

Aus dem Institut für Medizinische Informationsverarbeitung Biometrie und
Epidemiologie der Ludwig-Maximilians-Universität München

Vorstand: Prof. Dr. rer. nat. Ulrich Mansmann

Kausale Strategien in Beobachtungsstudien

Dissertation
zum Erwerb des Doktorgrades der Humanbiologie
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

Vorgelegt von
Hannes Buchner
aus München

Jahr
2017

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatter: Prof. Dr. rer. nat. Ulrich Mansmann

Mitberichterstatter: Prof. Dr. Jörg Schelling
Mitbetreuung durch den
Promovierten Mitarbeiter: PD Dr. Eva Annette Rehfuss
Dekan: Prof. Dr. med. dent. Reinhard HICKEL
Tag der mündlichen Prüfung: 29.03.2017

Eidesstattliche Versicherung

Ich, Hannes Buchner, erkläre hiermit an Eides statt,
dass ich die vorliegende Dissertation mit dem Thema
Kausale Strategien in Beobachtungsstudien

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München den,

Inhaltsverzeichnis

Inhaltsverzeichnis	4
1. Einleitung	5
1.1 Kurzdarstellung der Veröffentlichungen	9
1.1.1 Cooking and season as risk factors for acute lower respiratory infections in African children: a cross-sectional multi-country analysis	9
1.1.2 The vitamin E-binding protein afamin increases in maternal serum during pregnancy	11
1.2 Zusammenfassung.....	12
1.3 Summary	13
1.4 Literaturverzeichnis	15
2. Cooking and season as risk factors for acute lower respiratory infections in African children: a cross-sectional multi-country analysis.....	19
3. The vitamin E-binding protein afamin increases in maternal serum during pregnancy.....	20
Danksagung	22

1. Einleitung

Eine retrospektive Datenanalyse von Beobachtungsstudien sieht sich immer der Kritik ausgesetzt, eine statistische Optimierung der Ergebnisse vorgenommen zu haben [1]. Selbst in Ermangelung handfester Motive reicht allein die Möglichkeit der leichteren Publikation eindrucksvoller Ergebnisse aus (Stichwort Publikationsverzerrung [2]), um Auswertungen von Beobachtungsstudien zu kritisieren, da die hohen Qualitätsstandards prospektiv definierter klinischer Studienauswertungen meist nicht umsetzbar sind (bzgl. Statistikstandards der FDA vgl. ICH E9 [3]). Nicht zuletzt deshalb ist eine Analysestrategie für Beobachtungsstudien von immenser Bedeutung für die Qualitätssicherung und ist als solche zentraler Teil jeder empirischen Wissenschaft. Dies kann nicht ernst genug genommen werden, da eben auch die Glaubwürdigkeit empirischer Forschung in der öffentlichen Wahrnehmung davon abhängt.

In der Pharmabranche hat sich im Zusammenspiel mit dem Gesetzgeber in den letzten Dekaden ein hohes wissenschaftliches Niveau etabliert, speziell bzgl. vorspezifizierter Analysestrategien, aber auch Initiativen der FDA (clinicaltrials.gov) und der EU (EudraCT.ema.europa.eu) gegen Publikationsverzerrung sind erfolgreich auf den Weg gebracht. Beobachtungsstudien hingegen haben speziell durch die Anwendung kausaler Methoden ([4], [5]) und sogenannter „prospektiv-retrospektiver“ [6] Designs noch Verbesserungsspielräume. Während sich also in der Medikamentenentwicklung hohe Standards durchgesetzt haben, die manche sogar für Überregulierung halten [7], gilt es die Verbesserungsspielräume für Beobachtungsstudien noch auszuschöpfen.

Gerade die Nicht-Reproduzierbarkeit vieler Ergebnisse genetischer Analysen der letzten Jahrzehnte [8] sollte mahnendes Beispiel genug sein, auch wenn hier zusätzlich die Problematiken hochdimensionaler Daten auftreten. Allerdings kann auch die Kodierung von Variablen oder/und deren Kombination zu Scores durch die sich ergebenden unzähligen möglichen Varianten, Kombinationen und Verbindungsoptionen bereits als Problem hochdimensionaler Daten gelten [9].

