

Entwicklung zweier methodologischer Leitlinien zur Beurteilung von Blutungsmustern

**Eine Hilfe für
Zulassung und Verschreibung
auf der Basis einer empirischen Analyse**

von Diplom-Mathematiker
Christoph Gerlinger

von der Fakultät VIII - Wirtschaft und Management
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Gesundheitswissenschaften
- Dr. P.H. -

genehmigte Dissertation

Gutachter: Prof. Dr. M Harvey Brenner
Gutacher: PD Dr. Jens Wessel

Tag der wissenschaftlichen Aussprache: 19. Dezember 2005

Berlin, 2005

D 83

Danksagung

Mein Dank gilt Herrn Prof. Dr. M. Harvey Brenner und Herrn PD Dr. Jens Wessel für die Betreuung dieser Arbeit. Danken möchte ich Herrn Dr. Jürgen Dinger für die Unterstützung meines Vorhabens und die Erlaubnis, die Blutungskalender für diese Arbeit nutzen zu dürfen. Mein besonderer Dank gilt Herrn Dr. Jan Endrikat für die Anregung zu dieser Arbeit und für die regen Diskussionen über medizinische Fragen.

Meiner Familie danke ich recht herzlich für ihre geduldige Unterstützung.

Inhaltsverzeichnis

1	Einleitung	9
1.1	Historie	9
1.2	Das Blutungsmuster als Bestandteil der individuellen Lebensqualität.....	10
1.3	Natürliches Blutungsverhalten prämenopausaler Frauen.....	11
1.4	Natürliches Blutungsverhalten peri- und postmenopausaler Frauen.....	12
1.5	Relevanz des Blutungsmusters für die arzneimittelrechtliche Zulassung neuer Präparate	13
2	Ziel und Methodik	15
2.1	Bisherige Ansätze zur Analyse und Bewertung von Blutungsmustern.....	15
2.2	Ziel dieser Arbeit.....	16
2.3	Herangehensweise.....	16
2.3.1	Literaturrecherche	17
2.3.2	Statistische Datenanalyse	18
2.3.3	Entwicklung der Leitlinie zur Beurteilung von Blutungsmustern.....	19
3	Literaturrecherche	20
3.1	Methodik.....	20
3.1.1	Datenbankauswahl	20
3.1.2	PubMed Datenbankabfrage	21
3.1.3	Expertenbefragung	24
3.1.4	Literaturrecherche in interner Datenbank eines pharmazeutischen Herstellers	25
3.2	Ergebnisse der Literaturrecherche	26
3.3	Auswertung der Literatur	37
3.3.1	Blutungsstärke	37
3.3.1.1	WHO Definition	37
3.3.1.2	Schering AG	38
3.3.1.3	NV Organon	39
3.3.1.4	Ortho-McNeil Pharmaceuticals Inc.	40
3.3.1.5	Sonstige Definitionen.....	40

3.3.2	Blutungshäufigkeit	41
3.3.3	Blutungsmuster	43
3.3.3.1	Zyklus basierte Blutungsmuster-Methoden.....	44
3.3.3.2	Episoden basierte Blutungsmuster-Methoden.....	49
3.4	Zusammenfassung der Literaturrecherche	58
4	Empirische Untersuchung von Blutungsmustern	59
4.1	Datenbeschreibung	59
4.1.1	Fehlende Werte	62
4.2	Wahl des statistischen Verfahrens	66
4.3	Clusteranalyse	67
4.3.1	Überblick über Clusterverfahren	67
4.3.2	Abstandsbegriff	69
4.3.3	Hierarchisch agglomerative Verfahren zur Clusteranalyse	75
4.3.3.1	Single linkage Methode	77
4.3.3.2	Complete linkage Methode	77
4.3.3.3	Average linkage Methode	78
4.3.3.4	Methode von Ward	78
4.3.4	Dendrogramm	79
4.3.5	Wahl der Anzahl der Cluster	81
4.3.6	Software	83
4.4	Anwendung auf Blutungskalender	83
4.4.1	Abstandsdefinition zweier Blutungskalender	85
4.4.2	Ergebnisse der Clusteranalyse	89
4.4.2.1	Ergebnisse der Methode von Ward	91
4.4.2.1.1	Anzahl der Cluster	92
4.4.2.1.2	Lösung mit drei Clustern.....	97
4.4.2.1.3	Lösung mit vier Clustern.....	98
4.4.2.1.4	Lösung mit sechs Clustern.....	100
4.4.2.1.5	Lösung mit zwölf Clustern.....	101
4.4.2.2	Ergebnisse der anderen Cluster Methoden.....	105
4.4.2.2.1	Single Linkage.....	105
4.4.2.2.2	Complete Linkage.....	106

4.4.2.2.3 Average Linkage	108
4.5 Inhaltliche Interpretation der Ergebnisse	110
4.5.1 Lösung mit drei Clustern.....	111
4.5.2 Lösung mit vier Clustern.....	114
4.5.3 Lösung mit sechs Clustern.....	116
4.5.4 Lösung mit zwölf Clustern.....	119
4.5.5 Vergleich der Lösungen.....	123
4.6 Zusammenfassung der empirischen Untersuchung	124
5 Vergleich der Ergebnisse der Literaturrecherche und der empirischen Analyse	129
5.1 Methodik des Vergleichs	129
5.2 Ergebnis des Vergleichs	129
5.2.1 WHO Klassifikation.....	129
5.3 Entwicklung der neuen Methode zur Analyse der Zykluskontrolle	135
5.3.1 Ausgangsmethode.....	135
5.3.2 Erster Grenzfall der Ausgangsmethode.....	136
5.3.3 Zweiter Grenzfall der Ausgangsmethode	137
5.3.4 Dritter Grenzfall der Ausgangsmethode	138
5.3.5 Neue Methode zur Analyse der Zykluskontrolle	139
6 Klinische Relevanz und Beitrag zu Public Health	144
6.1 Klinische Relevanz	144
6.2 Beitrag zu Public Health	146
7 Leitlinien zur Verschreibung und Zulassung	149
7.1 Einleitung	149
7.2 Erhebung des Blutungsmusters	150
7.3 Leitlinie zur Verschreibung	151
7.3.1 Darstellung des Blutungsmusters für Einzelfälle	151
7.3.2 Bewertung des Blutungsmusters für Einzelfälle	153
7.3.3 Empfehlung zur Präparatauswahl und zur Dosistitration anhand des Blutungsmusters	154
7.4 Leitlinie zur Zulassung	155
7.4.1 Darstellung des Blutungsmusters	155

7.5	Bewertung des Blutungsmusters im Rahmen der Arzneimittelzulassung	161
8	Zusammenfassung und Ausblick	163
9	Abkürzungen	173
10	Literatur	175

Tabellenverzeichnis

Tabelle 1:	Anzahl der Treffer in PubMed	23
Tabelle 2:	Ergebnisse der Literaturrecherche	30
Tabelle 3:	Blutungsmusterindices nach WHO	55
Tabelle 4:	Klinisch relevante Blutungsmuster nach WHO	57
Tabelle 5:	Anzahl und Gesamtlänge der verfügbaren Blutungskalender nach Indikation.....	62
Tabelle 6:	Anzahl der verfügbaren Blutungskalender nach Länge und Indikation	63
Tabelle 7:	Kontingenztafel der Übereinstimmung von p binären Merkmalen....	72
Tabelle 8:	gebräuchliche hierarchisch agglomerative Clusterverfahren.....	77
Tabelle 9:	Anzahl Blutungskalender mit identischem Blutungsmuster	90
Tabelle 10:	Anzahl der Elemente pro Cluster - Lösung mit drei Clustern	98
Tabelle 11:	Anzahl der Elemente pro Cluster nach Indikation - Lösung mit drei Clustern	99
Tabelle 12:	Anzahl der Elemente pro Cluster - Lösung mit vier Clustern.....	100
Tabelle 13:	Anzahl der Elemente pro Cluster nach Indikation - Lösung mit vier Clustern	100
Tabelle 14:	Anzahl der Elemente pro Cluster - Lösung mit sechs Clustern.....	101
Tabelle 15:	Anzahl der Elemente pro Cluster nach Indikation - Lösung mit sechs Clustern	102
Tabelle 16:	Anzahl der Elemente pro Cluster - Lösung mit zwölf Clustern.....	104
Tabelle 17:	Anzahl der Elemente pro Cluster nach Indikation - Lösung mit zwölf Clustern	105
Tabelle 18:	Anzahl der Elemente pro Cluster - Complete-Linkage-Algorithmus	109
Tabelle 19:	Anzahl der Elemente pro Cluster - Average-Linkage-Algorithmus	111
Tabelle 20:	Klassifikation der Blutungskalender gemäß WHO	131

Tabelle 21:	Klassifikation der Blutungskalender gemäß WHO nach Indikationen.....	132
Tabelle 22	Klassifikation der Blutungskalender gemäß WHO nach Clustern - kontinuierlich verabreichte Präparate	133
Tabelle 23:	Klassifikation der Blutungskalender gemäß WHO nach Clustern - zyklisch verabreichte Präparate	135
Tabelle 24:	Blutungsmusterindices für alle Hormon-Präparate	158
Tabelle 25:	Blutungsmusterindices für alle zyklisch verabreichte Präparate	160
Tabelle 26	Hypothesen und Tests für Zykluskontrollstudien.....	161

Abbildungsverzeichnis

Abbildung 1:	PubMed Datenbankabfrage	23
Abbildung 2:	Beispiel eines Blutungskalenders zur klinischen Prüfung	42
Abbildung 3:	Beispiel Blutungskalender Individuum	42
Abbildung 4:	Beispiel Blutungssegment.....	51
Abbildung 6:	Beispiel für ein Dendrogramm.....	80
Abbildung 7:	Beispiel neue Blutungsepisode	86
Abbildung 8:	Beispiel verlängerte Blutungsepisode	87
Abbildung 9:	Beispiel verletzte Dreiecksungleichung bei Gewichtung	88
Abbildung 10:	Dendrogramm der Clusteranalyse nach Ward	92
Abbildung 11:	Semipartielles Bestimmtheitsmaß über der Anzahl der Cluster	93
Abbildung 12:	Pseudo-t-Statistik über der Anzahl der Cluster	94
Abbildung 13:	Kubisches Cluster-Kriterium über der Anzahl der Cluster	95
Abbildung 14:	Pseudo-F-Statistik über der Anzahl der Cluster	96
Abbildung 15:	Dendrogramm der Clusteranalyse - Single-Linkage-Algorithmus ...	106
Abbildung 16:	Dendrogramm der Clusteranalyse - Complete-Linkage-Algorithmus	107
Abbildung 17:	Dendrogramm der Clusteranalyse - Average-Linkage-Algorithmus	109
Abbildung 18:	Mittlere Blutungsintensität über die Zeit - drei Cluster	113
Abbildung 19:	Mittlere Blutungsintensität über die Zeit - vier Cluster	115
Abbildung 20:	Mittlere Blutungsintensität über die Zeit - sechs Cluster	117

Abbildung 21:	Mittlere Blutungsintensität über die Zeit - zwölf Cluster.....	119
Abbildung 22:	Mittlere Blutungsintensität über die Zeit - zwölf Cluster - 7 HRT dominierte Cluster	120
Abbildung 23:	Mittlere Blutungsintensität über die Zeit - zwölf Cluster- 4 OC dominierte Cluster	121
Abbildung 24:	Mittlere Blutungsintensität über die Zeit - zwölf Cluster- 1 nicht dominierter Cluster	122
Abbildung 25:	Mittlere Blutungsintensität über die Zeit - 4 kontinuierlich verabreichte Präparate	126
Abbildung 26:	Mittlere Blutungsintensität über die Zeit - zyklisch verabreichte Präparate	127
Abbildung 27:	Beispiel mehrerer Blutungsepisoden im Einnahme freien Intervall.	136
Abbildung 28:	vorfristige Entzugsepisode.....	138
Abbildung 29:	verspätete Entzugsepisode bei verkürzter einnahmefreier Zeit	139
Abbildung 30:	Extrem lange Blutungsepisode	142
Abbildung 31:	Graphische Darstellung des Blutungsmusters von Einzelfällen (Blutung schwarz und Spotting gestreift dargestellt)	152
Abbildung 32:	Graphische Darstellung des Blutungsmusters von Einzelfällen (Blutung und Spotting zusammengefasst)	153
Abbildung 33:	Graphische Darstellung der mittleren Blutungsintensität für eine Kohorte von OC Einnahmerinnen.....	156
Abbildung 34:	Dendrogramm der Clusteranalyse nach Ward mit optimaler Anzahl der Cluster	166
Abbildung 35:	Mittlere Blutungsintensität über die Zeit - 4 kontinuierlich verabreichte Präparate	167
Abbildung 36:	Mittlere Blutungsintensität über die Zeit - zyklisch verabreichte Präparate	168

1 Einleitung

1.1 Historie

Die Forschung nach einem zuverlässigen hormonalen Empfängnisverhütungsmittel begann in den zwanziger Jahren des vorigen Jahrhunderts mit Tierversuchen durch Ludwig Haberlandt. Er wies nach, dass nichtträchtige Tiere unfruchtbar werden, wenn ihnen Eierstöcke trächtiger Tiere eingepflanzt werden (1). Haberlandt schloss 1928 mit dem pharmazeutischen Hersteller Gideon Richter in Budapest einen Vertrag über die Herstellung eines hormonalen Empfängnisverhütungsmittels ab. 1930 wurde das Warenzeichen „Infecundin“ registriert. Das Verhütungsmittel wurde anscheinend getestet, diese Forschung führte aber zu keinem vermarkteten Produkt.

Diese sowie spätere Versuche von Gregory Pincus und Kollegen gingen davon aus, dass ein - aus heutiger Sicht sehr hoch dosiertes - Gestagen als Monosubstanz zu verwenden sei. Die gewünschte Ovulationshemmung wurde mit 300 mg Progesteron zwar erreicht, es kam jedoch zu erheblichen Nebenwirkungen in Form von unerwünschten Zwischenblutungen. Die Idee, ein Gestagen als Monosubstanz zu verwenden, wurde daraufhin verworfen. Zu Beginn der zweiten Hälfte des vorigen Jahrhunderts veranlasste Margret Sanger, die bereits 1916 die erste „Birth Control Clinic“ in Brooklyn, New York, gegründet hatte, Gregory Pincus zur Entwicklung eines Präparats für die hormonelle Kontrazeption auf der Basis der Kombination eines Gestagens mit einem Östrogen.

Gregory Pincus testete dazu die Kombination des Gestagens Norethynodrel in der Dosierung von 9,85 mg mit dem Östrogen Mestranol in der Dosierung von 150 µg. Die klinische Prüfung wurde im Jahre 1956 in San Juan, der Hauptstadt der Karibikinsel Puerto Rico, an 100 Frauen und einer Kontrollgruppe von 125 Frauen durchgeführt (2). Unter dem Handelsnamen „Enovid“ der Firma G. D. Searle & Co. wurde das von Pincus getestete Kombinationspräparat 1957 in den USA zugelassen. Die Indikation war jedoch Behebung von Menstruationsbeschwerden. Enovid wurde

1960 zusätzlich mit der Indikation Empfängnisverhütung zugelassen und war damit das weltweit erste hormonale Kontrazeptivum.

Die Schering AG brachte 1961 mit „Anovlar“ die erste „Antibabypille“ auf den europäischen Markt. Anovlar enthielt mit 4 mg Norethisteronacetat und 50 µg Ethinylestradiol bereits eine deutlich geringere Menge an Hormonen als Enovid (3).

Seit dieser Zeit werden neue Hormonpräparate zur Empfängnisverhütung im wesentlichen mit dem Ziel der weiteren Senkung der Dosierung entwickelt, um so die Nebenwirkungsraten kontinuierlich senken zu können. Dies ist insbesondere vor dem Hintergrund der breiten öffentlichen Debatte über die Nebenwirkungen der „Pille“ von Bedeutung.

Allerdings geht dieser Trend zu immer niedrigerer dosierten Präparaten durch die Senkung der Östrogendosis in der Regel mit einer Verschlechterung des Blutungsmusters einher, obwohl die kontrazeptive Wirksamkeit gewährleistet bleibt. Die Antibabypille auf dem europäischen Markt mit dem derzeit niedrigsten Östrogengehalt ist Mirelle® mit 15 µg Ethinylestradiol als Östrogen- und 60µg Gestoden als Gestagenkomponente. Auf dem nordamerikanischen Markt werden gegenwärtig nur orale Kontrazeptiva ab 20 µg Ethinylestradiol, wie zum Beispiel Levlite®, vertrieben.

1.2 Das Blutungsmuster als Bestandteil der individuellen Lebensqualität

Das Blutungsmuster ist ein wichtiges Kriterium für die Beurteilung sowohl von hormonalen Empfängnisverhütungsmitteln als auch von Präparaten zur Hormonersatztherapie während und nach den Wechseljahren, da die Einnahme eines solchen Präparates stets auch das Blutungsmuster beeinflusst (4). Die Einnahme eines sexualhormonhaltigen Präparates kann das natürliche Blutungsmuster verändern, da das Blutungsverhalten nicht mehr durch die

endogenen zyklischen Hormonspiegelschwankungen, sondern durch die exogene Hormoneinnahme bestimmt wird.

