und werden in dieser Studie durch die Magenumfangsmessung nicht als Risikofälle klassifiziert. Nach der oben definierten Risikogruppe würden demnach 13 Fälle (34,2%) der TÖF-Gruppe auffallen und 25 Fälle (65,8%) als falsch-negativ diagnostiziert.

Wird das Auftreten eines Polyhydramnions ebenfalls berücksichtigt, so würden weitere Fälle der Testgruppe auffällig. Wie bereits erwähnt, sind 14 (eingekreiste rote Dreiecke in Abbildung 22) von 38 Feten mit TÖF mit einem Polyhydramnion assoziiert. Von diesen 14 zeigen 9 (64,3%) einen Magenumfang, der oberhalb der 5. Perzentile liegt. Lediglich einer dieser 9 Polyhydramnien mit normalem Magenumfang hatte einen auffälligen Karyotypen (Trisomie 18). Mit einer Ausnahme waren alle diese Fälle mindestens in der 24. SSW. Kombiniert man die Suche von einem Magenumfang unter der 5. Perzentile mit dem Vorliegen eines Polyhydramnions (Magenumfang unter der 5. Perzentile oder Polyhydramnion), so würden 22 von 38 Feten mit TÖF auffällig werden. Dies entspricht einer Trefferquote von 57,8% nur bei den TÖF und 72,9% für die gesamte Testgruppe. 16 Fälle (von 38 der TÖF) zeigen weder einen Magenumfang unter der 5. Perzentile noch ein Polyhydramnion.

Zusätzlich zur Regressionsanalyse werden die oben berechneten Z-Werte der Stichproben in Abhängigkeit vom Abdomenumfang verglichen (Abbildung 23). Unter der Annahme von normalverteilten z-Werten der Kontrollgruppe lassen sich hier ebenfalls die 5. und 95. Perzentile (grüne Horizontalen in Abbildung 23) angeben. Betragsmäßig liegen die Perzentilen 1,65 Standardabweichungen ober- und unterhalb der Mittelwertlinie (Nulllinie). Erwartungsgemäß beträgt die Anzahl der Fälle der Testgruppe (ÖA mit und ohne Fistel) die unter die 5. Perzentile fallen 34 Fälle (57,63%). Im Einzelnen liegen 13 von 38 Fälle (34,2%) mit Fistel und 21 von 21 (100%) ohne Fistel unter der 5. Perzentile.

Zusätzlich wird hier die Perzentilenflüchtigkeit der ÖA ohne Fistel direkt sichtbar (ÖA ohne Fistel, schwarze Kreuze in Abbildung 23). Bei zunehmendem Abdomenumfang steigt der Abstand, gemessen in Standardabweichungen, zwischen der Subgruppe ÖA ohne Fistel und der 5. Perzentile.

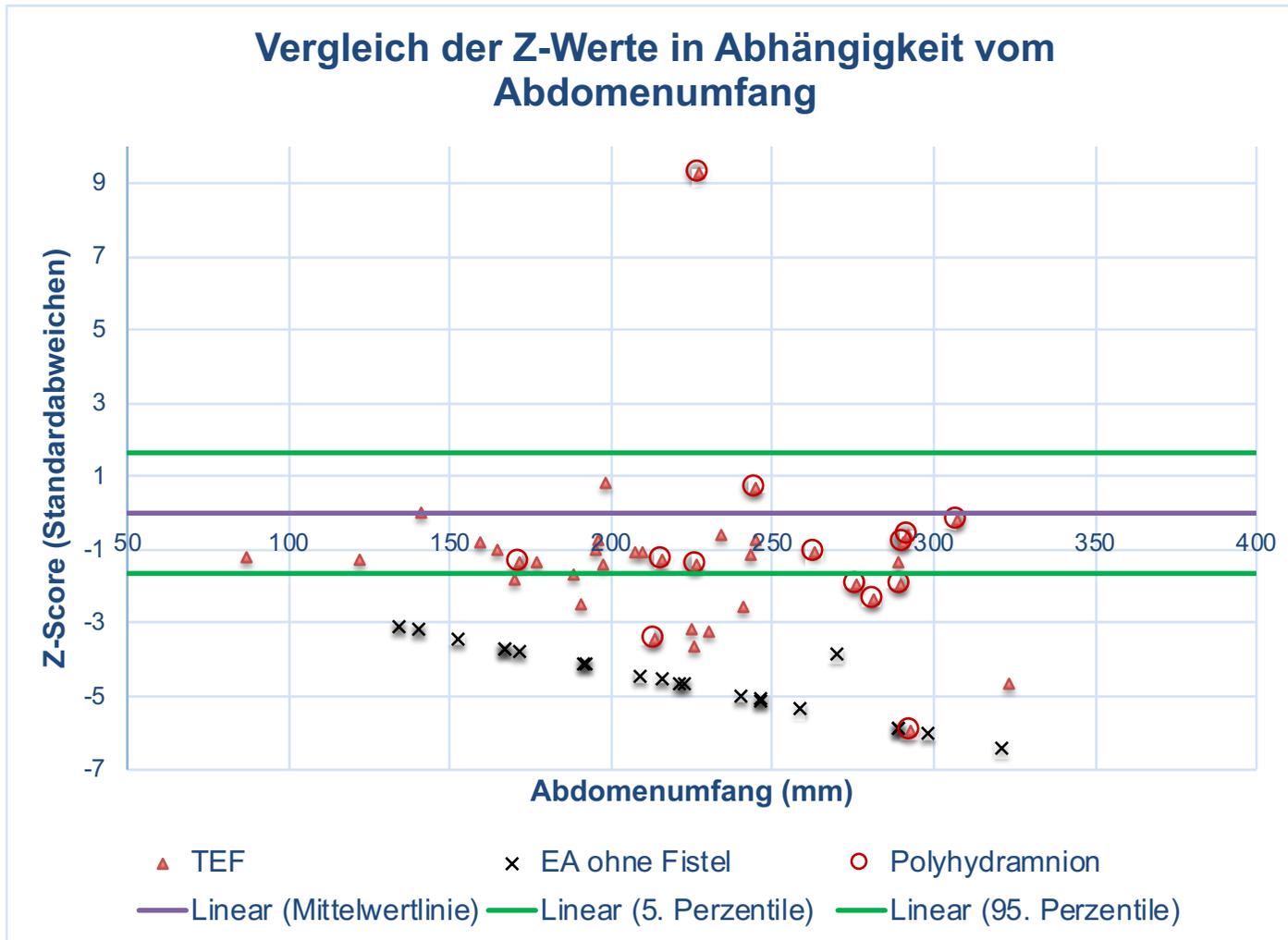


Abbildung 23 Vergleich der Z-Werte in Abhängigkeit vom Abdomenumfang: Mit einem Mittelwert von 0 liegt die Mittelwertlinie der Kontrollgruppe genau auf der x-Achse. Die 5. und 95. Perzentile werden durch die grünen Horizontalen dargestellt. Beobachtungen, die auf- oder bzw. unterhalb dieser Grenzlinien liegen, haben lediglich eine Wahrscheinlichkeit von 5% und weniger, noch zum gesunden Kollektiv zu gehören. Die ÖA ohne Fistel zeigen sich perzentilenflüchtig, da sie sich bei wachsendem Abdomenumfang weiter von der Kontrollgruppe entfernen.

Tabelle 7 fasst die Ergebnisse dieser Studie zusammen. Die objektivierte Messung des fetalen Magenumfangs in Abhängigkeit vom Abdomenumfang zeigt eine Sensitivität von 57,63%, mit einer Spezifität von 95,19%. Ausgedrückt in absoluten Zahlen würde der Magenumfang von 34 Feten mit ÖA pränatal auffallen. Daneben wären 18 Feten des gesunden Kollektivs falsch positiv, bei einer Spezifität von 95,19%.

Vier-Felder-Tafel

Testergebnis	ÖA	Gesund	Total
Pos.	34	18	52
Neg.	25	356	381
Total	59	374	
Sensitivität	57,63%		
Spezifität		95,19%	

Tabelle 7 Vier-Felder-Tafel der Regressionsanalyse: Mit der durchgeführten Statistik lässt sich die in der Tabelle gezeigte Sensitivität und die Spezifität berechnen.

3.3.4 Bewertung der Regressionsgeraden für die Kontrollgruppe

Die Qualität der Regressionsgeraden der Kontrollgruppe wird untersucht, da sie die Grundlage wesentlicher Aussagen dieser Studie bildet. Zur Bewertung der Regressionsgeraden aus Abbildung 17 und Abbildung 22 mit der Formel $AU = 0,1785 \cdot MU + 6,8087$ wird Tabelle 8 erstellt.

Regressions-Statistik

Multipler Korrelationskoeffizient	0,686
Bestimmtheitsmaß	0,470
Standard Error of Estimate (SEE = STABW Residuen)	10,000
SEE in % vom Mittelwert	23,23%
Beobachtungen	374

ANOVA

	Freiheitsgrade (df)	Prüfgröße (F)	F krit
Regression	1	330,157	0,000
Residuen	372		
Gesamt	373		

	Koeffizienten	t-Statistik	P-Wert	Untere 95%	Obere 95%
AU (mm)	0,1785	18,1702	0,000	0,1592	0,1979

Tabelle 8 Bewertung der Regressionsgeraden der Kontrollgruppe: In der Regressionsstatistik sind statistische Kenngrößen aufgeführt, die die Aussagekraft der Regressionsgeraden untersuchen. Der Multiple Korrelationskoeffizient gibt die Korrelation der beiden Testgrößen (Magenumfang und Abdomenumfang) an. Das Bestimmtheitsmaß als Angabe für die Erklärbarkeit der Veränderung des Magenumfangs (abhängige Variable) durch die Veränderung des Abdomenumfangs (unabhängige Variable) ist das Quadrat des Korrelationskoeffizienten. Der Standard Error of Estimate bzw Standardfehler der Schätzung ist die Standardabweichung der Residuen der Kontrollgruppe (Urban and Mayerl, 2011a). Um diesen Wert vergleichbar zu machen, wird er relativiert und in Prozent vom Mittelwert der abhängigen Variablen (Magenumfang) angegeben. Die ANOVA-Analyse dient zur Signifikanztestung des Bestimmtheitsmaßes anhand einer F-Statistik. Auch die Steigung wird auf ihre Signifikanz mit Hilfe eines T-Tests geprüft. Eine Steigung von 0 liegt in dieser Regressionsanalyse nicht im 95% - Konfidenzintervall der Steigung (Verwerfung der Nullhypothese).