Den schwarzen Peter aber nur den hochdimensionalen Daten zuzuschreiben greift sicherlich zu kurz, da es auch genug Beispiele aus klassischen Beobachtungsstudien gibt, deren Ergebnisse in randomisierten Studien nicht bestätigt werden konnten [10] und Nicht-Reproduzierbarkeit ist auch kein alleiniges Problem von

Beobachtungsstudien. So ist seit einigen Jahren in der Grundlagenforschung sogar von einer Reproduzierungskrise die Rede [11]. Genauso wenig kann aber allein eine solide statistische Analysestrategie alle Probleme lösen die mit Reproduzierbarkeit verbunden sind, auch wenn sicherlich in vielen Fällen schlicht irreführende statistische Methoden angewendet wurden (siehe z. B. [12]). Dies wird einem besonders klar, führt man sich vor Augen, wie erschreckend niedrig Reproduzierungsquoten in rigoros angelegten Untersuchungen sind (6 aus 53 in [13] und 14 aus 67 in [14]). Auch Fälschungen bzw. Fehler mögen einzelne Nichtreproduzierbarkeiten erklären [15], der Großteil dürfte aber sicherlich auf das Konto statistischer Trugschlüsse und allgemeiner Messungenauigkeiten der zugrundeliegenden Evaluierungstechniken gehen (wie z. B. Labortechnik [16] oder Medizingerätetechnik; siehe hierzu diverse FDA guidelines [17]).

Eine Vorspezifizierung der Analysestrategie für Beobachtungsstudien, wie der Prüfplan bzw. der Statistische Analyseplan in klinischen Studien [18], nach der Datenerhebung aber vor der Datenanalyse bzw. dem Zugang zu den Daten, wird für Biomarkerstudien in letzter Zeit unter dem Titel prospektiv-retrospektive Designs intensiv diskutiert [6]. Eine solche Vorspezifizierung bei einer unabhängigen dritten Stelle ist natürlich keine neu Idee [19], könnte aber auch auf viele andere Beobachtungsstudien angewendet werden, bei denen die Daten der jeweiligen Forschungsgruppe zu einem klar definierten Zeitpunkt zugänglich gemacht werden. So wird man beispielsweise in der 2014 gestarteten Offenlegungsoffensive der Pharmabranche seine Analysestrategie bestimmen müssen, bevor man Zugang zu den Daten erhält [20]. Es bleibt aber abzuwarten, inwieweit Publikationen, die durch diesen Zugang entstehen, der ursprünglich eingereichten Analysestrategie folgen. Die bisher sehr überschaubare Anzahl an Datenzugängen bzw. Forschungsanträgen kann sicher zum allergrößten Teil mangelnden Ressourcen zugeschrieben werden, mag aber auch mit dem Dilemma des Glaubwürdigkeitskonflikts zusammenhängen. Einerseits besteht Veröffentlichungszwang und andererseits bleibt die Schwierigkeit der Vorspezifizierung während deren Kontrolle in diesem Fall einfach möglich ist. Dies ist rein spekulativ, aber hätte sich sonst nicht schon längst eine unabhängige Institution zur Dokumentation von Vorspezifizierungen für Beobachtungsstudien entwickelt und durchgesetzt?

Ziel der vorliegenden Arbeit ist es, Analysestrategien für Beobachtungsstudien zu entwickeln, mit speziellem Fokus auf Modellbildung, Störfaktoren (Confounder) und kausaler Statistik. Hierzu wurden beispielhaft zwei Studien ausgewählt, zu deren Analyse komplexe statistische Methoden notwendig sind, um die zugrunde liegenden medizinischen bzw. epidemiologischen Fragestellungen zu beantworten. Anhand dieser Beobachtungsstudien werden Analysestrategien vorgestellt, die im ganzen oder in Teilen auf viele medizinische und epidemiologische Fragestellungen anwendbar sind.

Modellbildung

Zu den Schwierigkeiten der Modellbildung gehört zuallererst die Wahl der Modellklasse, welche meistens durch die Art der erhobenen Daten und inhaltliche Überlegungen bestimmt ist. Die meisten der mit dem statistischen Modell verbundenen Fragen können prospektiv beantwortet werden. Eine Vorspezifizierung anhängig von Modellannahmen, die erst anhand der tatsächlich erhobenen Daten, somit also retrospektiv, verifiziert werden können (vergleiche [21] für logistische Regressionen und [22] für gemischte Modelle) ist natürlich schwieriger, aber in den meisten Fällen auch möglich. Zumindest aber können in den allermeisten Fällen Sensitivitätsanalysen vorspezifiziert werden die kritische Modellannahmen aufweichen oder verändern. Eine weitere grundlegende Schwierigkeit ist die inhaltliche (Vor-)Auswahl von potenziell erklärenden Variablen sowie deren systematische Selektion anhand vordefinierter statistischer Kriterien. Ein weiterer für die Modellbildung wichtiger Punkt, der zwar nicht nur in Beobachtungsstudien, aber dort eben typischerweise vermehrt auftritt, sind fehlende Werte, wobei sich speziell breit angelegte epidemiologische Beobachtungsstudien oft mit einer großen Anzahl fehlender Werte auseinandersetzen müssen [23].