Unregelmäßige oder nicht vorhersehbare Blutungen schränken die Lebensqualität der Frau ein, da sie im täglichen Leben hinderlich sein können. In manchen Kulturen gehen die Einschränkungen des täglichen Lebens sogar soweit, dass sich Frauen während der Menstruationsblutung dem Verbot der Essenszubereitung oder des Geschlechtsverkehrs zu unterwerfen haben.

Die regelmäßige Blutung während der Einnahme eines hormonalen Präparates zur Empfängnisverhütung gilt als Zeichen, nicht schwanger zu sein. Blutungsprobleme sind einer der wichtigsten Gründe für die Entscheidung der Frau, ein Präparat zu wechseln, eine andere Form der Verhütung zu wählen oder eine Hormonersatztherapie ganz abubrechen. Im letzteren Falle kann der Verzicht auf die Osteoporoseprophylaxe zu einer erheblichen Erhöhung des Risikos führen, Knochenbrüche insbesondere der Wirbel und der Hüftknochen zu erleiden.

1.3 Natürliches Blutungsverhalten geschlechtsreifer Frauen

Das normale Blutungsverhalten geschlechtsreifer Frauen, die keine hormonhaltigen Präparate einnehmen, besteht in der Regel aus einer Blutungsepisode, der Menstruationsblutung, die regelmäßig etwa alle vier Wochen auftritt. Dabei zeigen sich jedoch sehr deutliche Unterschiede sowohl zwischen verschiedenen Frauen als auch im Zeitverlauf für die einzelne Frau. Belsey und Pinol (5) analysierten das Blutungsmuster von 1109 gesunden, geschlechtsreifen Frauen im Alter von 15 bis 49 Jahren, die keine Form der Empfängnisverhütung durch hormonale Kontrazeptiva oder Intrauterinspiralen anwandten. Die Frauen wurden im Durchschnitt 5,75 Jahre lang beobachtet. Belsey und Pinol fanden heraus, dass für 15-jährige Frauen der Median der Zykluslänge 32 Tage beträgt. Mit zunehmendem Lebensalter fällt der Median der Zykluslänge bis auf 27 Tage für 41-jährige Frauen. Nach dem 41. Lebensjahr steigt der Median der Zykluslänge wieder auf 29 Tage für 49-jährige Frauen an.

Die empirische Verteilung der Zykluslänge über das Lebensalter ist durch zwei markante Beobachtungen geprägt. Zum einen ist die Streuung der Werte zwischen den Frauen relativ gering. Der Interquartilsabstand der Zykluslänge beträgt im Alter von 21 bis 41 Jahren nur 3-4 Tage. Dies bedeutet, dass 50 % der Frauen eine Zykluslänge zwischen 27 und 30 Tagen haben, während je 25 % der Frauen eine Zykluslänge von unter 27 bzw. von über 30 Tagen haben. Im Alter von 15 bis 20 und zwischen 42 und 49 Jahren ist die Variabilität gegenüber dem Zeitraum dazwischen erhöht. Die zweite markante Beobachtung ist, dass es in allen Altersgruppen Frauen gibt, die einen im Vergleich zur Mehrheit der Frauen deutlich verlängerten Zyklus haben. In fast allen Altersgruppen zeigen 5 % der Frauen eine Zykluslänge von über 35 Tagen.

Die weitverbreitete Meinung, dass die natürliche Menstruationsblutung der Frau von den Mondphasen beeinflusst würde, entbehrt folglich einer verlässlichen empirischen Grundlage, zumal auch die durchschnittliche Zykluslänge weiblicher Schimpansen je nach Lebensalter bei 34 bis 40 Tagen liegt (6).

1.4 Natürliches Blutungsverhalten peri- und postmenopausaler Frauen

Das natürliche Blutungsverhalten geschlechtsreifer Frauen unterscheidet sich sehr deutlich vom natürlichen Blutungsverhalten peri- und postmenopausaler Frauen. Die Menopause ist ja gerade als die letzte spontane im Leben einer Frau auftretende Monatblutung definiert. Demzufolge tritt nach der Menopause keine natürliche Blutung mehr auf. Das Datum des Eintritts der Menopause kann daher nur retrospektiv bestimmt werden. Als Perimenopause wird der Zeitraum von einem Jahr vor der Menopause bis einem Jahr nach der Menopause bezeichnet.

Der Zeitraum vor der Menopause ist durch eine Zunahme der anovulatorischen Zyklen, d.h. Zyklen ohne Eisprung, gekennzeichnet. Der Grund für die anovulatorischen Zyklen liegt in der beginnenden Ovarialinsuffizienz und dem

relativen Östrogenüberschuss. Die Blutungen werden unregelmäßiger, es können Poly- und Oligomenorrhöen auftreten. Leider gibt es keine vergleichbare empirische Untersuchung des natürlichen Blutungsverhaltens perimenopausaler Frauen, wie sie Belsey und Pinol (5) für geschlechtsreife Frauen durchgeführt haben.

1.5 Relevanz des Blutungsmusters für die arzneimittelrechtliche Zulassung neuer Präparate

Für die arzneimittelrechtliche Zulassung neuer Präparate zur Empfängnisverhütung wird vom Committee for Proprietary Medicinal Products (CPMP) der europäischen Arzneimittelagentur (EMA) zusätzlich zum Nachweis der Sicherheit (Nebenwirkungsprofil) und der Wirksamkeit (Ovulationshemmung bzw. Schwangerschaftsverhütung) die Analyse des Blutungsmusters verlangt. Während die Anforderungen an die Nachweise für Sicherheit und Wirksamkeit (vgl. Gerlinger et al. 7) genau spezifiziert sind, sind die Anforderungen der europäischen Arzneimittelagentur an die Erhebung, die Auswertung und die Darstellung des Blutungsmusters jedoch recht vage. Obwohl sich das Blutungsmuster verschiedener Ovulationshemmer zum Teil deutlich unterscheidet, wird in der EMA Leitlinie zu den Zulassungsvoraussetzungen (8) lediglich gefordert, dass das Blutungsmuster in einer klinischen Studie mit einem aktiven Vergleichspräparat erhoben werden muss.

Die diesbezüglichen Anforderungen der EMA für die Zulassung von Präparaten zur Behandlung von Klimakteriumsbeschwerden sind nicht wesentlich präziser (9): „Bleeding data should include breakthrough bleeding or spotting, incidence of amenorrheic cycles and percentage of women with withdrawal bleeding where appropriate.“ Zusätzlich wird lediglich verlangt, dass die klinische Prüfung gegen ein aktives Vergleichspräparat durchgeführt wird, und dass die Behandlungsdauer mindestens ein Jahr sein sollte.

Auch für die arzneimittelrechtliche Zulassung von Therapeutika, die das Blutungsverhalten nachhaltig beeinflussen, werden von den einzelnen nationalen

Zulassungsbehörden der Mitgliedsstaaten der Europäischen Union Untersuchungen des Blutungsmusters verlangt. Die einzelstaatliche Regelung kommt dadurch zustande, dass es für Präparate - wie zum Beispiel zur Behandlung von Endometriose - derzeit keine harmonisierte Anforderung der europäischen Arzneimittelagentur gibt.

Die Bewertungskriterien, die von den Behörden der einzelnen Mitgliedsländer der Europäischen Union zur Beurteilung der von derartigen Therapeutika induzierten Blutungsmuster angewandt werden, sind auch hier recht unterschiedlich. Dies führt zu einem erheblichen Aufwand für den pharmazeutischen Unternehmer, da im Rahmen des europäischen Zulassungsverfahrens gemäß der „mutual recognition procedure“ die Anforderungen aller 27 beteiligten Länder¹ gleichzeitig erfüllt sein müssen.

Neben den Arzneimittelbehörden in der Europäischen Union bewerten natürlich die Arzneimittelbehörden anderer Länder, wie zum Beispiel die amerikanische Food and Drug Administration (FDA) oder das japanische Ministerium für Gesundheit, Arbeit und Wohlstand (MHLW), das Blutungsmuster im Rahmen der jeweiligen nationalen Zulassungsverfahren. Doch auch von diesen Behörden sind bisher keine Leitlinien zur Analyse und Bewertung von Blutungsmustern veröffentlicht worden.

¹ Dies sind die 25 Mitglieder der Europäischen Union (Belgien, Dänemark, Deutschland, Estland, Finnland, Frankreich, Griechenland, Irland, Italien, Lettland, Litauen, Luxemburg, Malta, Die Niederlande, Österreich, Polen, Portugal, Schweden, Slowakei, Slowenien, Spanien, Tschechische Republik, Ungarn, Vereinigtes Königreich und Zypern) sowie Island und Norwegen.

2 Ziel und Methodik

2.1 Bisherige Ansätze zur Analyse und Bewertung von Blutungsmustern

Es gibt eine fast 20 Jahre alte Empfehlung der Weltgesundheitsorganisation (WHO) für die Analyse von Blutungsdaten (10). Diese Empfehlung wurde entwickelt, um sehr verschiedene empfängnisverhütende Methoden wie zum Beispiel Intrauterinspiralen, orale Kontrazeptiva, Implantate oder Dreimonatsspritzen mit dem unbeeinflussten Blutungsmuster geschlechtsreifer Frauen vergleichen zu können.

Der von der europäischen Arzneimittelagentur und anderen nationalen Arzneimittelbehörden geforderte direkte Vergleich zweier vergleichbarer oraler Kontrazeptiva bzw. Hormonersatzpräparate ist mit dieser von der WHO empfohlenen Methode jedoch nur sehr bedingt möglich. Dies liegt insbesondere daran, dass durch die Anwendung eines oralen Kontrazeptivums bzw. eines sequentiellen Hormonersatzpräparates ein stabiler, regelmäßiger Zyklus von 28 Tagen erreicht werden soll, während die Menstruationszyklen von Frauen ohne Hormoneinnahme von variabler Länge sind.

Für die Analyse der Zykluskontrolle zweier Hormonpräparate ist es zudem erforderlich, die auftretenden Blutungsepisoden in Zwischenblutungs- und Entzugsblutungsepisoden zu klassifizieren. Diese Unterscheidung fehlt jedoch in der WHO-Empfehlung, vermutlich, weil sie für Präparate zur Langzeitanwendung nicht relevant ist.

Da es an einer einheitlichen Empfehlung für die Analyse von Blutungsmustern im Rahmen des Zulassungsverfahrens für Präparate zur hormonalen Kontrazeption und zur Hormonersatztherapie mangelt, werden derzeit Blutungsmuster je nach pharmazeutischem Unternehmen unterschiedlich definiert. Die Daten verschiedener

Unternehmen sind daher sowohl für die Zulassungsbehörde als auch für die Öffentlichkeit nur sehr bedingt vergleichbar.

2.2 Ziel dieser Arbeit

Ziel dieser Arbeit ist es, eine allgemeine methodische Leitlinie zur Beurteilung von Blutungsmustern zu entwickeln. Diese methodische Leitlinie soll einerseits den Vergleich zwischen den Auswertungen des Blutungsmusters verschiedenerer pharmazeutischer Hersteller im Rahmen der arzneimittelrechtlichen Zulassung neuer Präparate ermöglichen. Sie soll andererseits jedoch auch eine Hilfe bei der individuellen Verschreibung von hormonalen Kontrazeptiva und Hormonersatzpräparaten sein, da eine subjektiv optimierte Blutungskontrolle entscheidend zur Akzeptanz eines Präparates bei der individuellen Anwenderin beiträgt.

Unter einer optimalen Blutungskontrolle für ein sequenzielles Präparat versteht man im Allgemeinen das Ausbleiben von unerwünschten Zwischenblutungen und eine vorhersagbare, regelmäßige Entzugsblutung von kurzer Dauer und mäßiger Intensität. Für kontinuierlich verabreichte Präparate versteht man unter einer optimalen Blutungskontrolle das Ausbleiben von unerwünschten Zwischenblutungen.

Durch eine optimierte Blutungskontrolle kann einerseits die Gefahr einer Eisenmangelanämie verringert und andererseits die individuelle Lebensqualität der Anwenderin erhöht werden.

2.3 Herangehensweise

Die Leitlinie zur Datenerhebung, Analyse und Bewertung von Blutungsmustern wurde im wesentlichen in zwei Schritten erstellt. Zum einen wurde eine systematische Literaturrecherche durchgeführt um die derzeit verwendeten Methoden der pharmazeutischen Unternehmer und der von pharmazeutischen Herstellern unabhängigen Wissenschaftler zu erfassen und zu bewerten.

Zum anderen wurde eine empirische Untersuchung von insgesamt 4612 auswertbaren Blutungskalendern, die in klinischen Prüfungen von mehreren hormonhaltigen Präparaten in verschiedenen Indikationen geführt worden waren, durchgeführt. Ziel der empirischen Untersuchung war es, in den Blutungskalendern vorhandene Muster durch statistische Methoden zu erkennen und die so gefundenen Muster anschließend zu systematisieren und zu beschreiben.

Hervorzuheben ist die Tatsache, dass die beiden in der vorliegenden Arbeit verwandten Erkenntnisquellen zur Formulierung der Leitlinie zur Beurteilung von Blutungsmustern an zwei diametral entgegengesetzten Ausgangspunkten beginnen. Bei der systematischen Literatursauswertung wird publiziertes Wissen möglichst vollständig erfasst, während die empirische Analyse ein Clusterverfahren zur automatischen Klassifikation ohne jegliches Vorwissen verwendet.

Die Ergebnisse der Literaturrecherche und der empirischen Untersuchung wurden zur Aufstellung der Leitlinie zur Beurteilung von Blutungsmustern zusammengefasst und an Beispielen illustriert.

2.3.1 Literaturrecherche

Durch eine Literaturrecherche wurden die Definitionen für die Erfassung, Auswertung und Beurteilung des Blutungsmusters ermittelt, die von den auf dem Gebiet Fertilitätskontrolle und Hormontherapie forschenden pharmazeutischen Herstellern verwendet werden. Selbstverständlich wurden darüber hinaus auch die Methoden der von pharmazeutischen Herstellern unabhängigen Wissenschaftler gleichwertig berücksichtigt. Diese Literaturrecherche wurde in der öffentlich zugänglichen Datenbank medizinischer Literatur PubMed (11) durchgeführt. Zusätzlich zu der öffentlichen Quelle wurde in der internen Literaturdatenbank eines pharmazeutischen Herstellers, in der die gesamte interne und externe Literatur zu den Präparaten dieses pharmazeutischen Herstellers erfasst ist, recherchiert.

In Ergänzung der elektronischen Literaturrecherche wurde durch eine Befragung medizinischer Experten die für das Thema relevante Literatur ermittelt und mit den Resultaten der Literaturrecherchen in den Datenbanken verglichen. Die in der Literatur beschriebenen Erhebungs-, Auswertungs- und Bewertungsmethoden von Blutungsmustern wurden anschließend gegenübergestellt und zusammengefasst.

2.3.2 Statistische Datenanalyse

Neben der systematischen Auswertung der Literatur wurde eine sehr große Anzahl von Blutungskalendern empirisch ausgewertet. Für die vorliegende Arbeit konnten die Datenbanken eines auf den Indikationsgebieten der hormonellen Kontrazeption und der Hormonersatztherapie weltweit führenden pharmazeutischen Herstellers verwendet werden. Die untersuchten Datenbanken umfassen insgesamt 5602 Blutungskalender, die zum Teil über mehrere Jahre geführt wurden. Von den 5602 Blutungskalendern konnten 4612 in die empirische Auswertung einbezogen werden. Die empirisch analysierten Blutungskalender sind in Abschnitt 3.3.2 näher beschrieben. Sie wurden im Rahmen von klinischen Prüfungen zur Zulassung von unterschiedlichen Präparaten zur hormonellen Fertilitätskontrolle, zur Hormonersatztherapie sowie zur Behandlung der Endometriose erhoben.

Alle klinischen Prüfungen, in denen die in der vorliegenden Arbeit analysierten Blutungskalender erhoben wurden, wurden gemäß den Prinzipien der Deklaration von Helsinki (12) sowie den gesetzlichen Vorschriften des jeweiligen Prüflandes durchgeführt. Darüber hinaus wurden die „Good Clinical Practices“ (GCP) (13) in der jeweils gültigen Fassung beachtet und das Votum der jeweils lokal zuständigen Ethikkommission vor Beginn der klinischen Prüfung eingeholt. Darüber hinaus wurden für alle klinischen Studien und für die vorliegende Untersuchung sowohl die deutschen als auch die in den jeweiligen Prüfländern gültigen Datenschutzbestimmungen eingehalten.

Insgesamt 4612 der Blutungskalender in den zur Verfügung stehenden Datenbanken wurden mindestens drei Monate lang geführt. Die Daten aus diesen

Blutungstagebüchern wurden mit multivariaten statistischen Methoden empirisch untersucht. Ziel dieser Analyse war es herauszufinden, wie sich die derzeitigen Verfahren zur Klassifikation und Beurteilung des Blutungsmusters durch empirische statistische Methoden verbessern lassen.

2.3.3 Entwicklung der methodologischen Leitlinien zur Beurteilung von Blutungsmustern

Ausgehend von den Ergebnissen der Literaturrecherche und der statistischen Datenanalyse wurde eine methodologische Leitlinie für die Erfassung, Auswertung und Beurteilung von Blutungsmustern im Rahmen des Zulassungsverfahrens für Präparate zur hormonalen Kontrazeption und zur Hormonersatztherapie entwickelt. Dazu gehören auch Vorschläge für die zukünftige Erhebung einzelner Blutungsvorgänge mit Hilfe von optimierten Blutungskalendern. Zusätzlich wurde eine methodologische Leitlinie zur Berücksichtigung des Blutungsmusters bei der individuellen Verschreibung erarbeitet.