Tabelle 8 zeigt statistische Kenngrößen und Tests, die die Regressionsgerade bewerten. Ein F-Wert (F-krit, Tabelle 8) deutlich unter dem 0,05-Signifikanzniveau bestätigt die Signifikanz des Bestimmtheitsmaßes.

Dementsprechend lässt sich die Veränderung des Magenumfangs zu 47,02% durch die Veränderung des Abdomenumfangs erklären. Auch die Steigung (0,1785) dieser Regression ist auf einem Signifikanzniveau von 0.05 (P-Wert in Tabelle 8) hochsignifikant. Der Wert 0 liegt nicht im 95% Konfidenzintervall der Steigung der Regressionsgeraden. Der Standardfehler der Schätzung (Standard Error of Estimate bzw. SEE, Standardabweichung der Residuen der Kontrollgruppe) ist ein absolutes Maß der Streuung um die Regressionsgerade. Der Quotient aus SEE und Mittelwert der Zielvariable stellt eine relative Größe dar und besagt, dass die Residuen der Zielvariablen standardmäßig um 23,23% ihres Mittelwertes (hier Magenumfang von 43.045 mm im Mittel) schwanken. Dies ist ein Maß für die Prognosegenauigkeit der Regressionsgeraden für einen Magenumfang bei einem bestimmten Abdomenumfang. Außerdem erlaubt dieser Quotient durch seinen relativen Charakter einen Vergleich mit anderen Regressionsanalysen.

4 Diskussion

4.1 Der Magenumfang in Anhängigkeit vom Abdomenumfang als Indikator für eine Ösophagusatresie

Die vorliegende Dissertation konnte mithilfe einer Regressionsanalyse nachweisen, dass ein linearer Zusammenhang zwischen dem Magenumfang und dem Abdomenumfang besteht, gemessen in der abdominalen Biometrieebene und im Rahmen einer normalverlaufenden Schwangerschaft (Abbildung 17 und Tabelle 8). Es ist auf der Basis der Regressionsanalyse festzustellen, dass der Magenumfang von Feten mit ÖA statistisch signifikant kleiner ist als bei Feten der Kontrollgruppe (Tabelle 6 und Abbildung 21). Durch die Analyse der Residuen wurde gezeigt, dass der Magenumfang in Abhängigkeit vom Abdomenumfang ein Indikator für eine ÖA darstellt. Die Inferenzstatistik belegt die statistische Signifikanz dieser Aussagen (Tabelle 6 und Tabelle 8).

Für die gesamte Testgruppe ist mit der durchgeführten Regressionsanalyse eine Sensitivität von 57,63% mit einer Spezifität von 95,19% erreicht worden (Tabelle 7). Im Vergleich dazu präsentiert Bradshaw Sensitivitätsergebnisse von 57%, jedoch mit einer Spezifität von 99%. Durch Anpassung der Spezifität auf 99% in der eigenen Regressionsanalyse wird eine Sensitivität von lediglich 49,15% erreicht. Dadurch erscheinen die Werte von Bradshaw zunächst überlegen. Die Vergleichbarkeit ist jedoch kritisch zu betrachten, da die Sensitivitätsergebnisse stark von dem Anteil der Fälle ohne Fistel in der Testgruppe abhängt und die Magenumfangseinschätzung bei Bradshaw subjektiv erfolgte. Als Faustregel gilt dabei: Je höher die Anzahl der Fälle ohne Fistel, desto größer sind die Sensitivitäts- und Spezifitätswerte, da Fälle ohne Fistel offensichtlicher durch einen nicht-darstellbaren Magen erkennbar sind (Garabedian et al., 2015). Die Subjektivität schränkt die Verlässlichkeit der Spezifität sowie die Reproduzierbarkeit ein. Zudem wurde bei Bradshaw das Vorliegen eines

Polyhydramnions in den oben genannten Ergebnissen berücksichtigt, welches für die eigene Studie nicht einbezogen wurde (Bradshaw et al., 2016).

Erwartungsgemäß zeigt sich der deutlichste Unterschied zwischen der Kontrollgruppe und den Fällen ohne Fistel. In allen Fällen ohne Fistel (100%) liegt die Magenumfangsmessung in der Regressionsanalyse unter der 5. Perzentile und in nahezu allen Fällen ohne Fistel (20 von 21, 95,2%) war kein Magen darstellbar (Abbildung 22). Diese Ergebnisse bestätigen die aktuelle Studienlage (Houben and Curry, 2008, Brantberg et al., 2007, Garabedian et al., 2015): Demnach wäre ein Fall bei subjektiver Bewertung eventuell nicht erkannt worden (Abbildung 22), der mit der hiesigen Testmetrik aber auffällt. Falsch-negativ Raten sind bei Fällen ohne Fistel zwar selten, können jedoch auch dort vorkommen. Die Mukusproduktion der Magenschleimhaut kann zu einer Magenfüllung führen, die bei subjektiver Einschätzung zu falsch-negativen Befunden führen kann (Brantberg et al., 2007). Weiterhin fällt auf, dass der relative Anteil der Embryonen, die von Trisomie 18 betroffen sind, in der Gruppe ohne Fistel (52,4%) mehr als doppelt so hoch ist wie in der Gruppe mit Fistel (23,7%). Auch diese Beobachtung stimmt mit den bisherigen Studien überein (Depaepe et al., 1993, Felix et al., 2007, Pedersen et al., 2012). Folglich sollte besonders bei Befundkonstellationen (Polyhydramnion und keine Magenfüllung), die auf eine ÖA ohne Fistel hindeuten, nach weiteren Fehlbildungen gesucht werden, die für eine Trisomie 18 typisch sind.

Bei den Fällen mit Fistel liegen 34,2% unter der 5. Perzentile. Im Vergleich dazu präsentiert eine umfassende Studie von Garabedian Sensitivitätswerte von lediglich 17,9% bei den Fällen mit Fistel (Garabedian et al., 2015). Für einen genaueren und besseren Vergleich dieser Sensitivitätswerte wäre ein Spezifitätswert nötig, welcher in der Studie von Garabedian leider nicht angegeben wird. Die Sensitivitätsergebnisse aus der Studie von Garabedian basieren auf der Kombination aus pränatalem Ultraschall und fetalem MRT (Garabedian et al., 2015). Dementsprechend zeigt sich, dass sich die

durchgeführte Regressionsanalyse vorteilhaft auf die Sensitivität und die falsch-negativ Rate bei den Fällen mit Fistel auswirkt und der vergleichbaren Literatur überlegen ist. Allerdings wird in der hiesigen Studie eine 5% falsch-positiv Rate festgelegt.

4.2 Problematisch hohe „Falsch-Positiv Rate“

Ein nicht darstellbarer Magen oder ein subjektiv kleiner Magen und ein Polyhydramnion stellen die übliche Befundkonstellation dar, die zum Verdacht einer ÖA führt. Auf Grund der zum Teil subjektiven und indirekten Befundkonstellation sind hohe falsch-positiv Raten bekannt (Bradshaw et al., 2016, Houben and Curry, 2008, Kunisaki et al., 2014) – besonders bei der Einschätzung der Magengröße (Borsellino et al., 2006, Ethun et al., 2014). Die Subjektivität der Bewertung macht dabei genaue Angaben zur falsch-positiv Rate unmöglich, da diese untersucherabhängig ist. Eine festgelegte Spezifität von 95% in dieser Arbeit bedeutet, dass konstruktionsbedingt mit einer 5% falsch-positiv Rate bei Normalverteilung der Stichprobe zu rechnen ist. Am Beispiel der Geburten von 2015 in der Bundesrepublik Deutschland (Statistisches Bundesamt) wären das 36 900 falsch-positive Fälle. Die hohen falsch-positiv Raten sind also ein gemeinsames Problem der gängigen Befundkonstellation und dieser Studie.

Eine episodische Schluckaktivität, Testparameter indirekter Natur und eine niedrige Prävalenz der Erkrankung sind Hauptursachen für extrem hohe falsch-positiv Zahlen. Ein subjektiv sowie objektiv kleiner Magen kann häufig damit begründet werden, dass der Fetus in letzter Zeit nicht geschluckt hat (Choudhry et al., 2007, Shulman et al., 2002, Bradshaw et al., 2016). Weiterhin sind durch die subjektive Einschätzung der Magengröße die Reproduzierbarkeit und die Standardisierung der Ultraschallkriterien schwierig bis unmöglich. Diese Limitierung wurde versucht mit dieser Arbeit einzuschränken. Mit standardisierten Messungen könnten zusätzlich erhobene Messungen

miteinander verglichen und eine Entwicklung der Magengröße dokumentiert werden. Durch wiederholte Messungen während der Schwangerschaft könnte sich die falsch-positiv Rate weiter einschränken lassen.

Ein falsch positiver Befund ist problematisch in vielerlei Hinsicht. Verständlicherweise führt eine solche Vermutung/Diagnose zur Beunruhigung der Schwangeren und deren Umfeld, sodass hieraus eventuell eine starke psychische Belastung resultieren kann, durch die sogar im Extremfall eine vorzeitige Beendigung der Schwangerschaft ausgelöst werden kann (Kunisaki et al., 2014). Abgesehen von den zwischenmenschlichen und persönlichen Problemen, die durch eine Positivdiagnose entstehen können, belegen manche Studien Stress als Risikofaktor für eine Frühgeburt. Corticotropin-releasing Hormon (CRH), welches lokal von Dezidua, Plazenta und fetalen Membranen (Chorion und Amnion) produziert wird, induziert von selbigem Gewebe die Prostaglandinausschüttung, die zu Uteruskontraktionen führt (Lockwood, 1999). Obwohl das Frühgeburtsrisiko lediglich um das Zweifache ansteigt, muss eine Frühgeburt mit allen Konsequenzen bei der Äußerung eines Verdachts auf eine ÖA bedacht werden (Dole et al., 2003). Weiterhin sind die Risiken einer weiterführenden Diagnostik zu berücksichtigen. Der invasive Charakter einer Fruchtwasserpunktion birgt das Risiko von Infektionen, Verletzungen von mütterlichem und kindlichem Gewebe und der frühzeitigen Beendigung der Schwangerschaft (Kähler et al., 2013).

