

Diplomarbeit
im Studienfach Gesundheit

Evaluation einer Erhebung –
Prozess und Qualität einer krankenhausbasierten
Querschnittsstudie zur Stillförderung

Eingereicht bei: Prof. Dr. A. C. Seibt

Zweitprüferin: Prof. C. Deneke

Vorgelegt von:

Marlies Jöllenbeck,

Studiengang Gesundheit, Fachbereich Ökotrophologie,

Hochschule für Angewandte Wissenschaften Hamburg

Hamburg, März 2006

Zusammenfassung

Forschungsaufgabe: Die vorliegende Arbeit stellt, im Rahmen einer Zwischenauswertung, die Evaluation von Prozess- und Ergebnisqualität einer krankenhausbasierten Querschnittserhebung dar.

Untersuchungsobjekt: Das Evaluationsobjekt, die Studie QuaSti („Qualität im stillfreundlichen Krankenhaus“), ist eine Erhebung an einer „stillfreundlichen“ Hamburger Klinik. Neben der Versorgungsqualität dieser geburtshilflichen Abteilung bezüglich des Stillens werden die Stillprävalenz, Unterstützung bei Stillschwierigkeiten auf kommunaler und persönlicher Ebene und zahlreiche weitere quantitative und qualitative Variablen zu diesem Themenbereich erhoben. Der Studie QuaSti liegt mit der Querschnittserhebung ein epidemiologisches Untersuchungsdesign zu Grunde. Die Qualität epidemiologischer Studien ist im hohen Maße abhängig von der Art der Planung, der Durchführung und der Analyseverfahren. Um dies bewerten zu können, bedarf es einer systematischen Evaluation.

Zielsetzung: Ziel der Evaluation ist die Qualitätssicherung und Handlungsoptimierung im durchführungsmethodischen Vorgehen sowie eine Bewertung der Qualität der erhobenen Daten.

Hintergrund: Stillen ist von hoher Public-Health-Relevanz. Es ist mit zahlreichen, gesundheitlichen Vorteilen für Kind und Mutter verbunden. Diese Aussage ist gut kommuniziert und das Wissen darum verbreitet. Trotzdem ist die Stilldauer in Deutschland und vielen anderen Ländern kürzer als von der WHO empfohlen und insbesondere sozial benachteiligte Frauen werden von dieser Botschaft nicht erreicht. Gesundheitspolitische Initiativen wie die „Baby-Friendly Hospital Initiative“ (BFHI) der WHO/ UNICEF wollen weltweit das Stillen fördern. Da den Entbindungskliniken in diesem Zusammenhang eine Schlüsselrolle zugeschrieben wird, wurde dieses Setting als Ausgangspunkt für Stillförderung gewählt. Im Rahmen dieser Maßnahme zur Gesundheitsförderung werden Entbindungseinrichtungen nach definierten Kriterien („Zehn Schritte zum erfolgreichem Stillen“) von der WHO zertifiziert. Die Studie QuaSti ist die bundesweit erste, systematische Untersuchung eines stillfreundlichen Krankenhauses.

Methode: Die interne Evaluation erfolgt in Anlehnung an ein triangulatives Evaluationsverfahren mit Anwendung qualitativer und quantitativer Forschungsmethoden.

Ergebnis: Evaluationen können die Qualität epidemiologischer Untersuchungen maßgeblich sichern und verbessern.

Vorwort

„We are not no different from Gertrude Stein, who, on her deathbed, asked her lifelong companion, Alice B. Toklas, “What is the answer” And when Alice not could bring herself to speak, Gertrude asked, “in that case, what is the question.”” (Fontana, 668: 2000)

Mein Interesse an dem Thema Stillen begründet sich sowohl professionell als auch persönlich. Gesundheitsförderung stellt im Rahmen meines Studiums einen Schwerpunkt dar. Die Möglichkeit zur Mitarbeit an einer Querschnittserhebung, der Studie QuaSti, zum Thema Stillen interessiert mich, weil mich die Komplexität des Themas anspricht: Maßnahmen zur Stillförderung sind vorausschauend, weil damit die Gesundheit der heranwachsenden Generationen langfristig gefördert wird – und zwar in einem Umfang, wie nur wenige andere Maßnahmen dies vermögen. Doch welche Einflussfaktoren determinieren die individuelle Stillbereitschaft und -dauer? Wo und bezüglich welcher Zielgruppen besteht Bedarf an Stillförderung?

Meine Aufgabe lag darin, die Durchführung und das Ergebnis der Studie QuaSti im Rahmen einer Evaluation zu bewerten. Im Vordergrund der Arbeit stehen also nicht die inhaltlichen Ergebnisse der Studie QuaSti selbst, sondern deren methodische Qualität. Die Herausforderung dieser Aufgabe lag für mich darin, auf dieser methodischen Ebene zu bleiben. Durch die Mitarbeit und Evaluation der Studie QuaSti erhielt ich die Möglichkeit, theoretische Studieninhalte in der praktischen Umsetzung zu erleben, systematisch zu begleiten und methodische Techniken zu vertiefen. Dies ist hinsichtlich meiner beruflichen Fortentwicklung von großem Gewinn und Nutzen für mich.

Mit der Geburt meines Sohnes im Jahre 2004 hat das Thema Stillen für mich auch eine persönliche Dimension erhalten, die mich zur Auseinandersetzung mit der Thematik motivierte. Mein herzlicher Dank gilt Annette Seibt und Christiane Deneke, die Stillförderung ganz praktisch umsetzten, indem sie meine Lebenssituation mit Baby berücksichtigten und mir die notwendige Zeit und Unterstützung zur Erarbeitung der Diplomarbeit gewährten. Auch alle anderen Beteiligten der QuaSti-Studie trugen erheblich zum Entstehen dieser Arbeit bei, indem ich alle notwendigen Informationen und Daten zeitnah erhielt und großes Interesse wie auch Reflexionsbereitschaft erkennbar waren. Wesentliche Motivation für die Diplomarbeit war nicht zuletzt auch, dass die Mitarbeit im „QuaSti-Team“ einfach Spaß machte.

Inhaltsverzeichnis

Zusammenfassung	II
Vorwort	III
Inhaltsverzeichnis	IV
Tabellenverzeichnis	VI
Abkürzungsverzeichnis.....	VII
1 Einleitung.....	9
1.1 Problemstellung und Relevanz der vorliegenden Arbeit	9
1.2 Explikation der Forschungsaufgabe.....	11
1.3 Limitationen.....	12
1.4 Aufbau und Struktur der Arbeit.....	13
2 Grundlagen der Evaluationsforschung und Methodik.....	14
2.1 Grundlagen der Evaluationsforschung.....	14
2.2 Funktionen und Phasen einer Evaluation	15
2.3 Methodik der Evaluationsstudie	17
2.3.1 Triangulation.....	17
2.3.2 Qualitätskriterien für die vorliegende Evaluationsstudie.....	18
2.3.3 Grundlage für die Entwicklung der Evaluationskriterien	19
2.3.4 Der zeitliche Rahmen	19
3 Die Studie QuaSti– Hintergrund und Programmkontext.....	21
3.1 Public-Health-Relevanz des Stillens	21
3.2 Soziokultureller und gesundheitspolitischer Kontext.....	22
3.3 Hintergrund des WHO/ UNICEF Programms „Stillfreundliches Krankenhaus“	23
3.4 Internationale und nationale Trends zum Thema Stillen– die Datenlage...	24
3.5 Stand der Evaluation stillfördernder Einrichtungen- national und international.....	26
4 Die Studie QuaSti- Entwicklung und Vorgehensweise	27
4.1 Ausgangspunkt und Forschungsfrage	27
4.2 Studienumfang und Untersuchungsgang.....	29
4.3 Auswahl und Rekrutierung der Stichprobe.....	30
4.4 Das Projektteam der Studie QuaSti	31
4.5 Die Entwicklung der Erhebungsinstrumente	32

4.6	Auswahl und Schulung der InterviewerInnen.....	33
5	Darstellung der Prozessevaluation in der Studie QuaSti.....	35
5.1	Die Auswahl des Studiendesigns.....	35
5.2	Die Auswahl der Stichprobe.....	36
5.2.1	Die Stichprobenziehung.....	37
5.2.2	Zeitlicher Verlauf der Rekrutierung.....	38
5.3	Entwicklung und Einsatz der Erhebungsinstrumente.....	40
5.3.1	Die Entwicklung des Erhebungsinstruments mit Pilotstudie.....	40
5.3.2	Der Demographiebogen.....	41
5.3.3	Das Telefoninterview.....	42
5.4	Auswahl, Schulung und Kontrolle der InterviewerInnen.....	45
5.4.1	Die Auswahl der InterviewerInnen.....	45
5.4.2	InterviewerInnen- Schulung.....	45
5.4.3	Kontrolle in der Durchführung der Interviews.....	47
5.5	Rücklauf der Interviews.....	48
5.6	Qualitätssicherung.....	49
5.7	Datenhaltung- und Dokumentation.....	50
5.8	Datenschutz.....	51
5.9	Ethische Aspekte.....	52
6	Die Evaluation der Ergebnisse.....	54
6.1	Die Qualität der Stichprobe.....	54
6.1.1	Zusammensetzung der Stichprobe.....	55
6.1.2	Repräsentanz ausgewählter Gruppen.....	57
6.1.3	Die Response-Rate.....	59
6.1.3.1	Nichteilnehmerinnen-Analyse.....	61
6.1.3.2	Item-Nonresponse.....	62
6.2	Reliabilität.....	64
6.3	Objektivität.....	65
6.4	Kontextanalyse.....	66
6.5	Analyse des auswertungsmethodischen Verfahrens.....	67
6.6	Kommunikation und Public Health.....	68
7	Zusammenfassung und Diskussion.....	69
	Literaturverzeichnis.....	73
	Eidesstattliche Erklärung.....	80
	Anhang.....	81

Tabellenverzeichnis

Tab. 1: „Standards für Evaluation“ / DeGEval (Beywl et al., 2002: 9), Übersicht	18
Tab. 2: „Leitlinien und Empfehlungen Guter Epidemiologischer Praxis“, Übersicht	19
Tab. 3: Das Modell der Evaluationsstudie inkl. Bewertungskriterien	20
Tab. 4: Übersicht über den Rücklauf der Demographiebögen	39
Tab. 5: Rücklaufkontrolle Interviews; Übersicht.....	49
Tab. 6: Typisierung der Stichprobe.....	56
Tab. 7: Responseanalyse der Datenauswertung (vgl. Janßen et al., 1999: 175).....	60
Tab. 8: Gründe für Nichtteilnahme an den Interviews; Übersicht	61
Tab. 9: Typisierung der Nichtteilnehmerinnen	62

Abkürzungsverzeichnis

Abb.	Abbildung
Anm. d.	Anmerkung der
Bd.	Band
BF	Breastfeeding
BFHI	Baby-Friendly Hospital Initiative
bspw.	beispielsweise
BZgA	Bundeszentrale für gesundheitliche Aufklärung
bzw.	beziehungsweise
CATI	Computer Assisted Telephone Interview
DAE	Deutsche Arbeitsgemeinschaft für Epidemiologie
DeGEval	Deutsche Gesellschaft für Evaluation
d.h.	das heißt
ebd.	ebendort
erw.	erweitert(et)
et al.	et alii: und andere
etc.	et cetera
f.	folgende
ff.	fortfolgende
FKE	Forschungsinstitut für Kinderernährung (in Dortmund)
GEP	Gute Epidemiologische Praxis, Kurzform
ggf.	gegebenenfalls
HAW	Hochschule für Angewandte Wissenschaften
HAG	Hamburgische Arbeitsgemeinschaft für Gesundheitsförderung
HIS	Health Interview Survey
Hrg.	Herausgeber
i.e.	engl.: in example
Inc.	Incorporated
InGe	Institut für Gesundheitswissenschaften
Ink.	inklusive
Kap.	Kapitel
LBK	Landesbetrieb Krankenhäuser
QuaSti	Qualität im stillfreundlichen Krankenhaus
neubearb.	neubearbeitet(e)
Pkt.	Punkt
rev.	revidiert(e)
S.	Seite
Sic	lat. so: so lautet die Quelle
Sp.	Spalte
SPSS	Superior Performing Software Systems
Tab.	Tabelle
u.a.	und andere(s)
überarb.	überarbeitet(e)
UNICEF	United Nations International Children's Emergency Fund
URL	Uniform Resource Locator
Vol.	Volume
vollst.	Vollständig
vgl.	vergleiche
WHO	World Health Organisation
WHA	World Health Association
WWW	World Wide Web
zit. n.	zitiert nach
zugl.	zugleich

1 Einleitung

1.1 Problemstellung und Relevanz der vorliegenden Arbeit

“There is increasing interest in the development and application of formal “good epidemiology practices”. These not only consider issues of data quality, study design, and study conduct, but through their structured approach to the documentation of the study procedures, provide the potential for more rigorous reporting of the results in the scientific literature.” (Rushton, 2000:1)

Epidemiologische Studien stellen ein wesentliches Instrument zur Erforschung von Krankheitsursachen in definierten Populationen dar. (Hellmeier et al., 2003: 231) Gordis (2001: 3) beschreibt moderne Epidemiologie als Wissenschaft, die sich mit der Beobachtung, Untersuchung und Dokumentation der Verteilung und den Determinanten aller gesundheitsbezogenen Zustände und Ereignisse befasst.

Dabei wird die epidemiologische Datenbasis oftmals als Verteilungskriterium im Gesundheitswesen eingesetzt und dient der bedarfsgerechten Positionierung von Leistungen. (Schöffski et al., 2002: 2) Auch stellt sie die Grundlage einer Gesundheitsberichterstattung dar, die idealerweise Instrument einer gesundheitsfördernden Gesamtpolitik sein soll. (Trojan et al., 2001: 220) Dies verdeutlicht die Bedeutsamkeit dieses Wissenschaftsbereichs als gesundheitspolitisches Steuerungselement.

Epidemiologie stellt eine Basisdisziplin in vielen verschiedenen Fachrichtungen, wie in den Gesundheitswissenschaften, Pflegewissenschaften, Medizin, Psychologie und anderen Sozialwissenschaften dar. Dementsprechend groß ist die Bandbreite der Definitionen und Anwendungsbereiche. (Hellmeier et al., 2003: 232) Zahlreiche Aspekte beeinflussen die Qualität einer Studie maßgeblich, beginnend mit der Auswahl des probaten Studiendesigns bis hin zu den Prinzipien in der Durchführung statistischer Analysen. (Blettner et al., 2001: 97) Metaanalysen belegen, dass die Qualität epidemiologischer Studien variiert. Pocock et al. (2004: 329) weisen im Rahmen einer Analyse problematische Schwächen in der Planung, Durchführung und Analyse zahlreicher epidemiologischer Studien nach.

Aus der Wichtigkeit der Epidemiologie und der Problemstellung in diesem Feld lässt sich der Bedarf ableiten, die Qualität von Studien systematisch zu sichern und zu verbessern. Evaluationen stellen deshalb ein wesentliches Qualitätsmerkmal epidemiologischer Studien dar.

„Evaluation ist die systematische Untersuchung des Nutzens oder Wertes eines Gegenstandes. Solche Evaluationsgegenstände können z.B. Programme, Projekte, Produkte, Maßnahmen [...], Technologien oder Forschung sein.“ (Beywl et al., 2002: 16)

Evaluation dient durch kritische Reflexion des Untersuchungsobjekts einer wissenschaftlichen Weiterentwicklung und dem Erkenntnisgewinn. Darüber hinaus ist sie in ihrer praktischen Anwendung als Instrument der Qualitätssicherung relevant. Im Bereich des Gesundheitswesens erleben Evaluationen derzeit eine Konjunktur. Hier gewinnt der Nachweis von Wirksamkeit, Qualität und Wirtschaftlichkeit bestehender oder neu geplanter Maßnahmen angesichts begrenzter Ressourcen zunehmend an Bedeutung (Elkeles, 2004: 1) Evaluationen bilden die Grundlage evidenzbasierten¹ Handelns und können damit einer sinnvollen wie auch gerechteren Allokation der Ressourcen dienen.

Die vorliegende Arbeit evaluiert eine Querschnittserhebung. Das Evaluationsobjekt, die Studie „QuaSti“ („Qualität im stillfreundlichen Krankenhaus“), ist eine krankenhausbasierte Erhebung zum Thema Stillen. Erhoben werden die Stillprävalenz, sowie die Versorgungsqualität der geburtshilflichen Abteilung bezüglich des Stillens. Die Daten dienen der Einschätzung des Bedarfs von Angeboten, Unterstützung und Betreuung für stillende Frauen, auch auf kommunaler und privater Ebene. Die Studie QuaSti ist keine Studie im vielzitierten „Elfenbeinturm der Wissenschaft“ ohne praktischen Bezug, sondern vielmehr praxisnah mit dem Ziel angelegt, konkreten Handlungsbedarf für die Zielgruppe abzuleiten und die Situation stillender Frauen zu verbessern. Eine wesentliche Voraussetzung dafür stellt die Qualität der erhobenen Daten dar. Die Evaluation ist deshalb von praktischer Relevanz, um die Qualität der Studie zu sichern und zu verbessern. Sie macht Aussagen zur Prozessqualität und zur Qualität der erhobenen Daten und kann damit deren Evidenz bewerten und belegen.

Für den Themenbereich des Stillens und stillfreundlicher Krankenhäuser liegen in Deutschland keine publizierten Evaluationen epidemiologischer Erhebungen vor. Vor dem Hintergrund möchte die vorliegende Arbeit einen Beitrag in diesem Feld methodologischer Forschungsarbeit leisten.

¹ Evidenzbasierte Maßnahmen (z.B. in Medizin, Gesundheitsförderung) zielen darauf ab, die Praxis auf die jeweils „beste“ Evidenz zu stellen, d.h. es sollen nur solche Leistungen erbracht werden, deren Qualität und Wirksamkeit erwiesen sind. (Elkeles, 2004: 9)

1.2 Explikation der Forschungsaufgabe

„So it is important to be aware of how others' activities have been evaluated and to make your assessment of whether or not the evidence they provide is persuasive.“ (Naidoo et al., 2000: 374) Die Evaluation wird durchgeführt, um die Qualität der Studie QuaSti zu sichern und zu verbessern. Dafür wird eine Prozess- und Ergebnisevaluation durchgeführt.²

1. Prozessevaluation: Ziel der Bewertung im laufenden Forschungsprozess der Studie ist die Qualitätssicherung und Handlungsoptimierung, indem die Erkenntnisse direkt in den Forschungsprozess rückgekoppelt werden und zu Handlungskorrekturen führen können.³
2. Ergebnisevaluation: Diese erfolgt nach Abschluss der Feldphase oder, wie in der vorliegenden Arbeit, als Zwischenauswertung mit dem Vorliegen erster Daten.⁴ Sie stellt dabei eine Datenanalyse nach allgemein gültigen Gütekriterien für deren Qualität dar: Reliabilität, Validität und Objektivität. (Geyer, 2003: 35)

Die zentralen Forschungsfragen dieser Evaluation lauten: (vgl. dazu Kap. 7)

- Welche Qualität hat das erhebungsmethodische Vorgehen der Studie QuaSti?
- Welche Qualität haben die erhobenen Daten hinsichtlich ihrer Güte?
- Welchen Zielerreichungsgrad haben die Ergebnisse zur Beantwortung der Forschungsfrage der Studie QuaSti?

Der Evaluationsprozess mit Entwicklung eines adäquaten Untersuchungsdesigns zur Beantwortung dieser Fragen steht im Vordergrund der Darstellung. Die Evaluationsstudie hat grundsätzlich eine explorative Perspektive („goal-free“) und beschreibt und bewertet die auftretenden Effekte, Nebeneffekte, Zielverschiebungen, Konflikte und Wirkungszusammenhänge der Studie QuaSti. (vgl. Badura et al., 2003: 168) Aus Sicht der Evaluatorin ist es deshalb nicht sinnvoll, eine a priori Hypothesengenerierung vorzunehmen.

² Eine Evaluation ist, je nach ihrem zeitlichen Einsetzen, in verschiedene Phasen (Struktur-, Prozess- und Ergebnisevaluation) einteilbar. Eine detaillierte Definition der Phasen erfolgt in Abschnitt 2.2 (Anm. d. Autorin)

³ Die kontinuierliche Beobachtung des Programmverlaufs macht die Prozessevaluation mit zentralen Vorstellungen des „Total Quality Managements“ vergleichbar. (Christiansen, 1999: 54) Eine Vertiefung dieses Aspekts ist im Rahmen dieser Arbeit nicht möglich (Anm. d. Autorin)

⁴ Da die Studie QuaSti zum Zeitpunkt der Evaluation noch nicht beendet ist, stellt diese eine Zwischenauswertung dar. (Vgl. dazu Abschnitt 2.3.4)

1.3 Limitationen

Limitationen⁵ ergeben sich aus der Komplexität der Aufgabenstellung. *„Evaluation ist immer unvollständig, da es nicht möglich ist, jeden Aspekt einer Maßnahme zu bewerten.“* (Naidoo et al., 2003: 372) Die vorliegende Arbeit kann nicht die gesamte Bandbreite des Evaluandums⁶ erfassen und es muss deshalb eine Auswahl an Evaluationskriterien getroffen werden.

Eine weitere Beschränkung ergibt sich aus dem zeitlichen Start der Evaluation. Vereinbarungsgemäß führt die Evaluatorin⁷, im Rahmen einer Zwischenauswertung, eine Bewertung von Prozess und Ergebnis durch. (vgl. Abschnitt 2.3.4) Eine Strukturevaluation findet nicht statt. Dennoch werden einzelne strukturelle Aspekte aufgegriffen, die für den Gesamtverlauf von hoher Bedeutung sind.

Evaluationsforschung begleitet und bewertet Maßnahmen. Damit übernimmt sie Gestaltungs-, Kontroll-, Steuerungs-, und Bewertungsfunktion und steht potentiell im Spannungsverhältnis verschiedener Akteure mit unterschiedlichen Interessen. (Stockmann, 2004:19) Solche Interessenskonflikte können theoretisch einen limitierenden Faktor bei der Ausführung einer Evaluation darstellen. Die Projektleitung der Studie QuaSti hat explizit den Anstoß für die Evaluation gegeben. So besteht von dieser Seite aus eine hohe Bereitschaft zur Reflexion der eigenen Arbeit. Potentielle Konflikte durch die Evaluationsergebnisse wurden vorab erörtert und werden auch im laufenden Prozess konstruktiv kommuniziert.

Die Evaluation ist als interner Prozess angelegt, d.h. die Evaluatorin ist als Projektmitglied auch inhaltlich an der Studie beteiligt. Auch dieser Aspekt kann sich potentiell limitierend auswirken, weil es eine geringere Distanz und Neutralität gegenüber dem Evaluandum bedeuten kann. Die Evaluation wird deshalb nach Kriterien der Überprüfbarkeit und höchstmöglicher Transparenz gestaltet.

⁵ Ein wichtiges Kriterium für die Beurteilung der Evaluationsstudie ist die Transparenz und Nachvollziehbarkeit in der Auswahl der Kriterien. Um dies zu ermöglichen, erfolgt die Erläuterung der Limitationen vor Beginn des Evaluationsprozesses. (Anm. d. Autorin)

⁶ Evaluandum bezeichnet das zu evaluierende Objekt (vgl. Wiedmer, 2004: 84)

⁷ Die Begriffe „Autorin“ und „Evaluatorin“ werden in dieser Arbeit synonym verwendet (Anm. d. Autorin)

1.4 Aufbau und Struktur der Arbeit

Die vorliegende Arbeit umfasst insgesamt 7 Kapitel. Die vorangegangene Einleitung soll Aufschluss geben über die Relevanz von Evaluationen epidemiologischer Studien, über die Forschungsaufgabe sowie über die Limitationen, die sich ergeben.

Daran anschliessend werden im 2. Kapitel Grundlagen der Evaluationsforschung und die Methodik dieser Evaluationsstudie erläutert.

In Kapitel 3 und 4 werden das Evaluationsobjekt, Hintergrund und Programmkontext der Studie erläutert. Damit soll in die Thematik eingeführt und das Evaluationsobjekt vorgestellt werden. Die Darlegung des Programmkontextes ist unverzichtbarer Bestandteil des Evaluationsprozesses, weil dies die Voraussetzung für die Kontextanalyse darstellt.

Kapitel 5 und 6 stellen die Evaluation von Prozess- und Ergebnisqualität der Studie QuaSti anhand ausgewählter Kriterien dar. Dies demonstriert den methodologischen Forschungsprozess der Evaluation.

Im Kapitel 7 werden die Forschungsfragen der Evaluationsstudie zusammenfassend beantwortet und die Ergebnisse diskutiert. Abgeschlossen wird die Arbeit durch eine Reflexion des Evaluationsprozesses.

Anmerkungen zur Vorgehensweise

Evaluationsforschung hält eine Vielfalt synonym genutzter Begriffe bereit. Die Evaluatorin trifft deshalb zu Gunsten einer besseren Lesbarkeit eine Auswahl: Die vorliegende Arbeit wird als Evaluationsstudie bezeichnet, während das Evaluationsobjekt als „Evaluandum“ oder „Studie QuaSti“ bezeichnet wird. Im Zusammenhang mit der Studie QuaSti wird von „Vorgehen“ gesprochen, für die Evaluationsstudie wird der Begriff „Methode“ genutzt.

Da die Studie zum Zeitpunkt der Evaluation und der Erstellung dieser Arbeit noch nicht abgeschlossen ist, wird für Aspekte, die den noch laufenden Untersuchungsgang betreffen, die Präsensform gewählt. Ein Wechsel der Zeitformen ist daher teilweise nicht zu vermeiden.

Um den Lesefluss nicht zu behindern, finden sich im Text Formulierungen in maskuliner Form. Diese beziehen sich selbstverständlich stets auf beide Geschlechter.

2 Grundlagen der Evaluationsforschung und Methodik

2.1 Grundlagen der Evaluationsforschung

„Evaluation research is the systematic application of social research procedures in assessing the conceptualization and design, implementation, and utility of social intervention programs.“ (Rossi et al., 1993: 5 in Wottawa et al., 2003: 13)

Evaluation ist die systematische Bewertung von Sachverhalten, Objekten, Studien oder Personen anhand präzise definierter Kriterien. (Stockmann, 2004: 22) Die Evaluationsforschung ist ein Teilgebiet empirischer Sozialforschung. (ebd.) Definitiv wird zwischen Evaluationsforschung und Evaluation unterschieden. (Suchman, 1967, in Wottawa et al., 2003: 13) Evaluationsforschung meint die explizite Anwendung wissenschaftlicher Forschungsmethoden zur Durchführung einer Bewertung. Der Begriff Evaluation beschreibt den Prozess der Beurteilung eines Produktes, Prozesses oder Programms. (ebd.) Evaluationen werden in verschiedenen Bereichen (z.B. im Bildungswesen, Politik, Wirtschaft, Gesundheitswesen) mit sehr unterschiedlichen Erkenntnis- und Nutzwertinteressen eingesetzt. Zentrales Ziel stellt jedoch immer eine Bewertung der Qualität⁸ und der Wirksamkeit eines Objektes dar. (Badura et al., 2000: 574) Die Evaluation kann dabei als interner (Mitarbeiter des Projektes) oder externer (von außerhalb beauftragte Institution oder Person) Vorgang stattfinden.