Störfaktoren (Confounder) und kausale Statistik

Randomisierte Studien sind seit langem und zu Recht der Goldstandard zur Untersuchung kausaler Zusammenhänge und werden daher wann immer möglich in der Medikamentenentwicklung eingesetzt [3]. Sobald aber eine geeignete Randomisierung nicht mehr möglich ist, wie in den meisten Beobachtungsstudien,

benötigt man Methoden der kausalen Statistik sofern man sich nicht auf rein assoziierende Aussagen beschränken möchte. Natürlich kann es auch in randomisierten Studien nötig werden kausale Methoden anzuwenden. Dies kann beispielsweise in Subgruppen oder auch bzgl. Variablen nach denen nicht randomisiert wurde nötig werden. Zudem kann es natürlich immer spezielle Fragestellungen wie beispielsweise „Crossover“-Problematiken in onkologischen Studien geben, die die Anwendung kausaler Methoden nötig machen (compare [24] and [25]).

Eine wichtige Säule der kausalen Statistik ist, dass so genannte „Counterfactual thinking“ was die Möglichkeit sich Alternativen zur Realität vorzustellen bezeichnet. „Counterfactual thinking“ wurde als Philosophie in den 70er Jahren des letzten Jahrhunderts entwickelt und wurde 1974 erstmals von Rubin in der Statistik verwendet [26] und führte 1983 unter anderem zu der Entwicklung der „propensity scores“ von Rosenbaum und Rubin [27], der „potential outcome theory“ (Rubin 2004 [28]), „Inverse Probability of Treatment Weights“ (IPTW) [29] und nicht zuletzt auch der „Inverse Probability of Censoring Weights“ (IPCW) Methode [30]. Damit einhergehend machten die Verbesserung der mathematischen Sprache für kausale Modelle, insbesondere die Formalisierung von Zusammenhängen zwischen kausalen Graphen, „marginal structural models“, und der „potential outcome theory“ (vergleiche [31]) die kausale Inferenz und Graphen zu einem mächtigen Kommunikations- und Visualisierungs-Tool.

Während Pearl [32] noch zu Recht kritisiert, dass kausale Methoden noch nicht genug Aufmerksamkeit in medizinischen und epidemiologischen Anwendungen erfahren haben fanden in den letzten Jahren in der Epidemiologie und selbst in der Planung klinischer Studien immer mehr Methoden der kausalen Statistik, wie z. B. kausale Diagramme, Anwendung [33]. Kausale Diagramme ermöglichen eine systematische Evaluierung möglicher Einflußfaktoren in „graphischer Sprache“, welche von Medizinern und Statistikern gleichermaßen intuitiv verstanden werden kann [5]. Während das Design randomisierter Studien nach den wichtigsten Einflussfaktoren stratifizieren kann, bleibt der Beobachtungsstudie nichts anderes übrig, als diese zumindest zu „kontrollieren“. Der „Zeit“ als Störfaktor kommt hierbei eine zentrale Rolle zu, da abhängig von ihrer Rolle und Wichtigkeit ein longitudinales Design durchgeführt werden muss oder eine Querschnittstudie die sinnvollere Wahl

ist [34]. Auch eine Kombination beider kann sinnvoll sein [Kapitel 3], um einerseits den „Störfaktor“ Zeit und andererseits die örtliche Variabilität zu untersuchen, wobei immer die bereits existierende Evidenz aus vorhergehenden Studien mitberücksichtigt werden muss.

Während gut randomisierte Studien auf ausgeglichene Kovariablen hoffen können, selbst wenn diese nicht Teil der Randomisierungsstrategie sind [3], müssen Beobachtungsstudien grundsätzlich vom Gegenteil ausgehen und konsequenter Weise diese Daten erheben, um Ausgeglichenheit zu beweisen oder sie gegebenenfalls im finalen statistischen Modell zu berücksichtigen. Dabei ergeben sich durch die Einbindung von öffentlich verfügbaren Informationen, wie beispielsweise Wetterdaten [Kapitel 2] oder die Verbindung verschiedener Datenquellen (Register- und Gendatenbanken), immer größere Chancen für noch präzisere Modelle (siehe z. B. [35]).

1.1 Kurzdarstellung der Veröffentlichungen

Im Folgenden werden jeweils die Inhalte der einzelnen Veröffentlichungen kurz vorgestellt und der Beitrag des Doktoranden zu den verfassten Fachartikeln dargelegt.

1.1.1 Cooking and season as risk factors for acute lower respiratory infections in African children: a cross-sectional multi-country analysis

Hintergrund: Akute Infektionen der unteren Atemwege (ALRI) sind eine der Haupttodesursachen unter afrikanischen Kindern unter fünf Jahren. Ein erheblicher Teil davon entfällt auf Haushaltluftverschmutzung durch feste Brennstoffe.