Dementsprechend sollte die pränatale Diagnose bzw. eine weiterführende Diagnostik weder auf der subjektiven Beurteilung der Magengröße noch der einmaligen Messung des Magenumfangs beruhen. Allerdings kann der behandelnde Arzt durch eine objektive Messung vorzeitig auf den kleinen Magen des Fetus aufmerksam werden und seinen Beobachtungen entsprechend situationsadäquate Handlungsstrategien in Betracht ziehen.

4.3 Suche nach einem „Pouch-sign“ zur Verbesserung der pränatalen Detektionsrate

Zur Verbesserung der pränatalen Detektionsrate konzentriert sich ein Großteil der Forschung auf die sonographische Visualisierung des ösophagealen „Pouch-sign“. Definiert ist dieses Zeichen als Dilatation des proximalen Ösophagusstumpfes im Mediastinum und wurde 1983 zum ersten Mal beschrieben (Garabedian et al., 2014, Holland and Fitzgerald, 2010, Houben and Curry, 2008, Eyheremendy and Pfister, 1983). Typischerweise ist dieses Zeichen, ähnlich wie ein Polyhydramnion, erst im dritten Trimester darstellbar (Holland and Fitzgerald, 2010, Houben and Curry, 2008). In manchen Studien beläuft sich die Spezifität auf 100%, weshalb das „Pouch-sign“ oftmals als pathognomonisch für eine ÖA gilt (Garabedian et al., 2014, Brantberg et al., 2007). Allerdings ist dieses Zeichen in einem Großteil der Feten mit ÖA nicht zu sehen (Sensitivität ~ 40%) (Brantberg et al., 2007, Garabedian et al., 2014). Die Ursachen dafür liegen in der schwierigen Darstellung, die vom Gestationsalter, von der Position und Aktivität des Fetus sowie von der Schluckaktivität abhängig ist (Garabedian et al., 2014, Kunisaki et al., 2014, Brantberg et al., 2007). Zusätzlich ist besonders bei Feten mit TÖF das „Pouch-sign“ selten, da durch den Kurzschluss das Fruchtwasser in den Magen abfließen kann. Diese Eigenschaften machen die pränatale Visualisierung des „Pouch-sign“ schwierig und oftmals sehr zeitintensiv (20-30 Minuten) (Solt et al., 2010, Garabedian et al., 2014). Zusätzlich gibt es einzelne Studien, die das Vorkommen des „Pouch-signs“ als benignen und transienten Untersuchungsbefund in normalen Schwangerschaften beschreiben (Solt et al., 2010).

Auf Grund dieser Rahmenbedingungen ist das „Pouch-sign“ zur Bestimmung einer Risikogruppe für eine ÖA eher ungeeignet. Durch die hocheinzuschätzende Spezifität würde sie sich als komplementierende Untersuchung bei einer Risikogruppe eignen, besonders weil ein hoher positiver prädiktiver Wert (bis zu 100%) beschrieben wird (Garabedian et al., 2014)

4.4 Das Kernproblem: Ösophagusatresien mit tracheoösophagealer Fistel

In der Pränataldiagnostik stellen die Fälle mit Fistel das Kernproblem dar, da besonders diese Subgruppe unter den falsch-negativ Befunden vertreten ist und damit oftmals bis zur Geburt unerkannt bleiben (Holland and Fitzgerald, 2010, Hoopmann et al., 2015, Garabedian et al., 2015). Deutlich wird dies dadurch, dass von allen Feten, die erst postpartal mit einer ÖA diagnostiziert werden, über 90% eine Fistelanlage zeigen (Typ III nach Ladd, vergleichbar mit Typ IIIb nach Vogt) (Garabedian et al., 2015, Ladd and Swenson, 1947). Hierfür sind hauptsächlich drei Faktoren verantwortlich. Erstens kann die funktionelle Anatomie eine Magenfüllung ermöglichen, wenn ein Kurzschluss zwischen Trachea und distalem Ösophagusstumpf besteht. Das Ausmaß der Magenfüllung wird mit der Größe der TÖF assoziiert (Bradshaw et al., 2016, Houben and Curry, 2008, Pretorius et al., 1987, Rao et al., 1978). Dies erschwert die Detektion besonders, wenn die Beurteilung subjektiv erfolgt. Zweitens kommt epidemiologisch eine ÖA zu über 85% mit einer Fistel vor (Houben and Curry, 2008, Pinheiro et al., 2012). Das heißt neben der schwachen Detektion bilden die Fälle mit Fistel die überwältigende Mehrheit. Drittens sind begleitende Fehlbildungen bei ÖA mit einer Fistel seltener, die zu einer auffälligen Schwangerschaft führen würden (Holland and Fitzgerald, 2010, Pinheiro et al., 2012, Spitz, 2007). Diese Arbeit bestätigt die bereits publizierten Beobachtungen. Mit der objektiven Magenumfangsmessung ist es in der vorliegenden Arbeit gelungen, 34,2% (Spezifität 95,19%) der Feten mit Ösophagusatresie und einer TÖF aus der Test-Subgruppe TÖF pränatal zu erfassen. Die beschriebenen Ergebnisse sind im Vergleich zur Trefferquote von Garabedian (17,9%) als Fortschritt zu werten (Garabedian et al., 2015). Dennoch werden auch in dieser Arbeit Feten ohne Fistel erwartungsgemäß deutlich häufiger pränatal erfasst (100%).

Trotz einer generell besseren Prognose der Fälle mit Fistel, sind auch sie nicht komplikationslos. Unvorbereitetes medizinisches Personal, nichtsahnende Eltern, die Trennung von Mutter und Kind, eine Verzögerung der ersten chirurgischen Maßnahme sowie weitere medizinische Komplikationen sind Herausforderungen, die mit einer pränatalen Diagnose voraussichtlich besser zu organisieren und zu behandeln sind (Bradshaw et al., 2016, Brantberg et al., 2007, Garabedian et al., 2015, AWMF online, 2012). In diesem Sinne ist die objektive Messung des Magens, am Beispiel des Magenumfangs, zur Beurteilung der Magengröße gerade für die Fälle mit Fistel sinnvoll.

Die Suche nach geeigneten statistischen Screenings liefert einen möglichen Ansatz, um die pränatale Detektionsrate bei den Fällen mit Fistel zu verbessern. Abbildung 23 beschreibt zunächst einen Trend für die Fälle ohne Fistel, die sich bei zunehmendem Abdomenumfang perzentilenflüchtig zeigen. Hypothetisch könnte diese Eigenschaft auch auf die Subgruppe der TÖF zutreffen. Ein zunächst unauffälliger Magen (über der 5. Perzentile) könnte im Verlauf der Schwangerschaft unter die 5. Perzentile fallen. Auf dieser Basis wäre ein fetales Magenumfangsprofil potentiell sinnvoll, welches eine dynamische, im Zeitablauf der Schwangerschaft beobachtete und analysierte Entwicklung des Magenumfangs in Abhängigkeit vom Abdomenumfang beschreibt. Ähnliche Herangehensweisen finden bereits Anwendung in der täglichen Praxis, z.B. bei der Diagnostik der fetalen Wachstumsrestriktion (American College of Obstetricians and Gynecologists, 2013). Zudem wäre ein Magenumfangsprofil ein Lösungsansatz, um die hohen falsch-positiv und falsch-negativ Befunde zu reduzieren.

Der Querschnittscharakter der Studie erlaubt nur eine hypothetische Aussage zur Eigenschaft der Perzentilenflucht, die in weiteren Studien überprüft werden könnte. Allerdings unterstützen die Trendlinien in Abbildung 8 diese Vermutung, da die Testgruppe eine signifikant niedrigere Steigung aufweist als die Kontrollgruppe (Vergleich der Trendlinien der Kontrollgruppe vs. TÖF in

Abbildung 8). Die Trefferquote wäre dann abhängig vom Messzeitpunkt und dem statistischen Modell. Für eine zuverlässige Aussage über das trendmäßige Wachstum der TÖF-Gruppe ist das Bestimmtheitsmaß der TÖF-Gruppe zu schwach und die Fallzahl zu gering. Weiterhin befindet sich ein deutlicher Ausreißer in dieser Gruppe, der zu einer Verzerrung des Trends und der Verteilungseigenschaften der TÖF-Gruppe führt. Ohne den Ausreißer wäre ein stärkerer Trend zu erwarten. Deshalb sollte ein Fokus folgender Studien auf Ösophagusatresien mit Fisteln liegen, um diese Fragestellung zu analysieren.

4.5 Die Polyhydramnie als wichtiges Zusatzkriterium

Ein Polyhydramnion ist ein anerkanntes indirektes Zeichen für das Vorliegen einer ÖA (Houben and Curry, 2008, Pinheiro et al., 2012). Polyhydramnien traten nur in der Testgruppe auf, da sie ein Ausschlusskriterium für die Kontrollgruppe darstellten. In der Testgruppe kam es in 44,1% der Fälle zu einem Polyhydramnion. Laut Literatur wären Werte von bis zu 73% zu erwarten gewesen (Kunisaki et al., 2014). Die Häufigkeitsverteilung ist unterschiedlich und hängt vom Vorliegen einer Fistelanlage ab. Bei einer reinen Atresie (ÖA ohne Fistel) ist die Polyhydramnionanzahl erwartungsgemäß höher (57,1%) als bei Fällen mit Fistel (36,8%). Erklärt wird dies durch die vollständige Kontinuitätsunterbrechung des Gastrointestinaltraktes, welche die Passage von Fruchtwasser in den Magen unmöglich macht (Pretorius et al., 1987, Rao et al., 1978). Dennoch liegen beide Gruppen in den untersuchten Stichproben der Testgruppe unterhalb der in der Literatur zu findenden Werte.