Entsprechend der unterschiedlichen Einsatzgebiete besteht in der Evaluationsforschung ein ausgesprochener Methodenpluralismus. (Widmer, 2004: 86 ff.) Es kann die gesamte Bandbreite sozialwissenschaftlicher Forschungsmethoden genutzt werden. (ebd.) Der Einsatz quantitativer versus qualitativer Forschungsmethoden wird in der Evaluationsforschung, wie in anderen wissenschaftlichen Disziplinen auch, kontrovers diskutiert. (Naidoo et al., 2000: 373) Das Spektrum der Ansichten darüber ist breit, obgleich mittlerweile Annäherungen stattfinden: *„Neben dem „rein“ positivistischen (ergebnisorientiert/ quantitativen) und dem „rein“ phänomologischen (prozeßorientiert/ qualitativen) Ansatz gibt es in der neueren Evaluationsforschung Versuche, methodologische Kontroversen zu überwinden [...].“* (Badura et al., 2000: 582). Mittlerweile besteht Konsens darüber, dass quantitative und qualitative Methoden oft gut miteinander kombiniert werden können (Multimethodenansatz). (Stockmann, 2004: 21)

⁸ Qualität wird dabei definiert als: *„Die Gesamtheit von Eigenschaften und Merkmalen eines Produktes oder einer Dienstleistung, die sich auf deren Eignung zur Erfüllung festgelegter oder vorausgesetzter Erfordernisse bezieht.“* (Badura et al., 2000: 574)

Hintergrund und Entwicklung

Evaluationsforschung kann auf eine lange methodologische Tradition zurückblicken, deren Wurzeln sich wesentlich in einer politikwissenschaftlichen Diskussion über Wirkungsweise und Wirksamkeit des Staates begründen. Mit Entwicklung und Ausdifferenzierung der empirischen Sozialforschung hat sich die Evaluationsforschung zunehmend zu einem Instrument von pragmatischem Nutzen in den Sozialwissenschaften gewandelt. (vgl. Stockmann, 2004: 23 ff.; Mertens, 2004: 61 ff.) In Deutschland setzte die Evaluationsforschung Ende der sechziger Jahre des letzten Jahrhunderts, im Rahmen reform- und bildungspolitischer Programme ein. (Lübbert, 1999: 2) Diese Entwicklung folgte im internationalen Vergleich, insbesondere mit den USA, mit deutlicher Verzögerung.

Evaluationsforschung ist in Deutschland bis heute kein fest etablierter, eigenständiger Wissenschaftsbereich. (Wottawa et al., 2003: 13 ff.) Stockmann (2004; 33) konstatiert, dass die lange Geschichte der Evaluationsforschung in Deutschland in einem eigentümlichen Gegensatz zum Professionalisierungsgrad dieser Disziplin stehe. Als einer der Gründe führt er dazu an, dass die disziplinäre Segmentierung und die institutionelle Zersplitterung der Evaluation auf unterschiedliche Fachressorts der Weiterentwicklung von Theorien sowie der Wissensakkumulation in der Evaluationsforschung entgegenwirke. (vgl. ebd.) Dennoch hat die Professionalisierung der Evaluation in Deutschland in den letzten Jahren einen deutlichen Entwicklungsschub erfahren. Im Zuge dessen wurde im September 1997 die „Deutsche Gesellschaft für Evaluation“ (DeGEval) gegründet, um die Entwicklung, Akzeptanz und Nutzbarkeit dieser wissenschaftlichen Disziplin zu fördern und zu sichern. Die DeGEval veröffentlichte im Jahre 2001 die „Standards für Evaluation“ als Qualitätskriterien für Evaluationsforschung.⁹ (Beywl, 2002: 2ff.)

2.2 Funktionen und Phasen einer Evaluation

Generell können Evaluationen vier Funktionen (Aufgaben) zugeschrieben werden. Obgleich diese nicht trennscharf sind, sondern miteinander verbunden, ist es für die Ausrichtung der Evaluation von gravierender Bedeutung, welche Funktion im Vordergrund steht. (vgl. Stockmann, 2004: 17 ff.)

⁹ Die Entwicklung der Standards ist formuliert in Anlehnung an internationale Standards, insbesondere den „Standards for Evaluation of Educational Programs“, „Projects and Materials“ (Joint Committee, 1981) und der amerikanischen Fachgesellschaft „American Evaluation Association“ (Widmer, 2004: 92); (vgl. Abschnitt 2.5)

Dies sind die folgenden Funktionen: (ebd.)

Erkenntnisfunktion: Erkenntnisse über die Effektivität und Effizienz einer Maßnahme, Kausalzusammenhänge, Erreichbarkeit der Zielgruppe.

Kontrollfunktion: Die Beobachtung eines Programms führt zwangsläufig auch zu Informationen über die Arbeitsweise der Beteiligten. Daraus ergibt sich eine indirekte oder direkte Kontrollfunktion der Evaluation.

Dialog-/ Lernfunktion: Evaluationen liefern Informationen über die Zusammenarbeit verschiedener Akteure und bilanzieren, wie erfolgreich die Kooperation verläuft und wo Defizite auftreten. Dies bildet die Grundlage für einen Dialog und gemeinsames Lernen.

Legitimierungsfunktion: Evaluationen erlauben eine Einschätzung über die Effektivität und Effizienz einer Maßnahme. Damit kann der Umgang mit Finanzmitteln und der Wirkungsgrad einer Maßnahme belegt werden. Dies dient der Legitimation von Maßnahmen, auch auf gesundheitspolitischer Ebene.

Neben der generellen Ausrichtung einer Evaluation kann dabei in der praktischen Umsetzung nach drei verschiedenen Phasen des Evaluationsprozesses unterschieden werden: (nach Stockmann, 2004: 16 ff.)

1. *Strukturevaluation:* diese setzt bereits zum Zeitpunkt der Programmentwicklung ein und schließt die Konzeptualisierung und Ausarbeitung einer geplanten Intervention ein, um möglichst frühzeitig negative Effekte einer Maßnahme abschätzen zu können. Sie beleuchtet die strukturellen Rahmen- und Eingangsbedingungen (z.B. materielle, institutionelle, theoretische Aspekte) eines Vorhabens. Synonyme sind ‚ex-ante evaluation‘, ‚input-evaluation‘, ‚preformative evaluation‘ oder Programmevaluation.

2. *Prozessevaluation:* Diese setzt während der Implementationsphase der Maßnahme ein und übernimmt vorrangig Kontroll- und Beratungsfunktion während des Prozesses. Die Sammlung von Daten über den Verlauf dient als Steuerungselement und gibt Möglichkeiten zur Korrektur im laufendem Prozess („on going“). Synonym dieser Evaluationsform ist ‚formative Evaluation‘ oder auch ‚Begleitforschung‘. ‚Formativ‘ meint in diesem Zusammenhang aktiv- gestaltend, konstruktiv und kommunikationsfördernd.

3. *Ergebnisevaluation:* Diese Evaluation setzt ein, wenn die ersten Ergebnisse vorliegen. Mit Hilfe einer Wirkungsanalyse werden die Effekte einer Maßnahme überprüft und ‚im Lichte‘ der ex-ante formulierten Ziele geprüft. Diese Evaluationsform wird auch ‚ex-post-Evaluation‘ oder ‚summative Evaluation‘ genannt. ‚Summativ‘ impliziert eine mehr zusammenfassende, ergebnisorientierte und bilanzierende Perspektive.

2.3 Methodik der Evaluationsstudie

2.3.1 Triangulation

„Das Kernproblem sozialwissenschaftlicher Datengewinnung besteht, wie allgemein bekannt, in der begrenzten Überprüfbarkeit der Validität erfasster Informationen. Umso dringlicher erscheint es, dort, wo es die Anlage einer Studie erlaubt, das Verfahren der sogenannten Triangulation anzuwenden.“ (Siegrist et al., 2003: 274)

Für die vorliegende Evaluationsstudie wird ein triangulatives Evaluationsverfahren angewendet. Triangulation meint die Betrachtung eines Forschungsgegenstandes von (mindestens) zwei Punkten aus. Der Begriff kommt ursprünglich aus der Landvermessung und wurde von der qualitativen Sozialforschung übernommen. Der Sozialwissenschaftler Denzin prägte den Begriff und definierte ihn als: *„[...] the combination of methodologies in the study of the same phenomena.“* (1970: 297, zit. n. Flick, 2004: 13) Der integrative, methodische Einsatz ermöglicht eine Betrachtung eines Forschungsgegenstandes aus unterschiedlichen Perspektiven. Instrumente der Triangulation sind beispielsweise quantitative Messverfahren, qualitative Erhebungen durch Interviews, Analyse von Sekundärliteratur sowie deskriptive Elemente. *„Qualitative kann quantitative Forschung unterstützen und vice versa [...]“* (Bryman: 1992, zit. n. Flick, 2004: 68).

Ein triangulatives Evaluationsverfahren mit Einsatz unterschiedlicher Methoden ist aus Sicht der Autorin besonders geeignet, um die Komplexität des Evaluationsobjektes besser abzubilden. Diese Komplexität ergibt sich aus dem Gesamtkontext und wird nun in Kürze erläutert: Stillförderung ist ein Aspekt der Gesundheitsförderung. Die Etablierung von Maßnahmen zur Gesundheitsförderung in einem Krankenhaus ist immer noch aussergewöhnlich, da im „Medizinbetrieb“ traditionell die Behandlung von Krankheiten im Vordergrund steht.¹⁰ Dies erfordert ein Grundverständnis von Gesundheitsförderung der MitarbeiterInnen und steht möglicherweise im Gegensatz zum vorherrschenden, medizinisch geprägtem Selbstverständnis. (vgl. Naidoo et al., 2003: 5ff.)

Weiterhin werden mit der krankenhausbasierten Studie QuaSti vorrangig „Sozial- und Gesundheitsdaten“ vor dem Hintergrund eines Gesundheitsförderungsprozesses

¹⁰ In Krankenhäusern ist traditionell das klassische, biomedizinische Paradigma mit Fokussierung auf pathogenetische und pathologische Vorgänge (Krankheitsentstehung, Krankheit) etabliert. (vgl. Naidoo et al., 2003: 5 ff.) Dennoch sind innerhalb dieses Setting zunehmend Bestrebungen erkennbar, sich verstärkt auch der Erhaltung und Förderung der Gesundheit zuzuwenden. (ebd.) Ausdruck findet dies z.B. in der neuen Berufsbezeichnung für Krankenpflegekräfte („Gesundheits- und Krankenschwester/ -pfleger“). Eine Vertiefung dieses Aspektes ist im Rahmen dieser Arbeit nicht möglich. (Anm. d. Autorin)

erhoben, keine medizinischen Daten, wie beispielsweise in klinischen Studien üblich. Auch dies stellt eine Besonderheit dar. Daraus resultiert eine Vielschichtigkeit des Gesamtkontextes, der insbesondere die Prozessqualität der Studie QuaSti im hohen Maße beeinflusst. Vor diesem Hintergrund hat sich die Autorin für ein triangulatives Evaluationsverfahren entschieden.

2.3.2 Qualitätskriterien für die vorliegende Evaluationsstudie

„Die DeGEval- Standards sollen der Sicherung und Entwicklung der Qualität von Evaluationen dienen.“ (Beywl et al., 2002: 16)

Die „Standards für Evaluation“ sollen eine Orientierung bei der Durchführung und Bewertung von Evaluationen darstellen. Sie können auch die Planung und Umsetzung eines Evaluationsprojektes unterstützen. (Beywl et al., 2002: 16) Nach den Standards sollen Evaluationen vier grundlegende Eigenschaften aufweisen: Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit. (ebd.)

Diese sind wie folgt definiert:

<i>Nützlichkeit</i>
„Die Nützlichkeitsstandards sollen sicherstellen, dass die Evaluation sich an den geklärten Evaluationszielen sowie am Informationsbedarf der vorgesehenen Nutzer und Nutzerinnen ausrichtet“
<i>Durchführbarkeit</i>
„Die Durchführbarkeitsstandards sollen sicherstellen, dass eine Evaluation realistisch, gut durchdacht, diplomatisch und kostenbewusst geplant und ausgeführt wird“
<i>Fairness</i>
„Die Fairnessstandards sollen sicherstellen, dass in einer Evaluation respektvoll und fair mit betroffenen Gruppen und Personen umgegangen wird“
<i>Genauigkeit</i>
„ Die Genauigkeitsstandards sollen sicherstellen, dass eine Evaluation gültige Informationen und Ergebnisse zu dem jeweiligen Evaluationsgegenstand und den Evaluationsfragestellungen hervorbringt und vermittelt“

Tab. 1: „Standards für Evaluation“ / DeGEval (Beywl et al., 2002: 9), Übersicht

Die „Standards für Evaluation“ bilden die Grundlage dieser Evaluationsstudie. Ihre Anwendung dient der Transparenz und Qualitätssicherung in der Methodik. Weiterhin können sie als Bewertungsgrundlage herangezogen werden. (Widmer, 2004: 100) „Die Standards erlauben es dem Auftraggeber, die Evaluation fachlich fundiert zu bewerten.“ (Widmer, 2004: 103)

2.3.3 Grundlage für die Entwicklung der Evaluationskriterien

Für die systematische Bewertung der Studie QuaSti müssen Evaluationskriterien entwickelt werden. Da der Studie QuaSti ein epidemiologisches Studiendesign zugrunde liegt, werden dafür die „Leitlinien und Empfehlungen zur Sicherung von Guter Epidemiologischer Praxis“ (GEP) herangezogen. Diese Leitlinien wurden von der „Deutschen Arbeitsgemeinschaft für Epidemiologie“ (DAE) entwickelt, um einen Maßstab für gute wissenschaftliche Praxis zu setzen, gewissenhaftes Arbeiten zu sichern und einen vertrauensvollen Umgang unter Wissenschaftlern zu gewährleisten. Das Ziel der Leitlinien ist die Etablierung eines Qualitätsstandards für epidemiologische Forschung in Deutschland (vgl. Hoffmann et al., 2004: 3) Dabei sollen die Leitlinien kein „*kein enges und starres Handlungskorsett*“ (ebd.) darstellen, sondern möchten vielmehr als Handlungsempfehlung verstanden werden. Daher sind begründete Abweichungen möglich und stehen im Einklang mit „guter epidemiologischer Praxis“.

Die Arbeitsgruppe hat Leitlinien zu folgenden Themenbereichen aufgestellt:

Leitlinie 1: Ethik	Leitlinie 7: Auswertung
Leitlinie 2: Forschungsfrage	Leitlinie 8: Datenschutz
Leitlinie 3: Studienplan	Leitlinie 9: Vertragliche Rahmenbedingungen
Leitlinie 4: Probenbanken	Leitlinie 10: Interpretation
Leitlinie 5: Qualitätssicherung	Leitlinie 11: Kommunikation und Public Health
Leitlinie 6: Datenhaltung- und Dokumentation	

Tab. 2: „Leitlinien und Empfehlungen Guter Epidemiologischer Praxis“, Übersicht

2.3.4 Der zeitliche Rahmen

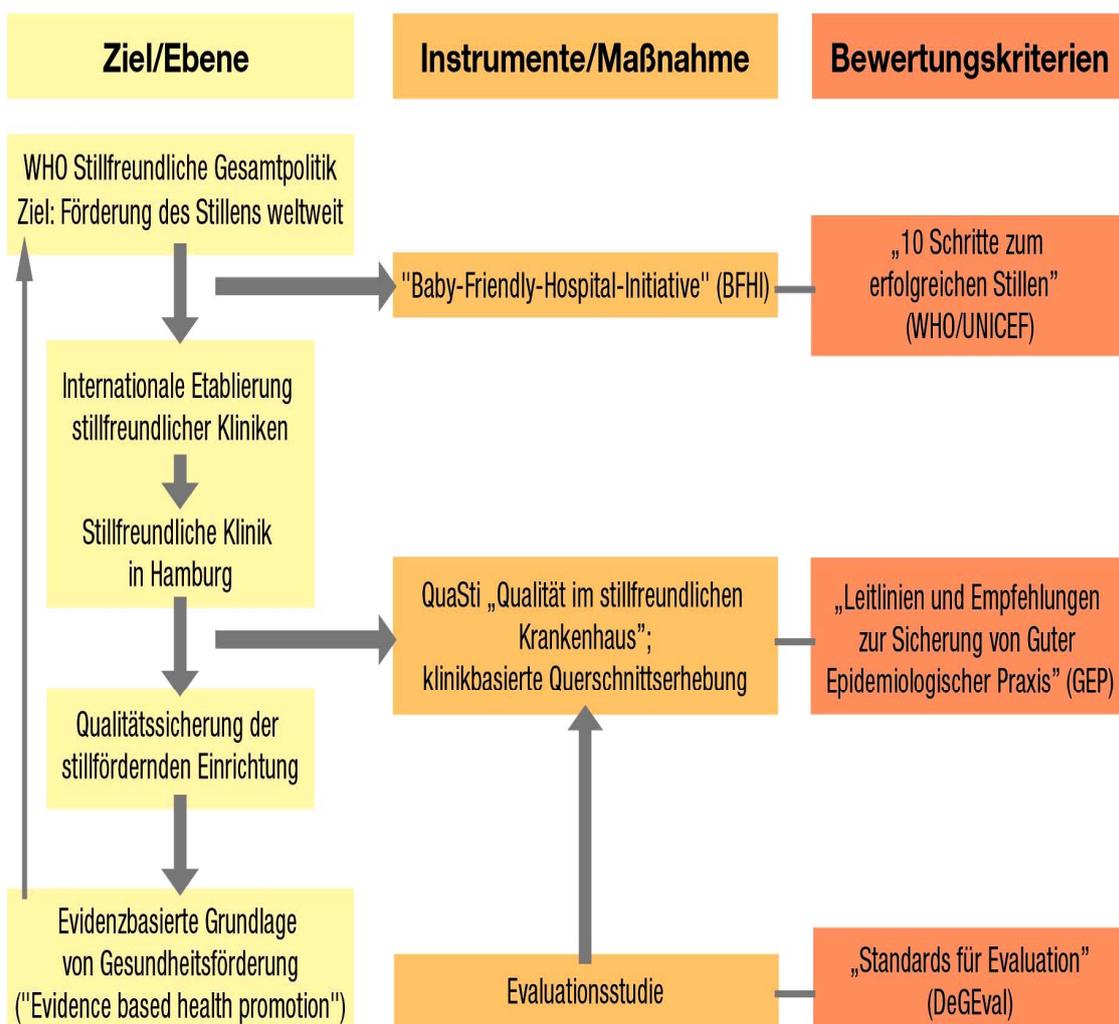
„*Evaluationsvorhaben sollen so rechtzeitig begonnen und abgeschlossen werden, dass ihre Ergebnisse in anstehende Entscheidungsprozesse bzw. Verbesserungsprozesse einfließen können.*“ (Beywl et al., 2002: 9)

Start der Evaluation war im März 2005 mit Hinzukommen der Evaluatorin zum Projektteam. Dies lag zeitnah zum Beginn der Hauptstudie am 01.01.2005 und dem Start der Implementationsphase. Die Erhebung der Daten sollte ursprünglich im Juli/ August 2005 abgeschlossen sein. Dies ließ sich aus verschiedenen Gründen jedoch nicht realisieren.¹¹ Da die Diplomarbeit zeitlichen Vorgaben unterlag, wurde zur Erstellung dieser Arbeit - im Sinne einer Zwischenauswertung - bei einem Stichprobenumfang von

¹¹ Die Gründe der zeitlichen Verzögerung sind Gegenstand der Evaluation und unter Abschnitt 5.2.2 erläutert.

n= 219 ein „Schnitt“ gemacht. Somit repräsentiert die Evaluation ein Zeitfenster der Studie über den Zeitraum vom 01.03. 2005 bis zum 31.11. 2005. Dies diente der Fertigstellung der Diplomarbeit innerhalb des geplanten Zeitraumes. Weiterhin stellt die Zwischenauswertung der Studie QuaSti ein wesentliches Element der internen Qualitätssicherung dar. Die Erkenntnisse der Zwischenauswertung haben Einfluss auf das weitere Vorgehen im Studienverlauf. Der Evaluationsprozess wird bis zum tatsächlichen Ende der Studie weitergeführt und in Form eines Evaluationsberichtes aufgearbeitet.

Der Aufbau und das Instrumentarium der Evaluationsstudie werden in der nachfolgenden Grafik nochmals zusammenfassend dargelegt:



Tab. 3: Das Modell der Evaluationsstudie inkl. Bewertungskriterien

3 Die Studie QuaSti– Hintergrund und Programmkontext

„Kontextanalyse: Der Kontext des Evaluationsgegenstandes soll ausreichend detailliert untersucht und analysiert werden.“ (Beywl, 2002: 10)

Vor der Darstellung des Evaluationsobjektes erfolgt nun eine Ausführung über den Hintergrund der Thematik und den Programmkontext. Dies dient der thematischen Übersicht sowie der Einordnung der Studie QuaSti in den Gesamtzusammenhang. Darüber hinaus ist eine Beschreibung des Programmkontextes, der sog. „Makro-Ebene“,¹² eine wesentliche Voraussetzung, um Kontexteinflüsse mit Wirkung auf das Evaluationsobjekt einschätzen zu können.

3.1 Public-Health-Relevanz des Stillens

„Schutz, Förderung und Unterstützung des Stillens ist in ganz Europa ein gesundheitspolitisches Thema ersten Ranges. Niedrige Stillraten und frühzeitiges Abstillen haben wesentliche, negative gesundheitliche und soziale Folgen für Frauen, Kinder, Gesellschaft und Umwelt.“ (Byrne, 2004: 10)

Stillen¹³ stellt die natürliche Ernährungsform des Säuglings dar, weil Muttermilch in ihrer Zusammensetzung ideal auf die Bedürfnisse des Kindes abgestimmt ist. Nach heutigem Kenntnisstand ist das Stillen mit zahlreichen gesundheitlichen Vorteilen für Mutter und Kind verbunden. Dass diese Art der Säuglingsernährung die kindliche Gesundheit, insbesondere die immunologische Abwehrlage und das Wachstum unmittelbar beeinflusst, ist seit langem bekannt. (Howie et al., 1990: 11) Gestillte Kinder erkranken seltener an Infektionen der oberen Luftwege, des Magendarmtrakts, an Sepsis, Mittelohrentzündungen und Harnwegsinfekten. (Abou-Dakn et al., 2003: 387) Durch Muttermilch wird auch ein wesentlicher Grundstein für die langfristige Gesundheit des Kindes gelegt. Studien belegen, dass das exklusive¹⁴ Stillen über mindestens 3-4 Monate die Risikorate, an einer atopischen Dermatitis und anderen Allergien zu erkranken, senkt. (Bergmann, 2001: 249) Weiterhin entwickeln gestillte Kinder in späteren Jahren seltener einen Diabetes Mellitus Typ 1, Morbus Crohn oder Adipositas. (Davis et al., 1988; Koletzko et al., 1989; Åkerblom et al., 1996, in Przyrembel, 2001: 4) Die emotionale und

¹² „Makro- Ebene“ meint regionale oder nationale Rahmenbedingungen, politische Einflüsse, soziales Klima (vgl. Christiansen, 1999: 50)

¹³ Stillen bezieht sich auf das Zuführen der Muttermilch direkt über die mütterliche Brust

¹⁴ Dies definiert die WHO als ausschließliches Stillen ohne Beigabe von Zusatzkost jeglicher Art (WHO, 2005)

kognitive Entwicklung des Säuglings wird positiv beeinflusst. (Widström, 2004: 12) Stillen wird als ein protektiver Faktor hinsichtlich einiger kindlicher Krebsformen diskutiert (Martin et al, 2005: 1) und stellt einen Schutzfaktor gegenüber dem plötzlichen Kindstod (SID = Sudden Infant Death) dar. (Jorch et al., 2003: 516) Doch auch für die Mütter ergeben sich gesundheitliche Vorteile. So zeigt sich, dass das sofortige Anlegen und Stillen die Rate der mütterlichen Wochenbettdepression senken kann. (Widström, 2004: 13) Stillen senkt das Risiko von Frauen, an Brustkrebs zu erkranken, signifikant. Bei einer Stillzeit von einem Jahr lässt sich eine individuelle Risikoreduktion des postmenopausalen Mamakarzinoms um 45% erreichen. (Scheele et al., 2003: 48) Auch das Risiko, an einem Ovarial- oder Uteruskarzinom zu erkranken, verringert sich. (Guise et al., 2003: 71)

Der signifikante Einfluss auf Morbiditäts- und Mortalitätsraten bestimmter Erkrankungen in diesen Bevölkerungspopulationen impliziert die hohe Public-Health-Relevanz. Damit verbunden sind auch zahlreiche ökonomische Aspekte von volkswirtschaftlicher Bedeutung. Direkte und indirekte Kosten¹⁵ durch das Auftreten akuter und chronischer Erkrankungen lassen sich durch stillfördernde Interventionen absenken. Stillen führt zu Einsparungen indirekter Kosten, wie beispielsweise psychosoziale Auswirkungen, die sich nicht eindeutig quantifizieren lassen. Dies macht eine Einschätzung des Gesamtnutzens stillfördernder Maßnahmen schwierig. (Scherbaum et al., 2004: 157)

3.2 Soziokultureller und gesundheitspolitischer Kontext

Stillen ist eng an soziokulturelle wie auch soziodemographische Bedingungen geknüpft und stark durch den gesellschaftlichen Wandel geprägt (Przyrembel, 2001: 13). Noch in den siebziger Jahren wurde dem Stillen sowohl medizinisch als auch gesellschaftlich keine hohe Bedeutung beigemessen. Mittlerweile haben sich die Erkenntnisse über die vielfältigen Vorteile durchgesetzt. Die Bereitschaft zum Stillen steht jedoch auch in Abhängigkeit vom soziodemographischen Kontext, denn „[...] *wer in Deutschland arm ist und weniger gut ausgebildet, der stillt weniger.*“ (Przyrembel, 2004: 4) Dies ist diskrepant, da besonders sozial benachteiligte Frauen und Familien durch das Stillen erheblich an Kosten für industrielle Fertigung einsparen könnten. Nicht alle sozialen Schichten werden offenbar gleichermaßen von stillfördernden Maßnahmen erreicht. (ebd.)