Auf der Basis der Einträge im Blutungskalender wurden klinisch relevante Indizes für die Komponenten des Blutungsmusters wie zum Beispiel die Entzugs- oder Zwischenblutung definiert. Diese Indizes wurden sowohl tabellarisch als auch graphisch detailliert beschrieben.

Für den Vergleich von Präparaten hinsichtlich des durch sie induzierten Blutungsmusters wurden geeignete statistische Tests vorgeschlagen. Es wurde sowohl der Fall der klassischen Überlegenheitshypothese (Nullhypothese: Das zu prüfende Präparat ist schlechter oder gleich dem Referenzpräparat) als auch der Nichtunterlegenheitshypothese ((Nullhypothese: Das zu prüfende Präparat ist wesentlich schlechter als das Referenzpräparat, sogenannte „non-inferiority“) behandelt.

3 Literaturrecherche

3.1 Methodik

Für die systematische Recherche der Literatur zum Thema Darstellung und Bewertung von Blutungsmustern wurden drei verschiedene Ansätze verfolgt. Es wurde sowohl eine öffentlich zugängliche medizinische Literaturdatenbank als auch eine interne Präparate- und Literaturdatenbank eines auf dem Gebiet der Empfängnisverhütung und Hormonersatztherapie führenden pharmazeutischen Herstellers durchsucht. Zusätzlich wurden zwei ausgewiesene klinische Experten auf dem Gebiet der hormonellen Kontrazeption sowie der peri- und postmenopausalen Hormonersatztherapie für Frauen nach relevanter Literatur befragt.

3.1.1 Datenbankauswahl

Für die Literaturrecherche wurde die Datenbank PubMed ausgewählt. PubMed ist eine Datenbank medizinischer Literatur, die von der US-amerikanischen National Library of Medicine betrieben wird. Die amerikanische National Library of Medicine ist Teil des National Instituts of Health (NIH). Die Datenbank ist über das Internet (11) allgemein und kostenlos zugänglich.

Neben PubMed gibt es noch eine Vielzahl anderer Datenbanken mit medizinischer Literatur. Die PubMed Datenbank enthält jedoch die Artikel sämtlicher wesentlicher biomedizinischer Zeitschriften. Die ausschließliche Verwendung der PubMed Datenbank ist dadurch gerechtfertigt, dass weder die Abfrage der internen Literaturdatenbank eines führenden pharmazeutischen Herstellers (vgl. Abschnitt 3.1.4) noch die Befragung von medizinischen Experten (vgl. Abschnitt 3.1.3) Literaturhinweise ergab, die über die Ergebnisse der PubMed Datenbank hinausgingen.

3.1.2 PubMed Datenbankabfrage

Im vierten Quartal 2002 wurden als erster Schritt sukzessive Datenbankabfragen in PubMed durchgeführt.

Zunächst wurde mit den beiden Stichworten „cycle“ und „control“ gesucht. Die beiden Stichwörter mussten dabei nur mindestens je einmal in einem beliebigen Feld der Datenbank für einen Artikel vorhanden sein, um einen Treffer zu erzeugen. Diese Datenbankabfrage ergab 25.555 Treffer. Als nächster Schritt wurde gefordert, dass die beiden Stichworte in geordneter Reihenfolge auftreten. Der Suchbegriff „cycle control“ verringerte die Anzahl der Treffer auf 1.938.

Eine weitere Suche wurde mit den beiden Stichworten „bleeding“ und „pattern“ durchgeführt. Wie in der Suche mit den Stichworten „cycle“ und „control“ mussten die beiden Stichwörter dabei nur mindestens je einmal in einem beliebigen Feld der Datenbank für einen Artikel vorhanden sein, um einen Treffer zu erzeugen. Diese Datenbankabfrage ergab 2.486 Treffer. Als nächsten Schritt wurde gefordert, dass jeweils die beiden Stichworte in geordneter Reihenfolge auftreten. Diese Einschränkung verringerte die Anzahl der Treffer auf 223 für den Suchbegriff „bleeding pattern“.

Da nicht auszuschließen ist, dass für eine Arbeit beide Suchbegriffe zu einem Treffer führen, wurden die beiden Suchbegriffe durch den booleschen „oder“ Operator (14) verknüpft. Dadurch werden alle Arbeiten, in denen mindestens einer der beiden Suchbegriffe „cycle control“ und/oder „bleeding pattern“ vorkommt, gefunden. Die Kombination der beiden Suchbegriffe ergab insgesamt noch 2.149 Treffer. Eine Verknüpfung der beiden Suchbegriffe durch den booleschen „und“ Operator, der verlangt, dass jeder der beiden Suchbegriffe vorkommen muss, erwies sich als zu restriktiv da nur zwölf Arbeiten dieses Kriterium erfüllten.

Eine erste stichprobenartige Sichtung der durch die Suchabfrage „cycle control“ oder „bleeding pattern“ erhaltenen 2.149 Treffer ergab, dass sich unter den Treffern viele

Arbeiten befanden, bei denen das Stichwort „cycle control“ in einem deutlich anderen Kontext als dem hier untersuchten gebraucht wurde, wie zum Beispiel in (15) in dem von der Kontrolle von Zellzyklen die Rede ist.

Die Literatursuche wurde daher durch die Verwendung von einschränkenden Nebenbedingungen weiter verfeinert. Die Einschränkung auf Artikel, die Frauen betrafen („human“ und „female“), verringerte die Anzahl der Treffer bereits auf 585. Abschließend wurde die Suche noch auf Artikel, die klinische Prüfungen betrafen, die nach 1995 erschienen und die in den Sprachen Englisch, Deutsch oder Französisch abgefasst waren, eingeschränkt.

Tabelle 1: Anzahl der Treffer in PubMed

Suchkriterium	Anzahl Treffer
„cycle“ und „control“	25.555
„cycle control“	1.938
„bleeding“ und „pattern“	2.486
„bleeding pattern“	223
„cycle control“ oder „bleeding pattern“	2.149
(„cycle control“ oder „bleeding pattern“) und <i>human</i> und <i>female</i>	585
finale Suchkriterium (vgl. Abbildung 1)	87

Die zeitliche Begrenzung (nach 1995) wurde in Anlehnung an die Gegebenheiten der Datenbasis des pharmazeutischen Herstellers, d.h. der Blutungskalender, die im weiteren empirisch untersucht wurden, gewählt. Bei diesem pharmazeutischen Hersteller wurden die Blutungskalender erst ab Mitte der neunziger Jahre vollständig

in elektronische Datenbanken eingegeben. Blutungsmusterdarstellungen früherer Jahre basierten auf bereits aggregiert erfassten Daten und können demzufolge mit den heutigen Methoden nicht verglichen werden.

Die Auswahl der Sprachen ist primär durch die Kenntnisse des Autors begründet. Es ist jedoch unbestritten, dass die relevante mathematische und medizinische Fachliteratur seit längerem in englischer Sprache publiziert wird. Auch die Expertenbefragung ergab keine Arbeiten in anderen als den gesuchten Sprachen. Daher wird davon ausgegangen, dass alle wesentlichen Konzepte der Literatur erfasst trotz der sprachlichen Einschränkung gefunden wurden.

Da sich der Inhalt der PubMed Datenbank ständig um neu erschienene Arbeiten erweitert und neu erschienene Arbeiten erst mit einer gewissen Zeitverzögerung in PubMed gelistet werden, wurden die Suche auf vor 2002 erschienene Artikel eingeschränkt. Nur so konnten reproduzierbare Resultate der Datenbankabfragen gewonnen werden. Die letztlich verwendete Datenbankabfrage in PubMed ist in Abbildung 1 dargestellt.

```
(((((("cycle control"[All Fields] OR  
"bleeding pattern"[All Fields]) AND  
Clinical Trial[ptyp]) AND  
(English[Lang] OR  
French[Lang] OR  
German[Lang])) AND  
"human"[MeSH Terms]) AND  
"female"[MeSH Terms]) AND  
("1996/01/01"[PDat] : "2001/12/31"[PDat]))
```

Abbildung 1: PubMed Datenbankabfrage

Diese Datenbankabfrage ergab im vierten Quartal 2002 insgesamt 86 Literaturstellen, die anschließend gesichtet und klassifiziert wurden. Im ersten

Quartal 2003 ergab diese Datenbankabfrage 87 Literaturstellen. Die nachträglich gefundene Arbeit (59) wurde in die Auswertung noch mit einbezogen.

3.1.3 Expertenbefragung

Um die Vollständigkeit der Ergebnisse der oben beschriebenen Datenbankabfrage in PubMed beurteilen zu können, wurde zusätzlich noch eine Befragung von klinischen Experten durchgeführt. Dazu wurden zwei dem Autor persönlich bekannte Experten, die sich seit längerem bei einem pharmazeutischen Hersteller mit klinischen Prüfungen des Blutungsmusters beschäftigen, nach für sie relevanter Literatur befragt. Der erste Experte ist als Gynäkologe auf dem Gebiet der Entwicklung von neuen Medikamenten zur oralen Empfängnisverhütung, die zweite auf dem Gebiet der Hormonersatztherapie für Frauen mit Klimakteriumsbeschwerden tätig.

Die befragten Experten nannten als die für sie relevante Literatur zur Methodik der Analyse und Bewertung von Blutungsmustern einerseits eine Auswahl aus den wissenschaftlichen Publikationen von Mitarbeitern des pharmazeutischen Herstellers, für den die beiden Experten tätig sind. Andererseits wurden die wegweisenden Arbeiten der WHO Task Force (10) als wesentliche Literatur genannt.

Darüber hinaus erbrachte die Befragung der Experten über die in der Literaturrecherche in PubMed gefundenen Artikel hinaus eine Vielzahl von Tagungsbänden, Produktbroschüren diverser pharmazeutischer Hersteller und sonstiger sogenannter „grauer“ Literatur. Bei der Sichtung des Materials ergaben sich jedoch keine neuen Erkenntnisse. Dies lag daran, dass zwar durchaus Angaben zum Blutungsmuster von Präparaten angegeben wurde, die genaue Beschreibung der Methodologie aber in der Regel in den „grauen“ Literaturstellen nicht angegeben war. Zudem wurden in den Produktbroschüren in aller Regel nur die Ergebnisse bereits publizierter Studien verkürzt wiedergegeben.

Die von den medizinischen Experten genannten wissenschaftlichen Publikationen von Mitarbeitern des pharmazeutischen Herstellers waren jedoch bereits sämtlich

auch in der biomedizinischen Datenbank PubMed gefunden worden. Die Expertenbefragung führte folglich zu keiner substanziellen Erweiterung der Literaturbasis für die vorliegende Arbeit. Es kann daher davon ausgegangen werden, dass die relevanten wissenschaftlichen Publikationen durch die PubMed Datenbankabfrage vollständig erfasst worden sind.

3.1.4 Literaturrecherche in interner Datenbank des pharmazeutischen Herstellers

Zur weiteren Kontrolle der in PubMed durchgeführten Literaturrecherche wurde die interne Datenbank des pharmazeutischen Herstellers über die Literatur zu den eigenen Produkten abgefragt. Hierbei wurde nur Literatur zu Präparaten, die für die Fragestellung dieser Arbeit von Relevanz sind, einbezogen. Es wurden sämtliche Präparate zur hormonellen Verhütung bei Frauen, zur Behandlung von Endometriose und zur Hormonersatztherapie bei Frauen in die Datenbankabfrage aufgenommen.

Diese interne Datenbank enthält für die erfassten Präparate deutlich mehr Zitate als PubMed. Bei der inhaltlichen Prüfung der Treffer der Suchabfrage stellte es sich jedoch heraus, dass der Unterschied in der umfangreicheren Definition der in die Datenbank aufzunehmenden Literatur besteht. So enthält die Datenbank des pharmazeutischen Herstellers neben zitierfähiger Literatur aus peer-reviewed Zeitschriften wie sie in PubMed enthalten sind, auch Abstracts von auf medizinischen Kongressen gehaltenen Vorträgen sowie unveröffentlichte interne Forschungsberichte.

Die Fundstellen aus der Datenbank des pharmazeutischen Herstellers erbrachten keine zusätzlichen, für die Fragestellung dieser Arbeit verwertbaren Artikel, da die zitierfähige Literatur aus den frauenheilkundlichen peer-reviewed Zeitschriften bereits vollständig in PubMed gefunden wurde. Die darüber hinaus gefundenen Abstracts der auf medizinischen Kongressen gehaltenen Vorträge konnten zu dieser Untersuchung jedoch nichts beitragen, da schon auf Grund ihrer Kürze die

Einzelheiten der Auswertung des Blutungsmusters nicht dargestellt wurden. Die internen Forschungsberichte waren entweder für die Fragestellung dieser Arbeit irrelevant oder sie betrafen klinische Prüfungen, deren die Auswertung des Blutungsmusters betreffende Methodologie bereits in peer-reviewed Zeitschriften publiziert worden war.

3.2 Ergebnisse der Literaturrecherche

Die Ergebnisse der Literaturrecherche sind in Tabelle 2 zusammengefasst. Einzelne in der Literaturrecherche gefundene Arbeiten konnten jedoch nicht für die Klassifikation der Methoden zur Analyse und Bewertung von Blutungsmustern verwendet werden. Dabei handelt es sich meistens um Arbeiten, die zwar die gesuchten Stichwörter enthalten, aber zur Beantwortung der Fragestellung dieser Arbeit nichts beitragen konnten, da die Stichwörter in einem anderen Kontext verwendet wurden. So untersuchten zum Beispiel Zhou und Xiao (25) Blutungen nach Einsetzen eines Intrauterinpessars zur postkoitalen Verhütung. Derartige Treffer der Literaturrecherche sind in Tabelle 2 mit

- NR entweder nicht relevant für die Fragestellung oder Blutungsmustermethodologie nicht ausreichend in der Arbeit beschrieben.

kenntlich gemacht worden. Darüber hinaus gab es zwei Arbeiten, von denen dem Autor trotz aller Bemühungen zur Literaturbeschaffung nur der Abstract zur Verfügung stand. Diese Arbeiten konnten folglich nicht in die genaue Untersuchung mit einbezogen werden, sie sind in Tabelle 2 mit

- NV Blutungsmustermethodologie nicht verfügbar

bezeichnet worden.

Für die Klassifikation der Literaturstellen nach Indikation des untersuchten Präparates und nach der verwendeten Methode zur Analyse und Bewertung des

Blutungsmusters der Literaturstellen wird das folgende Klassifikationsschema verwendet:

Die Indikation der in der betreffenden Arbeit untersuchten Präparate ist in die Kategorien

- HRT Hormonersatztherapie
- IC injizierbares Kontrazeptivum oder Implantatkontrazeptivum mit Langzeitwirkung
- IUD Intrauterinpessar
- OC orales Kontrazeptivum
- TC transdermales Kontrazeptivum
- VR Vaginalring zur Kontrazeption

eingeteilt. Die verschiedenen in der Literatur beschriebenen Methoden, das vaginale Blutungsmuster zu erheben und darzustellen, werden vom Autor zu den Kategorien

- EVRA Methodik der EVRA² Entwicklung
- Organon Methodik der Entwicklungsprojekte der Firma NV Organon
- Kalendermethode
- SAG alte Methodik der Schering AG
- WHO WHO Methodik

² EVRA ist ein transdermales Kontrazeptivum.

- Modifizierte WHO Methodik

zusammengefasst. Diese Kategorien werden im folgenden noch näher erläutert werden.