Methoden: Wir haben die Beziehung zwischen Kochpraktiken und ALRI in gepoolten Datensätzen von Demographic and Health Surveys, die zwischen 2000 und 2011 in afrikanischen Ländern südlich der Sahara durchgeführt wurden, untersucht. Die Auswirkungen des hauptsächlich benutzten Koch-Brennstoffs, die Lage der Kochstelle und deren Belüftung wurden in 18 (n = 56.437), bzw. in 9 (n = 23.139) und 6 Ländern (n = 14.561) untersucht. Dabei verwendeten wir ein kausales Diagramm

und multivariable logistische gemischte Modelle, um den Einfluss der Kovariaten auf individueller, regionaler und nationaler Ebene zu beurteilen.

Ergebnisse: Der hauptsächlich benutzte Koch-Brennstoff hatte einen statistisch signifikanten Einfluss auf das ALRI Risiko ($p < 0,0001$), wobei sich die Jahreszeit als Effekt-Modifizierer ($p = 0,034$) erwies. Während der Regenzeit wurde die Chance an ALRI zu leiden durch Kerosin (OR 1,64; CI: 0,99, 2,71), Kohle und Holzkohle (OR 1,54; CI: 1,21, 1,97), Holz (OR 1,20; CI: 0,95, 1,51) und minderwertigen Brennstoffe aus Biomasse (OR 1,49; CI: 0,93, 2,35) im Vergleich zu sauberen Brennstoffen erhöht. Im Gegensatz dazu waren in der Trockenzeit die entsprechenden Chancen für Kerosin (OR 1,23; CI: 0,77, 1,95), Kohle und Holzkohle (OR 1,35; CI: 1,06, 1,72) und minderwertigen Brennstoffe aus Biomasse (OR 1,07; CI: 0,69, 1,66) reduziert, aber für Holz (OR 1,32; CI: 1,04, 1,66) erhöht. Die Lage der Kochstelle erwies sich auch als jahreszeitabhängige statistisch signifikante ($p = 0,0070$) Einflussgröße für ALRI, insbesondere das Kochen in geschlossenen Räumen ohne separate Küche während der Regenzeit (OR 1,80; CI: 1,30, 2,50). Aufgrund der in Afrika seltenen Verwendung konnten wir jedoch keinen Einfluss von Herdbelüftungen nachweisen.

Schlussfolgerung: Wir fanden unterschiedliche und jahreszeitabhängige Risiken für verschiedene Arten von festen Brennstoffen und Kerosin, sowie der Lage der Kochstelle auf ALRI bei Kindern. Künftige Studien zu Haushaltluftverschmutzung sollten mögliche Wirkungs-Modifikationen von Koch-Brennstoffen durch die Jahreszeit berücksichtigen.

Eigenanteil: Der Doktorand war Erstautor dieser Arbeit und damit hauptverantwortlich für deren Konzeption, die Durchführung und Planung der Datenanalyse und der Formulierung des Manuskripts. Er hat die komplette statistische Auswertung, von der Wahl der Methoden und deren Beschreibung bis hin zur Durchführung der Auswertung übernommen. Er war für die Richtigkeit jeder einzelnen Zahl verantwortlich, aber auch für die globale Auswertungsstrategie. Im speziellen gelang es ihm durch das Einbinden öffentlich zugänglicher Wetterdaten über die Regenzeit in die kausale Auswertungsstrategie einen epidemiologischen Beitrag zu leisten. Diese Idee erwies sich in der Folge als sehr fruchtbar und stellt

letztendlich einen der wertvollsten epidemiologischen Erkenntnisgewinne dieser Arbeit dar.

1.1.2 The vitamin E-binding protein afamin increases in maternal serum during pregnancy

Hintergrund: Afamin ist ein aus der Leber stammendes Plasma-Glykoprotein mit Vitamin E-Bindungseigenschaften und einer mutmaßlichen Funktion auf die Fruchtbarkeit. In dieser Studie wurden Afaminserumkonzentrationen während und nach der Geburt von unkomplizierten Schwangerschaften und ein möglicher Zusammenhang zwischen Afaminserumkonzentrationen und dem Ausgang der Schwangerschaften untersucht.

Methoden: Afaminserumkonzentrationen wurden bei Frauen mit unkomplizierten Schwangerschaften in einer retrospektiven Kohorte (n = 466) in unterschiedlichen Schwangerschaftsstadien und in einer prospektiven Beobachtungsstudie (n = 76) in dem ersten, zweiten und dritten Trimester gemessen. Darüber hinaus wurde die Afaminserumkonzentration im ersten Trimester in einer Querschnittspilotstudie bestimmt, die Frauen mit Präeklampsie (PE), Schwangerschaft induzierte Hypertonie (PIH) und Frauen ohne Komplikationen während der Schwangerschaft (jeweils n = 13) eingeschlossen hat. Schließlich wurde die Expression von Afamin im menschlichen Plazentagewebe durch RT-PCR und Immunhistochemie untersucht.