¹⁵ Direkte Kosten beziehen sich auf Kosten, die direkt durch die Behandlung entstehen (z.B. Krankenhausaufenthalt, Medikamente). Indirekte Kosten umschreiben beispielsweise die Kosten durch Arbeitsunfähigkeit. (entnommen: Seminarskript zur Seminarreihe v. Prof. Dr. S. Busch, „Wirtschaftlichkeitsanalysen“, WS 2003/ 04, HAW Hamburg)

Die soziodemographischen und soziokulturellen Aspekte sind verknüpft mit einem gesundheitspolitischen Hintergrund. In diesem Kontext sind stillhindernde Entwicklungen, wie die Vermarktung von Muttermilchersatzprodukten bedeutsam. Die frühkindliche Ernährung ist durch ein reichhaltiges Angebot an artifizieller Fertigkost und Muttermilchersatzprodukten großer Unternehmen stark kommerzialisiert und durch Wirtschaftsinteressen gelenkt. Palmer (1988: 2ff.) führt aus, dass das Marketing und die Öffentlichkeitsarbeit dieser Unternehmen erheblichen Einfluss auf die Stillbereitschaft junger Mütter nehmen. Vor diesem Hintergrund verabschiedete die WHO im Mai 1981 den internationalen Kodex für die Vermarktung von Muttermilchersatzprodukten.¹⁶ Dieses verdeutlicht den gesundheitspolitischen Einfluss auf das Stillen. (Scherbaum et al., 2004: 150)

3.3 Hintergrund des WHO/ UNICEF Programms „Stillfreundliches Krankenhaus“

„Crucial for the dramatic increase of breastfeeding rates in Norway has been the fact that different Governments [...] increasingly integrated it into the health policy.“ (Nylander, 2004: 11)

Aufgrund der hohen Public-Health-Relevanz hat die Weltgesundheitsorganisation die Förderung des Stillens als einen zentralen Punkt von Gesundheitsförderung in ihre Agenda aufgenommen. Da die meisten Geburten in der westlichen Welt, national wie international, im Krankenhaus stattfinden, ist der Einfluss des „Setting“¹⁷ Krankenhaus auf das Stillverhalten hoch. Durch das frühe Anlegen, die Anleitung und Unterstützung bei Problemen stellen geburtshilfliche Kliniken die Weichen für den Stillbeginn und Stillerfolg. (Huch et al., 1996: 123) Um das Potential von Krankenhäusern als stillfördernde Einrichtung zu nutzen, riefen WHO/ UNICEF im Jahre 1991 das Programm „Baby-Friendly Hospital Initiative“ (BFHI) ins Leben.¹⁸ (UNICEF, 2005) Ziel dieser Kampagne ist, eine Rate des exklusiven Stillens von sechs Monaten zu erreichen. Es wurden „Zehn Schritte zum erfolgreichen Stillen“ definiert, deren Implementierung als Kriterium gilt, um

¹⁶ Der Kodex und die darauf folgenden Resolutionen der ‚World Health Association‘ (WHA) regeln die Grenzen des Marketings für Muttermilchersatzprodukte. Dieser Kodex versteht sich als weltweite Empfehlung zur Förderung des Stillens. (Scherbaum, 2004: 150)

¹⁷ Setting, engl.: *Umgebung, Schauplatz*, ist zum international etablierten Begriff in der Gesundheitsförderung geworden. Ein Setting wird als soziales System verstanden, in dem Menschen und ihre soziale, ökonomische und organisatorische Umwelt sowie die persönliche Lebensweise in starker wechselseitiger Beziehung zueinander stehen und dies maßgeblich die Bedingungen von Krankheit und Gesundheit bestimmen. (Grossmann et al, 2000: 100ff.)

¹⁸ Die in Deutschland etablierte Bezeichnung für die „Baby-Friendly Hospital Initiative“ lautet „Stillfreundliches Krankenhaus“.

„stillfreundliches Krankenhaus“ zu werden.¹⁹ Die Auswahl der Kriterien erfolgte nach evidenzbasierten Erkenntnissen hinsichtlich einer wirkungsvollen Stillförderung. (Byrne, 2004: 12) Guise et al. (2003: 70) belegen in einer Metaanalyse die Effektivität von stillfördernden Interventionen. In diesem Zusammenhang spielen insbesondere die Professionen des Gesundheitswesens eine wesentliche Rolle. *„The prevalence of BF [= Breastfeeding, Anm. d. Autorin] increases among the women that got health education and support from the health professionals.“* (Garcia et al., 2005: 295). Durch eine Verbesserung des Wissensstandes der jungen Mütter lässt sich eine höhere Initialrate an Stillenden sowie eine längere Stilldauer erreichen. (Palda et al., 2004: 967)

Die zertifizierten Krankenhäuser zeichnen sich dadurch aus, in besonderer Weise das Stillen zu fördern. Sie müssen sich alle 5 Jahre vor einem externen Gremium der UNICEF erneut zertifizieren, um diesen Status zu halten. (Abou-Dakn et al., 2003: 386 ff.)

Seit Start der Initiative sind mittlerweile mehr als 15 000 Einrichtungen in 134 Ländern als stillfreundliche Institution ausgezeichnet worden. In Deutschland existieren, bei einer Gesamtzahl von 1100 geburtshilflichen Abteilungen, derzeit 21 „Stillfreundliche Krankenhäuser“. (Cattaneo, 2004: 92) Drei davon befinden sich in Hamburg.²⁰ Im Jahre 2003 fanden 3% aller Geburten in Deutschland in einem der „stillfreundlichen Krankenhäuser“ statt. (ebd.)

3.4 Internationale und nationale Trends zum Thema Stillen– die Datenlage

„Less than 35% of infants worldwide are exclusively breastfed for even the first 4 months of life [...]“ (WHO, 2004)

Die „WHO Global Data Bank on Breastfeeding“ ermöglicht eine Übersicht über 94 Länder und umfasst damit circa 65% der Kinderpopulation (Alter: < 12 Monate) weltweit. (WHO, 2004) Obwohl die Stillrate innerhalb der letzten zwanzig Jahre national wie international gestiegen ist, liegen die Zahlen weiterhin unter dem von der WHO empfohlenem sechsmonatigen Zeitraum des exklusiven²¹ Stillens. (Palda et al., 2004: 170). Die insgesamt längsten Stillraten, bis zum zweitem Lebensjahr des Kindes und länger, finden sich in den Entwicklungsländern und osteuropäischen Staaten. Die Stillrate der einzelnen

¹⁹ Dies Kriterien umfassen personelle und strukturelle Bedingungen des Krankenhauses wie entsprechend geschultes Personal, ‚Rooming- In‘- Angebote, Leitlinien zur Stillförderung u.a.

²⁰ Dies sind folgende Hamburger Kliniken: Asklepios Klinik Nord/ Heidberg, Ev. Amalie-Sieveking-Krankenhaus gGmbH, Asklepios Klinik Harburg

²¹ Eine Resolution der WHO/ UNICEF empfiehlt das exklusive Stillen über 6 Monate, danach die, zeitlich und in der Auswahl, adäquate Einführung von ergänzender Kost. Eine Stillperiode über das erste Lebensjahr des Kindes hinweg wird empfohlen.(WHO, 2005)

Länder im Vergleich variiert, auch innerhalb Europas. Da europaweit kein standardisiertes Datenerfassungssystem zu Stilldaten vorliegt, ist die Datenlage unscharf. (Kersting, 2004: 5) *„In den meisten Ländern werden Daten gesammelt. Allerdings sind die erhobenen Stilldaten häufig widersprüchlich, manchmal ungenau und oftmals unvollständig.“* (Byrne, 2004: 11) Trotz der rudimentären Datenlage lässt sich ein genereller Anstieg der Stillrate weltweit erkennen.

In Deutschland liefert die 1997/98 bundesweit durchgeführte Studie SuSe²² („Stillen und Säuglingsernährung“) den bislang umfassendsten Einblick zum Thema Stillen in Deutschland. (Dulon et al., 2005: 1) Nach zwölf ausgewählten Kriterien der WHO und der Nationalen Stillkommission²³ wurden mittels einer Querschnitterhebung in 177 randomisiert ausgewählten Krankenhäusern Daten erhoben. Anschließend wurden in einer Längsschnitterhebung mit den dort entbundenen Frauen (n= 1717) Daten zum Stillen im ersten Lebensjahr generiert. Dieser integrierte Methodeneinsatz ist im Zusammenhang mit diesem Thema europaweit einzigartig. (Kersting, 2004: 6) Die Daten aus der SuSe- Studie zeigen, dass das initiale Stillen bei 91% liegt, dann jedoch rasch abfällt. So liegt die Rate des exklusiven Stillens nach 3 Monaten bei 33%, und nach 6 Monaten bei 10%. (ebd.) Weiterhin ließen sich Risikofaktoren (p= 0,01) für keine oder nur eine kurze Stillperiode identifizieren, wie beispielsweise eine geringe Schulbildung, Lebensalter der Mutter < 25 Jahre, alleinerziehende Mutter, keine Stillerefahrungen. (Dulon et al., 2005: 2) Eine Stratifizierung der Daten aus der SuSe- Studie zeigt, dass Mütter aus Ostdeutschland seltener und kürzer stillen als westdeutsche Frauen. (Dulon, 2001: 931)

Längerfristige Trends des Stillens werden seit 1985 im Rahmen der „DONALD“ - Studie („Dortmund Nutritional and Anthropometric Longitudinally Designed Study“) erfasst. In diese Kohortenstudie werden jährlich etwa 50 Säuglinge im Alter von 3 Monaten neu aufgenommen und bis zum Alter von 18 Jahren an insgesamt 29 Untersuchungsterminen begleitet. (Kersting, 2005: 2) Ziel der Studie am „Forschungsinstitut für Kinderernährung“ Dortmund (FKE) ist, Erkenntnisse zur Nährstoffzufuhr und dem Lebensmittelverzehr von Kindern und Jugendlichen zu gewinnen.

Seit Beginn der Studie ist der Anteil exklusiv gestillter Babys stetig gestiegen und entspricht den Daten der SuSe-Studie zum selben Zeitraum. (ebd.) Drei derzeit laufende

²² Ausgeschrieben vom „Bundesministerium für Gesundheit und Soziale Sicherung“ im Jahr 1997/98

²³ 1994 gegründet, um auf nationaler Ebene Stillförderung zu fördern und zu koordinieren (Przyrembel, 2001: 11)

Studien zum Thema Stillen in Deutschland lassen aktuelle Ergebnisse erwarten, die an die SuSe-Studie anknüpfen.²⁴

3.5 Stand der Evaluation stillfördernder Einrichtungen- national und international

Evaluationen der „Baby-Friendly Hospital Initiative“ aus dem Ausland belegen die Wirksamkeit dieses Programms zur Stillförderung. In den USA existieren derzeit 51 zertifizierte Einrichtungen. Am „Boston Medical Center“, Massachusetts, wurde eine Erhebung direkt vor, während und nach Implementation der Initiative durchgeführt. Dies zeigte innerhalb von vier Jahren einen Anstieg der initialen Stillrate von 58% auf 86,5%. Die Rate des exklusiven Stillens wuchs von 5,5% auf 33,5%. (Phillip et al., 2001: 677). In Brasilien lässt sich ein Anstieg der Stillprävalenz von 4% auf 42% seit Einführung der Initiative verzeichnen. (Lamounier, 1998: 19) Innerhalb Europas findet sich der höchste Zuwachs an Stillraten in den skandinavischen Ländern. Drei norwegische Surveys in einem Abstand von ca. 10 Jahren zeigen einen Anstieg des exklusiven Stillens von 1968 < 30% zu 1991 > 80% innerhalb der ersten zwölf Lebenswochen durch die Stillförderung. (Heiberg, 1995: 719) Ein vergleichbarer Trend ist auch in Schottland beobachtbar. (Broadfoot, 2005: 114) Eine qualitative Erhebung aus Wales demonstriert den Einfluss, den das Krankenhauspersonal auf die Stillbereitschaft nehmen kann. (Gould, 1994: 62 ff.)

Auch in Deutschland wird der Einfluss stillfreundlicher Kliniken auf das Stillen diskutiert. (vgl. Huch, 1996: 123- 132) Untersuchungen liegen hier darüber jedoch bislang nicht vor. Abou-Dakn et al. (2003: 387) fassen in einer Sekundäranalyse internationaler Studien die Einflussfaktoren der Entbindungsklinik auf den Stillbeginn zusammen. Weiterhin benennen sie positive Erfahrungen des Vivantes-Humboldt-Klinikums in Berlin mit den stillfördernden Richtlinien. Diese Klinik überprüft derzeit im Rahmen einer prospektiven Untersuchung den Langzeiterfolg auf die Vollzeit- und Gesamtstillquote. (ebd.)

Für die systematische Evaluation eines stillfreundlichen Krankenhauses in Deutschland konnte kein Beleg gefunden werden. Damit kann die Studie QuaSti einen wesentlichen Beitrag in diesem Forschungsfeld leisten.

²⁴ Dies sind folgende Studien:

- Lukassowitz et al., (2004) :„*Wie lange stillen Mütter in Deutschland*“, Beginn Juli 2004, Bundesinstitut für Risikobewertung (BfR)
- Müller et al., (2005): „*Stillen in Bayern*“, Beginn April 2005, Bayerisches Staatsministerium für Umwelt, Gesundheit und Verbraucherschutz
- Thefeld et al., (2002): „*Der Kinder- und Jugendgesundheitsurvey: Ermittlung des Gesundheitsverhaltens von Eltern und Kindern*“, Beginn 2003, Robert- Koch- Institut

4 Die Studie QuaSti- Entwicklung und Vorgehensweise

„Der Evaluationsgegenstand soll klar und genau beschrieben und dokumentiert werden, so dass er eindeutig identifiziert werden kann.“ [...] „Indem sie [=die Evaluation, Anm. d. Autorin] den Gegenstand systematisch beschreibt, bereitet sie dessen Bewertung vor.“ (Beywl, 2002: 10ff.)

Es erfolgt zunächst eine rein deskriptive Darstellung der Studie ohne Bewertung. Diese Ausführung stellt den Ausgangspunkt für die nachfolgende Evaluation dar.

Die Darstellung der Studie bezieht sich auf die ursprüngliche Studienplanung. Die Abweichungen, die sich bis zum jetzigen Zeitpunkt ergeben haben, sind Teil des Prozesses und damit Gegenstand der Evaluation. Sie werden deshalb zu einem späteren Zeitpunkt erläutert. (vgl. Kap. 5) Hintergrund und Programmkontext sind in Anlehnung an eine Checkliste²⁵ strukturiert, die von Evaluationsforschern zur Systematisierung von Kontexteinflüssen erstellt wurde. (Chen/ Rossi, 1983, in Christiansen, 1999: 50)

4.1 Ausgangspunkt und Forschungsfrage

Ausgangspunkt der Studie ist ein Lehrforschungsprojekt an der „Hochschule für Angewandte Wissenschaften“ (HAW), Hamburg, das im Jahre 2002/ 03 durchgeführt wurde.²⁶ In Fachkreisen ist der Trend beobachtbar, dass die Stilldauer insbesondere bei sozial benachteiligten Müttern sinkt. (Przyrembel, 2004: 4) Diese Beobachtung, auch auf Hamburger Ebene durch die „Hamburgische Arbeitsgemeinschaft für Gesundheitsförderung“ (HAG) bestätigt, gab den Anstoß zur Themenwahl. Studierende der Gesundheitswissenschaften haben sich über zwei Semester mit dem Thema Stillen und den Gründen des Nicht-Stillens beschäftigt und eine qualitativ angelegte Untersuchung dazu durchgeführt.²⁷ Über die Vorstellung der Ergebnisse entstand ein Kontakt zum Chefarzt der geburtshilflichen Abteilung der kooperierenden Klinik. In einem sich anschließendem Seminar an der HAW Hamburg entwickelten Studierende unter der

²⁵ Chen et al. (1983, 1990 in Christiansen, 1999: 50 ff.) unterscheiden zwischen Mikrokontext (z.B. Zielgruppe, Umstände der Leistungserbringung, Projektteam) und Makrokontext (z.B. nationale und internationale Einflüsse, politisches und soziales Klima).

²⁶ Das Projekt wurde von Prof. Dr. Annette C. Seibt und Prof. Christiane Deneke im Rahmen eines Fachprojektes angeboten und betreut.

²⁷ „Gründe und Hintergründe zum Thema Stillen und Nichtstillen- Ergebnisse einer qualitativen Untersuchung und Empfehlungen für stillfreundlichere Bedingungen und Unterstützung für sozial benachteiligte Mütter und Väter in Hamburg“, „Hochschule für Angewandte Wissenschaften“ und „Institut für Gesundheitswissenschaften“, Hamburg, Februar 2004

Leitung der Dozentinnen den Interviewbogen für die telefonische Befragung zum Thema Stillen. (vgl. Abschnitt 4.5)

Die Studie QuaSti ist ein Kooperationsprojekt zwischen einer Hamburger Klinik und dem „Institut für Gesundheitswissenschaften“ (InGe), ebenfalls Hamburg. Sie untersucht die Qualität des „stillfreundlichen“ Krankenhauses und zahlreiche weitere Variablen zu dieser Thematik. *„What is the breastfeeding prevalence and support seeking behavior of mothers having delivered in a baby-friendly hospital?“* (Seibt, Deneke, 2005)²⁸ Neben der Prävalenz des Stillens und Fragen zur Ernährung der Säuglinge werden zahlreiche umgebende Faktoren untersucht. Dies sind insbesondere die Qualität und Zugänglichkeit der Hilfestellung durch die MitarbeiterInnen, der Umgang bei Stillschwierigkeiten, sowie die Gründe für das Abstillen. Weiterhin werden Wissensstand und Einstellungen der Frauen bezüglich des Stillens erfasst. Auch die Nutzung von Hilfsangeboten in der Stillphase und deren Qualität auf kommunaler und persönlicher Ebene werden untersucht. Es werden Gesundheitsdaten (Körpergewicht und Nikotinkonsum) der Frauen, sowie ihr sozioökonomischer Status ermittelt.

Die Fragestellung der Studie QuaSti ist darauf angelegt, deskriptive Daten zu erheben. Ergänzend zur quantitativen Datenermittlung werden auch qualitative Daten, in Form offener Fragen, erhoben. Die Daten werden durch die direkte Befragung der Zielpopulation ermittelt. Damit kann die Studie eingeordnet werden in die Gruppe der „Health Interview Surveys“ (HIS).²⁹ (Geyer, 2003: 13) Methodisch liegt der Studie QuaSti ein epidemiologisches Querschnittsdesign zugrunde. Die Charakteristika einer Querschnittsstudie sind im folgenden Exkurs kurz dargelegt, weil sich daraus wichtige Bewertungskriterien ableiten lassen.

Exkurs: Charakteristika einer Querschnittsstudie

In der Epidemiologie wird zwischen verschiedenen Studienformen unterschieden, die je nach Fragestellung, Datenlage und Zielsetzung zum Einsatz kommen. Dabei wird differenziert zwischen deskriptiven, analytischen und experimentellen Studiendesigns. Querschnittstudien sind Observationsstudien und liefern deskriptive Aussagen über die

²⁸ Seibt, A., Deneke, C.: Abstract: „QuaSti- Breastfeeding- Promotion- Study“ für den ‚Health Promotion Section Workshop‘, EUPHA Conference (European Public Health Association) November 2005 Graz, Österreich, internes Dokument, Stand Juli 2005

²⁹ Neben der Darstellung statistischer Daten, beispielsweise in Form von Morbiditäts- und Mortalitätsstatistiken, gewinnt die Generierung von Daten über Interviews, den sogenannten „Health Interview Surveys“ zunehmende Bedeutung. (Geyer, 2003: 13) Diese umfassen z.B. den subjektiv eingeschätzten Gesundheitszustand auf Bevölkerungsebene, Einschätzung gesundheitlicher Belastungen oder die Ermittlung des Bedarfs von Angeboten für verschiedene Arten gesundheitsbezogener Versorgung oder Betreuung. (vgl. ebd.)

Verteilung bestimmter Krankheiten, Merkmale oder auch Verhaltensweisen innerhalb einer definierten Population. (Schäfer, 2004: 10) Sie werden häufig zur Ermittlung von Prävalenzen eingesetzt. Charakteristisch für dieses Studiendesign ist, dass Individualdaten erhoben werden, welche Exposition (Einflussfaktor) und Outcome (Ergebnis) innerhalb einer Population zum gleichen Zeitpunkt erfassen. Da dieses keine Aussage über den zeitlichen Ablauf der beobachteten Phänomene erlaubt, ist mit diesem Studiendesign ein kausaler Rückschluss grundsätzlich nicht möglich. (Gordis, 2001: 181)

Im Vergleich zu anderen Studiendesigns sind Querschnittsstudien relativ einfach durchzuführen und kostengünstig. Die Verteilung der interessierenden Merkmale wird entweder in der gesamten Bevölkerung oder in einer repräsentativen Stichprobe untersucht. Eine Kontrollgruppe ist nicht erforderlich. Das Studiendesign eignet sich zur Erhebung von Daten, für die keine Routinedaten (Sekundärdaten) vorliegen. (Schäfer, 2003: 22) Ihre Aussagen können dazu dienen, Hypothesen zu generieren (Geyer, 2003: 167) Querschnittsstudien werden insbesondere bei versorgungsepidemiologischen Fragestellungen, etwa zur bedarfs- und zielgruppenorientierten Planung im Gesundheits- und Sozialwesen, sowie bei Gesundheitsförderungsprogrammen eingesetzt. (Hellmeier et al., 2003: 244) Querschnittsstudien sind nicht geeignet, die Effektivität einer klinischen oder zielgruppenspezifischen Maßnahme zu testen.

4.2 Studienumfang und Untersuchungsgang

Start der Hauptstudie mit Verteilung der Demographiebögen war der 01.01.2005. Die Zielgröße für die Stichprobe wurde in der Planungsphase bei 500 mit den Frauen durchgeführten Interviews festgesetzt. Bei einer durchschnittlichen Fallzahl von etwa 2000 Geburten jährlich wurde für die Rekrutierung des geplanten Stichprobenumfangs von 500 Personen ein Zeitraum von 5- 6 Monaten berechnet. Die geschätzte Ausschöpfungsquote lag bei ca. 50- 75%. In der ursprünglichen Studienplanung wurde das voraussichtliche Ende der Datenerhebung mit Abschluss aller Interviews für August 2005 erwartet.

Die Erstbefragung mit Abfrage der demographischen Daten sowie der Einwilligung zu einem telefonischen Interview findet in der Klinik statt. Für das Ausfüllen des Demographiebogens resultieren daraus zwei unterschiedliche Zeitpunkte: wenn die Frauen sich kurz vor der Geburt zu einem ambulanten Vorstellungstermin einfinden oder aber während des stationären Aufenthaltes im Rahmen der Geburt. Dem Demographiebogen ist ein Datenschutzformular zur Einwilligung der Weitergabe der

persönlichen Daten beigefügt. Die Folgebefragung findet ca. 3 Monate nach Geburt mittels eines computergestützten Telefoninterviews (CATI) statt.³⁰ Die Frauen werden durch die Klinik vorab nochmals durch eine Postkarte an das in Kürze folgende Interview erinnert.

Die Projektmanagerin erhält den Rücklauf der Demographiebögen aus der Klinik. Diese werden von ihr nummeriert, für die Folgebefragung sortiert, auf Nutzbarkeit geprüft und in die Computer-Maske übertragen. Sie koordiniert die Verteilung und den Rücklauf der Interviews an die InterviewerInnen. Die Verteilung der Interviews an die InterviewerInnen erfolgt via E-Mail. Weiterhin steht die Projektmanagerin bei Schwierigkeiten und Fragen bezüglich technischer Probleme zur Verfügung. Für die InterviewerInnen wurde ein „Notfalltelefon“ (Projektleitung und Projektmanagement) etabliert, um bei unerwarteten Zwischenfällen eine zügige Rücksprache zu ermöglichen. In regelmäßigen Abständen finden Treffen des Projektteams statt. Die Sitzungen werden in Form von Prozessprotokollen dokumentiert. Weiterhin trifft sich das Projektteam regelmäßig mit den KlinikmitarbeiterInnen, um den jeweiligen Stand der Erhebungsphase sowie Probleme zu besprechen.

Ziel und Auftrag der Studie QuaSti ist es, der Klinik wesentliche Informationen über die derzeitige Qualität ihrer Arbeit als stillfreundliche Einrichtung zu liefern. Inwiefern die Daten Rückschluss auf die Effekte des gesundheitsfördernden Programms „Stillfreundliches Krankenhaus“ geben können, wird im Rahmen der Evaluation geprüft.

4.3 Auswahl und Rekrutierung der Stichprobe

Es liegt eine klare Definition der Zielpopulation vor. Diese konstituiert sich aus allen Frauen, die in der kooperierenden Klinik entbunden werden. Die Stichprobe wird, innerhalb eines definierten Zeitraums, ausschliesslich aus der Gruppe dieser Frauen gezogen. Die Rekrutierung der Stichprobe erfolgt durch die MitarbeiterInnen vor Ort. Diese Vorgehensweise ergibt sich aus den Rahmenbedingungen, da vor dem Hintergrund des Datenschutzes nur in der Klinik angestellte Personen primär Zugriff zu den Daten von PatientInnen der Klinik haben dürfen.

Eine Definition der Einschluss- und Ausschlusskriterien für Studienteilnehmerinnen liegt vor. Die Einschlusskriterien ergeben sich aus der oben genannten Beschreibung der Stichprobe. Um Einverständnis gebeten und interviewt werden ausschliesslich die entbundenen Frauen selbst. Als einziges Ausschlusskriterium wurde die Sprachkenntnis

³⁰ CATI= Computer Assisted Telephone Interview

gewählt: So werden vereinbarungsgemäß keine Frauen zur Teilnahme gebeten, die keine oder nur sehr geringe deutsche Sprachkenntnisse besitzen. Die Einschätzung darüber erfolgt ebenfalls ausschliesslich durch die MitarbeiterInnen vor Ort.

4.4 Das Projektteam der Studie QuaSti

Die Projektleitung besteht aus den Leiterinnen des „Instituts für Gesundheitswissenschaften“ (InGe) in Hamburg, Prof. Dr. A. C. Seibt und Prof. C. Deneke.³¹ Mit dem Projektmanagement ist Fr. U. Lüttmann, Studentin der Gesundheitswissenschaften im 6. Fachsemester betraut. Sie absolviert dies als unbezahltes Praktikum im Rahmen ihres Praxissemesters für den Zeitraum vom 01.01.2005 bis zum 31.08.2005 in Halbtags­tätigkeit. Hr. Dr. J. Pyrlík, ebenfalls „Institut für Gesundheitswissenschaften“, hat den Transfer der zuvor generierten Fragen in einen computergestützten Interviewleitfaden durchgeführt. Er übernimmt die Systemadministration der Software mit Umsetzung von Modifikationen und Verbesserungsprozessen des Fragebogens. Die Autorin dieser Evaluationsstudie, M. Jöllnbeck, ist Diplomandin der Gesundheitswissenschaften im 10. Semester. Ihre Aufgabe ist die Evaluation im Rahmen ihrer Diplomarbeit. Gleichzeitig ist sie als Mitglied des Projektteams auch inhaltlich in die Studienarbeit involviert. Als weitere MitarbeiterInnen kommen elf InterviewerInnen hinzu, die als „studentische Hilfskräfte“ die Interviews durchführen.³²

Die Zusammensetzung des Studienteams ist durch besondere Rollenkonstellationen geprägt. Die Projektleitung befindet sich sowohl für die Praktikantin, als auch für die Diplomandin in einer Anleitungssituation. Die Projektleiterinnen sind gleichzeitig auch Betreuerinnen der vorliegenden Diplomarbeit, die eine Reflexion ihres eigenen Projektes darstellt. Die Projektleitung selbst gab den Anstoß für die Evaluation der Studie und ist sich dieser „Doppelrolle“ bewusst.