Tabelle 2: Ergebnisse der Literaturrecherche

Autoren	Klassifikation
van de Weijer PH, Scholten PC, van der Mooren MJ, Barentsen R, Kenemans P (16)	HRT, WHO
Graser T, Koytchev R, Muller A, Oettel M (17)	HRT, SAG
Wren BG, McFarland K, Edwards L, O'Shea P, Sufi S, Gross B, Eden JA (18)	HRT, NR
Massai MR, Diaz S, Quinteros E, Reyes MV, Herreros C, Zepeda A, Croxatto HB, Moo-Young AJ (19)	OC, WHO
Petta CA, Hays M, Brache V, Massai R, Hua Y, Alvarez-Sanchez F, Salvatierra A, d'Arcangues C, Cook LA, Bahamondes L (20)	NR
al-Azzawi F, Wahab M, Thompson J, Pornel B, Hirvonen E, Ylikorkala O, van der Mooren MJ, Dillon J, Magaril C (21)	HRT, WHO
Bruhat M, Rudolf K, Vaheri R, Kainulainen P, Timonen U, Viitanen A (22)	HRT, WHO
Smallwood GH, Meador ML, Lenihan JP, Shangold GA, Fisher AC, Creasy GW (23)	TC, EVRA
Endrikat J, Cronin M, Gerlinger C, Ruebig A, Schmidt W, Dusterberg B (24)	OC, SAG

Autoren	Klassifikation
Zhou L, Xiao B (25)	NR
Endrikat J, Cronin M, Gerlinger C, Ruebig A, Schmidt W, Dusterberg B (26)	OC, SAG
Fruzzetti F, Genazzani AR, Ricci C, De Negri F, Bersi C, Carmassi F (27)	OC, Kalendermethode
Hampton RM, Short M, Bieber E, Bouchard C, Ayotte N, Shangold G, Fisher AC, Creasy GW (28)	OC, Kalendermethode
Endrikat J, Hite R, Bannemerschult R, Gerlinger C, Schmidt W (29)	OC, SAG
Jaisamrarn U, Reinprayoon D, Virutamasen P (30)	OC, NA
Ryan N, Rosner A (31)	HRT, NR
Vree ML, Schmidt J (32)	OC, Organon
Audet MC, Moreau M, Koltun WD, Waldbaum AS, Shangold G, Fisher AC, Creasy GW (33)	TC, EVRA
Petta CA, Hays M, Brache V, Massai R, Hua Y, Alvarez-Sanchez F, Croxatto H, d'Arcangues C, Cook LA, Bahamondes L (34)	NR
Garceau RJ, Wajszczuk CJ, Kaunitz AM (35)	IC, modifizierte WHO

Autoren	Klassifikation
Bassol S, Alvarado A, Celis C, Cravioto MC, Peralta O, Montano R, Novelli J, Albornoz H, Kesseru E, Soares A, Petracco A, Isaia B, Mendes J, Bahamondes L, de Melo NR, Reyes-Marquez R, Albrecht G (36)	OC, SAG
Boulay JL, Perruchoud AP, Reuter J, Bolliger C, Herrmann R, Rochlitz C (37)	NR
Zichella L, Sbrignadello C, Tomassini A, Di Lieto A, Montoneri C, Zarbo G, Mancone M, Pietro Battista P, Bertoli G, Perrone G (38)	OC, Kalendermethode
Foidart JM, Wuttke W, Bouw GM, Gerlinger C, Heithecker R (39)	OC, SAG
Bassol S, Cravioto MC, Durand M, Bailon R, Carranza S, Fugarolas J, Gaona R, Parada LM, Celis C, Santoyo S, Garza-Flores J, Vazquez L, Lopez C, Gurucharri C, Novelli J, Carneiro de Oliveira H, Mendez J, de Andrade ME, de Mello NR, de Melo K, Chada E, Yassle ME, Castaneda A, Gomez P, Arboleda C, Trujillo L, Bucheli R, Hidalgo I, Olavide R, Parejarios J, Succar J, Reyes-Marquez R, Albrecht G (40)	IC, WHO
Kaunitz AM (41)	OC, Kalendermethode
Ferguson H, Vree ML, Wilpshaar J, Eskes TK (42)	OC, Organon

Autoren	Klassifikation
Huber J, Foidart JM, Wuttke W, Merki-Feld GS, The HS, Gerlinger C, Schellschmidt I, Heithecker R (43)	OC, SAG
Bruni V, Croxatto H, De La Cruz J, Dhont M, Durlot F, Fernandes MT, Andrade RP, Weisberg E, Rhoa M (44)	OC, keine Details
Brouard R, Bossmar T, Fournie-Lloret D, Chassard D, Akerlund M (45)	NR
Letterie GS (46)	NR
Limpaphayom KK, Bunyavejchevin S (47)	NV
Heikkinen JE, Vaheiri RT, Ahomaki SM, Kainulainen PM, Viitanen AT, Timonen UM (48)	HRT, WHO
Endrikat J, Dusterberg B, Ruebig A, Gerlinger C, Strowitzki T (49)	OC, SAG
Rosenberg MJ, Meyers A, Roy V (50)	OC, sonstige Methode ³
Burch D, Bieshuevel E, Smith S, Fox H (51)	HRT, Kalendermethode
Saure A, Planellas J, Poulsen HK, Jaszczak P (52)	HRT
Belaisch J, Nappi C, Affinito P, De Gezelle H, Botelho M, Oliveira HM, Martinez-Oliveira J, Gonzaga F (53)	NA

³ verwendet einen Blutungsindex von 0 bis 100.

Autoren	Klassifikation
Kaunitz AM, Garceau RJ, Cromie MA (54)	OC, WHO
Reisman H, Martin D, Gast MJ (55)	OC, WHO und Kalendermethode
Archer DF, Maheux R, DelConte A, O'Brien FB (56)	OC, Kalendermethode
Zheng SR, Zheng HM, Qian SZ, Sang GW, Kaper RF (57)	IC, WHO
Bjarnason K, Cerin A, Lindgren R, Weber T (58)	HRT, modifizierte WHO
Weisberg E, Fraser IS, Mishell DR Jr, Lacarra M, Darney P, Jackanicz TM. (59)	VC, Kalendermethode
Ronnerdag M, Odland V (60)	IUD, WHO
Zheng SR, Zheng HM, Qian SZ, Sang GW, Kaper RF (61)	IC, WHO
Chavez A, DelConte A (62)	OC, modifizierte WHO
Yang TS, Tsan SH, Chen CR, Chang SP, Yuan CC (63)	NV
DelConte A, Loffer F, Grubb GS (64)	OC, modified WHO
Affandi B, Korver T, Geurts TB, Coelingh Bennink HJ (65)	IC, WHO
Sulak P, Lippman J, Siu C, Massaro J, Godwin A (66)	OC, Kalendermethode
da Silva MO, Costa MM (67)	IUD, NR

Autoren	Klassifikation
van de Weijer PH, Barentsen R, de Vries M, Kenemans P (68)	HRT, WHO
Swahn ML, Bygdeman M, Chen JK, Gemzell-Danielsson K, Song S, Yang QY, Yang PJ, Qian ML, Chang WF (69)	OC, modifizierte WHO
Serfaty D, Vree ML (70)	OC, Organon
Collaborative Study Group on the Desogestrel-containing Progestogen-only Pill (71)	OC, WHO, modifizierte WHO
Rowan JP (72)	OC, Kalendermethode
Cano A, Tarin JJ, Duenas JL (73)	HRT, sonstige Methode ⁴
Pal L, Lapensee L, Toth TL, Isaacson KB (74)	NR
Stanworth SJ, Bolton MJ, Hay CR, Shiach CR (75)	NR
Halbe HW, de Melo NR, Bahamondes L, Petracco A, Lemgruber M, de Andrade RP, da Cunha DC, Guazelli CA, Baracat EC (76)	OC, Organon
Vercellini P, Perino A, Consonni R, Oldani S, Parazzini F, Crosignani PG (77)	NR, sonstige Methode ⁵

⁴ verwendet relative Häufigkeit der Frauen mit Spotting pro Monat, Anzahl Tage mit Spotting pro Monat.

⁵ Bildlicher Score.

Autoren	Klassifikation
Martinez GH, Castaneda A, Correa JE (78)	IC, WHO
Chen JH, Wu SC, Shao WQ, Zou MH, Hu J, Cong L, Miao L, Wang C, Dong J, Gao J, Xiao BL (79)	IUD, VR, WHO
Carranza-Lira S, Martinez-Chequer JC, Santa Rita MT, Ortiz de la Pena A, Perez Y, Fernandez RL (80)	HRT, NR
Croxatto HB, Kovacs L, Massai R, Resch BA, Fuentealba B, Salvatierra AM, Croxatto HD, Zalanyi S, Viski S, Krenacs L (81)	OC, modifizierte WHO
Comparato MR, Yabur JA, Bajares M (82)	OC, modifizierte WHO
Nand SL, Webster MA, Baber R, O'Connor V (83)	HRT, sonstige Methode ⁶
Schipper I, Hop WC, Fauser BC (84)	NR
Graser T, Rossner P, Schubert K, Muller A, Bonisch U, Oettel M (85)	HRT, modifizierte WHO
Lubbert H, Nauert C (86)	HRT, sonstige Methode ⁶
Bannemerschult R, Hanker JP, Wunsch C, Fox P, Albring M, Brill K (87)	OC, Kalendermethode
Bergqvist A, Jacobson J, Harris S (88)	NR
Endrikat J, Muller U, Dusterberg B (89)	OC, SAG

⁶ Anzahl der Tage mit Blutung gezählt.

Autoren	Klassifikation
Csemiczky G, Dieben T, Coeling Bennink HJ, Landgren BM (90)	OC, Kalendermethode
AinMelk Y (91)	HRT, Kalendermethode
Kalogirou D, Antoniou G, Karakitsos P, Kalogirou O, Antoniou D, Giannikos L (92)	HRT, NR
Boonkasemsanti W, Reinprayoon D, Pruksananonda K, Niruttisard S, Triratanachat S, Leepipatpaiboon S, Wannakrairot P (93)	IC, WHO
Pansini F, De Paoli D, Albertazzi P, Bonaccorsi G, Campobasso C, Zanotti L, Pisati R, Giulini NA (94)	HRT, Kalendermethode
Saure A, Hirvonen E, Milsom I, Christensen A, Damber MG (95)	HRT, Kalendermethode
Boerrigter PJ, van de Weijer PH, Baak JP, Fox H, Haspels AA, Kenemans P (96)	HRT, Kalendermethode
Rice C, Killick S, Hickling D, Coelingh Bennink H (97)	OC, WHO
Mayagoitia SB, Hernandez-Morales C, Macias AM, Vega RR (98)	OC, NR
Messinis IE, Lolis D, Zikopoulos K, Milingos S, Kollios G, Seferiadis K, Templeton AA (99)	NR
Dusterberg B, Ellman H, Muller U, Rowe E, Muhe B (100)	OC, SAG

Autoren	Klassifikation
Rosenberg MJ, Waugh MS, Stevens CM (101)	OC, modifizierte WHO
Kunz J, Jaeger A (102)	OC, NR

3.3 Auswertung der Literatur

3.3.1 Blutungsstärke

Die in der Literatur beschriebenen Methoden zur Darstellung und Auswertung des Blutungsmusters lassen sich zunächst nach der unterschiedlichen Erhebung der Blutungsstärke unterscheiden.

3.3.1.1 WHO

Die Weltgesundheitsorganisation WHO (10) empfiehlt die Blutungsstärke in den Kategorien

- „spotting“
- „bleeding“

sowie

- „no bleeding or spotting“

zu erheben. Hierbei ist „spotting“ definiert als ein blutiger vaginaler Ausfluss, der jedoch keine Verwendung von Schutzmaßnahmen, wie Monatsbinden oder Tampons, erforderlich macht. Die Kategorie „bleeding“ ist hingegen definiert als ein blutiger vaginaler Ausfluss, der eine Verwendung von Schutzmaßnahmen

erforderlich macht. Tritt kein vaginaler Ausfluss auf, so wird dieses Ereignis als „no bleeding or spotting“ bezeichnet.

Bei der Definition der Blutungsstärke gemäß den WHO-Empfehlungen kommt es nicht darauf an, ob ein Monatsschutz, d.h. eine Binde oder ein Tampon, auch tatsächlich verwendet wurde. Maßgeblich für die Kategorisierung der Blutungsstärke ist lediglich, ob ein solcher Schutz – nach Ansicht der betroffenen Frau – erforderlich gewesen wäre.

3.3.1.2 Schering AG

Die Pharmafirma Schering AG verwendete noch bis Ende 1998 je nach klinischem Projekt verschiedene Begriffe für die Blutungsstärke, obwohl sich die Definitionen der Begriffe nur unwesentlich unterschieden. Das Konzept war, die Blutungsstärke nicht wie bei der Definition der Weltgesundheitsorganisation über die Verwendung von sanitären Produkten, sondern in Relation zu einer normalen, d.h. von einem Hormonpräparat unbeeinflussten, Regelblutung der Frau zu definieren. So wurden beispielsweise die Kategorien

- „keine Blutung“
- „weniger als normale Menses“
- „normale Menses“
- „mehr als normale Menses“

verwendet. Häufig wird die Blutungsstärke auch in die Kategorien

- „keine Blutung“
- „spotting“
- „leichte Blutung“

- „normale Blutung“
- „schwere Blutung“

erhoben. Für die anschließende biometrische Auswertung werden die Kategorien „leichte Blutung“, „normale Blutung“ und „schwere Blutung“ jedoch zur Kategorie „Blutung“ zusammengefasst. Der Grund für diese Zusammenfassung liegt darin, dass die Abgrenzung zwischen „spotting“ und den Blutungskategorien unabhängig davon ist, was die einzelne Frau als eine „normale Blutung“ ansieht. Dies hatte zur Folge, dass im zweiten Falle letztendlich die von der Weltgesundheitsorganisation empfohlene Definition der Blutungsstärke effektiv angewandt werden konnte.

3.3.1.3 NV Organon

In der Literatur finden sich noch weitere Definitionen der Blutungsstärke. So definiert die niederländische Pharmafirma NV Organon in den von ihr durchgeführten klinischen Prüfungen von Kontrazeptiva die Kategorien „spotting“ und „bleeding“ in Anlehnung an die Definitionen der Weltgesundheitsorganisation. Dabei wird jedoch die Grenze zwischen den beiden Kategorien „spotting“ und „bleeding“ anders als bei der WHO definiert.

Gemäß der von NV Organon verwandten Definition gilt ein blutiger vaginaler Ausfluss, der die Verwendung von maximal einer Schutzmaßnahme, wie einer Monatsbinde oder eines Tampons, erforderlich macht, noch als „spotting“. Die Kategorie „bleeding“ ist demzufolge als ein blutiger vaginaler Ausfluss, der die Verwendung von mindestens zwei Monatsbinden bzw. Tampons erforderlich macht, definiert.

Im Vergleich zur Definition der Blutungsstärke nach WHO führt die von NV Organon verwandte Definition zu weniger „bleeding“ und dafür mehr „spotting“.

3.3.1.4 Ortho-McNeil Pharmaceuticals Inc.

Die amerikanische Pharmafirma Ortho-McNeil Pharmaceuticals Inc. verwendete für die klinische Entwicklung ihres transdermalen kontrazeptiven Pflasters EVRA dieselbe Modifikation der WHO Definition der Blutungsstärke wie NV Organon: Ein Tag, an dem die Anwenderin genau ein Produkt für die Monatshygiene benötigte, wird bei Ortho-McNeil Pharmaceuticals noch als Tag mit „spotting“ gewertet, während die WHO den Tag als Tag mit „bleeding“ werten würde. Die übrigen Definitionen der Blutungsstärke von Ortho-McNeil Pharmaceuticals entsprechen der WHO-Empfehlung.

3.3.1.5 Sonstige Definitionen

Neben den beiden oben beschriebenen weit verbreiteten Ansätzen, die Blutungsstärke entweder über die Art und Menge der verwendeten Binden und Tampons oder relativ zu einer normalen Menses zu beschreiben, gibt es auch Ansätze (Vercellini et al. 77), die Blutungsstärke genauer über einen bildlichen Score (103) zu erfassen oder das Volumen des vaginalen Ausflusses in Millilitern zu erfassen (104).

Diese, nur scheinbar genaueren, Methoden zur Erhebung der Blutungsstärke haben sich jedoch in der Praxis nicht durchsetzen können. Sie werden nur vereinzelt eingesetzt. Der Grund hierfür liegt sicherlich zum einen darin, dass die tägliche Erhebung eines bildlichen Scores oder die Messung des Blutverlustes in Milliliter wesentlich aufwendiger ist, als die Erhebung einer Variablen auf einer dreistufigen Likert-Skala. Zum anderen dürfte der zusätzliche Erkenntnisgewinn klinisch nicht relevant sein, da das subjektive Empfinden der Anwenderin sicherlich eine wesentlichere Rolle bei der Bewertung ihrer individuellen Blutungsstärke spielt als die geblutete Menstruationsblutmenge in Millilitern. Die Bewertung ihrer individuellen Blutungsstärke ist natürlich auch abhängig von der Physis der Frau, was für die eine Frau wenig ist, kann für eine andere viel sein.

Die Erhebung der Blutungsstärke relativ zur normalen Menstruationsblutung ist bei klinischen Prüfungen von Präparaten zur Hormonersatztherapie postmenopausaler Frauen jedoch kritisch zu sehen. Dies liegt daran, dass die behandelten postmenopausalen Frauen zum Teil schon seit Jahren keine vaginale Blutung mehr hatten. Es kann also nicht unbedingt davon ausgegangen werden, dass die in die klinische Prüfung aufgenommen postmenopausalen Frauen noch über einen validen Referenzpunkt verfügen, um die Frage nach ihrer individuellen Blutungsstärke beantworten zu können. In diesem Fall sind die Definitionen der Blutungsstärke, die sich an der Art des benötigten Monatsschutzes orientieren, eindeutig zu bevorzugen.

3.3.2 Blutungshäufigkeit

Die Blutungshäufigkeit wird üblicherweise mit einem Blutungskalender erhoben. Unter einem Blutungskalender versteht man eine Tabelle in Form eines Kalenders, in den die Blutungsereignisse für jeden Tag eingetragen werden. Das menschliche Erinnerungsvermögen für derartige Ereignisse ist in der Regel schlecht. Deshalb hat es sich in der Praxis bewährt, die Einträge in die Blutungskalender täglich vorzunehmen, anstatt beispielsweise am Ende eines jeden Zyklus die Blutungsereignisse retrospektiv einzutragen.

Für klinische Prüfungen von Arzneimitteln zur hormonellen Kontrazeption oder zur Hormonersatztherapie werden heutzutage in der Regel Blutungskalender eingesetzt, welche für jeden Monat eine Seite verwenden (vgl. Abbildung 2). Früher wurden auch Blutungskalender verwendet, die an Stelle eines Kalendermonats einen vierwöchigen Zyklus, bestehend aus 21 Tagen Pilleneinnahme gefolgt von 7 einnahmefreien Tagen, abgebildet wurde.