Ergebnisse: Afaminserumkonzentrationen stiegen linear an und verdoppelten sich fast während der Schwangerschaft in beiden Studien (retrospektiven und der prospektiven Studie) bei Frauen ohne Komplikationen während der Schwangerschaft mit Median Afaminserumkonzentration von 61,9 mg / l, 79,6 mg / l und 98,6 mg / l im ersten, zweiten und dritten Trimester. Nach der Entbindung verringerte sich die mediane Afaminserumkonzentration auf die Ausgangswerte von 54,6 mg / l. In der Pilotstudie mit Komplikationen während der Schwangerschaft, zeigten Frauen mit PE deutlich höhere mediane Afaminserumkonzentrationen als Frauen mit unkomplizierter Schwangerschaft (70,0 mg / l vs. 55,4 mg / l, p = 0,007). Expressionsanalysen auf mRNA und Proteinebene in unkomplizierten Schwangerschaften ergaben keine Plazenta Afaminexpression.

Schlussfolgerung: Ein linearer Anstieg des mütterlich exprimierten Glykoproteins Afamin während der Schwangerschaft kann als grundlegende Referenz für nachfolgenden Untersuchungen von Afamin in schwangerschaftsbedingten Erkrankungen dienen.

Eigenanteil: Der Doktorand war Zweitautor dieser Veröffentlichung und hat die komplette statistische Auswertung, von der Wahl der Methoden und deren Beschreibung bis hin zur Durchführung der Auswertung übernommen. Er war für die Richtigkeit jeder einzelnen Zahl verantwortlich, aber auch für die globale Auswertungsstrategie. Darüber hinaus war er am Design der Forschungsfragestellung und der Erstellung des Manuskripts maßgeblich beteiligt.

1.2 Zusammenfassung

Bisher fanden kausale Methoden noch zu wenig Anwendung in Beobachtungsstudien. Die Vorspezifizierung der Analysestrategie und die damit verbundene Erhöhung der Glaubwürdigkeit, und sicher auch der Reproduzierbarkeit, wurden im Bereich der Beobachtungsstudien lange unterschätzt und stiefmütterlich behandelt. Diese Arbeit zeigt, wie kausale Methoden bei der Wahl des Studiendesigns (longitudinal vs. Querschnitt) und der Auswahl der Kovariablen benutzt werden können, selbst wenn das Vorwissen gering oder existierende kausale Diagramme noch Spielräume lassen.

Der erste Teil der Arbeit bewegt sich in einem Forschungsgebiet, in dem bereits relativ viel Literatur existiert und daher auf ein breites akzeptiertes Fachwissen zurückgegriffen werden kann, was die Benutzung kausaler Strategien möglich und meiner Meinung nach auch nötig macht. Da in diesem Forschungsfeld bereits longitudinale Studien durchgeführt wurden, bot die Verwendung von Querschnittsdaten eine perfekte Ergänzung, um speziell die örtliche Variation im Gegensatz zur zeitlichen zu untersuchen [34].

Im Gegensatz hierzu wird im zweiten Teil der Arbeit ein Biomarker erstmals bei Schwangeren gemessen. Während die globale Analysestrategie kausal ist, sind Teile der Analyse klassische reine Assoziationsstatistik bzw. sogar nur Hypothesen

generierender Natur, wie beispielsweise die Assoziationsanalysen des neuen Biomarkers mit diversen Lipiden. Da es sich bei dieser Forschungsarbeit um eine erstmalige Beschreibung eines Biomarkers handelt, wurde sein Verhalten sowohl im Querschnitt als auch über die Zeit gemessen, um den Einfluss der Zeit und in beschränktem Maß auch den des Ortes (3 Zentren) zu untersuchen.

Den Analysestrategien beider Teile ist gemein, dass zuerst das Design des jeweiligen gemischten Modells [36] (Longitudinales Lineares Gemischtes Modell und Nicht-lineares Multi-Level Logit Modell) mittels eines Kriteriums (BIC [37] und AIC [38]) bestimmt wird und anschließend alle *a priori* definierten Variablen mithilfe des selben Kriteriums ggf. in das Modell mit aufgenommen oder wieder aus dem Modell herausgenommen werden (Schrittweise-Selektion [39]). Die *a priori* Auswahl und Kodierung der Variablen erfolgte in beiden Fällen kausal, basierend auf vorhergehenden Veröffentlichungen und Standards bzw. deren Zusammenführung in einem kausalen Diagramm.