Ursachen hierfür können Chromosomenaberrationen und der Ultraschallzeitpunkt sein. Zunächst befinden sich 9 Feten in der Testgruppe noch vor der 24. SSW, sodass das Auftreten eines Polyhydramnions unwahrscheinlich ist (Kunisaki et al., 2014). Nur in einem Fall kam es zu einem Polyhydramnion vor der 24. SSW. Außerdem ist der Anteil an schwerwiegenden Chromosomenstörungen hoch, nämlich 33,9%. Eine relative Häufigkeit von 10%

wäre laut Literatur zu erwarten gewesen (Depaepe et al., 1993, Felix et al., 2007, Pedersen et al., 2012). Chromosomenstörungen haben meist zu einer frühzeitigen Schwangerschaftsbeendigung geführt (vor der 24. SSW), die die Entwicklung eines Polyhydramnions aus zeitlicher Sicht nicht erlaubte (Hoopmann et al., 2015). Diese Häufung könnte an der erhöhten Anzahl an Überweisungen von auffälligen Schwangerschaften an ein Perinatalzentrum Level I liegen. Die Häufigkeitsverteilung der Trisomien stimmt mit der Fachliteratur überein. Trisomie 18 stellt dabei die häufigste Chromosomenaberration unter den ÖA dar, gefolgt von Trisomie 21. Eher selten findet man eine Trisomie 13 (Pedersen et al., 2012, Depaepe et al., 1993, Felix et al., 2007).

Überraschend und interessant ist die Verteilung einer Polyhydramnie bei den Fällen mit Fistel, wenn man gleichzeitig den Magenumfang betrachtet (Abbildung 22). In dieser Arbeit zeigen eine Vielzahl der Fälle mit Fistel und einem Polyhydramnion einen unauffälligen Magenumfang (9 von 14 bzw. 64,3%, siehe Abbildung 22). Somit lässt sich die Trefferquote bei TÖF-Fällen mit diesem Zusatzkriterium deutlich verbessern, von 34,2% auf 57,8%. Diese Beobachtung widerspricht den Erwartungen, die von den theoretischen Überlegungen der Literatur geprägt sind. In der Theorie ist das Auftreten eines Polyhydramnions bei Fällen mit Fistel von der Größe der Fistel und der Schluckaktivität des Fetus abhängig (Houben and Curry, 2008). Große Fisteln erlauben die Passage ausreichender Mengen an Fruchtwasser, sodass kein Polyhydramnion entsteht und es zu einer Magenfüllung kommt (Pretorius et al., 1987, Rao et al., 1978, Brantberg et al., 2007). Diese Argumentation erklärt unter anderem die hohe falsch-negativ Rate bei der Beurteilung der Magengröße. Auf dieser Basis wäre ein Polyhydramnion besonders mit kleinem Magen zu erwarten. Insofern überraschen die Ergebnisse dieser Studie, da die meisten Polyhydramnien bei den Fällen mit Fistel einen unauffälligen Magenumfang zeigen (Abbildung 22). Zwei Feten mit Fistel, Polyhydramnion und normalem Magenumfang, weisen Chromosomenaberrationen auf. Lediglich eines der Feten hatte einen

Magenumfang unter der 5. Perzentile. Die Ursache, warum so viele Fälle mit Fistel und Polyhydramnion einen unauffälligen Magenumfang zeigen, ist mit den erhobenen Daten nicht ergründbar. Ob es sich bei dieser Auffälligkeit um eine Ausnahme in dieser Studie handelt oder reproduzierbar ist, muss in weiteren Untersuchungen geklärt werden. Da die Häufigkeit eines Polyhydramnions in der Testgruppe vergleichsweise niedrig ist, wird die Trefferquote von 57,8% bei den Fällen mit Fistel eventuell sogar unterschätzt (Kunisaki et al., 2014). Abschließend bestätigt diese Arbeit eine Polyhydramnie als wichtiges Zusatzkriterium, besonders beim Kernproblem der Fälle mit Fistel.

4.6 Diagnostische Alternativen zur Verbesserung der pränatalen Detektionsrate

Vielerlei diagnostische Alternativen stehen der Pränataldiagnostik zur Detektion einer fetalen ÖA zur Verfügung. Alle haben ihre Limitationen und Risiken, welche von zeitaufwändig über kostenintensiv bis invasiv rangieren. Die einzelnen Alternativen werden in diesem Kapitel kurz vorgestellt.

Da die Magengröße und ein Polyhydramnion lediglich als Indizien für eine ÖA anzusehen sind, verfolgen viele Forschungsansätze eine Suche nach direkten Merkmalen, um die pränatale Detektion zu verbessern. Dabei konzentriert man sich auf die direkte Visualisierung des fetalen Ösophagus, entweder mittels Ultraschall oder fetalem MRT.

Die sonographische Darstellung des fetalen Ösophagus gilt als große Herausforderung, weil der Ösophagus von Strukturen umgeben ist, die der eigenen sehr ähneln (Malingier et al., 2004). Eine ganzheitliche anatomische Darstellung des fetalen Ösophagus kann mit einem „high-resolution“ linearen Transducer gelingen. Dies funktioniert in 86,7% der Fälle zwischen der 19. und der 25. SSW. Zusätzlich konnte bei 92,3% innerhalb von fünf Minuten die

Passage von Fruchtwasser, im Rahmen des fetalen Schluckakts, nachgewiesen werden (Malingier et al., 2004). Eine Kasuistik beschreibt die pränatale, sonographische Diagnose einer ÖA mit einer TÖF durch Fokussierung auf den thorakalen Teil des Ösophagus (zwischen Trachea und Aorta) (Develay-Morice et al., 2007). Weitere Fortschritte erhofft man sich durch die 3-dimensionale Darstellung des Ösophagus. Objektive Kriterien, die einen gesunden Ösophagus im 3D-Ultraschall beschreiben, existieren allerdings noch nicht. Dementsprechend konzentriert sich die Forschung der 3-dimensionalen Darstellung auf eine verbesserte Visualisierung des „Pouch-sign“ (Achiron et al., 2008, Yagel et al., 2005). Eine detaillierte sonographische Darstellung des fetalen Ösophagus ist vielversprechend und technisch durchführbar. Allerdings wird dafür spezielle Hardware benötigt und die Untersuchung ist technisch anspruchsvoll und dadurch zeitaufwändig.

Neben dem pränatalen Ultraschall besteht die Möglichkeit eines fetalen MRT's. Die pränatale Diagnose mittels T2-Wichtung im MRT stützt sich auf den nicht-darstellbaren mediastinalen Teil des Ösophagus (Langer et al., 2001). ÖA typische Zeichen, wie ein nicht-darstellbaren Magen und das „Pouch-sign“, sind im MRT oftmals ebenfalls zu erkennen. Störfaktoren, wie die Position des Kindes und die Echogenität der Patientin auf Grund ihrer Körperkonstitution, treten bei einem fetalen MRT nicht auf (Garabedian et al., 2014, Langer et al., 2001). Allerdings führt die flache und damit unbequeme Rückenlage häufig zu mangelnder Adhärenz der Patientin (Yagel et al., 2005). Hiermit lassen sich Sensitivitäts- und Spezifitätswerte von 100% bzw. 80% erzielen, welche der eigenen Arbeit deutlich überlegen sind. Damit bietet das fetale MRT als weiterführende Diagnostik eine relevante Option (Langer et al., 2001). Allerdings ist ein fetales MRT sehr zeit- und kostenintensiv.

Als Ultima Ratio ist die invasive Diagnostik mittels Fruchtwasseruntersuchung anzusehen. Mit dem gewonnenen Fruchtwasser wird ein „amniotic fluid digestive enzyme assay“ durchgeführt. Auf Basis dieser biochemischen Untersuchung wird

ein „Esophageal atresia index“ errechnet, welcher das Produkt aus Alfafetoprotein (AFP) und gamma-glutamyl Transpeptidase (GGTP) angibt. Bei ÖA, unabhängig von ihrem Typ, stellte sich dieser Index signifikant erhöht dar und es ließen sich Sensitivitäts- und Spezifitätswerte von 98% und 100% erzielen (Czerkiewicz et al., 2011). Diese beeindruckenden Werte konnten in einer Studie von Garabedian et al. 2014 allerdings nicht wiederholt werden. Diese Gruppe präsentiert Sensitivitäts- und Spezifitätsergebnisse von 88,9% und 60% (Garabedian et al., 2014). Ähnlich zum 3-dimensionalen Ultraschall sind auch hier keine objektiven Grenzwerte für diese Untersuchung bekannt, sodass ein großflächiger Einsatz erschwert wird. Basierend auf diesen Schwankungen, kleinen Fallzahlen in den Studien und dem invasiven Charakter der Amniozentese sollte der „Esophageal Atresia Index“ als letzte Option in der pränatalen Diagnostik der ÖA angesehen werden.

Es bestehen umfangreiche pränataldiagnostische Möglichkeiten zusätzlich zur sonographischen Befundkonstellation aus kleinem/nicht-darstellbarem Magen und Polyhydramnion. Allerdings haben alle ihre Limitationen, sodass sie nicht in einem unselektierten Routinescreening zum Einsatz kommen können.

4.7 Nutzen einer pränatalen Diagnose

Der Nutzen der pränatalen Diagnose einer ÖA ist in der Literatur widersprüchlich, da in manchen Studien vorteilhafte Prognosen präsentiert werden, während andere Studien das Gegenteil belegen (De Jong et al., 2010a, Brantberg et al., 2007). Dabei wurden als Prognosekriterien die Länge des Krankenhausaufenthalts, die Komplikationsrate und die Mortalität einbezogen. Generell ist die Prognose stark davon abhängig, welcher Ösophagusatresietyp vorliegt. Bei pränatal-diagnostizierten Neugeborenen liegt zum großen Teil eine Atresie ohne Fistel vor, auf Grund der einfacheren Detektion mit den üblichen sonographischen Diagnoseparametern (Garabedian et al., 2015, Brantberg et al., 2007). Dieser Ösophagusatresietyp hat mehrere Eigenschaften, die seine

Prognose verschlechtern. Zunächst bestehen bei diesem Typ häufiger weitere Fehlbildungen oder Chromosomenaberrationen (Holland and Fitzgerald, 2010, Spitz, 2007). Weiterhin kommt es bei Ösophagusatresietypen I und II nach Vogt häufiger zu „long gap“ Atresien (Brantberg et al., 2007, Garabedian et al., 2015). Große Distanzen zwischen proximalem und distalem Stumpf erschweren das operative Verfahren und resultieren in einer erhöhten operativen und postoperativen Komplikationsrate (Garabedian et al., 2015). Diese Eigenschaften sind ursächlich für einen verlängerten Krankenhausaufenthalt, eine erhöhte Komplikationsrate und eine gestiegene Sterblichkeit (Brantberg et al., 2007, De Jong et al., 2010a, Garabedian et al., 2015). Studien, die eine schlechtere Prognose für pränatal-diagnostizierte Feten zeigen, weisen wahrscheinlich eine Überrepräsentation der Fälle ohne Fistel auf.