Kooperierende Teilnehmer

Die Querschnittsstudie findet in Kooperation mit der Hamburger Asklepios Klinik Nord statt. Das Klinikum ist seit dem Jahr 2000 als „stillfreundlich“ zertifiziert. Im November

³¹ Prof. Dr. Annette C. Seibt: PhD (USA), Dipl.- Päd., Prof. f. Public Health, Gesundheitsförderung und Prävention; Hochschule für Angewandte Wissenschaften,(HAW) Hamburg

Prof. C. Deneke, Dipl.-Soz., Schwerpunkt Gesundheitswissenschaften und Forschung, ebenfalls Hochschule für Angewandte Wissenschaften,(HAW) Hamburg

³² Dies sind: Martina Schlappner, Britta March, Dörte Heinsohn, Elke Forman, Anja Simmet, Petra Bödecker, Susann Aronsson, Marion Müller, Kirsten Andersen, Kay Seidl, Jan Mohr

2005 wurde es durch ein externes Gutachten rezertifiziert. Der leitende Chefarzt der Geburtshilfe und der gynäkologischen Abteilung, Hr. Dr. Scheele, hat den Anstoß für die Studie in seiner Abteilung gegeben. Als leitende Hebamme ist Fr. Juppe-Schütz involviert. Die Koordination der Studie von Seiten der Klinik wird durch Fr. Lange, Qualitätsmanagerin, durchgeführt. Hr. Dr. Scheele ist Mitglied der „Nationalen Stillkommission Deutschland“.

Die Klinik ist ein Krankenhaus der Schwerpunktversorgung im Verbund der Landesbetrieb Krankenhäuser (LBK). Der Landesbetrieb Krankenhäuser, bislang durch den Hamburger Senat als Träger finanziert, steht derzeit durch eine Teilprivatisierung vor erheblichen strukturellen Veränderungsprozessen.

4.5 Die Entwicklung der Erhebungsinstrumente

Die Entwicklung der Erhebungsinstrumente wurde primär im Rahmen eines studentischen Fachprojektes geleistet. (vgl. Abschnitt 4.1) Während eines Semesters wurden zwei Fragebögen entwickelt, der Demographiebogen und das Telefon-Interview. Die Klinik formulierte im Vorfeld die von ihnen gewünschten Fragenkomplexe. Diese wurden dann von den StudentInnen zu einem Interviewleitfaden aufgearbeitet. Ergänzend dazu wurden nach ausführlicher Literaturrecherche auch wissenschaftliche Erkenntnisse zum Thema Stillen aufgegriffen. Im Rahmen des Fachprojektes wurden 73 Frauen zum Thema Stillen befragt und die Daten ausgewertet. Diese umfassende erste Erhebung stellt die Pilotstudie für die Studie QuaSti dar.

Der Demographiebogen

Zur Erhebung des soziodemographischen Profils der Stichprobe wurden 14 Fragen zusammengestellt, die Folgendes umfassen:

- persönliche Daten der Frau (Name, Adresse, Telefonnummer, Alter)
- Name und Geburtstag des Neugeborenen
- Anzahl der Personen im Haushalt
- Höchster Schulabschluss, höchster Ausbildungsabschluss, momentane Tätigkeit
- Art und Höhe des Haushaltseinkommens
- Nationalität

Der Sozialindex generiert sich aus der Summe von Belastungen bzw. Unterversorgungslagen: Einkommen (< 400 € pro Person), Schulbildung (Hauptschule oder weniger), Berufsbildung (keine Ausbildung) und Migrationsstatus (nicht deutsch).

Das Telefon- Interview

Die Pilotstudie zeigte, welche Fragen modifiziert werden sollten, um für die Hauptstudie geeignet zu sein.³³ Der Fragebogen für das Telefoninterview wurde in Abstimmung mit der Klinik entsprechend überarbeitet und von 29 auf 32 Fragestellungen, mit > 250 Variablen, erweitert. Dann wurde dieser erneut im privaten Umfeld von Mitarbeitern des Projektteams getestet. Dabei wurden Methoden in Anlehnung an ein „Cognitive Debriefing“ angewendet, d.h. es wurden aktive kognitive Techniken genutzt in Form von „Probing“ (Hinterfragen von Antworten) und „Think aloud“ (die befragte Person wird aufgefordert, während oder nach der Beantwortung der Frage Überlegungen laut zu formulieren) (vgl. Kamtsiuris et al., 2002: S102). Weiterhin wurde eine Hebamme gebeten, den Fragebogen zu prüfen und gegebenenfalls zu ergänzen. Nach wiederholter Modifikation wurde die Papierform in einen „Computer Assisted Telephone Interview“ (CATI)–Leitfaden umgesetzt. Die Software–Grundlage dafür bildet das FileMaker Inc.-Programm „FileMaker Pro 7.0v3“®. Danach wurde die „FileMaker-Maske“ erneut innerhalb des Projektteams und im persönlichen Umfeld getestet. Die ersten Interviews mit Frauen aus der Stichprobe wurden durch die Projektmanagerin und Evaluatorsin durchgeführt, um die Handhabbarkeit der Instrumente vor Einsatz der InterviewerInnen zu prüfen.³⁴

4.6 Auswahl und Schulung der InterviewerInnen

Für die Interviews wurden zunächst vier weibliche und zwei männliche Studierende rekrutiert, die der Projektleitung persönlich bekannt waren. Einige Studierende besaßen Vorkenntnisse über die Thematik, da sie TeilnehmerInnen des Fachprojekts waren. Stillen stellt eine sensible Thematik mit Befragung sehr persönlicher Aspekte dar. Aus diesem Grund wurde der Einsatz der männlichen Interviewer wegen möglicher Vorbehalte, das Interview mit einem Mann zu führen, kritisch diskutiert. Es wurde eine Sonderauswertung dieser Interviews vereinbart, um einen diesbezüglichen Interviewer-Bias frühzeitig zu erkennen.

³³ Für die Modifikation des Fragebogens wurden sowohl die Ergebnisse der Pilotstudie als auch die Erfahrungen der InterviewerInnen einbezogen. Daraus ergab sich beispielsweise eine veränderte Reihenfolge der Fragen, so dass den Befragten durch die Voranstellung von Fragen zu ihrem Kind mehr Interesse entgegengebracht wird. Weiterhin wurden Items eingefügt, z.B. zur ambulanten Entbindungsart. Einige Fragen der Pilotstudie wurden für die Hauptstudie vorkategorisiert, also in quantitative Fragen umgewandelt, um die Auswertung zu vereinfachen.

³⁴ Demographie- und Fragebogen sind als Anlage beigefügt

Die InterviewerInnen wurden im Rahmen einer mehrstündigen Veranstaltung geschult. Die Schulung vermittelte Hintergrundwissen über Studie und Telefoninterviews, sowie den Umgang mit dem Fragebogen und der entsprechenden Software. Ergänzend wurde ein schriftlicher Interview-Leitfaden zusammengestellt, der nochmals Informationen zu Thematik und Hintergrundwissen gibt. Weiterhin legt dieser Leitfaden ein standardisiertes Vorgehen für die Durchführung der Interviews fest:

- Vorgaben für die Beantwortung von Fragen durch die Teilnehmerinnen (etwa wenn Teilnehmerinnen fragen, welches Ziel die Studie hat)
- Vorgabe der Anzahl der Anrufversuche (7 Versuche), um die Teilnehmerinnen zu erreichen
- Vorgabe des Besprechens des Anrufbeantworters (nach dem 3. Versuch)

Weiterhin enthält das Skript eine Referenzliste, damit die InterviewerInnen den Studienteilnehmerinnen bei Bedarf mit einer Adresse (z.B. Krisenintervention bei Frühgeborenen, Ambulanz für Schreikinder) weiterhelfen können.

5 Darstellung der Prozessevaluation in der Studie QuaSti

Die Prozessevaluation hat in und während der Implementationsphase vor allem eine Kontroll- und Beratungsfunktion zur Optimierung des Handlungsablaufs der Studie QuaSti. Die Bewertung des Prozesses soll der Steuerung des Ablaufs dienen und frühzeitig Korrekturmöglichkeiten aufzeigen. Im Folgenden werden auch Aspekte aufgegriffen (dies sind: Auswahl des Studiendesigns, Entwicklung der Erhebungsinstrumente, Schulung der InterviewerInnen), die grundsätzlich nicht Gegenstand von Prozessevaluation sind, da sie definitorisch in den Bereich der Planungsphase gehören. Diese Aspekte werden auf Grund der elementaren Bedeutsamkeit für den gesamten Forschungsprozess dennoch bearbeitet.

5.1 Die Auswahl des Studiendesigns

„Der Studientyp sollte beschrieben und seine Wahl angemessen begründet werden“ (Hoffmann et al., 2004: 9) Wesentliches Kriterien für die Auswahl des Studiendesigns ist die zu untersuchende Fragestellung, also die Zielsetzung der Erhebung. Hurrelmann et al. (2003: 243) führen aus, dass die explizite Formulierung der Forschungsfrage Grundlage und Voraussetzung für die Planung und Bewertung des Studiendesigns und der daraus abgeleiteten Erhebungsinstrumente darstellt. Sie formulieren neben diesem Hauptaspekt vier weitere Faktoren, die bei der Wahl eines Designs berücksichtigt werden sollten. Dies sind die zeitlichen und finanziellen Ressourcen, die Möglichkeiten der Datenerhebung, Art und Häufigkeit der zu beobachtenden Situation sowie die Literaturlage zum Thema. (ebd.)

Bewertung/ Einschätzung

Die Wahl des Studiendesigns ergibt sich aus der Forschungsfrage, die das methodische Vorgehen maßgeblich bestimmt. Die zentrale Fragestellung der Studie QuaSti ist darauf angelegt, deskriptive Daten zu sammeln. Sie zielt nicht darauf ab, analytische oder experimentelle Zusammenhänge zu erfassen. Sowohl zeitliche als auch finanzielle Ressourcen der Studie sind begrenzt. Vor diesem Hintergrund ist sie als Querschnittsstudie konzipiert. Dieses Studiendesign eignet sich besonders deshalb, weil es die Erhebung von Individualdaten der interessierenden Variablen erlaubt, relativ schnell durchführbar und kostengünstig ist. Bei der Untersuchung der Klinik ist insbesondere deren Status als „stillfreundliches Krankenhaus“ bedeutsam. Das stillfreundliche Programm stellt damit die Intervention dar, um Stillen zu fördern. Die Querschnittserhebung ist zweckdienlich, Hypothesen über die Effektivität der

stillfördernden Maßnahmen zu generieren. Einen kausalen Rückschluss auf deren Wirksamkeit lässt dieses Design jedoch nicht zu.

5.2 Die Auswahl der Stichprobe

„Sowohl die interne Validität als auch die Generalisierbarkeit der Ergebnisse sind im hohem Maße von [...] dem Auswahlverfahren der Studienteilnehmer abhängig.“
(Hoffmann et al., 2004: 9)

Die Stichprobe muss in ihrer Zusammensetzung charakteristisch für die Bevölkerung sein, auf die die Studienergebnisse bezogen werden. (Schäfer, 2004: 86) Vor Auswahl der Stichprobe bedarf es einer exakten Definition der Zielpopulation, um einen sog. „sampling frame“ (Auswahlrahmen) bestimmen zu können. Bei der Auswahl der Stichprobe ist zu beachten, dass sie in allgemeinen Merkmalen (wie z.B. Alter, Geschlecht) der Grundgesamtheit der definierten Zielpopulation ähnelt. (Schäfer, 2003: 24) Einschluss- und Ausschlusskriterien sollten bei der Planung definiert und begründet werden. (Hoffmann et al., 2004, 10)

Bewertung/ Einschätzung

Die Auswahl der QuaSti-Stichprobe ergibt sich aus dem Forschungsvorhaben. Eine exakte Definition der Zielpopulation liegt vor. (vgl. Abschnitt 4.3) Es wird ausschließlich aus dieser Gruppe eine Stichprobe gezogen. Die biologischen Merkmale, die sich aus dieser Vorgabe ergeben, (nämlich die Geschlechtsspezifität „weiblich“ und eine Altersbegrenzung auf das „fertile“ Alter), lassen hinsichtlich dieser Eigenschaften eine Repräsentanz der Ergebnisse, bezogen auf die Zielpopulation vermuten. Damit entfällt diesbezüglich die Notwendigkeit des Matching oder anderer Methoden zur Abwehr von Bias, die primär durch das Auswahlverfahren verursacht werden könnten (beispielsweise eine Stratifizierung nach Alter oder Geschlecht).

Stimmungen und das in einer Situation präsente Wissen determinieren das Antwortverhalten in einer Befragung. (Geyer, 2003: 91) Ein charakteristisches Merkmal der Zielpopulation ist, dass es bei der Befragung um ein besonderes Ereignis im Leben der Befragten geht: die Geburt ihres Kindes und die Stillzeit. Dieses wird in der Regel sehr gut erinnert. Die Interviews werden zeitnah durchgeführt. Deshalb sind hier keine bedeutsamen Verzerrungen durch ein eingeschränktes Erinnerungsvermögen, ein Recall-Bias, zu erwarten. Dies ist unabhängig davon, ob zum Zeitpunkt des Interviews noch gestillt wird oder bereits abgestellt wurde. Das charakteristische Merkmal findet sich auch in der Stichprobe wieder, so dass sich auch hier kein Potential für Verzerrungen ergibt.

Eine Definition der Ein- und Ausschlusskriterien für die Stichprobe liegt ebenfalls vor. Befragt werden ausschliesslich die entbundenen Frauen selbst. Eine abschließende Analyse der Stichprobe erfolgt im Rahmen der Ergebnisevaluation.

5.2.1 Die Stichprobenziehung

„Sind bei der Stichprobe bestimmte Bevölkerungsgruppen nicht angemessen vertreten, wird sich bei der Auswertung ein systematischer Fehler, ein sog. Selektions-Bias, einschleichen [...]“ (Schäfer, 2004: 26)

Bei der Rekrutierung der Stichprobe muss sichergestellt werden, dass sie zu einer repräsentativen Stichprobe in Bezug auf die Zielpopulation führt. (ebd.) Bei der Stichprobenauswahl sollte deshalb randomisiert werden. (Gordis, 2001: 181)

Bewertung/ Einschätzung

Aus dem Rekrutierungsverfahren, mit Auswahl der Stichprobe durch die KlinikmitarbeiterInnen, lassen sich vier relevante Einflussfaktoren ableiten:

1. Als Ausschlusskriterium zur Studienteilnahme gilt eine unzureichende Sprachkenntnis. So werden keine Frauen zur Teilnahme gebeten, die keine oder nur sehr geringfügige Kenntnis der deutschen Sprache besitzen. Die Auswahl der Studienteilnehmerinnen findet dabei nach einer persönlichen Einschätzung der auswählenden MitarbeiterInnen statt. Hierbei bestehen keine objektivierbaren Kriterien darüber, wie unzureichende Sprachkenntnisse exakt definiert sind. Daraus resultiert ein möglicher Ausschluss zur Befragung geeigneter Probandinnen. Im Umkehrschluss könnten Frauen trotz unzureichender Deutschkenntnisse in die Stichprobe gelangen. (vgl. Abschnitt 6.1.2)
2. Weiterhin ist zu erwähnen, dass die Klinik einen Nutzen aus „positiven Ergebnissen“ der Studie QuaSti zieht. Möglicherweise könnten daher besonders die Frauen rekrutiert werden, die offensichtlich bezüglich der Geburtsklinik oder des Stillens positiv gestimmt sind. Grundsätzlich ergibt sich aus den Zusammenhängen von Punkt 1 und 2 das Potential für Selektions-Bias.
3. Hinsichtlich des Selektions-Bias muss weiterhin bedacht werden, dass auch von Seiten der Studienteilnehmerinnen Verzerrungen entstehen können, nämlich durch ihre Bereitschaft zur Teilnahme. So ist denkbar, dass Frauen, für die Stillen ein sensibles Thema darstellt, bzw. Frauen, die nicht stillen möchten, eine Teilnahme ablehnen (Non-Responder). Im Umkehrschluss werden Frauen mit positiver Grundeinstellung zum Thema Stillen zur Teilnahme möglicherweise

besonders motiviert sein. Diese Form des Selektions-Bias wird Response-Bias genannt. (Schäfer, 2003: 24)

4. Auch Rahmenbedingungen wie beispielsweise Stress und zeitliche Überlastung nehmen Einfluss auf die Rekrutierung der Stichprobe. Da diese aber nicht beeinflussbare Faktoren darstellen, die Verzerrungen also zufällig sind, liegt hier kein systematischer Fehler vor.

Die Kontrolle darüber, ob die Stichprobe vollständig randomisiert erhoben wird, entzieht sich der Einflussnahme der Projektleitung. Auch eine Verzerrung durch Non- Responder und Response-Bias sind nicht beeinflussbar. Dies stellt ein großes Problem dar. Die Projektleitung ist sich dieser Problemstellungen bewusst und hat sie mit den MitarbeiterInnen der Klinik thematisiert. Es besteht Konsens zwischen den Kooperationspartnern, dass eine randomisierte Stichprobe ein wesentliches Gütekriterium der Studie QuaSti darstellt. Die Bewertung der Stichprobenziehung bezieht sich auf den laufenden Forschungsprozess. Eine abschließende Analyse dieser Problematik findet im Rahmen der Ergebnisevaluation statt. (vgl. Abschnitt 6.1)

5.2.2 Zeitlicher Verlauf der Rekrutierung

Ein wesentlicher Vorteil einer Querschnitterhebung ist der relativ geringe Ressourcenbedarf. Die Studiendauer ist dabei ein wesentlicher Faktor, da sie Einfluss nimmt auf die Kosten, die beispielsweise durch den personellen, den zeitlichen und den materiellen Aufwand entstehen. (Schäfer, 2004: 28)

Bewertung/ Einschätzung

Für die Rekrutierung der Stichprobe wurde ein Zeitraum von etwa sechs Monaten festgelegt. Die zuvor geschätzte Rücklaufquote der Demographiebögen, anhand der durchschnittlichen Geburtsrate ermittelt, erwies sich jedoch als nicht realistisch.

Rückfragen an die Klinik ergaben keine eindeutigen Rückmeldungen über die Problematik vor Ort, die zu der Verzögerung führten. So können lediglich Annahmen darüber gemacht werden. Gründe könnten beispielsweise personelle Überlastung durch hohen Arbeitsanfall oder auch mangelnde Motivation einzelner MitarbeiterInnen sein. Ein praktisches Problem vor Ort war weiterhin, das der Begleitbogen über den Datenschutz, ein betriebsinternes Dokument, zeitweilig nicht mehr zur Verfügung stand, so dass auch die Demographiebögen nicht verteilt werden konnten. Ferner wurde ein kleiner Anteil der Demographiebögen irrtümlich nicht den Patientenakten entnommen und fand sich deshalb im Archiv wieder.

Wochen(Monat)	2004 (Nov./ Dez.)	4 (Jan.)	8 (Feb.)	12 (März)	16 (April)	20 (Mai)	24 (Juni)	28 (Juli)	Gesamt
Ist/ Soll Rücklauf Demographiebögen									
Ist	27	10	20	54	23	42	19	24	219
Soll		100	100	100	100	100	—	—	500

Tab. 4: Übersicht über den Rücklauf der Demographiebögen

Für die Zwischenauswertung wurde der Datensatz am 31.10.2005 geschlossen. Zu diesem Zeitpunkt lagen 219 bearbeitete Interviews vor. Im Konsens zwischen Projektleitung und Klinik wird eine Stichprobe von 400 durchgeführten Interviews angestrebt. Bei einer erwarteten Ausschöpfungsquote von ca. 75% bedeutet dies, dass insgesamt ca. 600 Datensätze vorliegen müssen. Aus diesem Grund wurde der Zeitraum für die Rekrutierung der Stichprobe bis zum 31.11.2005 verlängert.

Aus der verlängerten Laufzeit ergibt sich während des laufenden Prozesses ein höherer Ressourcenverbrauch:

- Rekrutierung und Schulung einer weiteren Interviewerinnengruppe, da die erste Gruppe nicht mehr vollständig zur Verfügung steht
- Gewährleistung und Bezahlung des Projektmanagements bis zum Ende der Erhebungsphase Ende Februar 2006 (Tätigkeit der Projektmanagerin bis zum 31.08.2005 im Rahmen eines unbezahlten Praktikums)
- zeitliche Ressourcen der Projektleitung, da sich die Studie nun anders als geplant bis ins Wintersemester ausdehnt und mit anderen Terminplanungen des laufenden Semesters kollidiert
- zeitliche Verzögerung bis zum Vorliegen relevanter Daten für die Diplomandin
- verlängerter zeitlicher und personeller Aufwand für die KlinikmitarbeiterInnen, die Demographiebögen zu verteilen

Die zeitliche Verzögerung bei der Rekrutierung der Stichprobe führt zu einer Verlängerung der Untersuchungsphase. Zusammenfassend lässt sich feststellen, dass dies ungünstige Auswirkungen insbesondere auf den zeitlichen und finanziellen Ressourcenverbrauch hat.

5.3 Entwicklung und Einsatz der Erhebungsinstrumente

„Die Entwicklung eines Fragebogens gehört zu den wichtigsten Aufgaben im Rahmen eines Forschungsprozesses. Wer hier Fehler macht, gefährdet die zu realisierenden Untersuchungsziele, selbst wenn in allen anderen Phasen alles richtig gemacht wurde.“ (Elkeles, 2004: 40)

Erhebungsinstrumente müssen eine Reihe von Kriterien erfüllen, um die Informationen zu erhalten, die zur Beantwortung der Forschungsfragen erforderlich sind. Die Validität und Reliabilität der Erhebungsinstrumente entscheidet über die Datenqualität. (Geyer, 2003:91ff.) Aus diesem Grund sollten, sofern möglich, bereits validierte und erprobte Erhebungsinstrumente Einsatz finden. (Schäfer, 2004: 83) Bei der Entwicklung eines neuen Erhebungsinstrumentes müssen zahlreiche Aspekte berücksichtigt werden.

Die Evaluation der Erhebungsinstrumente setzt sich aus der Bewertung folgender Aspekte zusammen: der Entwicklung des Erhebungsinstruments mit Pilotstudie, des Demographiebogens und des Telefoninterviews.

5.3.1 Die Entwicklung des Erhebungsinstruments mit Pilotstudie

„In jeder epidemiologischen Untersuchung, bei der Primärdaten erhoben werden, ist zu prüfen, ob vor Beginn der Hauptstudie eine separate Pilotstudie erforderlich ist. Eine Pilotstudie wird in aller Regel erforderlich sein, wenn ein neues Erhebungsinstrument eingesetzt werden soll, [...]“ (Hoffmann et al., 2004: 15)

Die Erarbeitung eines neuen Erhebungsinstruments setzt zahlreiche Kriterien voraus. Es bedarf einer umfassenden Literaturrecherche zum Thema. Dies ist wichtig zur Einordnung in einen wissenschaftlichen Kontext und um Doppeluntersuchungen sowie unsinnige Fragestellungen zu vermeiden. (Schäfer, 2004: 82) Eine weitere Voraussetzung stellt die fachliche Qualifikation der MitarbeiterInnen dar, die an der Erarbeitung mitwirken. Weiterhin bedarf es zur Überprüfung des Instruments einer Pilotstudie, um damit Verständlichkeit, Handhabbarkeit und Konsistenz zu überprüfen. Dabei muss eine entsprechende Stichprobengröße beachtet werden, damit tatsächlich eine Validierung des Fragebogens möglich ist. (Geyer, 2003:135) Der Umfang der Pilotstudie hängt von der Zielsetzung wie auch Umfang der Hauptstudie ab, in der Regel werden 20- 50 Interviews als ausreichend eingeschätzt. (Siegrist, 2003: 272)

Bewertung/ Einschätzung

Für die Studie wurde ein neues Erhebungsinstrument generiert, da bereits erprobte und validierte Instrumente, die für den Forschungsauftrag der Studie QuaSti geeignet wären,

nicht zur Verfügung standen. Die Entwicklung des Erhebungsinstruments ist aus zwei sich ergänzenden Lehrforschungsprojekten an der „Hochschule für Angewandte Wissenschaften“ hervorgegangen. (vgl. Abschnitt 4.1) Die Dozentinnen des Fachprojekts verfügen über große Erfahrung im wissenschaftlichen Forschungsbereich. Die Rahmenbedingungen erlaubten allen Beteiligten eine ausführliche, thematische und methodische Auseinandersetzung mit umfassender Literaturrecherche. Durch die intensive Bearbeitung haben die Studierenden ein hohes Maß an methodischer und thematischer Kompetenz erlangt.³⁵ Damit sind wesentliche Gütekriterien zur Entwicklung des Erhebungsinstruments erfüllt. Im Rahmen der Pilotstudie wurden 73 Frauen befragt. Dies stellt einen sehr umfassenden Stichprobenumfang dar.

5.3.2 Der Demographiebogen

„Die Definition, durch die der Sozialstatus in dieser Studie bestimmt wird, muss im Vorfeld festgelegt werden und bestimmt maßgeblich das Design des Erhebungsinstrumentes mit.“
(Schäfer, 2004: 82)

Die Operationalisierung des Sozialstatus sollte auf der Basis allgemein anerkannter Definitionen für Sozialstatus formuliert werden, um eine Vergleichbarkeit mit anderen Studien zu ermöglichen. (ebd.) Die „Deutsche Arbeitsgruppe für Epidemiologie“ (DAE) hat in diesem Zusammenhang Empfehlungen zur „Messung und Quantifizierung soziographischer Merkmale in epidemiologischen Studien“ veröffentlicht.³⁶ (Jöckel et al., 1997: 1ff.)

Bewertung/ Einschätzung

Der Demographiebogen wurde, wie unter Abschnitt 4.5 ausführlich erläutert, auf der Basis von Indikatoren von Belastungen und Unterversorgungslagen (wie Netto-Haushaltseinkommen, Schulbildung und Familienstand) entwickelt. Die Auswahl der Indikatoren spiegelt die Empfehlungen der DAE detailliert und stellt ein adäquates Instrument zur Erstellung des soziographischen Profils der Studienpopulation dar.

³⁵ Die hohe Qualität der Pilotstudie ist besonders vor dem Hintergrund hervorzuheben, dass die Studierenden sich zum Start des Lehrforschungsprojektes im Grundstudium befanden, d.h. eine Reihe methodischer Grundlagen zunächst gelernt werden mussten.

³⁶ Epidemiologische Studien zeigen, dass der soziale Status eine wichtige einflussnehmende Determinante auf Krankheit, Expositionen und Risikofaktoren darstellt. Es liegen jedoch zahlreiche, theoretische Konstrukte und Operationalisierungen des Begriffs vor, so dass eine Vergleichbarkeit kaum möglich ist. Vor diesem Hintergrund hat die DAE Empfehlungen formuliert, um einen Konsens zu schaffen und Vergleichbarkeit herzustellen (Jöckel et al., 1997:1)

Aus dem Zeitpunkt des Verteilens des Demographiebogens resultiert ein Problem hinsichtlich des Items „Anzahl der Personen im Haushalt“: Der Demographiebogen wird zu zwei unterschiedlichen Zeitpunkten verteilt (entweder zum ambulanten Vorstellungstermin vor Geburt oder während des stationären Aufenthaltes nach Geburt). So ist das Neugeborene nicht stets mitberechnet bei der Angabe der Haushaltsgröße. Damit sind diese Angaben nicht valide und damit nicht nutzbar. Das Problem wurde von der Projektmanagerin bei ersten groben Prüfungen auf die Konsistenz der Angaben erkannt und die Vorgehensweise darauf hin korrigierend verändert. Für den weiteren Verlauf der Untersuchungsphase wird diese Frage im eigentlichen Interview erneut aufgegriffen und erfragt. Die bereits geführten und nicht mehr korrigierbaren Interviews sind bezüglich dieses Items separiert.