Beide Beispiele veranschaulichen Analysestrategien, die so weit als möglich kausale Methoden benutzen und Lösungsmöglichkeiten für typische Schwierigkeiten von Beobachtungsstudien geben. Die in mancher Hinsicht große Unterschiedlichkeit beider Beispiele (longitudinale vs. Querschnitts Daten, binärer vs. kontinuierlicher Endpunkt, epidemiologische und klinische Fragestellungen) ist hierbei ein immenser Vorteil und macht die Strategien beider Beispiele, oder zumindest Teile davon, für viele Fragestellungen anwendbar und zeigt eben auch, dass kausale Strategien in so gut wie allen Beobachtungsstudien anwendbar sind.

1.3 Summary

Until now, causal methods are not widely used in observational studies. In this field, pre-specification of the analysis strategy associated with an increase in credibility, and certainly also with an increase in reproducibility, were long underestimated and neglected. This work shows how causal methods can be used to choose the study design (longitudinal vs. cross-sectional) and select the covariates, even if previous knowledge is limited or existing causal diagrams are not fully defined.

The first part of the work is located in a field of research, where a relatively large amount of literature already exists, thus a large amount of accepted knowledge can be used. This makes the use of causal strategies possible and, in my opinion, necessary. Since in this field of research, longitudinal studies have already been conducted, the use of cross-sectional data offered a perfect complement to specifically examine the influence of local variation, in contrast to variation of time [34].

In the second part of the work, in contrast, for the first time a biomarker is measured in pregnant women. While the global strategy of analysis is causal, parts of the analysis are classical statistics of association or even only of hypotheses-generating nature, such as the analysis of association between the new biomarkers and various lipids. Since this is an initial characterisation of a biomarker, its behavior was measured both cross-sectional and over time, to investigate the influence of time, and of location to a limited extent (3 centers).

The analysis strategies of both parts have in common that firstly the design of the respective mixed model [36] (longitudinal linear mixed model and non-linear multilevel logit model) is determined by a criterion (BIC [37] and AIC [38]) and secondly the same criterion is used to decide if some of the priori defined variables should be included in the model or not (stepwise selection [39]). In both cases, the a priori selection and coding of the variables was performed causal, based on previous publications and standards and their extraction in a causal diagram.

Both examples illustrate analysis strategies that use as much as possible causal methods and provide solutions to typical problems of observational studies. The great diversity of both examples (longitudinal vs. cross-section data, binary vs. continuous endpoint, epidemiological and clinical question) is an immense advantage and makes the strategies of both examples, or at least parts thereof, applicable to many questions and shows furthermore that causal strategies are applicable in virtually all observational studies.

1.4 Literaturverzeichnis

- [1] Ioannidis JPA (2005) Why Most Published Research Findings Are False, PLoS Med. 2005 Aug; 2(8): e124., Published online 2005 Aug 30. doi: 10.1371/journal.pmed.0020124
- [2] Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. (2009) Publication bias in clinical trials due to statistical significance or direction of trial results. Cochrane Database Syst Rev. 2009 Jan 21;(1):MR000006. doi: 10.1002/14651858.MR000006.pub3.
- [3] U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) Center for Biologics Evaluation and Research (CBER), Guidance for Industry E9 Statistical Principles for Clinical Trials, September 1998 ICH
- [4] Robins JM, Hernan MA, Brumback B (2000) Marginal Structural Models and Causal Inference in Epidemiology, Epidemiology, Vol. 11, No. 5 (Sep., 2000), pp. 550-560 Published by: Lippincott Williams & Wilkins
- [5] Greenland S, Pearl J and Robins JM (1999) Causal Diagrams for Epidemiologic Research Epidemiology Vol. 10, No. 1 (Jan., 1999), pp. 37-48
- [6] Simon RM, Paik S, Hayes DF (2009) Use of Archived Specimens in Evaluation of Prognostic and Predictive Biomarkers, JNCI J Natl Cancer Inst (2009) 101(21): 1446-1452.doi: 10.1093/jnci/djp335 First published online October 8, 2009
- [7] Wirtschaftsverband der forschenden Pharma-Unternehmen in Deutschland, Politische Überregulierung gefährdet Wettbewerbsfähigkeit der Pharma-Industrie <http://www.vfa.de/de/wirtschaft-politik/artikel-wirtschaft-politik/politische-ueberregulierung-gefaehrdet-die-wettbewerbsfaehigkeit-der-pharmazeutischen-industrie.html>, aufgerufen am 20 Juni 2015
- [8] Ioannidis JPA, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG (2001) Replication validity of genetic association studies. Nat Genet. 2001;29:306–309.
- [9] Altman DG, Lausen B, Sauerbrei W and Schumacher M (1994). Dangers of using "optimal" cutpoints in the evaluation of prognostic factors. Journal of the National Cancer Institute 86(11), 829–835.