Diese Studie macht deutlich, dass eine höhere Trefferquote der TÖF im pränatalen Ultraschall möglich ist (57,8%, bei einer Befundkonstellation aus Magenumfang < 5. Perzentile in der Regressionsanalyse und/oder Polyhydramnion). Dadurch würde sich die Verteilung der Ösophagusatresietypen in den pränatal diagnostizierten Fällen verändern, mit einer größeren Repräsentation an Fällen mit Fistel. Aktuelle Studien zeigen bereits, dass das klinische Ergebnis von ÖA mit Fistel besser ist als das ohne Fistel (Garabedian et al., 2015). Erwartungsgemäß müsste sich die Prognose zu Gunsten der pränatal diagnostizierten Fälle verändern.

Im Allgemeinen muss davon ausgegangen werden, dass eine korrekte pränatale Diagnose im Interesse der Patientin, des Feten und der beteiligten medizinischen Einheiten ist. Für alle Beteiligten bedeutet eine pränatale Diagnose vor allem einen zeitlichen Gewinn.

So können sich die werdenden Eltern mit der Bedeutung sowie den Folgen der Diagnose beschäftigen und haben Zeit, sich vielseitig zu informieren und erforderliche Planungen zu realisieren. Die Chance, sich durch pränatale Diagnostik situationsentsprechend vorbereiten zu können, führt dazu, der

Herausforderung wohlüberlegt zu begegnen. Die pränatale Diagnose ist für den Fetus pränatal sowie postpartal von Vorteil. Pränatal sollte die Diagnose einer ÖA die Suche nach weiteren Fehlbildungen auslösen, besonders, wenn der Verdacht auf eine Atresie ohne Fistel besteht (Garabedian et al., 2015). Postpartal profitiert der Fetus vor allem dadurch, dass das medizinische Team bei der Geburt vorbereitet und damit handlungsfähig ist. In den letzten Jahrzehnten gelang bereits eine deutliche Verbesserung der anästhesiologischen, operativen und intensivmedizinischen Verfahren, die die Betreuung von Mutter und Kind optimierte (AWMF online, 2012). Mit einer erhöhten pränatalen Detektionsrate würden mehr Neugeborene und Mütter von diesen optimierten Verfahren profitieren. Zusätzlich müsste in weniger Fällen das Neugeborene von der Mutter getrennt werden, um in eine Spezialklinik überwiesen zu werden, was die Zeit bis zur ersten chirurgischen Intervention verlängert (Garabedian et al., 2015). Das medizinische Team kann die Zeit nutzen, um eine vollständige und ausgiebige Aufklärung durchzuführen und eine umfassende Betreuung anzubieten (Brantberg et al., 2007). Dies kann z.B. in Form einer psychologischen Betreuung der Eltern erfolgen. Durch die zusätzliche Zeit bis zur Geburt wird es dem interdisziplinären Team ermöglicht, nach weiteren Fehlbildungen zu suchen sowie sich auf die optimale Behandlung nach der Geburt vorzubereiten. Auch aus ökonomischer Sicht wäre eine pränatale Diagnose von Vorteil, da so Ressourcen gezielter und effektiver eingesetzt werden können.

4.8 Kritische Betrachtung dieser Arbeit

4.8.1 Keine Berücksichtigung der Dynamik der Magenfüllung

Auf Grund der Seltenheit der Erkrankung (siehe Kapitel 1.2) war die Wahl einer retrospektiven Fall-Kontrollstudie sinnvoll. Allerdings erlaubt der retrospektive Charakter der Studie keine Berücksichtigung der Dynamik der Magenfüllung. Diese wird unter anderem durch die Schluckaktivität des Feten und durch das

Gestationsalter beeinflusst (Houben and Curry, 2008, Malinger et al., 2004). Durch die Wahl einer bivariaten linearen Regression wird implizit unterstellt, dass die Größenzunahme des fetalen Magens zwischen der 19. und 40. SSW linear wächst, d.h. mit abnehmender Wachstumsrate. Studien zur Magenentwicklung zeigen jedoch, dass die Größenzunahme des Magens während der Schwangerschaft unterschiedlichen Phasen unterliegt (Goldstein et al., Nagata et al.). Weiterhin hängt die Magenfüllung von der Schluckaktivität der Feten ab (Houben and Curry, 2008). Diese Einschränkungen gelten allerdings für beide unabhängigen Stichproben, d.h. für die Kontroll- und Testgruppe gleichermaßen. Auch der Querschnittscharakter dieser Studie ist eher ungeeignet zur Beschreibung der Wachstumsdynamik des Magens. Um den Zeitablauf zu berücksichtigen, könnte man die Abhängigkeit von Magen- und Abdomenumfang vom Gestationsalter explizit in das statistische Modell aufnehmen.

4.8.2 Der Magenumfang als ein weniger fehleranfälliger

Parameter als die Magenfläche

Im Rahmen dieser Arbeit wurden zwei Parameter erhoben mit denen die Magengröße objektiv eingeschätzt werden sollte: zum einen der Magenumfang und zum anderen die Magenfläche. Bei der Auswertung der Daten zeigte sich, dass die erhobenen Daten zur Einschätzung der Magengröße deutlichen Schwankungen unterliegen. Besonders die Datenverteilung der Magenfläche ist deutlich breiter als die des Magenumfangs. Der entscheidende Unterschied besteht in der Quadrierung des Radius in der Flächenformel. Dadurch wird jede Ungenauigkeit bei der Vermessung des Magens quadriert („Error Maximization“). Deutlich wird dies an höheren Variationskoeffizienten und Standardfehlern des Mittelwertes. Die Präzision der Messung hat einen entscheidenden Einfluss auf die Verlässlichkeit des Regressionsmodells. Aus diesem Grund wurde für die objektive Einschätzung des Magens der Magenumfang gewählt.

4.8.3 Die Qualität der Schätzung ist entscheidend für die

Aussagekraft der Analyse

Die Regressionsanalyse der Kontrollgruppe bildet die Grundlage einer Vielzahl von statistischen Tests, die im Rahmen dieser Dissertation durchgeführt wurden. Dementsprechend ist eine kritische Bewertung der Regression nötig, um ihre Aussagekraft einordnen zu können. Zunächst müssen die folgenden statistischen Annahmen zur Verteilung der Regressionsresiduen gemacht werden (Urban and Mayerl, 2011c):

- a. Die Residuen haben bei jedem Abdomenumfang die gleichen Varianzen.
- b. Der Erwartungswert ist „0“.
- c. Die Residuen sind unabhängig vom Abdomenumfang.
- d. Die Residuen sind unabhängig voneinander verteilt.

Für die richtige Interpretation der Magenumfangswerte mithilfe der oben beschriebenen Regressionsanalyse ist ein normwertiger Abdomenumfang eine wichtige Voraussetzung. Das heißt, die gemessenen Abdomenumfangswerte müssen innerhalb der anerkannten, vorgegebenen Biometriewerte liegen (Papageorghiou et al., 2014). Bei einem zu kleinen Abdomen kann eine Magenumfangsmessung fälschlicherweise in dieser Regressionsanalyse als „unauffällig“ bewertet werden. Umgekehrt gilt, dass diese Regressionsanalyse einen Magenumfang fälschlicherweise als zu klein bewertet, wenn der Abdomenumfang zu groß ist.

Das nachweislich signifikante Bestimmtheitsmaß von 0,4702 der hier durchgeführten Regression (Tabelle 8, F-Statistik) gibt an, dass sich die Veränderungen des Magenumfangs zu 47,02% durch die Veränderungen des Abdomenumfangs erklären lassen. Die Steigung der Regression ist ebenfalls statistisch signifikant (Tabelle 8, T-Statistik) und enthält nicht den Wert „0“ in ihrem 95%-Konfidenzintervall. Dadurch wird das Regressionsmodell validiert

(Pindyck and Rubinfeld, 1981) und die Aussage gerechtfertigt, dass ein statistisch signifikanter Zusammenhang zwischen dem Magen- und dem Abdomenumfang besteht. Das Bestimmtheitsmaß könnte in diesem Fall leicht überschätzt werden, da beide Variablen (Ziel- und Einflussvariable) mit fortschreitender Zeit wachsen. Generell neigt jede Größe, die abhängig von der Zeit wächst, dazu, eine gewisse Erklärbarkeit zu zeigen gegenüber einer anderen Größe, die auch über Zeit wächst (Pindyck and Rubinfeld, 1981).

Nach Validierung des Regressionsmodells besteht nun die Frage, wie verlässlich die Vorhersage des Magenumfangs mit diesem statistischen Modell ist. Dies ist insofern wichtig, als dass die Genauigkeit der 5. Perzentile von der Präzision der Schätzung und den Normalverteilungseigenschaften der Residuen abhängt. Je genauer die Schätzung erfolgt, umso weniger streuen die Residuen um die Schätzgerade, umso kleiner ist ihr Abstand zur 5. Perzentile und umso stärker ist die Diskriminierungsfähigkeit des Modells, die Testgruppe von der Kontrollgruppe zu trennen. Da Mittelwert und Steigung der Kontrollgruppe signifikant unter den Werten der Testgruppe liegen, bewirkt eine niedrige Streuung der Residuen der Kontrollgruppe, dass relativ mehr Messungen der Testgruppe unter die 5. Perzentile fallen, weil der Abstand der 5. Perzentile zur Regressionsgeraden geringer wäre. In der hier vorgestellten Regression schwanken die Residuen standardmäßig um 10 mm um ihren Schätzwert (Tabelle 8, Standardfehler der Schätzung; SEE, Standard Error of Estimate). Diese Schwankungsbreite beträgt 23,23% vom Mittelwert des Magenumfangs der Kontrollgruppe (Quotient aus SEE und Mittelwert des Magenumfangs). Fraglich wäre somit, ob durch die Wahl des Schätzmodells der relative Schätzfehler von 23,23% gesenkt werden könnte. Dies könnte die Sensitivität weiter erhöhen oder es böte sich einen Spielraum, um die Spezifität zu steigern, indem die Risikogruppe nicht mit der 5. Perzentile definiert würde, sondern mit einer niedrigeren.