Auch bei dem Item „Netto-Haushaltseinkommen“ liessen sich im Demographiebogen zahlreiche Antwortlücken feststellen. Auch hier erfolgte eine Korrektur im laufenden Prozess mit erneuter Abfrage des Items im Telefoninterview. Bezüglich der Angaben über das Einkommen ist anzumerken, dass dies in Untersuchungen häufig eine problematische Frage darstellt, weil sie einerseits ungern beantwortet wird und andererseits oft fehlerhaft angegeben wird. (Jockel et al., 1997:9) Das Haushaltseinkommen wird häufig unterschätzt, da meist nur das Arbeitseinkommen oder die Rente ohne Berücksichtigung weiterer Einkommensarten (z.B. aus Vermögen, Mieteinkünften) benannt wird. (ebd.) Deshalb sind die Antworten meist nur als Trend interpretierbar. Diese Problematik wurde von der Projektleitung bereits formuliert. Eine Analyse des Items erfolgt in der Ergebnisevaluation. (vgl. Abschnitt 6.1.3.2)

5.3.3 Das Telefoninterview

„Alle interessierenden Variablen sollen präzise definiert und möglichst standardisiert operationalisiert werden. Für die Bestimmung sind möglichst valide und reliable Mess- und Erhebungsinstrumente einzusetzen.“ (Hoffmann et al., 2004:11)

Für die Frageformulierungen bestehen nach Elkeles (2004: 41ff.) eine Reihe von Kriterien, die bei der Operationalisierung der Fragen Berücksichtigung finden sollten:

- Begründungszusammenhänge für die Antworten, um die Ergebnisse interpretieren zu können
- Beantwortbarkeit der Fragen durch den Befragten einschätzen (Wissen, Erinnerungsfähigkeit, Akzeptanz, Informationsstand)
- Begriffe klar definieren (z.B. Haushaltsgröße, Entfernungen)

- Fachtermini vermeiden
- Mehrdimensionalität der Fragen vermeiden (keine und/ oder Verbindungen)
- Keine Suggestivfragen
- Auf soziale Erwünschtheit im Antwortverhalten achten
- Auf logisch und thematisch plausible Abfolge der Fragen achten
- „Keine Angabe“ und „Weiß nicht“ Kategorien vorgeben
- nur Fragen stellen, die unabdingbar sind für die Untersuchung
- Wissenschaftliche Erkenntnisse bei der Fragestellung berücksichtigen

Ist vor Beginn der Hauptstudie eine Pilotstudie erforderlich, sollte diese ausgewertet und dokumentiert vorliegen, um evtl. erforderliche Modifikationen für die Hauptstudie durchführen zu können. (Hoffmann et al., 2004: 15)

Weiterhin sollte bereits im Interview der Einsatz von Anreizen zur Teilnahme (sog. Incentives) Berücksichtigung finden, um die Teilnahme (Response) zu erhöhen. (Geyer, 2003: 64) Vor Durchführung der Interviews ist es vorteilhaft, den Anruf schriftlich anzukündigen, um zu vermeiden, dass eine hohe Verweigererquote durch misstrauische Befragte entsteht. (Siegrist et al., 2003: 264)

Bewertung/ Einschätzung

Die Entwicklung des Fragebogens geht aus einer qualitativ hochwertigen Pilotstudie hervor, die vor Beginn der Hauptstudie ausgewertet wurde. Damit liegt ein erprobtes Erhebungsinstrument vor. Der primäre Fragebogen wurde mehrfach modifiziert. Vor dem Einsatz im Feld wurde der Fragebogen mehrfach erprobt. Anregungen, die sich im Verlauf durch die Erfahrung der InterviewerInnen im Feld ergeben, wurden zeitnah aufgegriffen. Der Fragebogen unterliegt damit einer laufenden Modifikation und Qualitätskontrolle. Ein besonders positiver Aspekt stellt die Rücksprache und Austausch mit einer erfahrenen Hebamme vor dem Einsatz des Fragebogens im Feld dar. Durch diese Maßnahme gewinnt der Fragebogen an Qualität, weil Praxisbezug und Expertenwissen der Profession einbezogen wird. (vgl. Siegrist et al., 2003: 271)

Die oben aufgeführten Kriterien zur Operationalisierung von Fragen fanden bei der Fragebogenentwicklung Berücksichtigung. Es ist ein standardisierter Fragebogen, ergänzt durch offene Fragestellungen, entstanden. Die wissenschaftlichen Erkenntnisse der SuSe- Studie über wesentliche Risikofaktoren ($p=0,01$) für eine Stilldauer von weniger als 4 Monaten werden von der QuaSti- Erhebung inhaltlich explizit und umfassend

aufgegriffen. (vgl. Abschnitt 3.4) Auch die Frage nach Nikotinkonsum und mütterlichem Übergewicht wurde vor dem Hintergrund wissenschaftlicher Erkenntnisse aufgegriffen, da bekannt ist, dass diese Faktoren Einfluss auf die Stlldauer haben. (Horta et al., 2001: 304)

Der Incentive zur Teilnahme geschieht in Form eines ideellen Anreizes. In der Einleitung werden die Studienteilnehmerinnen auf die Bedeutsamkeit ihrer Angaben hingewiesen, und darauf, dass sie damit einen wichtigen Beitrag zur Verbesserung der Situation junger Familien leisten. Dies ist in diesem Kontext als Anreiz zur Teilnahme ausreichend. (vgl. Lavrakas, 1993: 133) Weiterhin erhielten Teilnehmerinnen vor dem Anruf eine Postkarte zur Vorankündigung. Dies ist vorteilhaft, um eine höhere Ausschöpfungsrate zu erreichen. Somit sind zahlreiche Kriterien zur Entwicklung eines validen und reliablen Erhebungsinstruments erfüllt.

Technische Aspekte des Fragebogens- Das CATI- System

Die Qualität von Telefoninterviews erhöht sich durch den Einsatz einer computergestützten Erhebung mit dem CATI- System erheblich. (Diekmann, 2000: 549) Neben dem Eingabekomfort für die InterviewerInnen bietet CATI zahlreiche Möglichkeiten, um fehlerhafte Eingaben zu minimieren. (z.B. durch Fehlersuchstrategien, Zurückweisung falscher Eingaben). (Siegrist et al., 2003: 264)

Bewertung/ Einschätzung

Der Einsatz des CATI- Systems in der Studie QuaSti hat positive Auswirkungen auf die Qualität. Dies begründet sich durch mehrere Aspekte:

- Hoher Komfort in der praktischen Durchführung steigert die Motivation der InterviewerInnen (Rücksprache mit einzelnen InterviewerInnen und teilnehmende Beobachtung während der Schulungen haben dies gezeigt)³⁷
- Laufende und zügige Verbesserung durch den Systemadministrator
- Fehlersuchstrategien, Filter bezüglich falscher/ungültiger Daten
- Automatische Dokumentation des Durchführungsdatums und der Interviewlänge
- Erleichtert Stichproben zur Kontrolle
- Erleichtert die Durchführung der Plausibilitätskontrolle
- Ermöglicht Datentransfer in das statistische Auswertungsprogramm SPSS®

³⁷ Die InterviewerInnen wurden im persönlichen und telefonischen Gespräch, in nicht standardisierter Form, von der Evaluatorin zur Handhabbarkeit des CATI- Systems befragt

5.4 Auswahl, Schulung und Kontrolle der InterviewerInnen

5.4.1 Die Auswahl der InterviewerInnen

„Das Datenerhebungspersonal ist sorgfältig auszuwählen und die soziale und fachliche Qualifikation sicherzustellen.“ (Hoffmann et al., 2004:15)

Die Auswahl der Personen für die Datenerhebung ist ein wesentlicher Aspekt im Rahmen der Datenerhebung. Neben der fachlichen Kompetenz sind auch auf soziale und persönliche Eignung zu achten. Dies ist insbesondere von Bedeutung, wenn Interviews zu sensiblen Themen durchgeführt werden. (z.B. Fragen zum Einkommen, zur Sexualität) Allen InterviewerInnen sollten gleichermaßen in der Lage sein, eine Vertrauensbasis mit den StudienteilnehmerInnen herstellen zu können. Ist dies nicht der Fall, kommt es zu sog. Inter-Observer-Verzerrungen. (Schäfer, 2004:90)

Bewertung/Einschätzung

Die Auswahl der InterviewerInnen erfolgte nach der persönlichen Einschätzung der Projektleitung. Alle Personen sind ihnen persönlich bekannt und teilweise TeilnehmerInnen des Lehrforschungsprojektes zum Thema Stillen. Die Teilnahme der männlichen Studenten wurde gesondert diskutiert und entschieden. Die Gefahr des Interviewer-Bias durch die männlichen Interviewer bei der Thematik Stillen wurde von der Projektleitung antizipatorisch aufgegriffen. Es wurde eine frühzeitige Zwischenauswertung dieser Interviews, insbesondere hinsichtlich der Ausschöpfungsquote, vereinbart und durchgeführt. (vgl. Abschnitt 5.4.3)

Die Auswahl der InterviewerInnen stützt sich maßgeblich auf die Einschätzung der Projektleitung über deren fachliche, persönliche und soziale Eignung. Diese Vorgehensweise stellt ein adäquates Verfahren bei der Auswahl geeigneter InterviewerInnen dar.

5.4.2 InterviewerInnen- Schulung

„Vor Beginn der Feldarbeit sollen die an der Datenerhebung beteiligten Personen ausführlich geschult und ausgebildet werden.“ (Hoffmann et al., 2004:15)

InterviewerInnen kommt im Rahmen der Umfrageforschung mittels Telefon eine zentrale Rolle für die Erhebung und Qualität der Daten zu. Neben dem Fragenstellen müssen die Antworten (korrekt) erfasst werden, entsprechend eingegeben und die Kooperationsbereitschaft des Gesprächspartners gesichert werden. Dies verdeutlicht, dass das Potential für Interviewer-Bias in solchen Untersuchungen hoch ist. (Lavrakas,

1993: 13) Hieraus lässt sich Reihe von Anforderungen an InterviewerInnen ableiten. “[...] *interviewers are the medium through which measurements are taken in [personal and telephone] surveys.*” (Groves, 1989, zit. n. Lavrakas, 1993: 124) Die InterviewerInnen sollten vor Beginn der Feldarbeit nach einem festgelegten Protokoll geschult werden. Dadurch wird ein standardisiertes Vorgehen bei der Durchführung der Interviews erreicht. Dies ist insbesondere von Bedeutung, wenn mehrere Personen an der Datenerhebung beteiligt (Beispiel: Interviewer A definiert „Abstillen“ ganz anders als Interviewer B) Mit einem standardisierten Verfahren wird Kongruenz in der Vorgehensweise sichergestellt. Dies minimiert die Gefahr von Interviewer-Bias und erhöht die Reliabilität. Regeln und Erläuterungen für die Durchführung der Erhebung sollen in Form eines Erhebungshandbuches schriftlich vorliegen und den InterviewerInnen zur Verfügung stehen. (Hoffmann et al., 2004:16) Die Interviewerschulung sollte so gestaltet werden, dass eine intrinsische Motivation und eine Identifikation mit dem Gesamtprojekt erreicht wird. Dies reduziert die Problematik gefälschter Dateneingaben. (Wottawa et al., 2003: 151)

Bewertung/ Einschätzung

In einer mehrstündigen Veranstaltung wurden die InterviewerInnen umfassend und in kongruenter Weise geschult, um einen einheitlichen Wissensstand aller an der Datenerhebung Beteiligten herzustellen. Dies ist eine wichtige Voraussetzung für Qualität der erhobenen Daten. Zusätzlich erhielten die InterviewerInnen ein Skript mit standardisierten Verfahrensweisen und Antworten. Die Betreuung der InterviewerInnen im laufenden Prozess wird nahezu ausschliesslich von der Projektmanagerin durchgeführt. Hier liegt ebenfalls eine Kongruenz hinsichtlich der Hilfestellung und des Informationsflusses vor, weil sie durch nur eine Person erfolgt. Die beigefügte Referenzliste wie auch die Einrichtung eines „Notfalltelefons“ erhöht die Qualität der Datenerhebung, da daraus eine ethische Verantwortung gegenüber den InterviewerInnen wie auch Studienteilnehmerinnen seitens der Projektleitung abgeleitet werden kann.

Trotz der Schulung ergaben sich technische Probleme der InterviewerInnen mit dem CATI-System. Der erläuternde Begleitbogen für den Einsatz des CATI-Systems wurde daraufhin von der Projektmanagerin mehrfach modifiziert und den Informationsbedürfnissen der InterviewerInnen angepasst.

Durch die Vorgehensweise in der Schulung der InterviewerInnen ist ein wesentliches Kriterium für eine qualitativ hochwertige Datenerhebung, insbesondere hinsichtlich der Minimierung von Interviewer-Bias, erfüllt worden.

5.4.3 Kontrolle in der Durchführung der Interviews

„Die wesentliche Voraussetzung für die Aufrechterhaltung der Qualität von Interviews liegt in einer gründlichen Einschulung und der kontinuierlichen Supervision des Interviewerstabes durch Identifikation von Schwachstellen und deren Beseitigung.“
(Geyer, 2003:47)

Eine Überprüfung der InterviewerInnen in der laufenden Untersuchungsphase wird in der Literatur kontrovers diskutiert. Insbesondere große Survey-Institute mit einer großen Gruppe von Datenerhebungspersonal führen stichprobenartig Kontrollen durch, indem Zielpersonen nochmals kontaktiert und zum Interview befragt werden. Informationen über das tatsächliche Ausmaß gefälschter Interviews sind rar. (Geyer, 2003: 49) Die Art der Fälschungen reicht dabei von Totalfälschung (Interviewer füllt Bogen selber aus ohne Kontakt zum Befragten), Teilfälschung (es werden nur Basisdaten erfragt, der Rest wird eigenmächtig ergänzt) bis hin zu Verletzung der Auswahlregeln und wissentlich fehlerhaftes Ausfüllen des Bogens, um das Interview zeitlich abzukürzen. Studien zeigen, dass Fälschungen insbesondere bei sehr umfangreichen Studien mit vielen InterviewerInnen anzutreffen sind. (ebd.)

Bewertung/ Einschätzung

Vor dem Hintergrund des Vertrauensverhältnisses der Projektleitung zu den InterviewerInnen wurden keine Kontrollen, beispielsweise durch Anrufe bei den Studienteilnehmerinnen, vereinbart. Deshalb entfällt eine systematische Kontrolle der InterviewerInnen. Aus organisatorischen Gründen war jedoch notwendig, einige Frauen wegen fehlender Daten nochmals anzurufen. Als Nebeneffekt zeigte sich, dass alle diese Frauen auch tatsächlich von den InterviewerInnen kontaktiert wurden.

Die Gruppe der InterviewerInnen ist in einer überschaubaren Größenordnung angelegt. Weiterhin wurden sie von der Projektleitung persönlich ausgewählt. Vor diesem Hintergrund ist es gerechtfertigt, keine systematischen Kontrollen während der Untersuchungsphase durchzuführen.

Kontrolle der männlichen Interviewer

Vereinbarungsgemäß erfolgte die Kontrolle der männlichen Interviewer nach Abschluss ihrer Interviews. Ein Vergleich zwischen männlichen und weiblichen Interviewern zeigt, dass die männlichen Interviewer eine Ausschöpfungsquote erreichen, die sich von denen der Interviewerinnen nicht unterscheidet. Dazu wurde die individuelle

Ausschöpfungsquote der InterviewerInnen ermittelt, die zwischen 70- 100% liegt.³⁸ Eine separate Analyse der Nicht- Teilnehmerinnen der männlichen Interviewer ergab, dass keine der Frauen das Interview abbrach, weil die Interviewer männlich waren. Weiterhin wünschte keine der Teilnehmerinnen zu Beginn des Gesprächs, von einer Frau interviewt zu werden. Das einige Teilnehmerinnen möglicherweise in der Befragung eine Interviewerin präferiert hätten, dies jedoch nicht offen darlegen mochten, erscheint theoretisch möglich, ist jedoch aus Sicht der Evaluatorin nicht als relevante Größe von Bedeutung. Zusammenfassend lässt sich daraus schliessen, dass der Einsatz männlicher Interviewer zu keinem Interviewer-Bias geführt hat. Die Projektleitung hat die Problematik in der Planungsphase der Studie antizipatorisch aufgegriffen und damit die Gefahr von diesbezüglichen Verzerrungen minimiert.

5.5 Rücklauf der Interviews

„Die telefonischen Umfragen [sic] erreichbaren Ausschöpfungsraten hängen neben der Zahl der Kontaktversuche, den Charakteristika der Befragten auch von den Anrufzeiten ab.“ (Geyer, 2003: 79)

Bei der Durchführung der Feldarbeit sollte eine Rücklaufkontrolle der Interviewbögen erfolgen, um die Ausschöpfungsquote im laufenden Prozess verfolgen zu können. (Elkeles, 2004: 50) Dies ermöglicht eine frühzeitige Korrektur, beispielsweise bei der Wahl der Anrufzeit, im laufenden Prozess. Wiseman & Mc Donald (1979, in Geyer, 2003: 70) generierten als wesentliche Faktoren für die Teilnahmebereitschaft:

- Zeitpunkt des ersten Kontakts, d.h. in welcher Verfassung der zu Befragende steht
- Ob ein Rückruf vereinbart wird, wenn das Interview nicht sofort erfolgen kann
- Der Aufwand, jemanden von der Teilnahme zu überzeugen

Grundsätzlich ist es notwendig, mehr als einen Kontaktversuch zu unternehmen, um eine zufrieden stellende Ausschöpfungsquote zu erreichen. (Geyer, 2003: 78) Die Angaben in der Literatur bezüglich eines akzeptablen Verhältnisses zwischen Versuchen und erfolgten Interviews sind uneindeutig. In einer Studie von Blasius et al. (1995, in Geyer, 2003: 78) wurde eine maximale Ausschöpfung erreicht, wenn zwischen drei und fünf Anrufe unternommen wurden. (vgl. ebd.)

³⁸ Diese Zahlen sind nur unter Vorbehalt verwertbar. Die tatsächliche Ausschöpfungsquote der einzelnen Interviewerinnen ist nur schwer ermittelbar, da die Projektmanagerin (noch) nicht durchgeführte Interviews erneut an andere Interviewerinnen verteilte, um eine möglichst hohe Gesamt-Ausschöpfungsquote zu erhalten. Diese Vorgehensweise ist in Hinblick auf das Gesamtziel notwendig und sinnvoll.

Bewertung/ Einschätzung

Es erfolge eine kontinuierliche Rücklaufkontrolle der Interviewbögen. So fiel frühzeitig auf, dass die Teilnehmerinnen während des Frühsommers deutlich schwieriger zu erreichen waren, auch wenn sehr unterschiedliche Zeitpunkte für die Anrufe gewählt wurden. Deshalb wurde vereinbart, die Anzahl der Kontaktversuche in den Urlaubsmonaten ggf. auf > 7 zu erhöhen und die Zeitpunkte für die Anrufe zu variieren. Dies ist mit einem hohen zeitlichen Ressourcenverbrauch für die InterviewerInnen verbunden, erscheint aber bei dem insgesamt überschaubaren Stichprobenumfang sinnvoll. Ein Erfahrungsaustausch der InterviewerInnen zeigte, dass die Frauen insbesondere zur Mittagszeit und in den Abendstunden gut erreichbar waren. Diese Erkenntnis wurde im weiteren Verlauf aufgegriffen. Weiterhin wurde bei Bedarf ein Termin vereinbart.

Monat	Anzahl der Interviews/ Rücklauf
Mai	3
Juni	38
Juli	54
August	24
September	50
Oktober	8
nicht durchführbar	42
Summe	219

Tab. 5: Rücklaufkontrolle Interviews; Übersicht

Die laufende Kontrolle des Rücklaufs mit zeitnaher und flexibler Korrektur in der Vorgehensweise führte zu einer Erhöhung der Ausschöpfungsquote. In der Feldphase wurden damit wesentliche Aspekte berücksichtigt, um eine möglichst hohe Response zu erreichen.

5.6 Qualitätssicherung

„Darüber hinaus sollten auch die Vorbereitungsschritte wie Interviewer-Schulung, die organisatorischen Maßnahmen der Qualitätssicherung und -kontrolle sowie die prozessbegleitende Evaluation beschrieben werden.“ (Hoffmann et al., 2004: 12)

Die interne Qualitätssicherung einer Untersuchung beinhaltet eine Reihe von Maßnahmen und ist unabdingbarer Bestandteil epidemiologischer Studien. Dies sind Maßnahmen wie beispielsweise interne Evaluation, Erstellung eines Studienplans, Dokumentationen des Untersuchungsgangs. Der Umfang der Qualitätssicherung bemisst sich dabei in Angemessenheit zum Gesamtumfang und Kosten der Studie. Zielvorgabe für die

Qualitätssicherung sind die im Studienplan³⁹ festgelegten Kriterien. (ebd.) Eine wichtige Grundlage zur Qualitätssicherung stellen die „Leitlinien Guter Epidemiologischer Praxis“ (GEP) dar. (Elkeles, 2004: 40)

Bewertung/ Einschätzung

Die Qualitätssicherung erfolgt in Form der vorliegenden Prozessevaluation, die sich an den „Leitlinien GEP“ orientiert. Im Rahmen dieser Evaluation erfolgt eine Analyse einzelner Aspekte, aus denen sich maßgeblich die Qualität der Studie konstituiert. Die Evaluation wurde von der Projektleitung veranlasst. Daraus lässt sich ein deutliches Interesse an Qualitätssicherung ableiten. Die Ergebnisse der vorliegenden Arbeit stellen eine Zwischenauswertung dar und dienen damit dem Studienmonitoring und der internen Qualitätssicherung, weil sich daraus Erkenntnisse für den weiteren Studienverlauf ableiten lassen. Damit liegt eine umfassende Qualitätssicherung vor.

Ein Studienplan für die Studie QuaSti liegt nicht vor. Nach den „Leitlinien Guter Epidemiologischer Praxis“ ist die Erstellung eines Studienplans wesentliche methodische Voraussetzung für die Qualität einer Studie. (Hoffmann et al., 2004: 9) Vor dem Hintergrund der überschaubaren Studiengröße ist es jedoch aus Sicht der Evaluatorin gerechtfertigt, die primär knapp bemessenen zeitlichen Ressourcen nicht für die Erstellung eines solchen Plans einzusetzen. Es liegt eine umfangreiche Dokumentation der Sitzungen in Form von Prozessprotokollen vor. Der Studienverlauf (beispielsweise Rücklauf der Bögen, Stand der Interviewdurchführungen) ist umfassend schriftlich dokumentiert, so dass eine Überprüfung der ursprünglichen Ziele und Vereinbarungen jederzeit möglich ist.

5.7 Datenhaltung- und Dokumentation

„Für die Erfassung und Haltung aller während der Studie erhobenen Daten sowie für die Aufbereitung, Plausibilitätsprüfung, Kodierung und Bereitstellung der Daten ist vorab ein detailliertes Konzept zu erstellen.“ (Hoffmann et al., 2004: 17)

Die erhobenen Daten sollten möglichst zeitnah in eine Datenbank überführt werden, die eine sichere Haltung und Erfassung der Daten gewährleistet. Damit wird eine zeitnahe Datenprüfung bereits parallel zur Felderhebung gewährleistet. (ebd.) Dies sollte durch entsprechend qualifiziertes Personal durchgeführt werden. *„Jede Auswertung ist nur so gut wie der ihr zugrunde liegende Datenbestand.“* (Schäfer, 2004: 90) Die

³⁹ Bestandteile des Studienplans sind beispielsweise: Fragestellung und Hypothese, Studientyp, Studienbasis, Studienumfang, Datenerfassungskonzeption. (Hoffmann et al., 2004: 9)

Plausibilitätskontrollen sollten auf der Basis eines Rohdatensatzes geprüft werden und Änderungen der Variablenwerte müssen einzeln dokumentiert werden. (Hoffmann et al., 2004: 17) Nicht jede Fehleingabe wird durch Plausibilitätsprüfungen erkannt. Eine sorgfältige Plausibilitätsprüfung reduziert jedoch die Anzahl der Fehler erheblich und ist damit wichtiges Element der Qualitätssicherung. (Schäfer, 2004: 90)

Bewertung/ Einschätzung

Alle erhobenen Daten werden zentral durch die Projektmanagerin verwaltet, die durch ihren ersten Beruf (vor Studiumsbeginn) als „medizinische Dokumentationsassistentin“ fundierte Vorkenntnisse und Qualifikationen mitbringt. Sie führt eine umfassende Sicherung aller Dokumente und Daten mit mehrfachen Sicherheitskopien durch. Mit Hilfe des CATI-Systems ist eine Übersicht der Daten in Form einer Datenbank jederzeit möglich. Dies ermöglicht eine zeitnahe und komfortable Überprüfung der Daten noch während der Erhebungsphase. Die Plausibilitätskontrolle wurde nach den „Leitlinien Guter Epidemiologischer Praxis“ durchgeführt und liegt in entsprechend dokumentierter Form vor. Alle Rohdatensätze wie auch die Einzelfassungen der geprüften und geänderten Versionen liegen vor.

Damit liegt eine adäquate und sorgfältige Datenhaltung und Dokumentation vor.

5.8 Datenschutz

„Bei der Planung und Durchführung epidemiologischer Studien ist auf die Einhaltung der geltenden Datenschutzvorschriften zum Schutz der informationellen Selbstbestimmung zu achten.“ (Hoffmann et al., 2004: 22) Bei wissenschaftlichen Untersuchungen muss dem Recht des Einzelnen auf informationelle Selbstbestimmung wie auch dem Recht auf Freiheit von Forschung und Wissenschaft Rechnung getragen werden. Die anonymisierten Daten sind zweckgebunden an wissenschaftliche Forschung. Zu Erfordernissen des Datenschutzes gehört nach Elkeles (2004: 49), dass:

- die durchführende Institution klar erkennbar ist
- der Befragte darüber informiert wird, wie er in die Stichprobe gelangte
- der Befragte auf die freiwillige Teilnahme hingewiesen wird
- die Angaben im Fragebogen niemals mit der Adresse verknüpft werden

Bewertung/ Einschätzung

Mit der umfassenden Information der Studienteilnehmerinnen sowie deren schriftlichem Einverständnis sind wesentliche Erfordernisse des Datenschutzes erfüllt.

Die Datensätze mit allen persönlichen Angaben der Teilnehmerinnen werden zur Durchführung als E-Mail-Anhang an die InterviewerInnen versendet. Nach ihrer Bearbeitung werden sie auf selbem Wege zurück gesendet. Dies vereinfacht das logistische Vorgehen bei der Verteilung der Interviews. Es stellt jedoch ein unsicheres Verfahren bei der Weitergabe von Daten dar, die den Datenschutzbestimmungen unterliegen, da die Angaben zu diesem Zeitpunkt noch mit der Adresse der Befragten verknüpft sind. Es liegen keine sicheren Angaben darüber vor, wie gut die jeweiligen Privatcomputer der InterviewerInnen, in Form von Sicherungssystemen, vor Fremdzugriff geschützt sind. Bei Dokumenten, die Datenschutzerfordernissen unterliegen, ist eine persönliche Übergabe oder eine postale Versendung zu bevorzugen.