- [10] Vandenbroucke JP (2004) When are observational studies as credible as randomised trials? *Lancet*.2004;363:1728–1731.
- [11] Llovera G, Hofmann K, Roth S, Salas-Pérdomo A, Ferrer-Ferrer M, Perego C et al. (2015) Results of a preclinical randomized controlled multicenter trial (pRCT): Anti-CD49d treatment for acute brain ischemia, *Science Translational Medicine* 05 Aug 2015: Vol. 7, Issue 299, pp. 299ra121, DOI: 10.1126/scitranslmed.aaa9853
- [12] Vul E, Harris C, Winkielman P & Pashler H (2009) Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition, *Perspectives on Psychological Science* May 2009 4: 274-290
- [13] Begley CG, Ellis LM (2012) Drug development: Raise standards for preclinical cancer research, *Nature* 483, 531–533 (29 March 2012), doi:10.1038/483531a, Published online 28 March 2012
- [14] Prinz F, Schlange T and Asadullah K (2011) Believe it or not: how much can we rely on published data on potential drug targets? *Nat Rev Drug Discov.* 2011 Aug 31; 10(9):712, doi: 10.1038/nrd3439-c1
- [15] Fang F, Markovac J, Roig M, Shafer S, Simmerling M, Smith R, Vaux D, Wager E, <http://retractionwatch.com>, aufgerufen am 20 Juni 2015
- [16] Guidance for Industry Bioanalytical Method Validation U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) Center for Veterinary Medicine (CVM) May 2001 BP: Bioanalytical Method Validation
- [17] U.S. Department of Health and Human Services Food and Drug Administration <http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/default.htm>, aufgerufen am 20 Juni 2015
- [18] Verordnung über die Anwendung der Guten Klinischen Praxis bei der Durchführung von klinischen Prüfungen mit Arzneimitteln zur Anwendung am Menschen (GCP-Verordnung - GCP-V), GCP-Verordnung vom 9. August 2004 (BGBl. I S. 2081), die zuletzt durch Artikel 8 des Gesetzes vom 19. Oktober 2012 (BGBl. I S. 2192) geändert worden ist"

- [19] Sauerbrei W, Abrahamowicz M, Altman DG, le Cessie S and Carpenter J on behalf of of STRATOS initiative (2014) STREngthening Analytical Thinking for Observational Studies: The STRATOS initiative. *Statistics in Medicine*, 2014. DOI: 10.1002/sim.6265
- [20] <https://www.clinicalstudydatarequest.com/>, aufgerufen am 20 Juni 2015
- [21] Hosmer DW, Lemeshow S and Sturdivant RX (2013) Model-Building Strategies and Methods for Logistic Regression, in *Applied Logistic Regression*, Third Edition, John Wiley & Sons, Inc., Hoboken, NJ, USA. doi: 10.1002/9781118548387.ch4
- [22] Fitzmaurice GM, Laird NM, Ware JH (2012) Residual Analyses and Diagnostics, in *Applied longitudinal analysis 2nd ed*, Vol. 998. John Wiley & Sons, 2012.
- [23] Rässler S, Rubin DB, and Zell ER (2008) Incomplete data in epidemiology and medical statistics. In Rao CR, Miller JP, and Rao DC (Eds.), *Handbook of Statistics 27: Epidemiology and Medical Statistics*, pp. 569–601. Elsevier.
- [24] European Medicines Agency, Oncology Working Party (2012) Guideline on the evaluation of anticancer medicinal products in man, 13 December 2012, EMA/CHMP/205/95/Rev.4
- [25] European Medicines Agency, Oncology Working Party (2012) Appendix 1 to the guideline on the evaluation of anticancer medicinal products in man, 13 December 2012, EMA/CHMP/27994/2008/Rev.1
- [26] Rubin DB (1974) Estimating causal effects of treatments in randomized and non-randomized studies," *Journal of Educational Psychology*. 1974;66:688–701. doi: 10.1037/h0037350.
- [27] Rosenbaum PR, Rubin DB (1983) The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70 (1): 41–55. doi:10.1093/biomet/70.1.41
- [28] Rubin DB (2004) Direct and indirect causal effects via potential outcomes," *Scandinavian Journal of Statistics*. 2004;31:161–170. doi: 10.1111/j.1467-9469.2004.02-123.x.
- [29] Robins JM , Hernán M. A, Brumback B, (2000). Marginal Structural Models and casual inference in epidemiology, *Epidemiology* Vol. 11 No.5