Die Ursachen für dieses Schwankungsausmaß sind vielseitig. Zum einen gestaltet sich die korrekte Messung des Magenumfangs als schwierig, da diese

stark von der Einstellung der Biometrieebene abhängig ist. Dieses Vorgehen erfordert viel Geschick und Erfahrung des Untersuchers. Zum anderen unterliegt der Magen im Verlauf der Schwangerschaft einem natürlichen Wachstum mit dynamischen Füllungsvolumina, auf Grund des episodisch erfolgenden fetalen Schluckakts (Sase et al., 2005, Goldstein et al., 1987, Nagata et al., 1990). Wachsende Größen zeigen typischerweise eine trichterförmige Punktwolke, die dadurch eine größere Schwankung bei zunehmendem Wachstumsfortschritt bedingt (Sase et al.). Wenn man also den Wachstumseffekt im Schätzmodell berücksichtigen könnte, ließe sich möglicherweise auch der relative Schätzfehler senken. Die episodische Schluckaktivität des Feten ist leider eine Variable, die schwer sein wird zu kontrollieren und die weiterhin einen Einfluss auf die Streubreite der Datenverteilung haben wird. Weiterhin wurde ein Großteil der Messungen im zweiten Trimenon durchgeführt (64,7%, Abbildung 5). Grund hierfür ist, dass die Pränataldiagnostik der Universitätsfrauenklinik Tübingen einen Schwerpunkt auf das zweite Screening setzt. Die Qualität der Regression würde bei gleichmäßiger Verteilung der Messzeitpunkte zunehmen. Alternativ oder zusätzlich könnte man separate Regressionsmodelle für das zweite und dritte Trimenon erstellen, um unterschiedliche Wachstumsmuster zu berücksichtigen (Goldstein et al., 1987, Nagata et al., 1990).

Auf Grund des zu erwartenden Schätzfehlers ist dieses Modell noch nicht reif für die Anwendung in der alltäglichen Praxis. Allerdings kann es statistisch signifikant zeigen, dass sich der Magenumfang mit dem Abdomenumfang schätzen lässt. Für die Anwendung eines Regressionsmodells in der klinischen Praxis sollte der Schätzfehler der Regression so klein wie möglich sein.

4.8.4 Ausreißer in der Subgruppe „TÖF“ des Testkollektivs

In mehreren Tabellen und Abbildungen ist ein besonders deutlicher Ausreißer zu erkennen, der der Subgruppe „TÖF“ des Testkollektivs zuzuordnen ist. Bereits in

Tabelle 2, welche die statistischen Kenngrößen der Rohdaten angibt, wird dieser Ausreißer deutlich. Dass es sich dabei um einen Einzelfall handelt, welcher neun Standardabweichungen oberhalb des Mittelwertes der Kontrollgruppe liegt (Tabelle 5), wird in Abbildung 21 und Abbildung 22 gezeigt.

Extreme Ausreißer sind für statistische Auswertungen oftmals schwer zu handhaben. Zunächst verzerren sie die statistischen Kenngrößen, besonders die Standardabweichung, den Mittelwert und den Variationskoeffizienten. Falls ein solcher Extremwert durch einen Messfehler verursacht worden wäre, würden die Charaktereigenschaften des Datensatzes verzerrt. Eine solche Verzerrung kann die Auswahl und die Aussagekraft der Ergebnisse von weiteren statistischen Tests beeinflussen.

Es gibt mehrere mögliche Ursachen für diesen Extremwert. Die einfachste Erklärung wäre natürlich ein Messfehler in der Datenerhebung oder ein Fehler in der Dokumentation. Leider ist dies im Nachhinein nicht sicher zu erschließen. Durch das Vorliegen einer großen distalen TÖF und eigenständiger Mukusproduktion der Magenschleimhaut wäre eine deutliche Magenfüllung erklärbar (Brantberg et al., 2007, Pretorius et al., 1987, Rao et al., 1978). Dass diese jedoch neun Standardabweichungen über dem Mittelwert der Kontrollgruppe liegt erscheint unwahrscheinlich.

Auf Grund der unmöglichen zweifelsfreien Aufklärung wurde der Extremwert im Datensatz belassen. Als Teil der Testgruppe hatte der Extremwert keinen Einfluss auf die Erstellung der Regressionsanalyse und die Berechnung der 5. Perzentile für die Kontrollgruppe.

4.8.5 Testkollektiv spiegelt nicht die Verteilung der

Grundgesamtheit wieder

Die Stichprobe der erkrankten Feten entspricht nicht der epidemiologischen Verteilung der Grundgesamtheit der Feten mit ÖA. Die Anzahl der Fälle ohne Fistel (35,6%) liegt prozentual in dieser Studie deutlich über der epidemiologischen Häufigkeit (< 15%) (Houben and Curry, 2008, Pinheiro et al., 2012). Ursache ist wahrscheinlich die bessere pränatale Erfassung der Fälle mit Fistel (Garabedian et al., 2015). Im Vergleich dazu liegt der Anteil der Fisteln in der Testgruppe lediglich bei 64,4%. Um die Epidemiologie wiederzugeben, wären Anteile von über 80% nötig (de Jong et al., 2010b, Ioannides et al., 2010, Pedersen et al., 2012, Pinheiro et al., 2012). Allerdings liegen wahrscheinlich weniger verwertbare pränatale Ultraschallbilder von Fällen mit Fistel vor, da diese deutlich schwerer zu erfassen sind (Garabedian et al., 2015). Durch die unterschiedliche Sensitivität und Spezifität zwischen beiden Subgruppen, wäre es erwägenswert, wenn zukünftige Diagnostikstudien diese Subgruppen differenziert betrachten würden.

4.8.6 Reproduzierbare Ergebnisse durch objektivierte Messung

Die Ergebnisse der objektiven Magenumfangsmessungen konnten intra- sowie interpersonell reproduziert werden. Überprüft wurde diese Aussage mithilfe von Bland-Altman Plots und Regressionsanalysen. Die zu diesem Zweck durchgeführten Tests attestieren jedoch, dass der eine Untersucher den Magenumfang im selben Ultraschallbild tendenziell etwas größer gemessen hat als der andere. Relativierend muss betont werden, dass der Gradient des statistischen Trends lediglich signifikant von 0 abweicht. Das Ausmaß der Abweichung ist allerdings sehr gering, erkennbar am geringen Korrelationskoeffizienten ($r = 0,26$) und der geringen Steigung (0,065). Verdeutlicht wird dies in Abbildung 15 und Abbildung 16, in denen optisch nur sehr schwer ein Unterschied in den Messungen abzulesen ist. Letztendlich wird

dieser Unterschied als so gering bewertet, dass er keinen relevanten Einfluss auf die Reproduzierbarkeit der Ergebnisse hat.

Generell kann man sagen, dass dieser Messunterschied nur dann zu einer Konsequenz führt, wenn sich dieser ausschließlich auf die Magenumfangsmessung beziehen würde. In einem solchen Fall wäre ein Kalibrierungsprozess nötig, da sich die Regressionsanalysen/-geraden (MU als Funktion von AU) zwischen den Untersuchern unterscheiden würden. Besonders deutlich würde dieser Unterschied am Regressionskoeffizienten „a“ (Schnittpunkt der y-Achse). Falls der eine Untersucher den Magen- und den Abdomenumfang gleichermaßen etwas größer vermisst, so hätte dies keinen Einfluss auf die Regressionsgerade und die Studienergebnisse wären dieselben. Da der Abdomenumfang im Rahmen dieser Studie nur durch einen Untersucher gemessen wurde, ist diese Fragestellung nicht abschließend zu klären.

Bezüglich der Beurteilung der sonographischen Magengröße sind reproduzierbare Ergebnisse in der Literatur selten und dementsprechend nicht vergleichbar. Dies kann allein mit der Subjektivität der Beurteilung in den meisten Studien begründet werden (Bradshaw et al., 2016, Brantberg et al., 2007, Holland and Fitzgerald, 2010). Subjektive und objektive Ergebnisse würden sich nur dann decken, wenn kein Magen erkennbar ist. Das gilt natürlich in erster Linie für die Fälle ohne Fistel, die auch in dieser Arbeit objektiv wie subjektiv alle erkannt wurden (Bradshaw et al., 2016, Stringer et al., 1995). Sobald es jedoch zu einer Magenfüllung kommt, ist die Qualität der subjektiven Einschätzung rein von der Expertise des Untersuchers abhängig. Damit wären alle Studienergebnisse zum Kernproblem der ÖA (Fälle mit Fistel) nicht vergleichbar. Um also Fortschritte in der sonographischen pränatalen Diagnostik der ÖA zu machen, erscheint die objektive Beurteilung der Magengröße sehr sinnvoll, damit sich die Studienergebnisse vergleichen lassen.

4.9 Fazit für die Praxis

Zunächst konnten alle objektiv erhobenen Ergebnisse dieser Arbeit reproduziert werden. Mit der Inferenzstatistik konnte schließlich statistisch signifikant gezeigt werden, dass der Magenumfang in Abhängigkeit vom Abdomenumfang einen Indikator für eine fetale ÖA darstellt. Mit der validierten Regressionsanalyse ist es gelungen, 100% der ÖA ohne TÖF und 34,2% der ÖA mit TÖF zu erkennen. Um diese Ergebnisse zu erzielen, musste mit einer 5% falsch-positiv Rate gearbeitet werden. Somit bleibt das bekannte Problem der erheblichen falsch-positiv Zahlen.

Basierend auf den hohen falsch-positiv Zahlen sollte diese Untersuchung nur in einer vorselektierten Patientengruppe durchgeführt werden. Eine solche Vorselektion könnte die bekannte Befundkonstellation aus nicht-darstellbarem/kleinem Magen und/oder Polyhydramnion sein oder das Vorkommen von Fehlbildungen, die mit einer ÖA assoziiert sind. Wenn bei einem bestehenden Verdacht diese Testmetrik im Anschluss auffällige Ergebnisse liefert, kann auf der Grundlage eine umfangreiche weiterführende Diagnostik des Ösophagus durchgeführt werden.