5.9 Ethische Aspekte

„Epidemiologische Untersuchungen müssen im Einklang mit ethischen Prinzipien durchgeführt werden und Menschenwürde sowie Menschenrechte respektieren.“ (Hoffmann et al., 2004: 6) Ethische Prinzipien ergeben sich aus nationalen wie internationalen Rechtsgrundlagen über Menschenrechte und noch spezieller formuliert, aus Patienten-, Probanden- und Forscherrechten. Dabei stellen diese ethischen Prinzipien auch dann eine wichtige Grundlage dar, wenn keine rechtliche Verpflichtung zu ihrer Beachtung besteht. (ebd.) Vor der Durchführung einer epidemiologischen Studie sollte die Stellungnahme einer Ethikkommission vorliegen. (ebd.) Vor diesem Hintergrund wurde von der „Deutschen Arbeitsgemeinschaft für Epidemiologie“ eine „Checkliste zur Begutachtung epidemiologischer Studien“ vorgelegt. (Wichmann et al., 2005: 1ff)

Bewertung/ Einschätzung

Vor Beginn der Studie regte die Projektleitung eine Vorlage des Studienvorhabens bei der Ethikkommission der Klinik an. Dies wurde vom Kooperationspartner aus Zeitgründen abgelehnt. Bei der vorliegenden Studie liegen keine interventionellen oder experimentellen Studienanordnungen vor, sondern es werden rein deskriptive Daten erhoben. Es liegt eine umfassende Probanden-Information sowie deren Einverständnis vor. (vgl. Wichmann et al., 2001: 3) Deshalb ist eine Einwilligung durch die Ethikkommission nicht zwingend erforderlich.

Die Studie wird vorrangig durch zwei Milchpumpenfirmen fremdfinanziert. Dies steht in keinem Gegensatz zu der Thematik der Studie im Kontext von Stillförderung. Es besteht kein Eigeninteresse der Geldgeber an bestimmten Ergebnissen. Damit ist die Unabhängigkeit gewahrt und die akquirierten Geldmittel stehen im Einklang mit einer ethischen Vorgehensweise.

Die Projektleitung hat eine sorgfältige Vorgehensweise für den Umgang mit krisenhaften Situationen sichergestellt, indem eine Referenzliste erstellt wurde. Damit erhalten die InterviewerInnen ein Instrument, um den Frauen bei Bedarf weiterhelfen zu können. Weiterhin wurden die InterviewerInnen im Rahmen der Schulung darauf vorbereitet, dass sensible Situationen am Telefon entstehen können. (beispielsweise eine schwere Erkrankung des Kindes, eine Krisensituation durch veränderte Lebensumstände der Frauen)

Damit werden zahlreiche, ethisch relevante Aspekte der an der Studie beteiligten Personengruppen berücksichtigt.

6 Die Evaluation der Ergebnisse

„Essentiell für die objektive Bestimmung des Zielerreichungsgrades ist der valide Nachweis eines Kausalzusammenhanges zwischen der zu beurteilenden Leistung und dem beobachteten Ergebnis [...]“ (Badura et al., 2003: 165)

Eine Ergebnisevaluation kann beginnen, wenn erste Ergebnisse vorliegen. Mit Hilfe einer Wirkungsanalyse werden die Effekte einer Maßnahme überprüft und ‚im Lichte‘ der ex-ante formulierten Ziele geprüft sowie Zusammenhänge beleuchtet. (Stockmann, 2004: 16) Die Ergebnisevaluation bewertet die Qualität der, im Rahmen der Zwischenauswertung, erhobenen Daten. Weiterhin wird der Zielerreichungsgrad der Studie QuaSti betrachtet, d.h. in welchem Maße die ex-ante formulierte Forschungsfrage beantwortet wird. Die Datenqualität wird anhand verschiedener Kriterien bewertet: *„Zu den Kriterien der Datenqualität zählen der Itemnonresponse, die Reliabilität und die Validität.“* (Elkeles,2004: 43) Geyer (2003: 38) ergänzt dies noch um die Kriterien der Objektivität.

6.1 Die Qualität der Stichprobe

“To evaluate this, authors should provide a comprehensive description of the study population and the sample population in terms of age distribution, place of residence, time period, nationality and ethnicity.“ (Blettner et al., 2001: 98) Die Repräsentativität der Stichprobe⁴⁰ stellt eine wesentliche Voraussetzung für die Validität der Daten dar. Validität bezieht sich auf die Anforderung, dass in einer Messung gemessen wird, was gemessen werden soll. (Geyer, 2003: 35) Die Aussagekraft der Studien hängt dabei sowohl von der methodischen Qualität (interne Validität) als auch von der Generalisierbarkeit der Daten auf die Zielpopulation (externe Validität) ab. (Latza, 2004: 327)

Die Bewertung der Repräsentativität erfolgt durch zwei wesentliche Faktoren:

- Zusammensetzung der Stichprobe (Vorliegen einer verzerrungsfreien Stichprobe)
- Response-Rate (Ausschöpfungsrate) mit differenzierter Analyse der Nichtteilnehmer-Rate und Item-Nonresponse (vgl. Latza, 2004: 328ff.)

⁴⁰ Der Begriff der „repräsentativen Stichprobe“ wird in der Literatur kontrovers diskutiert. Diekmann (2000: 368) konstatiert, dass eine Stichprobe niemals sämtliche Merkmalsverteilungen der Population repräsentieren kann. Aus seiner Sicht versteht sich der Begriff mehr als Metapher. Die Bezeichnung sei in der Statistik kein Fachbegriff. Treffender ist es Diekmanns Verständnis nach, von Zufallsstichproben oder einer Wahrscheinlichkeitsauswahl zu sprechen. (ebd.) Da die Bezeichnung der „repräsentativen Stichprobe“ im praktischem Sprachgebrauch vieler Studien fest etabliert ist, benutzt auch die Autorin diese weiterhin. Die von Diekmann angeführten Argumente werden jedoch im Rahmen der Bewertung berücksichtigt.

6.1.1 Zusammensetzung der Stichprobe

Die Repräsentativität von Stichproben ist ein Qualitätsmerkmal, welches beschreibt, dass die Struktur der Studienpopulation so beschaffen ist, dass die erhobenen Daten auf die Zielpopulation übertragbar sind. Sie trifft also Aussagen über die Generalisierbarkeit der Daten und damit der externen Validität. (Latza, 2004: 327) In einer Querschnittsstudie kann die Übertragbarkeit der Ergebnisse eingeschränkt sein, wenn zu viele Personen, die in der Stichprobe gezogen werden, später nicht an der Studie teilnehmen und damit zu Verzerrungen führen. (ebd.) Weiterhin ist bekannt, dass soziodemographische Merkmale Einfluss auf die Teilnahmebereitschaft nehmen. Eine überdurchschnittlich hohe Antwortbereitschaft findet sich in den Altersgruppen der 18- 29 Jährigen und 30- 40 Jährigen. Ein Geschlechtereffekt ist nicht beobachtbar. (Geyer, 2003: 79) Nicht teilnahmebereite Personen haben häufig eine höhere Wahrscheinlichkeit hinsichtlich bestimmter Lebensstilfaktoren oder sozial bedingter Risiken und Erkrankungen. (Latza, 2004: 328) Bei Personen mit niedrigem Bildungsstatus findet sich beispielsweise eine überproportional ablehnende Haltung gegenüber Befragungen. (Geyer, 2003: 79)

Bewertung/ Einschätzung

Die Typisierung der Stichprobe zeigt einen Vergleich der soziodemographischen Merkmale mit der definierten Zielpopulation. (vgl. Tab. 6) Die Auswahl der Vergleichsvariablen erfolgt vor dem Hintergrund, dass sozial benachteiligte Frauen kürzer stillen. Das soziodemographische Profil ist deshalb von besonderer Bedeutung hinsichtlich der Prävalenz des Stillens. Für die Zielpopulation sind zwei Datenquellen verfügbar, die ergänzend nebeneinander aufgestellt wurden. Als eine weitere Bezugsgröße wurde die Grundgesamtheit aller in Hamburg entbundenen Frauen des Jahres 2004 gewählt.⁴¹ Diese ist als Vergleichsgruppe besonders geeignet, da sie den Hamburger Durchschnitt aller entbundenen Frauen eines Jahres repräsentiert und in bestimmten Kriterien der Studienpopulation stark ähnelt (weibliches Geschlecht, Altersstruktur, kürzlich entbunden). Ergänzend wurden Daten des Mikrozensus⁴²

⁴¹ Dies bezieht sich auf die alle Entbindungen in einem der zwölf geburtshilflichen Abteilungen in Hamburg im Jahre 2004. Diese Daten stammen von der „Externen Qualitätssicherung“ (EQS) Hamburg. Die EQS Hamburg ist eine Arbeitsgemeinschaft der Hamburgischen Krankenhausgesellschaft (HKG) und der Krankenkassenverbände unter Einbeziehung der Ärztekammer Hamburg und des Medizinischen Dienstes der Krankenkassen. Zweck ist die Durchführung von Qualitätsvergleichen (Benchmark) zwischen den Krankenhäusern in Hamburg, um die Qualität dieser Versorgung kontinuierlich zu bestimmen. Die Teilnahme ist für die Krankenhäuser verpflichtend. (http://www.eqs.de/eqs/_index.php ; Stand 02.01.2006)

⁴² Der Mikrozensus ist eine amtliche Repräsentativstatistik, durchgeführt durch das statistische Bundesamt, und erhebt Daten über die Bevölkerungsstruktur, Arbeitsmarktsituation, wirtschaftliche

herangezogen, um in bestimmten Merkmalen (z.B. Netto-Haushaltseinkommen) zumindest einen Trend hinsichtlich des Profils der Stichprobe bestimmen zu können.

	Teilnehmerinnen* Studie QuaSti (n = 177) %	Zielpopulation** (n = 1895) %	[1]: Klinikqualitäts- gutachten*** (n = 68) %	Hamburg*** (n = 18191) %	Mikrozensus**** (n = 830000) %
Altersstruktur					
Unter 18 Jahre	0,5	0,4	[0]	0,6	-
18 - unter 30 Jahre	22,5	28,8	[22]	38,5	-
30 - 35 Jahre	40,6	34,3	[46]	31,7	-
35 - unter 40 Jahre	31,6	30,3	[26]	23,9	-
40 und > 40 Jahre	4,5	6,2	[6]	5,4	-
Familienstand					
Verheiratet	75,7	k. A.	[78]	k. A.	74
In Lebensgemeinschaft	19,2	k. A.	[18]	k. A.	6
Alleinstehend	5,1	k. A.	[4]	13,8	20
<small>(bezogen auf die Eltern- Kind-Gemeinschaften)</small>					
Schulbildung					
Abitur	62	k. A.	[78]	k. A.	21,2
Mittlere Reife	33	k. A.	[19]	k. A.	26,3
Hauptschule	5	k. A.	[3]	k. A.	44,1
Keinen Schulabschluss	0	k. A.	[k. A.]	k. A.	2,8
Beruflicher Abschluss					
Studium	31	47,9	[49]	19,6	10,6
Fachschule	18	k. A.	[k. A.]	k. A.	7,7
Lehre	46	21,5	[k. A.]	32,7	50,5
Keinen Abschluss	5	0,8	[k. A.]	2,6	28,4
Dt. Staatsangehörigkeit					
Dt. Staatsangehörigkeit	95	85,4	[k. A.]	71,6	81,7
Andere Länder	5	14,6	[k. A.]	28,4	18,3
Durchschnittliches Netto- Haushaltseinkommen					
Durchschnittliches Netto- Haushaltseinkommen	3000-3500 € (Median)	k. A.	[k. A.]	k. A.	1324 € (Median)

Legende

- * Bereinigte Netto- Stichprobe der Studie QuaSti
 - ** Daten zum Klinikum Nord, Hamburg, externe Qualitätssicherung Hamburg, (EQS) Jahresauswertung 2004
 - *** Daten aller 12 geburtshilflichen Abteilungen in Hamburg, EQS (ebd.)
 - **** Klinik-Qualitätsgutachten für das Klinikum Nord, Hamburg, Prof. Riegl & Partner, Gesundheitsforschungs- Institut, klinikinternes Dokument, 2003
- Anmerkung: Summenangaben unter 100% entstehen durch Non-Response zu diesem Item
k.A.: keine Angabe

Tab. 6: Typisierung der Stichprobe

und soziale Lage der Bevölkerung. (Breiholz et al., 2005: 89) Jährlich werden 1% aller Haushalte in Deutschland befragt, dies entspricht 390 000 Haushalte/ 830000 Personen (ebd.)

Die Analyse der Stichprobe weist eine Struktur auf, die der definierten Zielpopulation stark ähnelt. Die bislang erhobenen Daten sind übertragbar auf die Zielpopulation, die Stichprobe ist also repräsentativ hinsichtlich ihrer Zusammensetzung. Damit ist eine wesentliche Grundvoraussetzung für die externe Validität der Daten erfüllt. Daraus ergibt sich, dass die Ergebnisse der Studie QuaSti dem Krankenhaus klinikbezogen repräsentative Daten liefern.

Die Aufstellung der Hamburger Daten im Vergleich veranschaulicht das besondere soziodemographische Profil der Studienpopulation. Schon die Auswertung der Pilotstudie ergab eine deutlich höhere Quote an Frauen mit hohem Bildungsgrad als im Hamburger Durchschnitt. Obgleich bislang nur die Daten der Zwischenauswertung (n= 219) vorliegen, ist wahrscheinlich, dass sich die hohe soziale Selektion erneut darstellt. Hierfür ist eine Reihe von Gründen möglich. Ein wesentlicher Aspekt liegt in der Lage des Krankenhauses begründet, welches an wohlhabende Stadtrandgebiete, mit einer niedrigeren Rate an Migrantinnen angrenzt. Die Pilotstudie hat gezeigt, dass eine Entbindungsklinik häufig auf Grund der räumlichen Nähe gewählt wird, so dass hieraus eine Selektion resultiert. Die Stichprobe differiert in wesentlichen Merkmalen erheblich von den Hamburger Daten, insbesondere bezüglich der Schulbildung und Berufstätigkeit. Das durchschnittliche Netto-Haushaltseinkommen⁴³ ist mehr doppelt so hoch wie im bundesdeutschen Durchschnitt. Dies weist deutlich auf die sozial privilegierte Situation der Studienteilnehmerinnen hin. Eine Übertragung der Daten der Studie QuaSti auf den Hamburger Durchschnitt ist nicht zulässig, da die Studienpopulation diesbezüglich nicht repräsentativ ist. Dies war auch nicht Ziel der Studie. Der Vergleich mit Hamburg bzw. dem bundesdeutschen Durchschnitt wird nur ergänzend erwähnt, um das Profil der Stichprobe zu veranschaulichen.

6.1.2 Repräsentanz ausgewählter Gruppen

Obgleich die Stichprobe der Studie QuaSti die Zielpopulation grundsätzlich repräsentativ abbildet, können einzelne Gruppen über- oder unterrepräsentiert sein. Vor dem Hintergrund der Vorgehensweise bei der Rekrutierung der Stichprobe (im Verlauf es

⁴³ Dies wird nach dem Sozioökonomischen Panel (SOEP) definiert als: „Frage: *Wie hoch ist das monatliche Haushaltseinkommen aller Familienmitglieder heute? Bitte geben Sie den monatlichen Netto- Betrag an, also nach Abzug von Steuern und Sozialabgaben. Regelmäßige Zahlungen wie Renten, Wohngeld, Kindergeld, BAföG, Unterhaltszahlungen usw. rechnen sie bitte dazu!*“ (Frick, 2005: 2)

stationären Aufenthaltes nach Geburt) ist eine Analyse der ambulant⁴⁴ entbundenen Frauen notwendig, da diese Gruppe möglicherweise unterrepräsentiert ist.

Als zweite Gruppe wird die Teilnahme von Migranantinnen analysiert, da die Nationalität eine relevante Größe zur Einschätzung des Sozialstatus darstellt. Diese Gruppe wird vor dem Hintergrund näher beleuchtet, dass als Ausschlusskriterium für die Studie „unzureichende Deutschkenntnisse“ definiert wurde.

Drittens wird untersucht, wie hoch der Anteil der Frauen ist, die sich auf Grund der Stillfreundlichkeit für die Auftragsklinik entschieden haben. Diese Gruppe kann möglicherweise überrepräsentiert sein und damit die Ergebnisse hinsichtlich der Still-Prävalenz verzerren.

Bewertung/ Einschätzung

Ambulant entbundene Frauen: Insgesamt konnten im Rahmen der Erhebung (n= 219) 7 ambulant entbundene Frauen (= 3,1%) erfasst werden. Verglichen dazu liegt die durchschnittliche Rate in der Klinik bei 6,9%.⁴⁵ Daraus ergibt sich, dass die ambulant entbundenen Frauen unterrepräsentiert sind. Dies erklärt sich, wie bereits erläutert, in der Vorgehensweise der Rekrutierung und stellt einen systematischen Bias dar. Sowohl die KlinikmitarbeiterInnen als auch die Projektleitung haben mit dieser Problematik gerechnet. Eine Veränderung der Vorgehensweise ist aus logistischen Gründen nicht möglich. Der Einfluss die Unterrepräsentanz dieser, insgesamt kleinen, Gruppe auf das Gesamtergebnis kann derzeit nicht abschließend bewertet werden. Dafür bedarf es weiterer Untersuchungen darüber, in welchen, weiteren Merkmalen sich ambulant entbundene Frauen von Gruppe der stationär entbundenen Frauen unterscheiden.

Teilnehmerinnen mit Migrationshintergrund: Insgesamt wurden mit den Studienteilnehmerinnen (n= 219) 15 Frauen nicht-deutscher Nationalität (= 6,8%) erfasst. Die durchschnittliche Rate liegt in der Klinik bei 14,6%. Von den 15 Interviews konnten 6 Interviews aus verschiedenen Gründen nicht durchgeführt werden. Eine Einzelfallanalyse der durchgeführten Interviews ergab, dass bei 2 Interviews durch eine dritte Person gedolmetscht wurde, ein Interview auf Englisch geführt wurde und bei einem weiteren Sprachschwierigkeiten mit der Teilnehmerin angemerkt wurde.

⁴⁴ Definition „ambulant“: dies wird definiert als Aufenthalt von insgesamt bis zu 24h inklusive des vorgeburtlichen Aufenthaltes.

⁴⁵ Daten aus der ‚Externen Qualitätssicherung‘ Hamburg, (EQS), Asklepios Klinik Nord, Hamburg, Jahresauswertung 2004, klinikinternes Dokument 2005

Die Verwertbarkeit dieser Interviews muss in Frage gestellt werden. Zusammenfassend liegt damit in der Stichprobe eine Unterrepräsentanz von Frauen mit Migrationshintergrund vor.

Die Einschätzung darüber, ob bei den Teilnehmerinnen das Ausschlusskriterium vorliegt, wird durch die KlinikmitarbeiterInnen vor Ort getroffen. Die Analyse dieser Gruppe zeigt, dass das Ausschlusskriterium offenbar unzureichend definiert wurde, da Frauen trotz geringer oder keiner Deutschkenntnisse in die Stichprobe gelangten. (vgl. Abschnitt 5.2.1) Diese Problematik wurde von der Projektleitung frühzeitig aufgegriffen. Es wurde trotz der Problematik vereinbart, die Vorgehensweise in der Klinik nicht zu verändern. Englischsprachige Interviews wurden toleriert und ausschliesslich durch InterviewerInnen mit sicheren Englischkenntnissen durchgeführt. Die Interviews werden auch zukünftig im Rahmen der Plausibilitätskontrolle einer Einzelfallanalyse unterzogen. Insgesamt stellt dies eine sinnvolle Vorgehensweise dar, um eine möglichst hohe Teilnahmequote dieser Personengruppe zu erreichen.

Stillfreundlichkeit: 7,3 % aller Frauen (n= 177) entschieden sich wegen der Stillfreundlichkeit für die Klinik. Häufig wurde dies zusammen mit anderen Gründen erwähnt. Im Vergleich dazu entschieden sich 36,7 % eindeutig auf Grund der räumlichen Nähe für die Klinik. Dies deckt sich mit den Ergebnissen der Pilotstudie. Eine Verzerrung der Stichprobe bezüglich einer Überrepräsentanz von Frauen, die die Klinik auf Grund der Stillfreundlichkeit wählten, liegt nicht vor.

6.1.3 Die Response-Rate

„The distinction between the study population (target population, i.e. population for which results are wanted), the sample population (brut sample, i.e. population selected to be included in the study) and the population on which the analysis is based (net population, i.e. sample population minus subjects which are eligible but do not participate) is helpful in evaluation for which population the results are valid and to what degree they can be generalised.“ (Blettner et al., 2001: 98)

Durch eine randomisierte Stichprobe aus einer definierten Zielpopulation kann eine Repräsentativität erreicht werden. Diese ist jedoch nur dann gewährleistet, wenn die so ausgewählten Probanden auch teilnehmen. Die Response (Ausschöpfungsquote) stellt deshalb ein wesentliches Kriterium der Stichprobenqualität und deren Validität dar. Der Begriff Response bezieht sich auf die Bereitschaft von Probanden aus einer definierten Grundgesamtheit zur Teilnahme an einer Studie. Generell sollte eine möglichst hohe Response angestrebt werden. Eine niedrigere Response kann jedoch nicht pauschal mit

geringer Qualität einer Studie gleichgesetzt werden. (vgl. Latza et al., 2004: 334) Bei telefonischen Befragungen ist eine durchschnittliche Response-Rate von 50- 70% zu erwarten. (Diekmann, 2000: 359) Diekmann (ebd.) schlägt zur Berechnung der Ausschöpfungsquote (A) folgende vor:

$$A = \text{Anzahl ausgewerteter Interviews} / \text{bereinigter Bruttostichprobenumfang} \times 100$$

Bewertung/ Einschätzung

Die Ausschöpfungsquote der Zwischenauswertung liegt nach der Berechnungsgrundlage von Diekmann (ebd.) bei 80,8 %. Damit wurde eine überdurchschnittlich hohe Ausschöpfungsquote realisiert. Die nachfolgende Tabelle stellt eine detaillierte Response-Analyse mit Erhebung der Brutto- und Nettostichprobe dar. Dies dient der Übersicht und erlaubt einen Rückschluss auf die Zusammensetzung der Gesamt-Stichprobe im Verhältnis zu der Population, deren Daten dann erhoben werden konnten.

	Absolut	Bruttostichprobe	Nettostichprobe	Bereinigte Nettostichprobe	Erreichte Personen	Teilnahmebereit	Abgeschlossene Interviews
Bruttostichprobe, davon:	219	100 %					
Gesperrte Datensätze, z.B. nicht ermittelbare Tel.-Nr.	7	3,2%					
Nettostichprobe, davon:	212	96,8 %	100 %				
Keinen Anschluß	5	2,3 %	2,4 %				
Verzogen	1	0,45 %	0,5 %				
Verstorben*	2	0,9 %	0,9 %				
Kommunikationsprobleme	0	0 %	0 %				
Falsche Tel.-Nr.	2	0,9 %	0,9 %				
Bereinigte Nettostichprobe davon beim letzten Versuch:	202	92,3 %	95,3 %	100%			
Anrufbeantworter	3	1,4 %	1,4 %	1,5 %			
Besetzt	1	0,45 %	0,5 %	0,5 %			
Keine Antwort	9	4,1 %	4,2 %	4,5 %			
Erreichte Personen, davon:	189	86,3 %	89,2 %	93,5 %	100 %		
Harte Verweigerung	12	5,5 %	5,6 %	5,9 %	6,3 %		
Teilnahmebereit, davon:	177	80,8 %	83,2 %	87,6 %	93,7 %	100 %	
Abgebrochene Interviews	0	0 %	0 %	0 %	0 %	0 %	
Technische Störungen	0	0 %	0 %	0 %	0 %	0 %	
Abgeschlossene Interviews	177	80,8 %	83,2 %	87,6 %	93,7 %	100 %	100 %

* Kinder sind verstorben. Interviews wurden deshalb nicht durchgeführt.

Tab. 7: Responseanalyse der Datenauswertung (vgl. Janßen et al., 1999: 175)

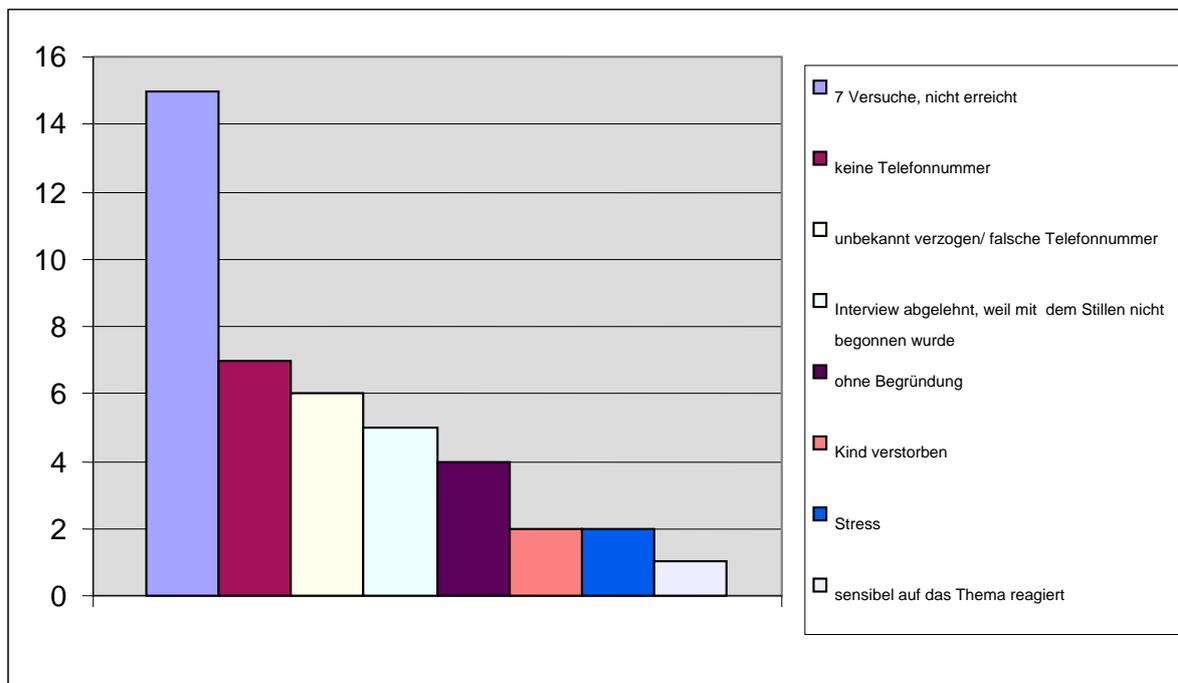
6.1.3.1 Nichtteilnehmerinnen-Analyse

Bei Beurteilungen der Response ist zu berücksichtigen, dass sich die Gruppe der NichtteilnehmerInnen systematisch von teilnehmenden Probanden unterscheiden können. Um Auswirkungen potentieller Selektionseffekte (sog. Nonresponse-Bias) schätzen zu können, ist eine möglichst genaue Analyse der NichtteilnehmerInnen vorzunehmen. (Latza, 2004: 329) Für den Vergleich werden Variablen ausgewählt, die im Rahmen der Studie inhaltlich von Bedeutung sind. Voraussetzung ist, dass die Verteilung dieser Variablen in der Grundgesamtheit bekannt ist. (Elkeles, 2003: 50). Grundsätzlich erfolgt bei der Beurteilung der Nichtteilnehmer-Rate eine Differenzierung der Ausfälle in:

- *nicht stichprobenneutral*: dies sind Verweigerungen der Kontakt- oder Zielpersonen, Interview-Abbrüche und alle personenbezogenen Gründe (z.B. Alter, mangelnde Sprachkenntnisse, gesundheitliche Gründe, nicht eingehaltene Termine) sowie die Nichterreichbarkeit der Teilnehmer. (Stögbauer, 2000, in Geyer, 2003: 77)
- *stichprobenneutral*: in Telefoninterviews sind dies Nummern, die zu Fax- und Datenendgeräten gehören oder Nummern, die an Firmen und Behörden vergeben sind (ebd.)