- [30] Robins JM, Finkelstein DM (2000) Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics*. 2000 Sep;56(3):779-88.
- [31] J. Pearl (1995) "Causal Diagrams for Empirical Research" *Biometrika*, 82(4), 669--710, December 1995.
- [32] J. Pearl (2009) "Causal inference in statistics: An overview," *Statistics Surveys*, 3:96--146, 2009.
- [33] Kalisch M, Fellinghauer BA, Grill E, Maathuis MH, Mansmann U, Buhlmann P, and Stucki G (2010) Understanding human functioning using graphical models. *BMC Med Res Methodol*, 10:14, 2010.
- [34] Markovitz AR, Goldstick JE, Levy K, Cevallos W, Mukherjee B, Trostle JA, et al. (2012) Where science meets policy: comparing longitudinal and cross-sectional designs to address diarrheal disease burden in the developing world. *Int J Epidemiol*, 41(2):504-13. Epub ahead of print 2012 Jan 9.
- [35] Nasseh D, Engel J, Mansmann U, Tretter W, Stausberg J. (2014) Matching study to registry data: maintaining data privacy in a study on family based colorectal cancer. *Stud Health Technol Inform*. 2014;205:808-12.
- [36] Laird NM, Ware JH (1982) Random-Effects Models for Longitudinal Data. In: *International Biometric Society (Hrsg.): Biometrics*. 38, Nr. 4, 1982, S. 963–974. doi:10.2307/2529876. PMID 7168798
- [37] Schwarz GE (1978) "Estimating the dimension of a model", *Annals of Statistics* 6 (2): 461–464, doi:10.1214/aos/1176344136, MR 468014.
- [38] Akaike H (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6): 716–723. doi:10.1109/TAC.1974.1100705. MR0423716
- [39] Harrell FE (2001) "Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis," Springer-Verlag, New York.

2. Cooking and season as risk factors for acute lower respiratory infections in African children: a cross-sectional multi-country analysis

Hannes Buchner¹, Eva A. Rehfuess¹

¹*Institute for Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-University, Munich, Germany*

This chapter was published as

Buchner H, Rehfuess EA (2015) **Cooking and Season as Risk Factors for Acute Lower Respiratory Infections in African Children: A CrossSectional Multi-Country Analysis**. PLoS ONE 10(6): e0128933. [doi:10.1371/journal.pone.0128933](https://doi.org/10.1371/journal.pone.0128933)

and can be found here

<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0128933>

3. The vitamin E-binding protein afamin increases in maternal serum during pregnancy

Michael Hubalek [a](#), Hannes Buchner [b](#), Manfred G. Mörtl [c](#), Dietmar Schlembach [c,d](#), Berthold Huppertz [e](#), Branka Firulovic [a](#), Wolfgang Köhler [f](#), Erich Hafner [g](#), Benjamin Dieplinger [f](#), Ludwig Wildt [h](#), Hans Dieplinger [i,j](#)

a Department of Gynecology and Obstetrics, Innsbruck Medical University, Innsbruck, Austria

b Institute of Medical Informatics, Biometry and Epidemiology, Medical School, Ludwig Maximilian University of Munich, Germany

c Department of Gynecology and Obstetrics, Medical University of Graz, Austria

d Department of Gynecology and Obstetrics, Jena University, Jena, Germany

e Institute of Cell Biology, Histology and Embryology, Medical University of Graz, Austria

f Department of Laboratory Medicine, Konventhospital Barmherzige Brüder, Linz, Austria

g Department of Obstetrics and Gynecology, Danube Hospital/SMZ-Ost, Vienna, Austria

h Department of Gynecological Endocrinology and Reproductive Medicine, Innsbruck Medical University, Innsbruck, Austria

i Division of Genetic Epidemiology, Department of Medical Genetics, Clinical and Molecular Pharmacology, Innsbruck Medical University, Innsbruck, Austria

j Vitateq Biotechnology GmbH, Innsbruck, Austria

This chapter was published as

Michael Hubalek, Hannes Buchner, Manfred G. Mörtl, Dietmar Schlembach, Berthold Huppertz, Branka Firulovic, Wolfgang Köhler, Erich Hafner, Benjamin Dieplinger, Ludwig Wildt, Hans Dieplinger: **The vitamin E-binding protein afamin increases in maternal serum during pregnancy** Clinica Chimica Acta 2014, 434 (2014) 41–47, <http://dx.doi.org/10.1016/j.cca.2014.03.036>

and can be found here

<http://www.sciencedirect.com/science/article/pii/S000989811400165X>

Danksagung

Mein Dank gilt meinen Eltern die mich immer machen ließen, die mich studieren ließen, meiner Oma und Opa die mich unterstützten und natürlich Bernadette, was soll ich sagen für alles. Meine Schwester und meine Freunde will ich auch nicht vergessen: Danke Euch.

Euer Hannes