Die vorgelegte Dissertationsschrift gibt Auskunft über die absolute Anzahl der Fälle der Testgruppe, die in die hiesige Risikogruppe fallen würde. Es wurde allerdings noch kein statistisches Modell erarbeitet, um eine Prognose zu erstellen, die eine Detektionsrate für die Grundgesamtheit der Testgruppe, insbesondere für die TÖF-Gruppe, vorhersagt. Hierfür könnte eine neue retrospektive statistische Studie angelegt werden, die sich auf die Fälle mit tracheoösophagealer Fistel konzentriert. Eine solche Studie könnte die Grundlage einer prospektiven Studie bilden, die die erarbeitete Prognose überprüft.

5 Referenzen

ABELE, H., STARZ, S., HOOPMANN, M., YAZDI, B., RALL, K. & KAGAN, K. O. 2012. Idiopathic polyhydramnios and postnatal abnormalities. *Fetal Diagn Ther*, 32, 251-5.

ACHIRON, R., GINDES, L., ZALEL, Y., LIPITZ, S. & WEISZ, B. 2008. Three- and four-dimensional ultrasound: new methods for evaluating fetal thoracic anomalies. *Ultrasound Obstet Gynecol*, 32, 36-43.

AMERICAN COLLEGE OF OBSTETRICIANS AND GYNECOLOGISTS 2013. ACOG Practice bulletin no. 134: fetal growth restriction. *Obstet Gynecol*, 121, 1122-33.

AWMF ONLINE. 2012. *Leitlinie der Deutschen Gesellschaft für Kinderchirurgie: Kurzstreckige Ösophagusatresie* [Online]. Available: http://www.awmf.org/uploads/tx_szleitlinien/006-045l_S2k_%C3%96sophagusatresie_kurzstreckig_2012-02_verlaengert.pdf [Accessed 2016].

BLAAS, H.-G. K., HASENÖHRL, G. & STAUDACH, A. 2013. *Ultraschalldiagnostik in Geburtshilfe und Gynäkologie*, Springer-Verlag.

BORSELLINO, A., ZACCARA, A., NAHOM, A., TRUCCHI, A., AITE, L., GIORLANDINO, C. & BAGOLAN, P. 2006. False-positive rate in prenatal diagnosis of surgical anomalies. *J Pediatr Surg*, 41, 826-9.

BRADSHAW, C. J., THAKKAR, H., KNUTZEN, L., MARSH, R., PACILLI, M., IMPEY, L. & LAKHOO, K. 2016. Accuracy of prenatal detection of tracheoesophageal fistula and oesophageal atresia. *J Pediatr Surg*, 51, 1268-72.

BRANTBERG, A., BLAAS, H. G., HAUGEN, S. E. & EIK-NES, S. H. 2007. Esophageal obstruction-prenatal detection rate and outcome. *Ultrasound Obstet Gynecol*, 30, 180-7.

BRUMFIELD, C. G., DAVIS, R. O., OWEN, J., WENSTROM, K. & KYNERD, P. M. 1998. Pregnancy outcomes following sonographic nonvisualization of the fetal stomach. *Obstet Gynecol*, 91, 905-8.

CHOUDHRY, M., BOYD, P. A., CHAMBERLAIN, P. F. & LAKHOO, K. 2007. Prenatal diagnosis of tracheo-oesophageal fistula and oesophageal atresia. *Prenat Diagn*, 27, 608-10.

CRISERA, C. A., GRAU, J. B., MALDONADO, T. S., KADISON, A. S., LONGAKER, M. T. & GITTES, G. K. 2000. Defective epithelial-mesenchymal

interactions dictate the organogenesis of tracheoesophageal fistula. *Pediatr Surg Int*, 16, 256-61.

CZERKIEWICZ, I., DREUX, S., BECKMEZIAN, A., BENACHI, A., SALOMON, L. J., SCHMITZ, T., BONNARD, A., KHEN-DUNLOP, N. & MULLER, F. 2011. Biochemical amniotic fluid pattern for prenatal diagnosis of esophageal atresia. *Pediatr Res*, 70, 199-202.

DASHE, J. S., MCINTIRE, D. D., RAMUS, R. M., SANTOS-RAMOS, R. & TWICKLER, D. M. 2002. Hydramnios: anomaly prevalence and sonographic detection. *Obstet Gynecol*, 100, 134-9.

DE JONG, E., DE HAAN, M., GISCHLER, S., HOP, W., COHEN-OVERBEEK, T., BAX, N., DE KLEIN, A., TIBBOEL, D. & GRIJSEELS, E. 2010a. Pre- and postnatal diagnosis and outcome of fetuses and neonates with esophageal atresia and tracheoesophageal fistula. *Prenatal diagnosis*, 30, 274-279.

DE JONG, E. M., FELIX, J. F., DE KLEIN, A. & TIBBOEL, D. 2010b. Etiology of esophageal atresia and tracheoesophageal fistula: "mind the gap". *Curr Gastroenterol Rep*, 12, 215-22.

DEPAEPE, A., DOLK, H. & LECHAT, M. F. 1993. The epidemiology of tracheoesophageal fistula and oesophageal atresia in Europe. EUROCAT Working Group. *Arch Dis Child*, 68, 743-8.

DEVELAY-MORICE, J., RATHAT, G., DUYME, M., HOFFET, M., FREDOUILLE, C., COUTURE, A., ALLAL, H., DESCHAMPS, F., FRANDJI-BARBIER, N. & MARÈS, P. 2007. [Ultrasonography of fetal esophagus: healthy appearance and prenatal diagnosis of a case of esophagus atresia with esotracheal fistula]. *Gynecologie, obstetrique & fertilité*, 35, 249-257.

DOLE, N., SAVITZ, D. A., HERTZ-PICCIOTTO, I., SIEGA-RIZ, A. M., MCMAHON, M. J. & BUEKENS, P. 2003. Maternal stress and preterm birth. *Am J Epidemiol*, 157, 14-24.

ETHUN, C. G., FALLON, S. C., CASSADY, C. I., MEHOLLIN-RAY, A. R., OLUTOYE, O. O., ZAMORA, I. J., LEE, T. C., WELTY, S. E. & CASS, D. L. 2014. Fetal MRI improves diagnostic accuracy in patients referred to a fetal center for suspected esophageal atresia. *J Pediatr Surg*, 49, 712-5.

EYHEREMENDY, E. & PFISTER, M. 1983. Antenatal real-time diagnosis of esophageal atresias. *J Clin Ultrasound*, 11, 395-7.

FAHRMEIR, L., KÜNSTLER, R., PIGEOT, I. & TUTZ, G. 2011a. *Statistik: Der Weg zur Datenanalyse*, Springer Berlin Heidelberg.

FAHRMEIR, L., PIGEOT, I., KÜNSTLER, R. & TUTZ, G. 2011b. *Statistik. Der Weg zur Datenanalyse*, Berlin: Springer (Springer-Lehrbuch).

FAUSETT, S. R. & KLINGENSMITH, J. 2012. Compartmentalization of the foregut tube: developmental origins of the trachea and esophagus. *Wiley Interdiscip Rev Dev Biol*, 1, 184-202.

FELIX, J. F., TIBBOEL, D. & DE KLEIN, A. 2007. Chromosomal anomalies in the aetiology of oesophageal atresia and tracheo-oesophageal fistula. *Eur J Med Genet*, 50, 163-75.

GARABEDIAN, C., SFEIR, R., LANGLOIS, C., BONNARD, A., KHEN-DUNLOP, N., GELAS, T., MICHAUD, L., AUBER, F., GOTTRAND, F. & HOUFFLIN-DEBARGE, V. 2015. Does prenatal diagnosis modify neonatal treatment and early outcome of children with esophageal atresia? *Am J Obstet Gynecol*, 212, 340.e1-7.

GARABEDIAN, C., VERPILLAT, P., CZERKIEWICZ, I., LANGLOIS, C., MULLER, F., AVNI, F., BIGOT, J., SFEIR, R., VAAST, P., COULON, C., SUBTIL, D. & HOUFFLIN-DEBARGE, V. 2014. Does a combination of ultrasound, MRI, and biochemical amniotic fluid analysis improve prenatal diagnosis of esophageal atresia? *Prenat Diagn*.

GOLDSTEIN, I., REECE, E. A., YARKONI, S., WAN, M., GREEN, J. & HOBBS, J. C. 1987. Growth of the fetal stomach in normal pregnancies. *Obstetrics and gynecology*, 70, 641-644.

GOTO, S., IHARA, Y., URATA, Y., IZUMI, S., ABE, K., KOJI, T. & KONDO, T. 2001. Doxorubicin-induced DNA intercalation and scavenging by nuclear glutathione S-transferase pi. *FASEB J*, 15, 2702-14.

GOYAL, A., JONES, M. O., COURIEL, J. M. & LOSTY, P. D. 2006. Oesophageal atresia and tracheo-oesophageal fistula. *Arch Dis Child Fetal Neonatal Ed*, 91, F381-4.

HILL, L. M., BRECKLE, R., THOMAS, M. L. & FRIES, J. K. 1987. Polyhydramnios: ultrasonically detected prevalence and neonatal outcome. *Obstet Gynecol*, 69, 21-5.

HOLLAND, A. J. & FITZGERALD, D. A. 2010. Oesophageal atresia and tracheo-oesophageal fistula: current management strategies and complications. *Paediatr Respir Rev*, 11, 100-6; quiz 106-7.

HOOPMANN, M., KAGAN, K. O., BORGMEIER, F., SEITZ, G., ARAND, J. & WAGNER, P. 2015. Measurement of Gastric Circumference in Foetuses with Oesophageal Atresia. *Geburtshilfe Frauenheilkd*, 75, 1148-1152.

HOUBEN, C. H. & CURRY, J. I. 2008. Current status of prenatal diagnosis, operative management and outcome of esophageal atresia/tracheo-esophageal fistula. *Prenat Diagn*, 28, 667-75.

IOANNIDES, A. S. & COPP, A. J. 2009. Embryology of oesophageal atresia. *Semin Pediatr Surg*, 18, 2-11.

IOANNIDES, A. S., HENDERSON, D. J., SPITZ, L. & COPP, A. J. 2003. Role of Sonic hedgehog in the development of the trachea and oesophagus. *J Pediatr Surg*, 38, 29-36; discussion 29-36.

IOANNIDES, A. S., MASSA, V., FERRARO, E., CECCONI, F., SPITZ, L., HENDERSON, D. J. & COPP, A. J. 2010. Foregut separation and tracheo-oesophageal malformations: The role of tracheal outgrowth, dorso-ventral patterning and programmed cell death. *Developmental Biology*, 337, 351-362.