Es stellte sich jedoch heraus, dass Blutungskalender, die einen vollen Monat abbilden, weit weniger anfällig für fehlerhafte Einträge der Pillennutzerin sind, als Blutungskalender die nur einen 28-tägigen Zyklus abbilden. So sind insbesondere Verwechslungen des Datums beim Übergang von einem Zyklus zum nächsten bei einem Blutungskalender mit einer monatlichen Darstellung ausgeschlossen.

Month/year		Study-no. 305220		Volunteer-no.		Volunteer initials _____ 1st name / surname																										
Please tick		S	M	T	W	T	F	S	S	M	T	M	T	F	S	S	M	T	W	T	F	S	S	M	T	M	T	F	S	S	M	T
Start of a new package		☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
No. of tablets taken:		0	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
		1	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
		2	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Bleedings:		none	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
		spotting	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
		light	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
		normal	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
		heavy	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Check The diary was checked with the volunteer		Date				Signature of investigator												☐ checked by CRA														

Abbildung 2: Beispiel eines Blutungskalenders zur klinischen Prüfung (105)

Für die jeweilige Anwenderin ist ein über mehrere Monate reichender Blutungskalender, wie er beispielsweise von der Kaisers Betriebskrankenkasse im Internet zur Verfügung gestellt wird (vgl. Abbildung 3), übersichtlicher.

Januar	Februar	März	April	Mai	Juni	Juli	August	September	Oktober	November	Dezember
01	01	01	01	01	01	01	01	01	01	01	01
02	02	02	02	02	02	02	02	02	02	02	02
03	03	03	03	03	03	03	03	03	03	03	03
04	04	04	04	04	04	04	04	04	04	04	04
05	05	05	05	05	05	05	05	05	05	05	05
06	06	06	06	06	06	06	06	06	06	06	06
07	07	07	07	07	07	07	07	07	07	07	07
08	08	08	08	08	08	08	08	08	08	08	08
09	09	09	09	09	09	09	09	09	09	09	09
10	10	10	10	10	10	10	10	10	10	10	10
11	11	11	11	11	11	11	11	11	11	11	11
12	12	12	12	12	12	12	12	12	12	12	12
13	13	13	13	13	13	13	13	13	13	13	13
14	14	14	14	14	14	14	14	14	14	14	14
15	15	15	15	15	15	15	15	15	15	15	15
16	16	16	16	16	16	16	16	16	16	16	16
17	17	17	17	17	17	17	17	17	17	17	17
18	18	18	18	18	18	18	18	18	18	18	18
19	19	19	19	19	19	19	19	19	19	19	19
20	20	20	20	20	20	20	20	20	20	20	20
21	21	21	21	21	21	21	21	21	21	21	21
22	22	22	22	22	22	22	22	22	22	22	22
23	23	23	23	23	23	23	23	23	23	23	23
24	24	24	24	24	24	24	24	24	24	24	24
25	25	25	25	25	25	25	25	25	25	25	25
26	26	26	26	26	26	26	26	26	26	26	26
27	27	27	27	27	27	27	27	27	27	27	27
28	28	28	28	28	28	28	28	28	28	28	28
29	29	29	29	29	29	29	29	29	29	29	29
30	30	30	30	30	30	30	30	30	30	30	30
31	31	31	31	31	31	31	31	31	31	31	31

Regelblutung
 schwach ☐ ☐ ☐
 mittel ☐ ☐ ☐
 stark ☐ ☐ ☐

Abbildung 3: Beispiel Blutungskalender Individuum (106)

In klinischen Prüfungen wird jedoch ein Blutungskalender über einen Monat pro Seite einem Blutungskalender mit mehreren Monaten pro Seite vorgezogen. In diesen Kalender werden nicht nur die Intensität der Blutungen, sondern auch die tägliche Pilleneinnahme eingetragen. Ferner wird es dadurch möglich, einzelne Monate des Blutungskalenders schon vor dem Ende der Studie einzusammeln und in die Studiendatenbank einzugeben. Durch dieses Vorgehen lässt sich die nötige Zeit zwischen dem Ende der Behandlung der letzten Studienteilnehmerin und dem Vorliegen des Studienberichts deutlich verringern. Dies kann im Endeffekt zu einer schnelleren klinischen Entwicklung und dadurch zu einer früheren Zulassung des geprüften Präparates durch die Arzneimittelbehörden führen.

3.3.3 Blutungsmuster

Das meist zyklische Blutungsverhalten geschlechtsreifer Frauen wird in der Literatur unter dem übergreifenden Begriff „Blutungsmuster“ zusammengefasst. Der Begriff wurde aber auch auf das Blutungsverhalten von Frauen nach der Menopause oder von Frauen unter der Einnahme von sexualhormonhaltigen Kontrazeptiva ausgedehnt.

Bei der bisherigen Entwicklung von sexualhormonhaltigen Präparaten, die das Blutungsverhalten der Anwenderinnen nachhaltig beeinflussen, wurde bisher stets darauf Wert gelegt, dass das Blutungsmuster unter der Einnahme möglichst dem natürlichen Blutungsmuster geschlechtsreifer Frauen weitgehend entspricht. Erst im Jahre 2003 wurde mit dem Kontrazeptivum Seasonale® der amerikanischen Pharmafirma Barr Laboratories Inc. die weltweit erste Zulassung eines oralen Kontrazeptivums, das einen auf 84 Tage verlängerten Zyklus erreichen soll, erteilt.

Die in der Literatur beschriebenen Methoden zur Erfassung und Beschreibung des Blutungsmusters lassen sich in zwei Klassen, in die Episoden-basierten und in die Zyklus-basierten Methoden, einteilen. Bei den Episoden-basierten Methoden werden zunächst zusammenhängende Tage mit Blutungen zu Blutungsepisoden

zusammengefasst. Die Blutungsepisoden werden dann für die Analyse des Blutungsmusters als Einheiten betrachtet.

Demgegenüber stehen die Zyklus-basierten Methoden. Hier werden Tage mit Blutungen in bestimmten Zeiträumen des Zyklus, wie zum Beispiel den sieben Tagen ohne Hormoneinnahme beim klassischen Pillenregime (bestehend aus 21 Tagen mit Pilleneinnahme gefolgt von 7 Tagen ohne Pilleneinnahme) zusammengefasst. In diesem Fall spielt es keine Rolle, ob die gezählten Blutungen an aufeinanderfolgenden Tagen stattfanden, oder ob sie singuläre Ereignisse waren.

3.3.3.1 Zyklus-basierte Blutungsmuster-Methoden

Die Zyklus-basierten Blutungsmuster-Methoden wurden überwiegend bei der Beschreibung des Blutungsverhaltens unter der Einnahme von Präparaten zur hormonellen Kontrazeption verwendet. Der Einnahmezyklus, der für die allermeisten hormonellen Kontrazeptiva gilt, besteht aus 21 Tagen mit Pilleneinnahme, gefolgt von sieben Tagen ohne Pilleneinnahme. Insgesamt ergibt sich damit ein Einnahmezyklus von 28 Tagen, d.h. von genau vier Wochen. Auch für sehr niedrig dosierte neuere Präparate wie zum Beispiel Meliane (26), für die eine verlängerte Einnahmeperiode von 23 Tagen gewählt wurde, wurde die vierwöchige Dauer des gesamten Pillenzyklus beibehalten. Dies geschah, indem ein Einnahmeschema von 23 Tagen mit Wirkstoffeinnahme gefolgt von 5 Tagen ohne Wirkstoffeinnahme gewählt wurde. Um die Compliance bei einem derartigen ungewöhnlichen Einnahmemuster zu erhöhen, wurden der Packung vier wirkstofffreie, d.h. sogenannte Placebopillen, hinzugefügt. Damit ist eine tägliche Einnahme des Präparates ohne Unterbrechungen möglich. In den USA ist dieses Einnahmeschema auch für klassische Pillen mit 21-tägiger Wirkstoffgabe üblich (107), d.h. 21 wirkstoffhaltige Pillen gefolgt von 7 Placebopillen.

Die Idee der Zyklus-basierten Blutungsmuster-Methoden besteht darin, den vierwöchigen Pilleneinnahmezyklus in feste vordefinierte Abschnitte zu unterteilen. Die Tage mit Blutungen bzw. Spotting, die innerhalb dieser Abschnitte auftreten,

werden dann jeweils getrennt betrachtet. Dieses Vorgehen ermöglicht eine einfache Unterscheidung der erwünschten Hormonentzugsblutungen von den unerwünschten Zwischenblutungen.

Exemplarisch sei hier die Vorgehensweise der Schering AG bei der klinischen Entwicklung des dropirenonhaltigen monophasischen Kontrazeptivums Yasmin® in der Mitte der Neunziger Jahre des letzten Jahrhunderts detailliert dargestellt (39, 43). Alle beobachteten Pillenzyklen mit Ausnahme des ersten Pillenzyklus wurden in drei Abschnitte eingeteilt:

- Abschnitt 1: Tag 1 bis 3 des Pillenzyklus
- Abschnitt 2: Tag 4 bis 21 des Pillenzyklus
- Abschnitt 3: Tag 22 bis 28 des Pillenzyklus

Der erste Pillenzyklus wurde in leicht geänderte Abschnitte eingeteilt. Dieses Vorgehen war notwendig, da bei Frauen, die mit der hormonellen Kontrazeption beginnen, im ersten Zyklus der Übergang vom natürlichen Menstruationszyklus auf den durch die Hormoneinnahme gesteuerten Pillenzyklus stattfindet. Laut Patienteninformation war der erste Tag der Pilleneinnahme der erste Tag der Menstruationsblutung. Idealtypischerweise erfolgt die Hormonentzugsblutung unmittelbar nach dem Beenden des 21-tägigen Einnahmeintervalls zu Beginn der einwöchigen pillenfreien Zeit. Dies hatte jedoch zur Folge, dass bei der Umstellung der normalerweise vierwöchige Zeitraum zwischen zwei aufeinanderfolgenden Monatsblutungen auf etwa drei Wochen verkürzt wurde. Die Einteilung der Abschnitte für den ersten Zyklus war wie folgt:

- Abschnitt 1: Tag 1 bis 7 des Pillenzyklus
- Abschnitt 2: Tag 8 bis 21 des Pillenzyklus
- Abschnitt 3: Tag 22 bis 28 des Pillenzyklus

Im Vergleich zur Einteilung der Abschnitte für die nachfolgenden Pillenzyklen wurde der erste Abschnitt von drei auf sieben Tage verlängert und der zweite Abschnitt wurde entsprechend von 18 auf 14 Tage verkürzt. Der dritte Abschnitt blieb mit sieben Tagen unverändert, er entspricht also in allen Zyklen der pilleneinnahmefreien Periode.

Für die Analyse des Blutungsmusters wurden zunächst die Blutungstage, die in dem jeweiligen Abschnitt auftraten, summiert und direkt in den Berichtsbogen, engl. case record form (CRF), eingetragen. Diese Aggregation der Blutungstage erfolgte getrennt nach den erhobenen Blutungsstärkekategorien. Ein nicht zu unterschätzender Vorteil der Zyklus-basierten Methode ist, dass nur bereits aggregierte Daten, nämlich die Anzahl der Blutungstage pro Blutungsstärkekategorie und pro Zyklusabschnitt, in die Datenbank eingegeben werden müssen. Im Gegensatz dazu erfordern Episoden-basierten Blutungsmuster-Methoden eine vollständige Eingabe der Blutungskalender in die Datenbank.

Im Rahmen der Beschreibung des Blutungsmusters werden bei den Zyklus-basierten Methoden die Blutungstage in den einzelnen Abschnitten des Zyklus entweder als Entzugsblutung oder als Zwischenblutung definiert. Im beschriebenen Falle der Yasmin®-Entwicklung wurden alle Blutungstage im Abschnitt 2 als Zwischenblutungstage definiert. Dagegen wurden Blutungstage im Abschnitt 3 des Zyklus und Blutungstage im Abschnitt 1 des Folgezyklus als Entzugsblutung gewertet. Der erste Abschnitt des ersten Zyklus wurde getrennt behandelt. Da für Frauen, die mit der hormonellen Kontrazeption beginnen, der erste Einnahmetag der erste Blutungstag einer Menstruationsblutung sein sollte, wurden Blutungen im ersten Abschnitt des ersten Zyklus nicht als Zwischenblutung gewertet.

Die quantitative Beschreibung der Anzahl der Blutungstage erfolgte getrennt nach Zwischenblutungen und Entzugsblutungen mit Hilfe der beschreibenden Statistik. Neben der Anzahl der vorhandenen Blutungsinformationen wurden das arithmetische Mittel, die Standardabweichung, das Minimum, das Maximum, der Median, sowie das 25%- und das 75%-Perzentil der Anzahl der Blutungstage pro

Blutungsstärkekategorie (z.B. „weniger als normale Blutung“, „normale Blutung“, „mehr als normale Blutung“) angegeben. Zudem wurden alle Blutungstage unabhängig von ihrer jeweiligen Blutungsstärke addiert und ebenfalls mit beschreibender Statistik dargestellt. Hierbei wurden dieselben statistischen Kenngrößen wie für die nach Blutungsstärke differenzierte Betrachtung verwendet.

Die Betrachtung des Blutungsmusters für hormonale Kontrazeptiva ist jedoch nicht auf die Anzahl der Blutungstage der Entzugsblutung und der Zwischenblutungen beschränkt. Vielmehr richtet sich das Hauptaugenmerk auf die dichotomisierten Zielgrößen des unerwünschten Auftretens einer oder mehrerer Zwischenblutungen in einem Zyklus sowie des Ausbleibens der erwünschten Entzugsblutung in einem Zyklus.

Daneben wurde eine globale Zielgröße zusätzlich zu diesen innerhalb des jeweiligen Zyklus aggregierten Zielgrößen definiert. Hierbei handelt es sich um das Auftreten einer oder mehrerer Zwischenblutungen in einem oder mehreren der Zyklen während der gesamten Studiendauer. Der erste Zyklus blieb hierbei unberücksichtigt, da es im ersten Zyklus durch die Umstellung vom natürlichen Menstruationszyklus zum hormoninduzierten Pillenzyklus zu relativ häufigen Zwischenblutungen kommt. D.h. es wurden die Zyklen zwei bis sechs im Falle einer halbjährigen Studie bzw. die Zyklen zwei bis dreizehn im Falle einer ganzjährigen Studie betrachtet. Diese globale dichotome Zielgröße des Auftretens von Zwischenblutungen in der Studie wurde für die Blutungsmusterbewertung als Hauptzielgröße definiert. Die dichotomisierten Zielgrößen wurden ebenfalls mit beschreibender Statistik analysiert. Hierbei wurden die statistischen Kenngrößen der absoluten und relativen Häufigkeiten verwendet. In vielen Arbeiten wird das Ausbleiben einer Entzugsblutung in einem Zyklus bereits als Amenorrhö bezeichnet, obwohl die Weltgesundheitsorganisation Amenorrhö als das Ausbleiben jedweder vaginaler Blutung für einen Zeitraum von 90 Tagen definiert.

Die Zyklus-basierten Blutungsmuster-Methoden der verschiedenen pharmazeutischen Unternehmer unterscheiden sich im Wesentlichen in der

unterschiedlichen Definition der Abschnitte, in die der Pilleneinnahmezyklus unterteilt wird. Eine gängige Unterteilung ist auch das Einnahmeschema selbst:

- Abschnitt 1: Tag 1 bis 21 des Pillenzyklus
- Abschnitt 2: Tag 22 bis 28 des Pillenzyklus

Eine Anpassung der Definitionen der Abschnitte ist auch dann erforderlich, wenn das Einnahmeschema der Präparate modifiziert wird. So wurden beispielsweise von dem pharmazeutischen Unternehmen Schering AG für die klinische Prüfung des 23 Tage-Regimes von Meliane (26) die Definitionen

- Abschnitt 1: Tag 1 bis 7 des Pillenzyklus
- Abschnitt 2: Tag 8 bis 23 des Pillenzyklus
- Abschnitt 3: Tag 24 bis 28 des Pillenzyklus

für den ersten Zyklus und

- Abschnitt 1: Tag 1 bis 5 des Pillenzyklus
- Abschnitt 2: Tag 6 bis 23 des Pillenzyklus
- Abschnitt 3: Tag 24 bis 28 des Pillenzyklus

für die auf den ersten Zyklus folgenden Zyklen angewandt.

Bei der klinischen Entwicklung des hormonhaltigen Pflasters Ortho-Evra® zur Kontrazeption wurde eine kombinierte Zyklus- und Episoden-basierte Blutungsmuster-Methodologie angewandt. Alle Blutungstage im Abschnitt Tag 1 bis 21 wurden als Zwischenblutung betrachtet, es sei denn, die Blutung war mit einer Entzugsblutung im Abschnitt von Tag 22 bis 28 zusammenhängend.

3.3.3.2 Episoden-basierte Blutungsmuster-Methoden

Die Episoden-basierten Blutungsmuster-Methoden gehen auf die bahnbrechende Publikation von Rodriguez et al. (108) aus dem Jahr 1976 zurück. Anstatt wie bei den Zyklus-basierten Methoden die Anzahl der Blutungstage in einem bestimmten Kalenderabschnitt als Untersuchungsgröße zu betrachten, werden die einzelnen Blutungstage zunächst zu Blutungsepisoden zusammengefasst. Die Blutungsepisoden werden dann als Untersuchungsgröße angesehen. Ein weiterer Unterschied zu den Zyklus-basierten Methoden ist die Definition der Untersuchungseinheit. Bei den Zyklus-basierten Methoden ist der Pilleneinnahme- bzw. der Menstruationszyklus die Untersuchungseinheit, während bei den Episoden-basierten Blutungsmuster-Methoden die Frau die Untersuchungseinheit darstellt. Letzteres ist aus methodologischen Gründen natürlich eindeutig vorzuziehen, da in diesem Falle alle Frauen gleich gewichtet werden und die Gewichtung nicht von der Anzahl der Zyklen abhängt.