Bewertung/ Einschätzung

Die Nichtteilnehmerinnen stellen sich ausschliesslich aus nicht stichprobenneutralen Ausfällen verschiedener Ursache zusammen.



Tab. 8: Gründe für Nichtteilnahme an den Interviews; Übersicht

Es erfolgte eine Analyse, ob sich die Nichtteilnehmerinnen in ihrer Zusammenstellung von der Stichprobe unterscheiden. Dafür wurde eine Typisierung dieser Gruppe nach soziodemographischen Merkmalen (vgl. Tab. 6) vorgenommen und in Bezug zu den Daten der Teilnehmerinnen gesetzt. In einzelnen Items bestehen Unterschiede. Das Netto-Haushaltseinkommen der Nichtteilnehmerinnen liegt eine Kategorie unter dem der Teilnehmerinnen mit einer deutlich höheren Rate an Antwortverweigerern. Insgesamt besteht große Ähnlichkeit zwischen den beiden Gruppen. Damit liegt hier keine systematische Selektion der Nichtteilnehmerinnen vor, die zu Verzerrungen der Ergebnisse führen können.

	Teilnehmerinnen-Rate (n = 177) %	Nichtteilnehmerinnen-Rate (n = 42) %
Verheiratet	76 %	64 %
Schulbildung		
Abitur	62 %	67%
Mittlere Reife	33 %	22 %
Hauptschule	5 %	7 %
Keinen Schulabschluss (Keine Angabe)	0 % (0 %)	2 % (2 %)
Beruflicher Abschluss		
Studium	31 %	36 %
Fachschule	18 %	24 %
Lehre	46 %	19 %
Keinen Abschluss (Keine Angabe)	5 % (0 %)	12 % (9 %)
Deutsche Staatsangehörigkeit	95 %	86 %
Durchschn. Netto-Haushaltseinkommen (Median) (Keine Angabe)	3000-3500 € (12,5 %)	2500-3000 € (26 %)
Art des Einkommens		
Sozialhilfe	3 %	7 %
Asylbewerber	0 %	0 %
Arbeitslosengeld (Keine Angabe)	6 % (3 %)	2 % (17 %)

Tab. 9: Typisierung der Nichtteilnehmerinnen

6.1.3.2 Item-Nonresponse

Die Item-Nonresponse stellt den prozentualen Anteil derer dar, die auf eine Frage mit „Weiss nicht“, „Keine Angabe“ antworten oder diese gänzlich offen lassen. Anteile, die höher als 5% liegen, indizieren, dass diese Frage Probleme aufweist, beispielsweise nicht richtig verstanden wird bzw. den Befragten unangenehm ist. (Elkeles, 2003: 43) Auch bei

einer ausreichend hohen Gesamt-Response kann ein Item-Nonresponse-Problem entstehen, weil einzelne Fragen nicht von allen Studienteilnehmern beantwortet werden (beispielsweise Fragen zu dem Haushalts-Nettoeinkommen). Die Item-Nonresponse ist ein entscheidendes Kriterium für die interne und externe Validität von Studien. (Latzka, 2004: 329)

Bewertung/ Einschätzung

Bei der Analyse der Daten wurden besonders sensible Fragen ausgewählt (wie der Sozialstatus), die für die Aussagekraft der Studie von Bedeutung sind

- *Netto-Haushaltseinkommen*: Diese Frage wurde von 12,5 % der Teilnehmerinnen nicht beantwortet. (22 von n=177).⁴⁶ Damit liegt die Item-Response für diese Frage bei 87,5 %. Bei dieser Frage konnte aus verschiedenen Gründen nicht mehr nachgefasst werden (entweder kein Nachfassen möglich wegen fehlender Rufnummer, oder aber trotz Nachfrage Angabe verweigert). Bereits während der Feldphase wurde deutlich, dass eine Reihe von Frauen diese Frage nicht beantwortet, so dass sie im Interview nochmals aufgegriffen wurde. Da die Quote der Nichtbeantwortung deutlich über 5% liegt, stellt dies einen Item-Nonresponse-Bias dar.
- *Schulbildung und Ausbildung*: hier besteht kein Problem, da diese Fragen nahezu vollständig beantwortet wurden. Die Kontrolle dieses Items ist wichtig, weil der Bildungsstatus ein Parameter zur Generierung des Sozialindex darstellt, der Rückschluss auf den Sozialstatus gibt.
- *Rauchen*: Dieses Item stellt eine sensible Frage hinsichtlich der „sozialen Erwünschtheit“ des Verhaltens dar. Eine Analyse ergab, dass allen Frauen, denen diese Frage gestellt wurde⁴⁷, diese auch beantworteten. Somit liegt diesbezüglich kein Item-Nonresponse-Problem vor. Der Aspekt der sozialen Erwünschtheit⁴⁸ sollte jedoch Berücksichtigung finden bei der Auswertung dieses Items.

⁴⁶ Bei den Nichtteilnehmerinnen machten 26,2% keine Angabe dazu. (11 von n=42)

⁴⁷ Keine Antwort bezüglich des Rauchens liegt nur bei den Frauen vor, die zum Zeitpunkt des Interviews nicht mehr stillten und deshalb nicht bis zum Ende des Fragebogens gelangten, also dazu nicht befragt wurden

⁴⁸ Der Begriff der „sozialen Erwünschtheit“ meint Antwortverzerrungen bei Befragten, die durch das Bestreben entstehen, bei den Adressaten einen positiven Eindruck zu machen. (vgl. Geyer, 2003: 128)

6.2 Reliabilität

Reliabilität bezeichnet das Ausmaß, in dem mehrfache Messungen die gleichen Ergebnisse bzw. Messwerte liefern; also in welchem Maße das Testergebnis reproduzierbar ist. Dabei ist für die Reliabilität ist nicht von Bedeutung, ob das Instrument auch valide ist. (Geyer, 2003: 37 ff.) Zur Überprüfung der Reliabilität stehen verschiedene Methoden zur Verfügung, beispielsweise ein sog. Test-Retest-Verfahren.⁴⁹ Dabei wird das Instrument zweimal oder häufiger an der gleichen Personengruppe angewendet und auf individuenbezogene und gruppenbezogene Invarianzen geprüft, d.h. ob die ermittelten Werte über mehrere Zeitpunkte unverändert bleiben. Die Reliabilität kann jedoch dann verfälscht sein, wenn die wiederholt Befragten sich erinnern und um Konsistenz im Antwortverhalten bemühen. (ebd.) Auf Grund der einflussnehmenden Erinnerungseffekte ist die Wiederholung solcher Befragungen häufig nur theoretisch möglich und in der Praxis kaum sinnvoll durchführbar. (ebd.) Bei telefonischen Interviews ergibt sich aus der kürzeren Latenzzeit zwischen Frage und Antwort potentiell eine geringere Reliabilität, weil das Antwortverhalten zügiger und die Exposition mit dem Interviewmaterial geringer ist als beim persönlichen Gespräch. (Geyer, 2003: 72)

Einschätzung/ Bewertung

Eine Reliabilitätsprüfung durch eine Wiederholung von Interviews wird von der Evaluatorin als nicht sinnvoll eingeschätzt, da die Beeinflussung durch Erinnerungseffekte zu hoch wäre. Weiterhin unterliegen viele der untersuchten Merkmale zeitbezogenen Einflüssen (beispielsweise Stildauer, Stillkomplaktionen), so dass keine Konsistenz im Antwortverhalten zu erwarten ist. Zudem erscheint ein solches Prüfverfahren vor dem Hintergrund des Studiumfanges (insbesondere des zeitlichen Rahmens) als unrealistisch.

Das Erhebungsinstrument der vorliegenden Studie ist im Rahmen der Pilotstudie umfassend geprüft und ausgewertet worden. Ein Vergleich der Daten aus der Pilotstudie mit denen aus der jetzigen Zwischenauswertung zeigt eine Konsistenz im Antwortverhalten mit gruppenbezogen reproduzierbaren Antwortmustern auf, sofern dies im Rahmen einer Befragung zu individuellen Lebenssituationen überhaupt vergleichbar ist. Die Konsistenz und Plausibilität des Antwortverhaltens wurde auch im Rahmen der Plausibilitätskontrolle geprüft. Es konnten keine gruppenbezogenen Invarianzen im

⁴⁹ Weitere Verfahren sind die Split-Half-Methode (das Interview wird so gesplittet, dass zwei Teile mit der halben Zahl der Items entstehen) und die Paralleltestmethode (zwei Instrumente werden eingesetzt, die das gleiche messen sollen), die Reliabilität zeigt sich dann in den Korrelationsberechnungen der beiden Messungen (Geyer, 2003: 38)

Antwortmuster festgestellt werden. Die Fragen des Interviews sind aus Sicht der Evaluatorin so strukturiert, dass zuverlässige Ergebnisse zu erwarten sind. Durch die Vorgehensweise in der InterviewerInnen-Schulung wurde ein standardisiertes Vorgehen bei der Durchführung der Interviews erreicht. Dies erhöht die Reliabilität der Erhebung maßgeblich. (vgl. Abschnitt 5.4.2)

Vor diesem Hintergrund wird das Erhebungsinstrument als grundsätzlich reliabel eingeschätzt.

6.3 Objektivität

„Wenn eine Datenerhebung durchgeführt wird, sollten die erhaltenen Ergebnisse möglichst unabhängig von den situativen Bedingungen sein.“ (Geyer, 2003: 38ff). Dies meint insbesondere Interviewer- und Unterseuchereffekte. Bei der Objektivität werden drei unterschiedliche Qualitäten unterschieden:

- *Durchführungsobjektivität*: dies bezieht sich auf die Unabhängigkeit der Ergebnisse von den untersuchenden Personen, beispielsweise die Beeinflussung durch die InterviewerInnen (Interviewer-Bias). (ebd.)
- *Auswertungsobjektivität*: dies meint die Unabhängigkeit der Ergebnisse von den auswertenden Personen. Die Auswertungsobjektivität wird durch eine stark strukturierte und standardisierte Vorgehensweise bei der Auswertung erreicht. (ebd.)
- *Interpretationsobjektivität*: dies meint eine objektive Interpretation der Ergebnisse nach standardisierten Interpretationsregeln. (ebd.)

Ein weiteres wesentliches Merkmal der Objektivität stellt die „Unabhängigkeit der Forschung“ dar, die sicherstellt, dass keine Beeinflussung der ForscherInnen in ihrer Tätigkeit, beispielsweise durch vorteilhafte Begünstigungen angesichts eines bestimmten Ergebnisses, vorliegt. (Hoffmann et al., 2004: 23)

Einschätzung/ Bewertung

Die Durchführungsobjektivität wurde bereits im Rahmen der Prozessevaluation ausführlich mit Analyse der potentiell einflussnehmenden Faktoren (beispielsweise der InterviewerInnen-Schulung) behandelt. (vgl. Abschnitt 5.4.2)

Die Auswertung und Interpretation der Ergebnisse stehen mit Abschluss der Hauptstudie noch aus und können deshalb zum derzeitigen Zeitpunkt nicht beurteilt werden. Eine erste Zwischenauswertung und Interpretation der Ergebnisse, von der Studienleitung im

Rahmen eines Kongresses⁵⁰ präsentiert, weist ein standardisiertes und transparentes Auswertungsverfahren mit schlüssiger Interpretation der Ergebnisse auf.

Bezüglich der Forschungstätigkeit liegt vollständige Unabhängigkeit der Studienleitung zum wissenschaftlichen Forschungsprojekt vor. Unabhängigkeit in der Forschung ist ein wesentliches Güte Merkmal wissenschaftlicher Arbeit. Damit sind zahlreiche Kriterien der Objektivität erfüllt.

6.4 Kontextanalyse

Die Kontextanalyse erlaubt eine Bewertung der einflussnehmenden Kontextfaktoren, d.h. in welchem Maße die soziale Realität auf Durchführung und Ergebnisse der Studie Einfluss nimmt. (Christiansen, 1999: 50) Dafür wird eine Systematisierung der Kontexteinflüsse durchgeführt. (vgl. Kap. 4) Kontexteinflüsse beziehen sich sowohl auf die Bedingungen im direkten Bezug zur Studie (beispielsweise MitarbeiterInnen des Projektteams, Kooperationspartner) als auch auf den Makro-Kontext (beispielsweise die politische Situation oder medienwirksame Ereignisse, die Einfluss auf das Probandenverhalten nehmen könnten).

Bewertung/ Einschätzung

Es erfolgt eine Kontextanalyse bezüglich der Rekrutierung der Stichprobe, da der verzögerte zeitliche Verlauf erheblichen Einfluss auf den Studienverlauf nimmt. Als ein wesentlicher Einflussfaktor des Kontextes ergeben sich die besonderen Bedingungen der Klinik, die derzeit privatisiert wird. Die Privatisierung geschieht vor dem Hintergrund einer politischen Entscheidung. Der „Landesbetrieb Krankenhäuser“ (LBK) dessen Träger bislang die Freie und Hansestadt Hamburg war, wurde 2004 zu 74,9% an den privaten Investor „Asklepios“ verkauft.⁵¹ Dies löst erhebliche Unsicherheiten, wie beispielsweise Sorge um veränderte Arbeitsverträge und Angst vor Verlust des Arbeitsplatzes bei den MitarbeiterInnen der Klinik aus. Persönliche Gespräche mit Betroffenen bestätigen dies. Weiterhin besteht nach Berichten der KlinikmitarbeiterInnen in der geburtshilflichen Abteilung eine hohe Arbeitsdichte mit stresshaftem Arbeitsalltag. Vor dem Hintergrund dieser Gesamtsituation und Stimmungslage ist es möglich, dass die Motivation und Bereitschaft zur (zusätzlichen) Mitarbeit an der Studie möglicherweise leidet. Weiterhin ist

⁵⁰ EUPHA Conference (European Public Health Association), Health Promotion Section Workshop, Graz, Österreich, November 2005

⁵¹ <http://www.lbk-hh.de/>

die hohe Arbeitsdichte zu berücksichtigen. Dieses wären mögliche Gründe für den stark verzögerten Rücklauf der Demographiebögen.

Eine eingehende Lektüre der Tagespresse während des Studienverlaufs zeigt keine relevanten medienwirksamen Ereignisse mit Bezug zum Thema Stillen auf, die einen signifikanten Einfluss auf das Studienergebnis vermuten lassen. Im Jahre 2003 wurde der Fall eines inadäquat zusammengesetzten Fertigmilchprodukts weltweit bekannt, der in Israel zum Tod zweier Säuglinge führte.⁵² Nach Einschätzung der Evaluatorin liegt dieses Ereignis zu weit zurück, um Einfluss auf das derzeitige Stillverhalten von Frauen in Hamburg zu nehmen.

6.5 Analyse des auswertungsmethodischen Verfahrens

Die Auswertung epidemiologischer Studien soll [...] zügig, valide, transparent und jederzeit für Dritte nachvollziehbar unter Verwendung adäquater Methoden und ohne unangemessene Verzögerung erfolgen. (Hoffmann, et al., 2004: 20)

Da die Analyse bei gesundheitsrelevanten Fragestellungen von hohem Interesse für den Auftraggeber wie auch für die Öffentlichkeit ist, sollte die Datenauswertung möglichst zeitnah nach der Dateneingabe erfolgen und für Dritte jederzeit nachprüfbar gestaltet werden. (ebd.) Die Auswertungsprinzipien einschliesslich der statistischen Methoden sollten vorab festgelegt werden. (Schäfer, 2004: 92) Für die Auswertung sollte genügend Zeit und fachlich qualifiziertes Personal zur Verfügung stehen, um große „Datenfriedhöfe“ zu vermeiden. (ebd.)

Bewertung/ Einschätzung

Da die Gesamterhebung zum derzeitigen Zeitpunkt noch nicht abgeschlossen ist, kann dieser Aspekt nur begrenzt bewertet werden. Die Auswertung der Daten erfolgt mit dem statistischem Auswertungsprogramm SPSS®. Dies ist ein in den Sozialwissenschaften fest etabliertes Instrument und im vorliegendem Falle besonders geeignet, da es mit dem CATI-System kompatibel ist. So sind Daten transferierbar und müssen nicht erneut manuell eingegeben werden. Dies trägt erheblich zur Minimierung von Eingabefehlern bei.

Die Projektleitung hat in einer ersten Auswertung der Daten der Zwischenauswertung das Konzept der „Sozialen Benachteiligung“ gewählt. Dies erscheint besonders geeignet, da die Stilldauer maßgeblich von soziodemographischen Merkmalen abhängt. Die geringe Fallzahl sozial benachteiligter Frauen durch die hohe soziale Selektion wurde dabei von

⁵² <http://www.medizin-forum.de/afp/medizin/031117153122.a67uw966.html> Stand 12.12. 2005

der Projektleitung offen dargelegt. Die Berücksichtigung dieses Aspektes ist auch für die Endauswertung von Bedeutung, um entsprechende Zusammenhänge richtig interpretieren zu können. Die zeitlichen und personellen Ressourcen für die Auswertung sind voraussichtlich knapp bemessen. Dies begründet sich vor allem durch die zeitliche Verzögerung in der Feldphase. Die Projektleitung plant eine Vergabe des umfangreichen qualitativen Anteils der Erhebung als Diplomarbeit. Dies ermöglicht eine umfassende Bearbeitung und Auswertung und vermeidet einen „Datenfriedhof“. Sofern vor Abschluss der gesamten Auswertung beurteilbar, erfüllt das auswertungsmethodische Vorgehen zahlreiche Kriterien „Guter Epidemiologischer Praxis.“

6.6 Kommunikation und Public Health

“The quality of an epidemiological study goes beyond technique. Of equal importance is its actual benefit to the health of the population: do the findings contribute towards prevention or improved management of cases?” (Blettner et al., 2001:100)

Nach den „Leitlinien Guter Epidemiologischer Praxis“ sollten Forschungsergebnisse, deren Anliegen die Umsetzung von Ergebnissen in gesundheitswirksame Maßnahmen sind, der interessierten Öffentlichkeit und betroffenen Bevölkerungsgruppen zeitnah präsentiert und angemessen kommuniziert werden, beispielsweise in Form einer Empfehlung. (vgl. Hoffmann et al., 2004:12) Die Instrumente der Studie sollten dabei transparent dargelegt werden. (ebd.)

Bewertung/ Einschätzung

Obgleich die Gesamtdaten mit Interpretation noch ausstehen, lassen sich bereits jetzt zahlreiche Implikationen der Ergebnisse mit Public-Health-Relevanz ableiten. Dies können beispielsweise Handlungsempfehlungen für die Stillbetreuung sozial benachteiligter Frauen sein, um einem frühzeitigen Abstillen vorzubeugen. Die Projektleitung präsentierte erste Zwischenergebnisse direkt nach ihrer Erhebung einer Fachöffentlichkeit. Weitere Publikationen, seitens der Klinik und der Projektleitung, sind geplant. Es bestehen generell Ideen für weitere Forschungsvorhaben. Die Kooperationspartner werden die klinikbezogenen Aspekte aufgreifen, um die stillfördernde Qualität vor Ort weiter zu entwickeln und zu verbessern. Direkt nach Abschluss der Studie ist eine Präsentation der Ergebnisse mit interessierten MitarbeiterInnen der Klinik, externem Fachpersonal sowie allen am Projekt Beteiligten geplant.

Die Ergebnisse der Studie werden damit umfassend in der Öffentlichkeit kommuniziert und dienen als Grundlage, um daraus zukünftige, praxisbezogene Handlungsfelder für Stillförderung abzuleiten und Empfehlungen zu formulieren.

7 Zusammenfassung und Diskussion

Die Bewertung der Erhebung, durchgeführt in Anlehnung an die „Leitlinien Guter Epidemiologischer Praxis“, zeigt generell eine hohe Prozess- und Ergebnisqualität der Studie QuaSti. Für die Zusammenfassung und Diskussion werden nun die Forschungsfragen der Evaluation (vgl. Abschnitt 1.2) erneut aufgegriffen:

1. Welche Qualität hat das erhebungsmethodische Vorgehen der Studie QuaSti?

Als besondere Stärke im Rahmen der Prozessevaluation lassen sich die sorgfältige Entwicklung der Erhebungsinstrumente und die Pilotstudie, die Schulung und Betreuung der InterviewerInnen, die Datendokumentation, sowie die Berücksichtigung ethischer Aspekte hervorheben. Während der Feldphase wurden Probleme, insbesondere durch das Engagement der Projektmanagerin, flexibel und zeitnah bearbeitet und korrigiert. Beispielsweise ergaben sich aus dem Verteilungsmodus der Demographiebögen zu zwei unterschiedlichen Zeitpunkten (vor oder nach der Geburt) nicht valide Angaben zum Item „Anzahl der Personen im Haushalt“. Dieses Problem wurde frühzeitig erkannt und im laufenden Prozess korrigiert. (vgl. Abschnitt 5.3.2) Eine solche Vorgehensweise erhöht die erhebungsmethodische Qualität maßgeblich, weil damit eine Handlungsoptimierung im laufenden Prozess mit positiven Auswirkungen auf das Gesamtergebnis erreicht wird.

Als Problemfeld, mit Potential für verschiedene Formen des Selektions-Bias, erweist sich die Vorgehensweise in der Stichprobenziehung. (vgl. Abschnitt 5.2.1) Dieses Potential lässt sich nicht vollkommen ausschliessen und sollte deshalb im Rahmen der Gesamtauswertung deskriptiv aufgegriffen werden. Als größtes Problem während der Feldphase zeigte sich die zeitliche Verzögerung bei der Rekrutierung der Stichprobe. Die KlinikmitarbeiterInnen und die Projektleitung stellten im Konsens fest, dass eine Vergabe der Rekrutierung an eine „ausenstehende“ Person möglicherweise effektiver gewesen wäre und den Ablauf beschleunigt hätte.

2. Welche Qualität haben die erhobenen Daten hinsichtlich ihrer Güte?

Die Evaluation der Zwischenergebnisse mit detaillierter Analyse der Stichprobenqualität zeigt eine Repräsentativität der Stichprobe, bezogen auf die Zielpopulation. Damit liefert die Studie der Klinik repräsentative Daten, wenngleich einzelne Gruppen innerhalb der Stichprobe unterrepräsentiert sind (z.B. ambulant entbundene Frauen; vgl. Abschnitt 6.1.2). Im Rahmen der Zwischenauswertung konnte mit 80,8% eine hohe Ausschöpfungsquote realisiert werden. Die Analyse der Item-Nonresponse zeigt, dass bestimmten Items (z.B. dem Netto-Haushaltseinkommen) bei der Endauswertung besondere Aufmerksamkeit zukommen sollte. Die Bewertung der Validität, Reliabilität und

Objektivität ergibt, dass die Studie wesentliche Gütekriterien wissenschaftlicher Forschungsarbeit erfüllt.

Zahlreiche potentielle Probleme wurden von der Projektleitung bereits im Vorfeld antizipatorisch aufgegriffen. Dies ist wesentlich für die Qualität der Studie. Nicht alle Verzerrungen und Fehler innerhalb eines wissenschaftlichen Forschungsprozesses können eliminiert werden. (vgl. Blettner et al., 2001: 97ff.) Wesentlich ist, dass das Potential erkannt und gegebenenfalls benannt wird. (ebd.) Da die Evaluation eine Zwischenauswertung darstellt, konnten die identifizierten Problemfelder noch im laufenden Prozess korrigierend aufgegriffen werden.

Auch die freundliche und kooperative Zusammenarbeit zwischen Projektteam und KlinikmitarbeiterInnen wirkte sich positiv auf den Studienablauf aus.

3. Welchen Zielerreichungsgrad haben die Ergebnisse zur Beantwortung der Forschungsfrage der Studie QuaSti?

Zusammenfassend betrachtet hat die Studie einen hohen Zielerreichungsgrad und ist zur Beantwortung der zentralen Forschungsfrage (Prävalenz des Stillens; Versorgungsqualität der geburtshilflichen Abteilung) der Studie geeignet. Generell lässt sich eine deutliche Zufriedenheit der Frauen mit der Qualität und Stillunterstützung der Klinik ableiten. Die Zielpopulation weist ein besonderes soziodemographisches Profil mit einer großen Gruppe sozial privilegierter Frauen auf.

Ein Vergleich der Zwischenergebnisse mit Daten der Studie SuSe zeigt, dass die Klinik eine höhere Stillquote erreicht als der bundesweite Durchschnitt.⁵³ Die Daten lassen jedoch keinen Rückschluss darüber zu, warum dies so ist (beispielsweise eine steigende Stilldauer als gesellschaftliche Entwicklung). Der Vergleich ist allenfalls als Trend verwertbar, da sowohl die Studienpopulationen als auch die Kliniken (Vorhandensein stillfördernder Programme oder nicht) in ihren Konstellationen differieren. Die Daten der Studie QuaSti sind ausschließlich für die Zielpopulation repräsentativ.

In der SuSe-Studie wurden Risikofaktoren für kein oder nur kurzes Stillen generiert, die durch strukturelle und organisatorische Bedingungen im Krankenhaus entstehen (z.B. „Rooming-in“-Praktiken, Zeitpunkt des ersten Stillens nach Geburt). (vgl. Abschnitt 3.4) Die Erhebung zeigt, dass die Klinik diese Risikofaktoren erfolgreich reduziert.

⁵³ Seibt, A., Deneke, C. (2005): Präsentation der Ergebnisse/ Studie QuaSti : EUPHA Conference (European Public Health Association), Health Promotion Section Workshop, Graz, Österreich, November 2005

Mit der Studie QuaSti erfolgt die bundesweit erste, systematische Untersuchung eines „Stillfreundlichen Krankenhauses.“ Die Ergebnisse der Studie sind dazu geeignet, Hypothesen zu generieren, beispielsweise hinsichtlich Sozialstatus und stillfördernder Intervention. (Beispiel: *„Die stillfördernde Intervention „stillfreundliches Krankenhaus“ erhöht die Stilldauer bei sozial benachteiligten Frauen.“*) Diese Zusammenhänge sind von besonderem Interesse, da sich daraus der Handlungsbedarf für vulnerable Zielgruppen ableiten lässt. Die umfassende Datenlage der Studie QuaSti kann als Ausgangspunkt dienen, um Verbesserungs- und Interventionsbedarf für stillende Frauen wesentlich zielgruppenspezifischer aufzuzeigen. Ein Rückschluss auf die Effektivität der Intervention „stillfreundliches Krankenhaus“ ist jedoch nicht zulässig, da mit der Querschnittserhebung keine kausalen Zusammenhänge erfasst werden können.