KÄHLER, C., GEMBRUCH, U., HELING, K.-S., HENRICH, W. & SCHRAMM, T. 2013. Empfehlungen der DEGUM zur Durchführung von Amniozentese und Chorionzottenbiopsie. *Ultraschall in der Medizin-European Journal of Ultrasound*, 34, 435-440.

KLUTH, D. & FIEGEL, H. 2003. The embryology of the foregut. *Semin Pediatr Surg*, 12, 3-9.

KUNISAKI, S. M., BRUCH, S. W., HIRSCHL, R. B., MYCHALISKA, G. B., TREADWELL, M. C. & CORAN, A. G. 2014. The diagnosis of fetal esophageal atresia and its implications on perinatal outcome. *Pediatr Surg Int*.

LADD, W. E. & SWENSON, O. 1947. Esophageal Atresia and Tracheo-esophageal Fistula. *Ann Surg*, 125, 23-40.

LANGER, J. C., HUSSAIN, H., KHAN, A., MINKES, R. K., GRAY, D., SIEGEL, M. & RYAN, G. 2001. Prenatal diagnosis of esophageal atresia using sonography and magnetic resonance imaging. *J Pediatr Surg*, 36, 804-7.

LEE, K. D., OKAZAKI, T., KATO, Y., LANE, G. J. & YAMATAKA, A. 2008. Esophageal atresia and tracheo-esophageal fistula associated with coarctation of the aorta, CHARGE association, and DiGeorge syndrome: a case report and literature review. *Pediatr Surg Int*, 24, 1153-6.

LOCKWOOD, C. J. 1999. Stress-associated preterm delivery: the role of corticotropin-releasing hormone. *Am J Obstet Gynecol*, 180, S264-6.

MALINGER, G., LEVINE, A. & ROTMENSCH, S. 2004. The fetal esophagus: anatomical and physiological ultrasonographic characterization using a high-resolution linear transducer. *Ultrasound Obstet Gynecol*, 24, 500-5.

MCKENNA, K. M., GOLDSTEIN, R. B. & STRINGER, M. D. 1995. Small or absent fetal stomach: prognostic significance. *Radiology*, 197, 729-33.

NAGATA, S., KOYANAGI, T., HORIMOTO, N., SATOH, S. & NAKANO, H. 1990. Chronological development of the fetal stomach assessed using real-time ultrasound. *Early human development*, 22, 15-22.

PAPAGEORGHIOU, A. T., OHUMA, E. O., ALTMAN, D. G., TODROS, T., CHEIKH ISMAIL, L., LAMBERT, A., JAFFER, Y. A., BERTINO, E., GRAVETT, M. G., PURWAR, M., NOBLE, J. A., PANG, R., VICTORA, C. G., BARROS, F. C., CARVALHO, M., SALOMON, L. J., BHUTTA, Z. A., KENNEDY, S. H. & VILLAR, J. 2014. International standards for fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. *Lancet*, 384, 869-79.

PEDERSEN, R. N., CALZOLARI, E., HUSBY, S. & GARNE, E. 2012. Oesophageal atresia: prevalence, prenatal diagnosis and associated anomalies in 23 European regions. *Arch Dis Child*, 97, 227-32.

PINDYCK, R. S. & RUBINFELD, D. L. 1981. Analysis of Variance and Correlation. In: LIEBERMAN, B. E. & NORTON, M. S. (eds.) *Econometric Models and Economic Forecasts*. McGraw-Hill International Editions.

PINHEIRO, P. F., SIMOES E SILVA, A. C. & PEREIRA, R. M. 2012. Current knowledge on esophageal atresia. *World J Gastroenterol*, 18, 3662-72.

PRETORIUS, D. H., DROSE, J. A., DENNIS, M. A., MANCHESTER, D. K. & MANCO-JOHNSON, M. L. 1987. Tracheoesophageal fistula in utero. Twenty-two cases. *J Ultrasound Med*, 6, 509-13.

QI, B., DIEZ-PARDO, J. A., NAVARRO, C. & TOVAR, J. A. 1996. Narrowing the embryologic window of the adriamycin-induced fetal rat model of esophageal atresia and tracheoesophageal fistula. *Pediatr Surg Int*, 11, 444-7.

RAO, S. B., SLOVIS, T. L., CRADOCK, T. V. & SHAH, T. 1978. Visualization of the intestinal tract by amniography in a fetus with esophageal atresia. *Pediatr Radiol*, 7, 241-2.

REDDY, U. M., ABUHAMAD, A. Z., LEVINE, D. & SAADE, G. R. 2014. Fetal imaging: executive summary of a joint Eunice Kennedy Shriver National Institute of Child Health and Human Development, Society for Maternal-Fetal Medicine, American Institute of Ultrasound in Medicine, American College of Obstetricians and Gynecologists, American College of Radiology, Society for Pediatric Radiology, and Society of Radiologists in Ultrasound Fetal Imaging workshop. *Obstet Gynecol*, 123, 1070-82.

ROBERTS, W., ZURADA, A., ZURADA-ZIELI, N. A., GIELECKI, J. & LOUKAS, M. 2016. Anatomy of trisomy 18. *Clin Anat*, 29, 628-32.

SAKIYAMA, J., YAMAGISHI, A. & KUROIWA, A. 2003. Tbx4-Fgf10 system controls lung bud formation during chicken embryonic development. *Development*, 130, 1225-34.

SASAKI, T., KUSAFUKA, T. & OKADA, A. 2001. Analysis of the development of normal foregut and tracheoesophageal fistula in an adriamycin rat model using three-dimensional image reconstruction. *Surg Today*, 31, 133-9.

SASE, M., ASADA, H., OKUDA, M. & KATO, H. 2002. Fetal gastric size in normal and abnormal pregnancies. *Ultrasound Obstet Gynecol*, 19, 467-70.

SASE, M., MIWA, I., SUMIE, M., NAKATA, M., SUGINO, N., OKADA, K., OSA, A., MIKE, H. & ROSS, M. G. 2005. Gastric emptying cycles in the human fetus. *Am J Obstet Gynecol*, 193, 1000-4.

SCHNEIDER, A., HOMMEL, G. & BLETTNER, M. 2010. Linear Regression Analysis. *Dtsch Arztebl Int*, 107, 776-782.

SHULMAN, A., MAZKERETH, R., ZALEL, Y., KUINT, J., LIPITZ, S., AVIGAD, I. & ACHIRON, R. 2002. Prenatal identification of esophageal atresia: the role of ultrasonography for evaluation of functional anatomy. *Prenat Diagn*, 22, 669-74.

SOLOMON, B. D. 2011. VACTERL/VATER Association. *Orphanet J Rare Dis*, 6, 56.

SOLT, I., ROTMENSCH, S. & BRONSHTEIN, M. 2010. The esophageal 'pouch sign': a benign transient finding. *Prenat Diagn*, 30, 845-8.

SPITZ, L. 2007. Oesophageal atresia. *Orphanet J Rare Dis*, 2, 24.

STATISTISCHES BUNDESAMT. 2016. *Geburten* [Online]. Available: <https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Bevoelkerung/Geburten/Geburten.html> [Accessed 2016].

STOLL, C. G., ALEMBIK, Y. & DOTT, B. 1991. Study of 156 cases of polyhydramnios and congenital malformations in a series of 118,265 consecutive births. *Am J Obstet Gynecol*, 165, 586-90.

STRINGER, M. D., MCKENNA, K. M., GOLDSTEIN, R. B., FILLY, R. A., ADZICK, N. S. & HARRISON, M. R. 1995. Prenatal diagnosis of esophageal atresia. *J Pediatr Surg*, 30, 1258-63.

URBAN, D. & MAYERL, J. 2011a. Das Grundmodell der Regressionsanalyse: Regression im bivariaten Modell. *Regressionsanalyse: Theorie, Technik und Anwendung*. Wiesbaden: VS Verlag für Sozialwissenschaften.

URBAN, D. & MAYERL, J. 2011b. Entdeckung und Beseitigung von Modellverstößen. *Regressionsanalyse: Theorie, Technik und Anwendung*. Springer.

URBAN, D. & MAYERL, J. 2011c. Inferenzstatistik in der Regressionsanalyse. *Regressionsanalyse: Theorie, Technik und Anwendung*. Wiesbaden: VS Verlag für Sozialwissenschaften.

URBAN, D. & MAYERL, J. 2011d. Testverfahren im bivariaten Regressionsmodell. *Regressionsanalyse: Theorie, Technik und Anwendung*. Wiesbaden: VS Verlag für Sozialwissenschaften.

WONG, M. T., SCHOLVINCK, E. H., LAMBECK, A. J. & VAN RAVENSWAAIJ-ARTS, C. M. 2015. CHARGE syndrome: a review of the immunological aspects. *Eur J Hum Genet*, 23, 1451-9.

WYLLIE, J. P., WRIGHT, M. J., BURN, J. & HUNTER, S. 1994. Natural history of trisomy 13. *Arch Dis Child*, 71, 343-5.

YAGEL, S., SONIGO, P., ROUSSEAU, V., SARNACKI, S., COHEN, S. & BENACHI, A. 2005. Esophageal atresia diagnosed with three-dimensional ultrasonography. *Ultrasound Obstet Gynecol*, 26, 307-8.

6 Anhang

Erklärung zum Eigenanteil der Dissertationsschrift:

Die Doktorarbeit wurde an der Universitätsfrauenklinik Tübingen unter der Betreuung von Herrn Prof. Dr. med. Markus Hoopmann durchgeführt. Die Datenerhebung erfolgte zusammen mit Herrn Patrick Zanker, zu der Zeit Student der Humanmedizin im klinischen Abschnitt. Die statistische Auswertung erfolgte eigenständig durch mich, nach Beratung mit Herrn Prof. Dr. med. Karl Oliver Kagan.

Ich versichere, dass ich die beigefügte Dissertation selbständig verfasst habe und keine weiteren als die von mir angegebenen Quellen oder Hilfsmittel verwendet habe. Zudem versichere ich, dass die beigefügte Dissertation in keinem anderen Promotionsverfahren eingereicht wurde und kein endgültig gescheitertes Promotionsverfahren diesem vorausging.

Düsseldorf, den

Felix Borgmeier