Rodriguez et al. verwandten den Begriff „bleeding episode“ jedoch in anderer Weise als dies heute allgemein üblich ist. Sie definierten eine Folge von zusammenhängenden Tagen, an denen vaginale Blutungen auftraten, als „bleeding run“ und die sich daran anschließende Folge von zusammenhängenden blutungsfreien Tagen als „non-bleeding interval“. Ein „bleeding run“ und das sich unmittelbar anschließende „non-bleeding interval“ wurden zusammen als „bleeding episode“ bezeichnet. Auf der Grundlage dieser Blutungsepisoden wurden von Rodriguez et al. Blutungsmusterindices definiert.

Um eine longitudinale Darstellung des Blutungsmusters zu ermöglichen, wurden die Blutungsmusterindices für definierte Zeitabschnitte, sogenannte Referenzperioden, berechnet. Die Länge einer Referenzperiode kann an sich willkürlich gewählt werden, sie sollte aber die natürliche Periodizität der Menstruationsblutung von ungefähr vier Wochen nicht unterschreiten. Eine zu lange Referenzperiode hingegen bedingt eine zu starke Verdichtung der Daten über die Zeit. Rodriguez et al. (108) schlugen in

ihrer Publikation eine Referenzperiodenlänge von 90 Tagen vor, da diese Wahl für die übliche Länge von Blutungsmusterstudien, die zwischen 6 Monaten und ein bis zwei Jahre beträgt, geeignet ist.

Die von Rodriguez et al. (108) geprägten Begriffe und Methoden wurden von der Weltgesundheitsorganisation aufgegriffen und modifiziert(10, 5). Da die ursprüngliche Nomenklatur von Rodriguez et al. heute nicht mehr verwendet wird, wurde in dieser Arbeit auf eine ausführliche Darstellung der überholten Begrifflichkeiten verzichtet.

Definition Blutungsepisode

In einem Blutungskalender werden zusammenhängende Tage, an denen eine vaginale Blutung stattfand, zu einer Blutungsepisode (engl. „bleeding episode“) zusammengefasst. Hierbei ist zu beachten, dass einzelne Tage ohne Blutung eine Blutungsepisode nicht beenden. Eine Blutungsepisode ist folglich eine Folge von einem oder mehreren Tagen, an denen geblutet wurde, die auf beiden Seiten von jeweils mindestens zwei blutungsfreien Tagen begrenzt wird. Das blutungsfreie Intervall (engl. „bleeding-free interval“) ist dementsprechend eine Folge von mindestens zwei Tagen, an denen keine vaginale Blutung auftrat. Eine Blutungsepisode und das sich unmittelbar an sie anschließende blutungsfreie Intervall werden zu einem Blutungssegment (engl. „bleeding segment“) zusammengefasst.

Am besten lassen sich diese Definitionen an einem Beispiel wie in Abbildung 4 erläutern.

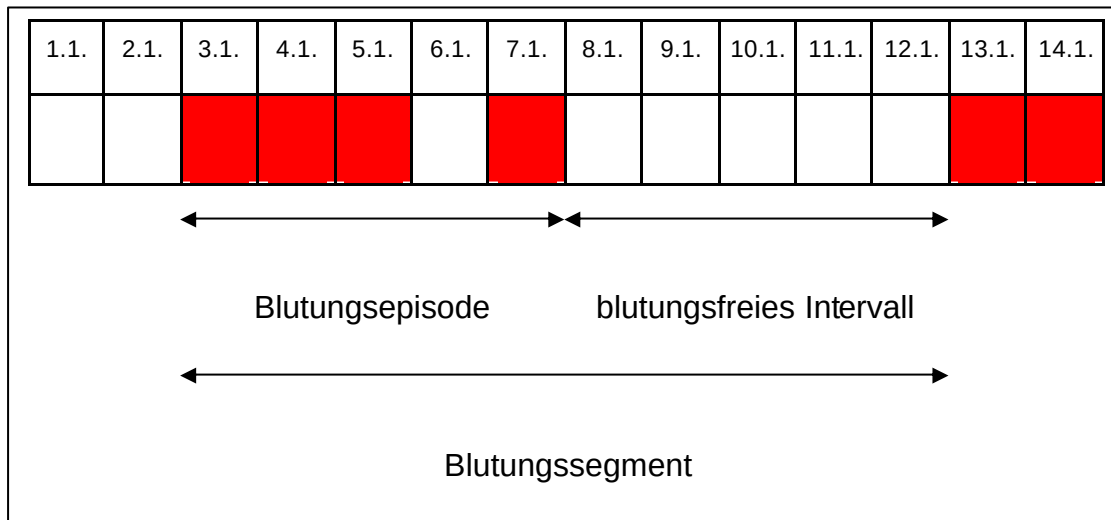


Abbildung 4: Beispiel Blutungssegment

Die Abbildung 4 zeigt einen zweiwöchigen Ausschnitt aus einem Blutungskalender vom 1. bis zum 14. Januar. Für den 3., 4., 5. und 7. Januar sind Blutungen eingetragen, ebenso für den 13. und 14. Januar. Da der 1. und der 2. Januar blutungsfrei waren, begann die erste Blutungsepisode am 3. Januar. Sie endete am 7. Januar, da der einzelne blutungsfreie Tag am 6. Januar die Blutungsepisode laut Definition nicht beendet. Dies geschieht erst am 8. Januar, da sowohl der 8. als auch der 9. Januar blutungsfrei waren.

Das sich an diese Blutungsepisode anschließende blutungsfreie Intervall dauerte vom 8. bis zum 12. Januar. Mit dem ersten Blutungstag nach mindestens zwei blutungsfreien Tagen begann am 13. Januar eine weitere Blutungsepisode. Das erste Blutungssegment dauerte vom 3. bis zum 12. Januar. Die Länge des Blutungssegments entspricht also der Dauer vom Beginn einer Blutungsepisode bis

zum Beginn der nächsten Blutungsepisode. Die Blutungsstärke wurde der Einfachheit halber bei diesem Beispiel nicht berücksichtigt.

Diese Blutungsereignisse werden den zu analysierenden Referenzperioden zugeordnet. Unter Referenzperioden versteht man aufeinander folgende Zeitabschnitte gleicher Länge, in welche die Studiendauer unterteilt wird. Durch die Einteilung der Studiendauer in die Referenzperioden lassen sich Änderungen des Blutungsverhaltens im Zeitverlauf darstellen. Üblicherweise wird eine Referenzperiode von 90 Tagen gewählt. Diese Länge entspricht in etwa einem Vierteljahr.

Eine Besonderheit gibt es bei der Zuweisung der Blutungsereignisse zu den einzelnen Referenzperioden zu beachten. Es kann vorkommen, dass ein Blutungsereignis in einer Referenzperiode beginnt, aber erst in der darauffolgenden Referenzperiode endet. In diesem Falle wird das betreffende Blutungsereignis nur derjenigen Referenzperiode, in der es beginnt, zugeordnet, um Doppelzählungen zu vermeiden.

Ausgehend von den obigen Definitionen der Blutungsereignisse „Blutungsepisode“, „blutungsfreies Intervall“ und „Blutungssegment“ werden nun für jede untersuchte Frau und für jede der Referenzperioden die Blutungsmusterindices definiert, die in Tabelle 3 dargestellt sind.

Hierbei ist zu beachten, dass selbst bei vollständiger Datenerhebung die von der WHO definierten Blutungsmusterindices nicht in jedem Falle berechnet werden können. Die Spannweite, d.h. die Differenz zwischen Maximum und Minimum, kann nur berechnet werden, wenn mindestens zwei der betreffenden Blutungsereignisse, wie zum Beispiel die Blutungsepisode, in der Referenzperiode begonnen haben. Die Berechnung der übrigen Indices, mit Ausnahme der Anzahl der Blutungsereignisse, erfordert den Beginn von mindestens einem Blutungsereignis in der nächsten Referenzperiode.

Bei der Untersuchung des Blutungsmusters von normal menstruierenden Frauen tritt das Problem der eben beschriebenen mangelnden Berechenbarkeit einzelner Blutungsmusterindices kaum auf, da ja in etwa drei Blutungsepisoden pro Quartal erwartet werden. Es tritt jedoch gehäuft auf bei der Untersuchung von unbehandelten Frauen während der Wechseljahre und von Frauen, die kontinuierlich kombinierte Hormonersatzpräparate einnehmen, da die Anzahl der Blutungsereignisse in diesen Populationen generell zurückgehen. Die Ursache dafür liegt im Falle der unbehandelten Frauen während der Wechseljahre an den gehäuft auftretenden spontanen Blutungsunregelmäßigkeiten und im Falle der Frauen, die kontinuierlich kombinierte Hormonersatzpräparate einnehmen, an den potentiell erheblichen Blutungsunregelmäßigkeiten in den ersten Monaten nach Einnahmebeginn.

In diesen beiden Populationen ist das mehrmonatige Ausbleiben einer Blutung weder ein seltenes noch ein unerwünschtes Ereignis. Dies stellt den Untersucher vor die Wahl, entweder sich auf die Auswertung der Anzahl der Blutungsepisoden zu beschränken, oder die ursprünglich gewählte Länge der Referenzperiode zu erhöhen.

Tabelle 3: Blutungsmusterindices nach WHO

Blutungsmusterindices pro Subjekt und Referenzperiode
• Anzahl der Blutungsepisoden • Anzahl der Blutungstage • arithmetisches Mittel der Länge der Blutungsepisoden • Minimum der Länge der Blutungsepisoden • Maximum der Länge der Blutungsepisoden • Spannweite der Länge der Blutungsepisoden
• arithmetisches Mittel der Länge der blutungsfreien Intervalle • Minimum der Länge der blutungsfreien Intervalle • Maximum der Länge der blutungsfreien Intervalle • Spannweite der Länge der blutungsfreien Intervalle
• arithmetisches Mittel der Länge der Blutungssegmente • Minimum der Länge der Blutungssegmente • Maximum der Länge der Blutungssegmente • Spannweite der Länge der Blutungssegmente

Die für die einzelne Frau und die einzelnen Referenzperioden abgeleiteten Blutungsmusterindices werden dann mittels deskriptiver Statistik, d.h. Anzahl der nicht fehlenden Werte, arithmetischer Mittelwert, Standardabweichung, Minimum, erstes Quartil, Median, drittes Quartil und Maximum ausgewertet. Je nach Fragestellung können aber auch andere deskriptive Statistiken verwendet werden.

Darüber hinaus definiert die WHO auf der Grundlage der in Tabelle 3 beschriebenen Blutungsmusterindices sogenannte „klinisch relevante“ Blutungsmuster. Diese sind in Tabelle 4 beschrieben. Hierbei ist zu beachten, dass diese Definitionen nur für eine Referenzperiodenlänge von 90 Tagen sinnvoll sind, da die Anzahl von Blutungsereignissen in einer Referenzperiode natürlich von der Länge der Referenzperiode abhängt.

Tabelle 4: Klinisch relevante Blutungsmuster nach WHO

klinisch relevante Blutungsmuster nach WHO (für 90-tägige Referenzperioden)	
1. Amenorrhö	kleine Blutung während der gesamten Referenzperiode
2. seltene Blutung	weniger als drei Blutungsepisoden während der Referenzperiode
3. häufige Blutung	mehr als fünf Blutungsepisoden während der Referenzperiode
4. unregelmäßige Blutung	- zwischen drei und fünf Blutungsepisoden während der Referenzperiode - weniger als drei blutungsfreie Intervalle mit einer Länge von jeweils mindestens 14 Tagen während der Referenzperiode
5. verlängerte Blutung	mindestens eine Blutungsepisode mit einer Länge von mindestens 14 Tagen während der Referenzperiode
6. „normales“ Blutungsmuster	keines der klinisch relevanten Blutungsmuster 1. bis 5.

Die in Tabelle 4 beschriebene Klassifikation klinisch relevanter Blutungsmuster der WHO ermöglicht leider keine eindeutige Klassifikation der Blutungsmuster, da die

Definition der Kategorie „5. verlängerte Blutung“ gleichzeitig mit einer der Definitionen der Kategorien „2. seltene Blutung“, „3. häufige Blutung“ oder „4. unregelmäßige Blutung“ erfüllt sein kann. Die übrigen Kategorien der WHO Klassifikation schließen sich jedoch gegenseitig aus, so wie dies von einem korrekten Klassifikationsschema erwartet werden kann.

Die Referenzperiodenmethode von Rodriguez et al. (108) und die darauf aufbauende WHO Klassifikation (10) wurden ursprünglich entwickelt, um das Blutungsmuster von Frauen, die empfängnisverhütende Mittel einnehmen, mit dem Blutungsmuster von nicht verhütenden geschlechtsreifen Frauen vergleichen zu können. Sie eignen sich insbesondere zur Darstellung des Blutungsmusters von hormonellen Langzeitpräparaten wie Spiralen, Implantaten, Intrauterinsystemen oder Dreimonatsdepotpräparaten.

Die Referenzperiodenmethode eignet sich jedoch weniger gut zur Beurteilung der Zykluskontrolle sequenzieller hormonaler Kontrazeptiva und von sequenziellen Hormonersatzpräparaten, durch die ein bestimmtes Blutungsmuster durch eine regelmäßige Hormongabe, gefolgt von regelmäßigem Hormonentzug, induziert werden soll.

Zur Analyse der Blutungsmuster bei zyklischen Hormonanwendungen wurden ursprünglich die Zyklus-basierten Methoden entwickelt (vgl. Abschnitt 3.3.3.1). Diese eignen sich jedoch nicht für Langzeitpräparate, die keinen Zyklus vorgeben. Darüber hinaus haben die Zyklus-basierten Methoden den Nachteil, dass sie keine Rücksicht auf zusammenhängende Blutungsepisoden nehmen können. So kann es vorkommen, dass eine einzelne Blutungsepisode bei der Anwendung einer Zyklus-basierten Methode gleichzeitig als Zwischenblutung als auch als Entzugsblutung gezählt wird. Dieser Fall tritt immer dann auf, wenn sich die Blutungsepisode über mehr als einen der festen Kalenderabschnitte erstreckt.

Eine neue Methode, die Episoden-basierten und die Zyklus-basierten Methoden zu kombinieren, ist im Abschnitt 5.3 beschrieben.

3.4 Zusammenfassung der Literaturrecherche

Zusammenfassend lässt sich sagen, dass die Zyklus-basierte Blutungsmuster-Methode den Vorteil der einfachen Definitionen und der im Vergleich zu den Episoden-basierten Blutungsmuster-Methoden unkomplizierteren Datenerfassung hat. Der wesentliche Nachteil der Zyklus-basierten Blutungsmuster-Methoden liegt hingegen darin, dass nicht unterschieden wird, wie häufig unerwünschte Blutungsereignisse auftreten. Insbesondere im Falle von Durchbruchblutungen, d.h. von Zwischenblutungen, welche die Verwendung eines sanitären Produktes unbedingt erfordern, kommt es schon darauf an, ob es sich um ein Ereignis von dreitägiger Dauer handelt oder ob es sich um drei verschiedene Ereignisse handelt. Diese Unterscheidung wird nur bei den Episoden-basierten Methoden getroffen. Ein weiterer Nachteil der Zyklus-basierten Methoden ist die Tatsache, dass die Definition der Grenzen der einzelnen Abschnitte eines Zyklus nicht einer gewissen Willkür entbehrt. Insbesondere wird bei den Zyklus-basierten Blutungsmuster-Methoden eine Blutung, die vom letzten Tag eines Abschnitts bis zum ersten Tag des folgenden Abschnitts andauert, sowohl als Zwischenblutung als auch als Entzugsblutung gezählt werden, obwohl es sich eigentlich nur um eine einzige Blutungsepisode handelt.

Die Zyklus-basierten Blutungsmuster-Methoden lassen sich grundsätzlich auch auf die Analyse des Blutungsmusters von sequenziellen Präparaten zur Hormonersatztherapie, bei denen die Gestagenkomponente einem dem natürlichen 28-Tage Menstruationsblutungszyklus folgt, anwenden. Sie sind jedoch für die hauptsächlich verschriebenen kontinuierlich kombinierten Hormonersatzpräparate, die täglich die gleiche Menge sowohl an Östrogen als auch an Gestagen enthalten, nicht geeignet, da diese Präparate keinen zyklischen Hormonentzug beinhalten. Demzufolge wären sämtliche Blutungsereignisse als unerwünschte Zwischenblutungen zu werten. Letzteres trifft auch z.B. auf Gestagenmonopräparate zur Behandlung der Endometriose zu, die ohne Unterbrechung in gleicher Stärke täglich verabreicht werden.