Um die Wirksamkeit der Intervention „Stillfreundliches Krankenhaus“ oder auch die aus der Studie QuaSti generierbaren Hypothesen zu untersuchen, ist eine weiterführende Studienanordnung mit einer geeigneten Vergleichsgruppe notwendig. Tiefergehende Aspekte, die das Stillen beeinflussen, wie beispielsweise Scham, sexualisierte Gewalt und andere psychologische Faktoren können mit der Studie QuaSti nicht erfasst werden. Auch hier ergibt sich Potential für weitere Untersuchungen.

Durch die Studie QuaSti ist ein validiertes und evaluiertes Erhebungsinstrument entstanden. Dies stellt eine geeignete Grundlage für sich anschließende Forschungsvorhaben dar.

Reflexion des Evaluationsprozesses

„We have learned that the choice of methods (and measures and instruments and data) depends much more on the type of question being asked than on the qualities of any particular method.“ (Chelminsky, 1995:6)

Aus der vorliegenden Arbeit lassen sich folgende Implikationen zusammenfassend ableiten: Die Evaluation der Studie QuaSti demonstriert, dass eine systematische Begleitung und Bewertung des Studienverlaufs die Qualität epidemiologischer Untersuchungen maßgeblich sichern und verbessern kann. Die Evaluation macht den Forschungsprozess transparent und untermauert die empirische Evidenz der wissenschaftlichen Erkenntnisse. Dies impliziert ihre hohe Relevanz als Instrument der Qualitätssicherung von Studien.

Es entsteht ein kontinuierlicher Reflexionsprozess mit direkter Rückkopplung zum praktischen Handlungsfeld. Dabei ist die Beobachtung und Bewertung der Prozessqualität von besonderer Wichtigkeit, weil diese maßgeblich die Qualität der Ergebnisse bestimmt.

Aus diesem Grund hat sich der frühzeitige Beginn der Evaluation bei der Studie QuaSti, kurz vor Beginn der Prozessphase, als sehr vorteilhaft erwiesen.

Für die Autorin lag die Herausforderung der Evaluationsstudie in der Auswahl der angemessenen Fragen bzw. Kriterien, um eine Bewertung vorzunehmen, welche die Studie QuaSti adäquat abbildet. Das Bemühen lag darin, die Grundlagen der Bewertung transparent und objektiv zu gestalten. Doch auch Evaluationen selbst müssen qualitätsgesichert durchgeführt und kritisch reflektiert werden. Als Grundlage dieser Arbeit wurden deshalb die „Standards für Evaluation“ gewählt. Diese boten eine wertvolle Orientierungshilfe bei der Planung und Durchführung der Evaluation. Weiterhin sind sie eine geeignete Grundlage, um diese kritisch zu bewerten. Eine Anwendung der „Standards für Evaluation“ dient damit der Qualitätssicherung und Weiterentwicklung dieser Evaluationsstudie als auch, ganz generell, des Wissenschaftsbereiches der Evaluationsforschung.

Evaluationsforschung muss in der Regel komplexe Zielkriterien beachten: objektive und subjektive gesundheitliche Ereignisse, Zufriedenheit der Klienten und Stakeholder oder auch Kosten einer Maßnahme. (Füllgraff, 1999: 35) Um dieser Komplexität gerecht zu werden, eignet sich eine triangulative Vorgehensweise. Triangulation bedeutet die Anwendung verschiedener qualitativer und quantitativer Messverfahren, sowie Bewertungen, durchgeführt auf der Grundlage von Sekundärliteratur. Die Autorin hat den „Methoden-Mix“ in sinnvoller Ergänzung seiner Komponenten zueinander erlebt. Hierdurch wurde es erleichtert, die Vielschichtigkeit des Evaluationsobjektes zu erfassen und darzustellen.

Die Autorin hat während des gesamten Evaluationsprozesses größtmögliche Unterstützung aller Beteiligten erfahren. Auch dieser Aspekt ist von wesentlicher Bedeutung für die Qualität einer Evaluation.

Literaturverzeichnis

- Abou-Dakn, M., Strecker, J.R., (2003): Einflussfaktoren der Entbindungsklinken auf den Stillbeginn, Zentralbibliothek Gynäkologie, Stuttgart: Georg Thieme Verlag KG, 386-392
- Badura, B., Strodtholz, P. (2003): Soziologische Grundlagen der Gesundheitswissenschaften: Handbuch Gesundheitswissenschaften, München: Juventa-Verlag, 3. Auflage, 145- 168
- Badura, B., Strodtholz, P. (2000): Qualitätsförderung, Qualitätsforschung und Evaluation im Gesundheitswesen: Das Public Health Buch, München: Urban& Fischer Verlag, 574- 584
- Badura, B., Siegrist, J., (Hrg.) (1999): Einleitung: Evaluation im Gesundheitswesen, Ansätze und Ergebnisse, München: Juventa-Verlag, 9-12
- Beaglehole, R., Bonita, R., Kjellström, T. (1997): Einführung: Einführung in die Epidemiologie, Bern: Hans Huber Verlag, 14
- Bergmann, R., Bundeszentrale für gesundheitliche Aufklärung (Hrg.) (2001): Muttermilch und Allergie: Stillen und Muttermilchernährung, Gesundheitsförderung Konkret, Band 3, Köln: BZgA, 249- 253
- Beywl, W. (2002): Standards für Evaluation, Deutsche Gesellschaft für Evaluation e.V. (DeGEval): <http://www.degeval.de/calimero/tools/proxy.php?id=227> Stand 14. 07.2005
- Blettner, M., Heuer, C., Razum, O., (2001): Critical reading of epidemiological papers. A guide: The European Journal of Public Health, London: Oxford University Press, 97-101
- Breiholz, H., Duschek, KJ., Hansch, E., Nöthen, M. (2005): Leben und Arbeiten in Deutschland: Ergebnisse des Mikrozensus 2004, Wiesbaden: Statistisches Bundesamt www.destatis.de Stand 12.11.2005
- Broadfoot, M., Britten, J., Tappin, DM., Mac Kenzie, JM (2005): The Baby-Friendly Hospital Initiative and the breast feeding rates in Scotland, British Medical Journal, online: <http://fn.bmjournals.com/cgi/content/abstract/90/2/F114> Stand 19.07. 2005
- Byrne, D.: Schutz, Förderung und Unterstützung des Stillens in Europa: Ein Aktionsplan, Dokument der EU- Konferenz zur Förderung des Stillens in Europa: <http://www.ilca.org/liasion/European%20Blueprint%20PDF/Blueprint%20German.pdf> Stand 15.08.2005
- Cattaneo, A., (2004): Breastfeeding in Europe: a blueprint for action: Journal of Public Health, Berlin: Springer-Verlag, 89- 96
- Chelminsky, E. (1995): New dimensions in evaluation: World Bank Operations Evaluations Department (OED): Evaluation and Development: proceedings of the World Bank Conference Washington D.C., 3-11

- Christiansen, G., Bundeszentrale für gesundheitliche Aufklärung (Hrg.) (1999): Die Aufgabenstellung: Evaluation- Ein Instrument zur Qualitätssicherung in der Gesundheitsförderung, Band 8, Köln: BZgA, 14- 17
- Christiansen, G., Bundeszentrale für gesundheitliche Aufklärung (Hrg.) (1999): Monitoring von Implementationsprozessen: Evaluation- Ein Instrument zur Qualitätssicherung in der Gesundheitsförderung, Band 8, Köln: BZgA, 48- 55
- Dulon, M., Kersting, M., Schach, S. (2001): Duration of breastfeeding and associated factors in Western and Eastern Germany: Acta Paediatrica, London: Taylor& Francis, Vol. 90, 931- 935
- Dulon, M., Kersting, M., (2005): Erfassung des Stillverhaltens in der Bundesrepublik Deutschland: die SuSe-Studie, Bundesministerium für Gesundheit und Soziale Sicherung: www.bmgs.bund.de/deu/gra/themen/forschung/2305_2371.php Stand 28.03.2005
- Diekmann, A., (2000): Telephonische Befragung [sic] : Empirische Sozialforschung, Grundlagen, Methoden, Anwendungen, 6. Auflage, Hamburg: rororo, 429- 439
- Diekmann, A., (2000): Zufallsstichproben in der Praxis : Empirische Sozialforschung, Grundlagen, Methoden, Anwendungen, 6. Auflage, Hamburg: rororo, 355- 369
- Elkeles, T. (2004): Evaluation: Definitionen, Abgrenzungen und Prinzipien: Evaluation im Gesundheitswesen, e-learning-Seminar der „Hochschule für Angewandte Wissenschaften“ Hamburg, Wintersemester 2004, 4- 15
- Elkeles, T. (2004): Voraussetzungen und Methoden der Evaluationsforschung: Evaluation im Gesundheitswesen, e-learning-Seminar der „Hochschule für Angewandte Wissenschaften“ Hamburg, Wintersemester 2004, 33- 64
- Flick, U. (2004): Triangulation qualitativer und quantitativer Forschung: Triangulation, Wiesbaden: Verlag für Sozialwissenschaften, 67- 84
- Fontana, A., Frey, J. (2000): From Structured Questions to Negotiated Text: Handbook of Qualitative Research, 2. Auflage, London: Sage Publications, 645
- Fülgraff, G. (1999): Perspektiven der Public-Health-Forschung: Public-Health-Forschung in Deutschland, Bern: Verlag Hans Huber, 33- 36
- Frick, JR., (2005): Deutsches Institut für Wirtschaftsforschung: Zeitreihen zur Entwicklung von Indikatoren zu zentralen Lebensbereichen, SOEP-Monitor, www.diw.de/deutsch/spo/service/soepmonitor/soepmonitor_person2004.pdf Stand 01.12. 2005
- Garcia, CM., Garcia, CS., Juan MP., Ruiz ME., Parellada E. (2005): Breast-feeding: can staff positively affect its duration (Article in Spanish), Atencion Primaria, 35, 295- 300 http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=15826512&dopt=Citation Stand 12.07.2005
- Geyer, S. (2003): Einführung: Forschungsmethoden in den Gesundheitswissenschaften, Weinheim, München: Juventa- Verlag, 13- 22

- Geyer, S. (2003): Daten und ihre Eigenschaften: Forschungsmethoden in den Gesundheitswissenschaften, Weinheim, München: Juventa- Verlag, 23- 41
- Geyer, S. (2003): Arten der Befragung: Forschungsmethoden in den Gesundheitswissenschaften, Weinheim, München: Juventa- Verlag, 43-89
- Geyer, S. (2003): Auswahlverfahren und die Wahl eines Untersuchungsdesigns: Forschungsmethoden in den Gesundheitswissenschaften, Weinheim, München: Juventa- Verlag, 157- 180
- Geyer, S. (2003): Fragen und Antworten: Forschungsmethoden in den Gesundheitswissenschaften, Weinheim, München: Juventa- Verlag, 91- 132
- Gordis, L. (2001) : Einleitung: Epidemiologie, Marburg: Kilian- Verlag, 3- 14
- Gordis, L. (2001): Fall- Kontroll- Studien und Querschnittsstudien: Epidemiologie, Marburg: Kilian- Verlag, 165- 185
- Gould, EA. (1994): Breastfeeding Support: The Contribution of Health Professions, University of Wales: Dissertation for the degree of Master in Public Health
- Grossmann, R, Scala, K., Bundeszentrale für gesundheitliche Aufklärung (Hrg.) (2000): Setting- Ansatz in der Gesundheitsförderung: Leitbegriffe der Gesundheitsförderung, Reihe „Blickpunkt Gesundheit“ 3, 3. Auflage, Schwabenheim: Verlag Peter Sabo, 100-101
- Guise, MD., Palda V., Westhoff C., (2003): The Effectiveness of Primary Care-Based Interventions to Promote Breastfeeding: Systematic evidence Review and Meta-Analysis for the US Preventive Services Task Force: Annals of Family Medicine. <http://www.Annafammed.org> Stand 15.07.2005
- Heiberg, E., Helsing, E.: Changes in Breastfeeding practices in Norwegian maternity wards: national surveys 1973, 1982 and 1991: Acta Paediatrica, London: Taylor& Francis: 719-724
- Hellmeier, W., Brand, H., Laaser, U., Hort, A., Hurrelmann, K. (Hrg.) (2003): Epidemiologische Verfahren in den Gesundheitswissenschaften: Handbuch Gesundheitswissenschaften, 3. Auflage, München: Juventa- Verlag, 231- 257
- Hoffmann, W., Latza, U., Terschüren, C. (2004): Leitlinien und Empfehlungen zur Sicherung von Guter Epidemiologischer Praxis (GEP)- Langversion- überarbeitete Fassung nach Evaluation <http://medweb.uni-muenster.de/institute/epi/dae/Empfehlungen.doc> Stand 12.06.2005
- Hoffmann, ,U., Terschüren, C., Holle, R., Kamtsiuris, P.,Bergmann, M, Kroke, A., Sauer, S., Stang, A., Latza, U. (2004a): Zum Problem der Response in epidemiologischen Studien in Deutschland (Teil 2): Das Gesundheitswesen, Stuttgart: Georg Thieme Verlag KG, 482- 491
- Horta, BL., Kramer, MS., Platt, RW. (2001): Maternal smoking and the risk of early weaning: a meta- analysis: American Journal of Public Health, New York: Oxford University Press, Vol. 91, 304- 307

- Howie PW, Forsyth J., Ogston SA. (1990): Protective Effect of breast feeding against infection, British Medical Journal, London: BMJ Publishing Group, 11-15
- Huch, R., Bergmann, R., Kersting, M., Schneider, H., Schubinger, G. (1996): Stillen- nur wegen des stillfreundlichen Krankenhauses, PerinatalMedizin, Berlin: Springer- Verlag: Vol. 8, Nr. 4, 123- 132
- Hurrelmann, K., Laaser, U. (Hrsg.), (2003): Sozialwissenschaftliche Verfahren in den Gesundheitswissenschaften: Handbuch Gesundheitswissenschaften, 3. Auflage, München: Juventa-Verlag, 231- 257
- Janßen, C., Crispin, A., Weitkunat, R., Abel, T. (1999): Lebensstile und Gesundheit- Eine empirische Studie zum Verständnis spezifischer Verhaltensmuster bei Erwachsenen im zweiten und dritten Lebensabschnitt-: Public-Health-Forschung in Deutschland, Bern: Verlag Hans Huber, 171- 179
- Jorch, G, Fischer, D., Beyer, U. (2003): Prävention des plötzlichen Säuglingstodes: Monatsschrift Kinderheilkunde, Hamburg: Springer- Verlag, Heft 151, 5, 514- 519
- Jöckel, K.-H., Babitsch, B., Bellach, B.-M., Bloomfield, K., Hoffmeyer-Zlotnik, J., Winkler, J., Wolf, C. (1997): Messung und Quantifizierung soziographischer Merkmale in epidemiologischen Studien: Empfehlungen der Deutschen Arbeitsgemeinschaft für Epidemiologie (DAE) der GMDS und DGMS: www.daeppi.de Stand 01.12.2005
- Kamtsiuris, P., Bergmann, K.E., Dippelhofer, A., Hölling, H., Kurth, B.-M.,Thefeld, W. (2002): Der Pretest der Kinder und Jugendgesundheitsurveys: Methodische Aspekte und Durchführung: Das Gesundheitswesen, 64, Sonderheft 1, Stuttgart: Georg-Thieme Verlag, S99- S106
- Kersting, M. (2005): Pizza und Süßes statt Gemüse? Was essen Kinder und Jugendliche heute? Forschungsinstitut für Kinderernährung (FKE) Dortmund: <http://www.kinderumweltgesundheit.de/KUG/index2/pdf/aktuelles/essen.pdf> Stand 06.19.2005
- Kersting, M. (2004): Stillen in Deutschland- Datenlage: Zehn Jahre Nationale Stillkommission in Deutschland, Internationales Symposium, Bundesinstitut für Risikobewertung (BfR), Berlin: <http://www.bfr.bund.de/cm/251/Abstractband.pdf> Stand 07.08.2005
- Latza, U., Stang, A., Bergmann, M. (2004): Zum Problem der Response in epidemiologischen Studien in Deutschland (Teil 2): Das Gesundheitswesen; Stuttgart: Georg Thieme Verlag 66, 326- 336
- Lamounier, JA. (1998): Experience of the Baby Friendly Hospital Initiative: Rev Assoc Med Bras., 44, 319-24, Pub Med online: www.ncbi.nlm.nih.gov/entrez/query. Stand 27.07. 2005
- Lavrakas, P. (1993): Telephone Survey Methods, 2. Auflage, Vol. 7, London: Sage Publications, 13
- Lee, B. (2005): Theories of Evaluation: Evaluationsforschung, Opladen: Leske & Budrich, 135- 173

- Lübbert, D. (1999): Zusammenfassung verschiedener Texte zur Evaluationsforschung, Universität Ruhr- Bochum, Fakultät für Sozialwissenschaften, online: <http://www.luebbert.net/uni/methoden/eval/index.php> Stand 14.06.2005
- Lukassowitz, I. (2004): „Wie lange stillen Mütter in Deutschland“, Studie des Bundesinstituts für Risikobewertung (BfR): www.idw-online.de/pages/de/news?print=1&id=83506 Stand 12.06.2005
- Machan, C. (2004): Triangulative Evaluation regionaler Befundnetzwerke aus Sicht niedergelassener Ärzte: <http://www.egms.de/en/meetings/gmds2004/04gmds049.shtml> Stand 10.07.2005
- Martin, RM., Gunnell, D. (2005): Breast-feeding and childhood cancer: A systematic Review, International Journal of Cancer, Heidelberg: Wiley InterScience, 1020- 1031
- Mertens, D. (2004): Institutionalizing Evaluation in the United States of America: Evaluationsforschung, Opladen: Leske & Budrich, 45- 60
- Mielk, A. (2000): Ansätze zur Verringerung der gesundheitlichen Ungleichheit: Soziale Ungleichheit und Gesundheit, Bern: Hans- Huber- Verlag, 367- 397
- Müller, E. (2005): Studie „Stillen in Bayern“, Bayerisches Staatsministerium für Umwelt, Gesundheit und Verbraucherschutz, www.stmugv.bayern.de/de/gesundheit/qiba/sons-proj/stillverh.htm Stand: 15.06.2005
- Naidoo, J., Wills, J. (2000): Evaluation in health promotion, Health Promotion; Foundations for Practice, London: Bailliere Tindall, 2. Auflage, 369- 388
- Naidoo, J., Wills, J. (2003) :Sichtweisen und Verständnis der Gesundheit: Lehrbuch der Gesundheitsförderung, Köln: Bundeszentrale für gesundheitliche Aufklärung, 5- 26
- Nylander, G. (2004): Warum stillen in Norwegen fast alle? Warum gibt es in Europa so große Unterschiede? Zehn Jahre Nationale Stillkommission in Deutschland, Internationales Symposium, Bundesinstitut für Risikobewertung (BfR), Berlin: <http://www.bfr.bund.de/cm/251/Abstractband.pdf> Stand 07.08.2005
- Palda, V., Guise, J.-M., Wathen, C. (2004): Interventions to promote breast- feeding: applying the evidence in clinical practice, http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=15023925 Stand 06.08.2005
- Palmer, G. (1988): Why breastfeeding is political: The Politics of Breastfeeding, London: Pandora Press, 1-11
- Pfaff, H., Bentz, J. (2000): Subjektive Daten- Objektive Analyse: Das Public Health Buch, 1. Auflage, München: Urban & Fischer, 310- 328
- Philipp, BL., Merewood, A., Miller LW., Chawla N., Murphy- Smith, MM., Gomes JS., Cimo S., Cook, JT. (2001): Baby- friendly hospital initiative improves breastfeeding initiation rates in a US hospital setting: Pediatrics, American Academy of Paediatrics, 108 (3): 766- 768
- Pocock, S., Collier, TJ., Dandreo, KJ., de Stavola, B., Goldman, MB., Kalish, LA., Kasten, L., McCormack, V., (2004): Issues in the reporting of epidemiological studies: a survey

- of recent practice, British Medical Journal, online:
<http://bmj.bmjournals.com/cgi/content/abstract/329/7471/883> Stand 12.09.2005
- Przyrembel, H. (2004): Warum braucht Deutschland eine Nationale Stillkommission, Zehn Jahre Nationale Stillkommission in Deutschland, Internationales Symposium, Bundesinstitut für Risikobewertung (BfR), Berlin:
<http://www.bfr.bund.de/cm/251/Abstractband.pdf> Stand 07.08.2005
- Przyrembel, H., Bundeszentrale für gesundheitliche Aufklärung (Hrg.), (2001): Die Vorteile der Muttermilch: Stillen und Muttermilchernährung, Gesundheitsförderung Konkret, Band 3, Köln: BZgA, 13- 24
- Przyrembel, H., Bundeszentrale für gesundheitliche Aufklärung (Hrg.), (2001): Vorwort: Stillen und Muttermilchernährung, Gesundheitsförderung Konkret, Band 3, Köln: BZgA, 11
- Rushton, L., (2000): Reporting of occupational and environmental research: use and misuse of statistical and epidemiological methods: Occupational Environmental Medicine; New York: Lippincott Williams & Wilkins, 57:1-9
- Rychetnik, L., Frommer, M., Hawe, P., Shiell, A. (2002): Criteria for evaluating evidence on public health interventions: Journal of Epidemiology and Community Health; London: BMJ Publishing group, 56:119-127
- Schäfer, T. (2004): Einführung: Projektdurchführung und Studiendesign, Skript der Hochschule Magdeburg-Stendal (FH), Verbundprojekt „Hochschulen für Gesundheit“, 10 <http://www.hochges.de> Stand 15.05. 2005
- Schäfer, T. (2004): Deskriptive Studien: Projektdurchführung und Studiendesign, Skript der Hochschule Magdeburg-Stendal (FH), Verbundprojekt „Hochschulen für Gesundheit“, 20-30 <http://www.hochges.de> Stand 15.05. 2005
- Schäfer, T. (2004): Studiendurchführung: Projektdurchführung und Studiendesign, Skript der Hochschule Magdeburg-Stendal (FH), Verbundprojekt „Hochschulen für Gesundheit“, 81- 94 <http://www.hochges.de> Stand 15.05. 2005
- Scheele, M., Abou-Dakn, M., Strecker, J-R. (2003): Does breast-feeding prevent breast cancer? Stillen als Brustkrebsprävention- eine Übersicht: Zentralbibliothek Gynäkologie, Stuttgart: Georg Thieme Verlag KG, 48- 52
- Schöffski, O., Graf v.d. Schulenburg, J.-M., (Hrg.) (2002): Vorwort zur ersten Auflage, Gesundheitsökonomische Evaluation, Berlin: Springer- Verlag, 1-3
- Seibt, A., Deneke, C. (2005): Abstract: „QuaSti- Breastfeeding- Promotion- Study“: Health Promotion Section Workshop', EUPHA Conference (European Public Health Association) November 2005, Graz, Österreich, internes Dokument, Stand Juli 2005
- Siegrist, J., Badura, B. (Hrsg.) (1999): Chancen und Grenzen sozialwissenschaftlicher Evaluationsforschung: Evaluation im Gesundheitswesen, Ansätze und Ergebnisse, München: Juventa- Verlag, 43- 51

- Siegrist, J., Geyer, S. (2003): Sozialwissenschaftliche Verfahren in den Gesundheitswissenschaften: Handbuch Gesundheitswissenschaften, 3. Auflage, München: Juventa- Verlag, 259- 279
- Scherbaum, V., Perl, FM., Kretschmer, U. (2004): Stillen. Frühkindliche Ernährung und reproduktive Gesundheit, Köln: Deutscher Ärzte- Verlag, 135- 155
- Stockmann, R. (Hrsg.) (2004): Evaluation in Deutschland: Evaluationsforschung, Opladen: Leske& Budrich, 13- 43
- Thefeld, W. (2002) „Der Kinder- und Jugendgesundheitsurvey: Ermittlung des Gesundheitsverhaltens von Eltern und Kindern“, Robert-Koch-Institut, Berlin: <http://www.kiggs.de/experten/downloads/dokumente/Sonderheft036.pdf> Stand 07.08.2005
- Trojan, A., Siegrist, J., Deutsche Gesellschaft für Public Health (Hrsg.) (1999): Ziel: Verbesserung der Gesundheit durch Prävention und Gesundheitsförderung: Public-Health- Forschung in Deutschland, Bern: Hans- Huber- Verlag, 149- 151
- Trojan, A., Legewie, H. (2001): Politische Strategien und Strukturen: Nachhaltige Gesundheit und Entwicklung, Frankfurt: Verlag für Akademische Schriften (VAS), 215- 289
- UNICEF (2005): The Baby- Friendly Hospital Initiative: www.unicef.org/programme/breastfeeding/baby.htm Stand 15.07.2005
- WHO (2004): Infant feeding, Nutrition and food security: www.Euro.who.int/eprise/main/WHO/Progs/NUT/Infant/20020730_1 Stand 19.08.2005
- WHO (2005) Monitoring and Evaluating child health programmes: WHO Department of Child and Adolescent Health and Development: http://www.who.int/child-adolescent-health/Overview/Child_health/child Stand 19.08.2005
- Wichmann, HE., Jöckel, KH., Raspe, HH., Kohlmann T. (2005): Entwurf einer Checkliste zur ethischen Begutachtung epidemiologischer Studien.“ <http://daepi.de> Stand 17.08.2005
- Widmer, T., Stockmann, R. (Hrsg.), (2004): Qualität der Evaluation- Wenn Wissenschaft zur praktischen Kunst wird: Evaluationsforschung, Opladen: Leske& Budrich, 83- 109
- Widström, A. (2004): Closeness, breastfeeding and bonding. The importance of the first weeks: Zehn Jahre Nationale Stillkommission in Deutschland, Internationales Symposium, Bundesinstitut für Risikobewertung (BfR), Berlin: <http://www.bfr.bund.de/cm/251/Abstractband.pdf> Stand 07.08.2005
- Wottawa, H., Thierau, H. (2003): Wissenschaft, Moral und die Grenzen der Planbarkeit: Lehrbuch Evaluation, 3. Auflage, Bern: Hans Huber Verlag, 13-22
- Wottawa, H., Thierau, H. (2003): Durchführung von Evaluationsstudien: Lehrbuch Evaluation, 3. Auflage, Bern: Hans Huber Verlag, 13-22

Eidesstattliche Erklärung

Ich versichere, dass ich die vorliegende Arbeit ohne fremde Hilfe selbstständig verfasst und nur die angegebenen Hilfsmittel benutzt habe. Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen sind unter Angabe der Quelle kenntlich gemacht.

Hamburg, der 20. März 2006

Marlies Jöllnbeck

Anhang

Demographiebogen.....	A-1
Fragebogen, Seite 1.....	A-2
Fragebogen, Seite 2.....	A-3
Fragebogen, Seite 3.....	A-4
Fragebogen, Seite 4.....	A-5
Fragebogen, Seite 5.....	A-6