4 Empirische Untersuchung von Blutungsmustern

Parallel zur Auswertung der medizinischen und biometrischen Fachliteratur zum Thema Blutungsmuster wurde eine empirische Untersuchung von Blutungskalendern, die aus klinischen Studien mit Hormonpräparaten stammten, durchgeführt. Die gesamte Datenbasis mit Blutungskalendern eines auf den Gebieten der hormonellen Kontrazeption und Hormonersatztherapie weltweit führenden pharmazeutischen Herstellers wurde für diese Untersuchung verwendet. Die Datenbasis umfasst die Blutungskalender von insgesamt mehr als fünftausend Frauen.

4.1 Datenbeschreibung

Die für diese Arbeit verwendeten Blutungskalenderdaten stammen alle aus klinischen Prüfungen, die für die arzneimittelrechtliche Zulassung sexualhormonhaltiger Präparate durchgeführt wurden.

Alle Daten, die im Rahmen der klinischen Studien erhoben wurden, sind in eine elektronische Studiendatenbank eingelesen worden. Das bedeutet, dass alle Blutungsdaten von zwei verschiedenen Datentypisten unabhängig voneinander eingegeben wurden. Diese beiden Eingaben wurden anschließend von einem Programm miteinander verglichen, um somit Einlesefehler der Datentypisten nahezu auszuschließen. Darüber hinaus wurden die Einträge im Blutungskalender mit den übrigen Daten der Studienteilnehmerin zusammen auf ihre Plausibilität geprüft. Falls bei der Plausibilitätsprüfung Unstimmigkeiten der Daten erkannt werden, löst der für die Studie verantwortliche Datenmanager diese auf. Dabei wird wie folgt vorgegangen: Bei einfachen Tippfehlern legt der verantwortliche Datenmanager den korrekten Wert fest, indem er den korrekten Wert aus den Berichtsbögen manuell in die Datenbank einträgt. Im Falle von nicht plausiblen Daten wird beim Prüfarzt unter Verwendung eines Datenkorrekturbogens nachgefragt und der korrigierte Wert wird, nach erneuter Plausibilitätsprüfung, vom Datenmanager in die Datenbank eingestellt.

Um die Erarbeitung der studienübergreifenden Dokumente eines Zulassungsantrags für ein neues Medikament zu ermöglichen, werden die Studiendatenbanken der einzelnen im Zulassungsantrag enthaltenen klinischen Prüfungen in eine Projektdatenbank integriert. Zum Teil wird statt mehrerer Projektdatenbanken eine Substanzdatenbank erstellt. Diese enthält dann alle Studien mit der Prüfsubstanz unabhängig von der jeweiligen Indikation des Zulassungsantrags. Die Zusammenfassung mehrerer Projektdatenbanken in eine Substanzdatenbank ist insbesondere dann von Vorteil, wenn einzelne Studien eines Entwicklungsprojektes für die Zulassungsanträge mehrerer Indikationen verwendet werden können. Dies trifft zum Beispiel auf Studien zur Pharmakokinetik der Substanz oder auf Studien der Sicherheit der Substanz in Patienten mit eingeschränkter Nierenfunktion zu.

Für diese Arbeit wurden Daten aus vier verschiedenen Datenbanken des pharmazeutischen Herstellers, der die Daten zur Verfügung stellte, ausgelesen. Hierbei handelt es sich um zwei Projektdatenbanken sowie zwei Substanzdatenbanken. Jede der beiden Projektdatenbanken enthielt jeweils die Daten der klinischen Prüfungen eines Präparates zur hormonellen Kontrazeption. Die erste Substanzdatenbank enthielt alle Studien eines synthetischen Gestagens in den Indikationen Fertilitätskontrolle und Hormonersatztherapie. Die zweite Substanzdatenbank enthielt alle Studien eines anderen synthetischen Gestagens in den Indikationen Fertilitätskontrolle, Endometriose und Hormonersatztherapie.

Tabelle 5 gibt einen Überblick, über die Anzahl der Frauen, deren Blutungskalender in den genannten Datenbanken verfügbar sind und in welcher Indikation die klinischen Prüfungen, in denen sie geführt wurden, durchgeführt wurden. Die überwiegende Mehrzahl dieser Blutungskalender wurde für mehr als drei Monate geführt (vgl. Tabelle 6).

Mehrere Projektdatenbanken von älteren Projekten des pharmazeutischen Herstellers konnten leider nicht in diese Untersuchung mit einbezogen werden. Der Grund dafür ist, dass die Blutungskalender früher auf Grund von Beschränkungen

von Speicherplatz und Rechnerleistung nur in bereits zusammengefasster Form in die Studiendatenbanken eingegeben wurden (vgl. Abschnitt 3.3.1.2).

Tabelle 5: Anzahl und Gesamtlänge der verfügbaren Blutungskalender nach Indikation

Indikation	Anzahl der Blutungskalender	Gesamtlänge aller Blutungskalender in Jahren
Endometriose	321	120.3
Hormonersatztherapie	3246	2389.8
Hormonale Kontrazeption	2035	917.1
Summe	5602	3427.2

Die Datenbasis der vorliegenden Untersuchung umfasst somit alle bis Ende 2003 bei dem pharmazeutischen Hersteller verfügbaren Blutungskalender, für die tägliche Eintragungen in den Datenbanken vorliegen.

Tabelle 6: Anzahl der verfügbaren Blutungskalender nach Länge und Indikation

Indikation	Länge der Blutungskalender in Tagen								
	1-90		91-180		181-360		>360		alle
	n	%	n	%	n	%	n	%	n
Endometriose	88	27.4	210	65.4	23	7.2	0	0.0	321
Hormonersatz- therapie	454	14.0	369	11.4	1601	49.3	822	25.3	3246
hormonelle Kontrazeption	241	11.8	939	46.1	855	42.0	0	0.0	2035
Summe	783	14.0	1518	27.1	2479	44.3	822	14.7	5602

4.1.1 Fehlende Werte

Bei vielen empirischen Untersuchungen tritt das Problem fehlender Werte auf. In klinischen Prüfungen geführte Blutungskalender bilden darin leider keine Ausnahme. Da fehlende Werte je nach ihrer Ursache zu einer Verzerrung des Auswertungsergebnisses führen können, ist eine genaue Untersuchung der möglichen Auswirkungen fehlender Werte in empirischen Studien geboten.

Für die Untersuchung von Blutungskalendern ist das Problem der fehlenden Werte jedoch nicht von großer Bedeutung, da die Zuverlässigkeit der Studienteilnehmerinnen insbesondere in Studien von oralen Kontrazeptiva als sehr gut zu bezeichnen ist.

Man unterscheidet in Abhängigkeit des statistischen Modells bei der Generierung fehlender Werte drei Arten:

- „completely missing at random“: Die geplante Beobachtung eines Wertes w fehlt rein zufällig. Die dichotome Zufallsvariable X mit den Ausprägungen „Wert w wurde beobachtet“ und Ausprägungen „Wert w wurde nicht beobachtet“ ist statistisch unabhängig sowohl von w als auch von allen anderen im Rahmen der Untersuchung erhobenen Variablen.
Beispiel: Ein Blutungskalender, der beispielsweise auf dem Postwege verloren geht, würde als „completely missing at random“ gelten.
- „missing at random“: Die geplante Beobachtung eines Wertes w fehlt zufällig. Die dichotome Zufallsvariable X mit den Ausprägungen „Wert w wurde beobachtet“ und Ausprägungen „Wert w wurde nicht beobachtet“ ist statistisch unabhängig von w , sie kann aber von anderen im Rahmen der Untersuchung erhobenen Variablen abhängen, mit Ausnahme der Beobachtungen, die zu einem späteren Zeitpunkt als die Beobachtung von w erhoben wurden.
Beispiel: Studentinnen ohne Kinder ziehen öfter um als Angestellte mit Kindern, daher würden durch einen vorzeitigen Studienabbruch wegen Umzugs bedingte fehlende Werte in einem Blutungskalender als „missing at random“ gelten.
- „not missing at random“: Die geplante Beobachtung eines Wertes w fehlt nicht zufällig. Die dichotome Zufallsvariable X mit den Ausprägungen „Wert w wurde beobachtet“ und Ausprägungen „Wert w wurde nicht beobachtet“ ist statistisch abhängig von w .
Beispiel: Typischerweise tritt diese Art fehlender Werte in klinischen Prüfungen auf, wenn ein Patient zu krank ist, um den fehlenden Wert beobachten zu können. Bei Blutungskalendern von klinischen Studien im Rahmen der Hormonersatztherapie, der Behandlung der Endometriose oder der hormonellen Empfängnisverhütung ist ein derartiger Fall aber höchst unwahrscheinlich.
Streng genommen müsste man auch den Fall, dass Frauen nur Blutungen in

ihren Blutungskalender eintragen und an Tagen ohne Blutung keine Einträge vornehmen, unter die Kategorie „not missing at random“ einordnen. In der praktischen Durchführung der Studien wird aber ein derartiges Eintrageverhalten vom Prüfarzt bei der gemeinsamen Durchsicht des Blutungskalenders mit der Studienteilnehmerin sicher erkannt. Die fehlenden Einträge werden daraufhin ergänzt, so dass letztlich in diesem Falle die Beobachtungen vollständig in der Datenbank erfasst werden können.

Das größte Problem bei der Berücksichtigung fehlender Werte in einer empirischen Untersuchung ist das informative Fehlen „not missing at random“. Wie oben ausgeführt, tritt dieses Problem jedoch bei den in dieser Arbeit untersuchten Blutungskalendern, wenn überhaupt, dann höchstens in einem für die Fragestellung dieser Arbeit vernachlässigbaren Anteil auf.

Das Problem fehlender Werte in den Blutungskalendern wurde dadurch gelöst, dass einzelne fehlende Werte in den Blutungskalendern ersetzt wurden. In der Praxis werden zwei Ersetzungsalgorithmen für fehlende Werte häufig angewandt. Zum einen ist dies der sogenannte „Last Value Carried Forward“ (LOCF) Algorithmus, bei dem ein fehlender Wert durch den letzten beobachteten Wert ersetzt wird. Zum anderen ist dies der „worst case“ Ansatz, bei dem ein fehlender Wert durch den ungünstigsten Wert ersetzt wird.

Beide Standardverfahren sind für die Ersetzung fehlender Werte in Blutungskalendern jedoch unbefriedigend. Beim LOCF Algorithmus würde eine Lücke im Kalender zur Nichtblutung, wenn am letzten Tag vor der Lücke nicht geblutet worden war, und dies selbst dann, wenn die erste Eintragung nach der Lücke eine Blutung ist. Beim „worst case“ Algorithmus würde jede Lücke, unabhängig von ihrem Kontext, als Blutung gezählt.

Daher wird für diese Arbeit ein modifizierter „worst case“ Ersetzungsalgorithmus verwendet. Eine Lücke in den Beobachtungen im Blutungskalender wird durch das Maximum der Blutungsstärken am Tag unmittelbar vor sowie am Tag unmittelbar

nach der Lücke ersetzt. Hierbei wurde die Ersetzung auf Lücken von bis zu sieben Tagen begrenzt, um einen zu großen Einfluss des Ersetzungsalgorithmus für fehlende Werte auf das Ergebnis zu vermeiden. Blutungskalender, die trotz dieser Ersetzungsregel noch fehlende Werte aufwiesen, wurden aus der Auswertung ausgeschlossen. In aller Regel betraf dies die Blutungskalender von Frauen, welche ihre Studieneilnahme vorzeitig abgebrochen hatten. Kurze Blutungskalender wegen vorzeitigem Studienabbruch bedeuten für diese Untersuchung jedoch keinen Erkenntnisgewinn, da es sich hier um ein nicht informatives Fehlen der Blutungsdaten handelt.

Es kann daher davon ausgegangen werden, dass fehlende Eintragungen in den untersuchten Blutungskalendern nicht zu einer notablen Verzerrung der Ergebnisse dieser empirischen Untersuchung führen können.

4.1.2 Beschränkung der Datenbasis

Für die empirische Untersuchung der Blutungskalender stellte sich die Frage nach der Beschränkung der Datenbasis. Einerseits sollten möglichst viele Blutungskalender in die empirische Auswertung mit einbezogen werden, andererseits müssen die zu untersuchenden Blutungskalender auch über eine gewisse Mindestlänge verfügen. Wie aus Tabelle 6 ersichtlich ist, waren die Längen der Blutungskalender im vorliegenden Datenbestand sehr heterogen. Der Hauptgrund für die unterschiedliche Länge der Blutungskalender ist durch das Studiendesign bedingt, da die in den Datenbanken erfassten Studien von unterschiedlicher Dauer waren. Darüber hinaus führte auch die vorzeitige Beendigung der Studie durch Studienteilnehmerinnen zu verkürzten Blutungskalendern.

Es wurden daher nur Blutungskalender in die empirische Analyse einbezogen, die eine Mindestlänge von 90 Tagen aufwiesen. Die 90 Tage wurden in Anlehnung an die Definitionen der WHO (10) gewählt. Durch diese Definition konnten über 85% der Blutungskalender einbezogen werden. Für Blutungskalender, die länger als 90 Tage geführt worden waren, wurden nur die ersten 90 Tage in Betracht gezogen. Diese

Einschränkung war für die vorliegende Arbeit unerheblich, da sich das Blutungsverhalten unter der Einnahme von Hormonpräparaten nach dem zweiten Monat nicht mehr wesentlich ändert.

4.2 Wahl des statistischen Verfahrens

Bei den Einträgen in einen Blutungskalender handelt es sich im Prinzip um Daten in Form einer diskreten Zeitreihe. Für jedes Subjekt i und zu jedem Zeitpunkt t wird eine Blutungsstärke $X_{i,t}$ beobachtet. Für Zeitreihen mit Wiederholungen gibt es aber noch keine statistischen Verfahren, die auch in den gängigen Softwarepaketen zur statistischen Datenanalyse wie SAS (109), SPSS (110), oder S-Plus (111) implementiert sind.

Daher konnte sich die vorliegende Arbeit nur auf die klassischen Methoden der statistischen Analyse multivariater Daten stützen. Eine gute Übersicht dieser Methoden findet sich zum Beispiel in den Monographien von Anderson (112) oder Dillon und Goldstein (113).

Da das Ziel der vorliegenden Arbeit die Unterscheidung von verschiedenen Blutungsmustern ist, kommen von den klassischen Verfahren nur die Diskriminanzanalyse und die Clusteranalyse in Betracht. Beide Verfahren zerlegen eine Menge von Beobachtungen in disjunkte Teilmengen.

Eine Diskriminanzanalyse setzt jedoch voraus, dass die Grundgesamtheit der zu klassifizierenden Objekte in eine vorab bekannte Anzahl von echten Teilmengen zerlegt wird. Die Diskriminanzanalyse setzt weiter voraus, dass eine Teststichprobe mit Objekten, von denen die Zugehörigkeit zu den Teilmengen bereits bekannt ist, existiert, anhand derer die verbleibenden Objekte klassifiziert werden. Beide dieser Voraussetzungen sind jedoch für das Blutungsmusterproblem nicht gegeben. Weder sind die verschiedenen Blutungsmuster a priori bekannt, noch existiert eine Lernstichprobe mit bereits bestimmten Blutungsmustern zugeordneten Blutungskalendern.

Die Blutungskalenderdaten wurden daher mit Hilfe der Clusteranalyse empirisch untersucht. Die Methode der Clusteranalyse wurde für die vorliegende Arbeit ausgewählt, weil sie gerade kein Vorwissen über das Blutungsmuster verlangt. Das in den Blutungskalendern vermutete Blutungsmuster soll durch „unsupervised pattern recognition“, um den Sprachgebrauch der Literatur zur künstlichen Intelligenz (114) zu verwenden, durch die Clusteranalyse erkannt werden.

4.3 Clusteranalyse

Die Clusteranalyse gehört zu den wichtigsten Methoden der multivariaten Datenanalyse und Mustererkennung für numerische Daten. Das Ziel der Clusteranalyse ist es, eine Grundgesamtheit von Objekten, die durch mehrere Variablen (Merkmale) beschrieben werden, in Teilmengen (sogenannte Cluster oder Klassen) zu zerlegen. Informationen zu möglichen Beziehungen der einzelnen Variablen (Merkmale) untereinander sind für die Clusteranalyse nicht nötig, d.h. eine spezifische Kenntnis des Datenraumes ist nicht explizit nötig.

Ziel der Clusteranalyse ist es, dass die Objekte, die zu demselben Cluster gehören, möglichst ähnlich (interne Homogenität) sein sollen, während die Objekte, die verschiedenen Clustern angehören, dagegen möglichst unterschiedlich (externe Heterogenität) sein sollen. Die gebildeten Cluster ermöglichen dann eine Typologisierung der zu untersuchenden Objekte in „natürliche“ Gruppen, bei denen jedes Objekt den anderen Objekten seiner Gruppe ähnlicher ist, als den Objekten, die anderen Gruppen zugeordnet wurden. Diese Gruppen sind bei der Clusteranalyse nicht von vornherein festgelegt, sie werden vielmehr als Muster im zu untersuchenden Datenraum erkannt.

4.3.1 Überblick über Clusterverfahren

Unter dem Begriff der Clusteranalyse werden unterschiedliche statistische Verfahren zusammengefasst. Grundsätzlich lassen sich die Methoden zur Clusteranalyse zunächst in disjunktive und nicht disjunktive Verfahren unterscheiden. Bei einer

disjunktiven Methode wird gefordert, dass sich die gebildeten Cluster nicht überschneiden. Jede Untersuchungseinheit wird folglich genau einem Cluster zugeordnet. Im Gegensatz dazu können die Untersuchungseinheiten bei einer nicht disjunktiven Methode auch mehreren Clustern gleichzeitig angehören. Zu den nicht disjunktiven Methoden gehören auch die probabilistischen Clusterverfahren, bei denen die Zugehörigkeit eines Objektes zu einem Cluster durch eine Wahrscheinlichkeit beschrieben wird. Im Rahmen dieser Arbeit wurden nur disjunktive Clusterverfahren angewandt, da in der klinischen Forschung Ambivalenzen möglichst vermieden werden. Zudem erfüllt ein nicht disjunktives Verfahren die Anforderung nach externer Heterogenität der Cluster in geringerem Maße als ein disjunktives Verfahren.

Eine weitere wichtige Unterscheidung der Methoden zur Clusteranalyse besteht zwischen sogenannten hierarchischen und partitionierenden Verfahren. Bei den hierarchischen Verfahren wird schrittweise eine Folge von Einteilungen der Grundgesamtheit in k Cluster vorgenommen, wobei k zwischen 1 und der Anzahl n der Objekte im Datensatz liegt. Bei hierarchischen Verfahren werden einmal gebildete Cluster nicht mehr verändert. Sie werden lediglich im nächsten Schritt des hierarchischen Algorithmus ggf. bei agglomerativen Verfahren mit einem anderen Cluster fusioniert bzw. bei divisiven Verfahren in zwei Cluster unterteilt. Es ergibt sich also eine Folge von Einteilungen der Grundgesamtheit in 2, 3, 4, ..., n Cluster. Abschließend kann der Untersucher die ihm a priori unbekannte optimale Anzahl an Clustern festzulegen.

Dagegen werden bei partitionierenden Verfahren die Untersuchungseinheiten in eine vom Untersucher fest vorgegebene Anzahl k von Clustern eingeteilt. Für die in dieser Arbeit vorgesehene empirische Untersuchung der Blutungskalender kommen jedoch nur hierarchische Verfahren in Betracht, da die Anzahl der für die Fragestellung relevanten verschiedenen Blutungsmuster a priori unbekannt ist.

Die hierarchischen Verfahren zur Clusteranalyse lassen sich wiederum in agglomerative und in divisive Clusterverfahren unterscheiden. Bei einem hierarchisch

agglomerativen Verfahren wird im ersten Schritt jede Untersuchungseinheit als ein einelementiger Cluster aufgefasst. In jedem sukzessiven Gruppierungsschritt werden die beiden Cluster mit dem geringsten Abstand zueinander jeweils zu einem neuen Cluster zusammengefasst. So entsteht eine Folge von Einteilungen in Cluster bis im letzten Schritt alle Untersuchungseinheiten zu einem einzigen Cluster zusammengefasst worden sind, der wiederum genau die Grundgesamtheit bildet. Dagegen wird bei einem hierarchisch divisiven Clusterverfahren mit einem einzigen Cluster, der Grundgesamtheit, begonnen. In jedem der sukzessiven Schritte des hierarchisch divisiven Clusterverfahrens wird dann einer der bestehenden Cluster in zwei kleinere Cluster zerlegt. Im Gegensatz zu agglomerativen Verfahren sind divisive Verfahren jedoch algorithmisch sehr aufwendig (vgl. 113 Abschnitt 5.3.1.2). Dies liegt daran, dass ein Cluster mit n Elementen auf insgesamt $2^{n-1} - 1$ Arten in zwei Cluster zerlegt werden kann und der Algorithmus unter allen diesen Möglichkeiten die optimale Zerlegung finden soll. Aus den genannten Gründen kommen daher für diese Arbeit nur hierarchisch agglomerativen Verfahren in Frage.

Überblicksarbeiten zu den verschiedenen weiteren Methoden der Clusteranalyse, die hier nicht behandelt werden, sind zum Beispiel Everitt, Landau und Leese (115) sowie Bacher (116).

4.3.2 Abstandsbegriff

Das Kernstück jeder Clusteranalyse ist, neben der Auswahl des Clusteralgorithmus, die Definition der Ähnlichkeit der einzelnen Objekte, die in Cluster eingeteilt werden sollen. Wie zu erwarten, gibt es in der Literatur eine Fülle von Vorschlägen, die Ähnlichkeit zweier Objekte zu definieren. Dabei ist zu beachten, dass die Wahl der Ähnlichkeitsdefinition natürlich nicht willkürlich sein kann. Vielmehr muss die Definition der Ähnlichkeit zweier Objekte substanzwissenschaftlich interpretierbar sein.

Für binäre Daten wird die Ähnlichkeit in der Regel auf einer Skale von 0 bis 1 angegeben. Eine Ähnlichkeit von 0 bedeutet hier, dass sich zwei Objekte in allen

binären Merkmalen unterscheiden, während eine Ähnlichkeit von 1 bedeutet, dass die beiden Objekte in allen Merkmalen übereinstimmen, d.h. dass sie gleich sind.

Bei der Definition der Ähnlichkeit ist zu unterscheiden, ob die Abwesenheit eines bestimmten Merkmals in beiden Objekten zur Information über die Ähnlichkeit der beiden Objekte beiträgt oder nicht. Everitt, Landau und Leese (1985, Seite 36) nennen als Beispiel die Taxonomie von Tierarten, bei denen das übereinstimmende Fehlen von Flügeln bei zwei Tierarten noch keine Aussage über die Ähnlichkeit der beiden Tierarten erlaubt. Im Gegensatz dazu trägt das gemeinsame Fehlen einer Blutung bei der Betrachtung der Ähnlichkeit zweier Blutungskalender durchaus zur Information bei.

Werden die Übereinstimmungen und die Nichtübereinstimmungen der p binären Merkmale von zwei Objekten in einer Kontingenztafel der Form von Tabelle 7 dargestellt, so lässt sich die Ähnlichkeit zweier Objekte anhand der Einträge in der Kontingenztafel einfach beschreiben.

Tabelle 7: Kontingenztabelle der Übereinstimmung von p binären Merkmalen

	Objekt j			
Objekt i	Merkmal	vorhanden	nicht vorhanden	Summe
	vorhanden	a	b	a+b
	nicht vorhanden	c	d	c+d
Summe		a+c	b+d	a+b+c+d=p

Wenn die Abwesenheit eines bestimmten Merkmals in beiden Objekten zur Information über die Ähnlichkeit der beiden Objekte nicht beträgt, so wird häufig der Jaccard'sche Koeffizient

$$s_{ij} = \frac{a}{a+b+c}$$

als Maß für die Ähnlichkeit der beiden Objekte i und j verwendet. Die Anzahl der Merkmale, die für die beiden Objekte i und j gleichzeitig nicht vorhanden sind, geht nicht in die Berechnung des Jaccard'schen Koeffizienten ein.

Wenn hingegen die Abwesenheit eines bestimmten Merkmals in beiden Objekten zur Information über die Ähnlichkeit der beiden Objekte beträgt, so wird häufig der einfache Übereinstimmungskoeffizient

$$s_{ij} = \frac{a+d}{a+b+c+d}$$

als Maß für die Ähnlichkeit der beiden Objekte i und j verwendet. Der einfache

Übereinstimmungskoeffizient gibt die relative Häufigkeit der Übereinstimmung der Merkmalsausprägungen der beiden Objekte i und j an.

Für nominale Merkmale mit mehr als zwei Ausprägungen schlagen Everitt, Landau und Leese (115, Seite 38 bis 39) zunächst deren Zerlegung in mehrere binäre Merkmale vor. Die Autoren schreiben jedoch selbst, dass dies eine unbefriedigende Lösung sei. Sie schlagen anschließend vor, für nominale Merkmale den einfachen Übereinstimmungskoeffizienten zu verwenden, indem die relative Häufigkeit der Übereinstimmung der nominalen Merkmale zwischen den beiden Objekten als Maß für deren Ähnlichkeit gewählt wird. Dieses Vorgehen ist jedoch für ordinal skalierte Daten unbefriedigend, da die inhärente Ordnung der Merkmalsausprägungen nicht bei der Erfassung der Ähnlichkeit zweier Objekte berücksichtigt wird. Im Falle der ordinal skalierten Blutungsstärken kann durchaus davon ausgegangen werden, dass die Ähnlichkeit von „keine Blutung“ und „Spotting“ echt größer ist als die Ähnlichkeit von „keine Blutung“ und „Blutung“.

Bacher (116, Seite 213ff) schlägt für ordinal skalierte Merkmale die Verwendung von Abstandsmaßen, die für metrisch skalierte Merkmale definiert sind, vor, da diese auch eine „ordinale Interpretation“ haben. Abstandsmaße für metrisch skalierte Merkmale werden im folgenden erläutert. Dazu ist es jedoch zuerst notwendig, vom Begriff der Ähnlichkeit zweier Objekte auf den Begriff der Verschiedenheit, d.h. des Abstandes zweier Objekte, überzuleiten.

An Stelle einer Definition der Ähnlichkeit zweier Objekte kann auch eine Definition des Abstands zweier Objekte voneinander für die Clusteranalyse verwendet werden. In der Regel lässt sich eine Definition der Ähnlichkeit auch in eine äquivalente Definition des Abstands transformieren. So lassen sich beispielsweise der Jaccard'schen Koeffizient und der einfache Übereinstimmungskoeffizient durch die Definition

$$d_{ij} = 1 - s_{ij}$$

in einen Abstand überführen. Dieser Abstand ist wie die Ähnlichkeit auf das Intervall $[0; 1]$ normiert.

Die Software (vgl. Abschnitt 4.3.6), die für die vorliegende Arbeit verwendet wurde, erfordert die Eingabe des Abstandes der Objekte. Daher wird im folgenden nur auf die Beschreibung von Abständen und nicht weiter auf die Beschreibung von Ähnlichkeiten eingegangen.

Der Abstand zweier Objekte wird üblicherweise mit Hilfe eines Abstandsmaßes beschrieben. An ein Abstandsmaß $d(x,y)$ für zwei Elemente x und y aus einer Grundgesamtheit M werden die folgenden Anforderungen gestellt:

1. $d(x,x)=0 \quad \forall x \in M$
2. $d(x,y) \geq 0 \quad \forall x \in M, \forall y \in M, x \neq y$
3. $d(x,y)=d(y,x) \quad \forall x \in M, \forall y \in M$
4. $d(x,y)+d(y,z) \geq d(x,z) \quad \forall x \in M, \forall y \in M, \forall z \in M$

Die ersten beiden Bedingungen sagen aus, dass der Abstand zwischen zwei Objekten stets nicht negativ ist und dass der Abstand eines Objekts zu sich selbst gleich null ist. Die dritte Bedingung verlangt die Symmetrie des Abstands. Die vierte Bedingung ist die Dreiecksungleichung, die besagt, dass der direkte Weg (d.h. Abstand) zwischen zwei Punkten im Datenraum nie länger als ein Umweg über einen beliebigen dritten Punkt ist.

Für metrisch skalierte Merkmale gibt es in der Literatur ebenso wie für die nominal skalierten Merkmale eine Fülle von Vorschlägen zur Abstandsdefinition. Die meisten dieser Vorschläge sind Spezialformen der verallgemeinerten Minkowski-Metrik. Die verallgemeinerte Minkowski-Metrik mit den Parametern r und q definiert den Abstand zweier Punkte x und y im p -dimensionalen Raum durch

$$d_{(r,q)}(x, y) = \left(\sum_{i=1}^p |x_i - y_i|^r \right)^{\frac{1}{q}}$$

wobei x_i die i -te Koordinate von x und y_i die i -te Koordinate von y bezeichnet. Wird r gleich q gesetzt, so ergibt sich die eigentliche Minkowski-Metrik,

$$d_r(x, y) = \left(\sum_{i=1}^p |x_i - y_i|^r \right)^{\frac{1}{r}}$$

die in der mathematischen Literatur auch unter dem Begriff L_r -Norm bekannt ist. Hierbei muss zusätzlich noch gelten, dass $r \geq 1$ ist, damit die vorgenannten Anforderungen für ein Abstandsmaß erfüllt sind.

Um eine Minkowski-Metrik als Definition des Abstands zweier Objekte für eine Clusteranalyse anwenden zu können, muss der Wert des Parameters r festgelegt werden. Der Parameter r gibt an, wie stark kleine Unterschiede in vielen Merkmalen im Verhältnis zu großen Unterschieden in wenigen Merkmalen gewichtet werden sollen. Bei einem kleinen Wert von r werden kleine Unterschiede in vielen Merkmalen deutlich stärker gewichtet als bei einem großen Wert von r . Im Gegensatz dazu werden große Unterschiede in wenigen Merkmalen bei einem kleinen Wert von r weniger stark gewichtet als bei einem großen Wert von r .

In der Regel werden jedoch für die Clusteranalyse die Definitionen $r = 1$ oder $r = 2$ verwendet. Mit der Festlegung auf $r = 1$ ergibt sich die L_1 -Norm,

$$d_1(x, y) = \sum_{i=1}^p |x_i - y_i|$$

die auch mit dem Namen City-Block-Metrik bezeichnet wird. Der Abstand der beiden Objekte x und y ist hier durch die Summe der Beträge der Abstände der einzelnen Koordinaten x_i und y_i definiert. Dadurch werden die absoluten Abweichungen aller Merkmale der Objekte x und y gleich gewichtet.

Wird als Parameter der Minkowski-Metrik $r = 2$ verwendet, so ergibt sich der euklidische Abstand

$$d_2(x, y) = \left(\sum_{i=1}^p |x_i - y_i|^2 \right)^{\frac{1}{2}} = \sqrt{\sum_{i=1}^p (x_i - y_i)^2}$$

der beiden Objekte x und y , der auch als L_2 -Norm bezeichnet wird. Der euklidische Abstand stimmt mit der üblichen geometrischen Entfernung im Raum überein. Er ist auch die Grundlage der Methode der kleinsten Quadrate, wie sie beispielsweise in der Statistik für die Schätzung einer Regressionsgerade angewandt wird.

Von allen Abstandsdefinitionen wird der euklidische Abstand lt. Everitt, Landau und Leese (115, Seite 40) für die Clusteranalyse am meisten verwendet. Die für diese Arbeit getroffene Wahl der Abstandsdefinition für Blutungskalender wird im Abschnitt 4.4.1 näher erläutert werden.

4.3.3 Hierarchisch agglomerative Verfahren zur Clusteranalyse

Unter der Vielzahl der inzwischen veröffentlichten hierarchisch agglomerativen Verfahren zur Clusteranalyse bezeichnen Everitt, Landau und Leese (115, Seite 62) sechs Verfahren als Standardverfahren. Im einzelnen nennen sie single linkage, complete linkage, average linkage, centroid linkage, median linkage sowie das Verfahren von Ward.

Demgegenüber nennen Dillon und Goldstein (113, Abschnitt 5.3.1.1) nur single linkage, complete linkage, Average-Linkage-sowie das Verfahren von Ward als wesentliche Verfahren zur hierarchischen Clusteranalyse. Diese gebräuchlichen Verfahren zur hierarchischen agglomerativen Clusteranalyse sind in Tabelle 8 zusammengefasst und werden anschließend näher erläutert.

Tabelle 8: gebräuchliche hierarchisch agglomerative Clusterverfahren

Methode	Abstandsdefinition zwischen Clustern
single linkage	kleinster paarweiser Abstand zwischen einem Element des einen und einem Element des anderen Clusters
complete linkage	größter paarweiser Abstand zwischen einem Element des einen und einem Element des anderen Clusters
average linkage	durchschnittlicher paarweiser Abstand zwischen einem Element des einen und einem Element des anderen Clusters
Ward	Zunahme der Quadratsummen zwischen den Clustern über alle Merkmale summiert

Für die Auswahl einer Methode zur hierarchischen agglomerativen Clusteranalyse gibt es kein uniformes Optimalitätskriterium. Je nach Art der in den Daten enthaltenen Cluster erzeugen die verschiedenen Methoden verschieden gute Ergebnisse. So ergibt beispielsweise die Methode von Ward nach einer Simulationsstudie von Hands und Everitt (117) gute Ergebnisse, wenn die in den Daten vorhandenen Cluster ungefähr die gleiche Anzahl von Elementen besitzen, und nicht so gute Ergebnisse, wenn die Anzahl von Elementen in den einzelnen Clustern grob unterschiedlich ist.

Everitt, Landau und Leese (115, Seite 67) schließen ihren Vergleich der verschiedenen Methoden mit der Bemerkung, dass die Single-Linkage-Methode in praktischen Anwendungen oftmals wegen ihrer Neigung zur Verkettung keine brauchbaren Ergebnisse liefert. Dagegen würde die Methode von Ward in der Praxis

gut funktionieren, selbst wenn die eigentlich von ihr geforderte geometrische Interpretation nicht gewährleistet sei.

Der methodische Ansatz der vorliegenden Arbeit ist es, mehrere der genannten Clusterverfahren zur empirischen Erkennung von Blutungsmustern anzuwenden, um anschließend deren Ergebnisse auf substanzwissenschaftliche Interpretierbarkeit zu untersuchen.

4.3.3.1 Single-Linkage-Methode

Die Single-Linkage-Methode verwendet als Definition des Abstandes zweier Cluster den kleinsten paarweisen Abstand zwischen einem Element des einen und einem Element des anderen Clusters. Dies bedeutet, dass der Abstand zweier Cluster C_i und C_j durch

$$d(C_i, C_j) = \min \{d(x, y)\} \text{ mit } x \in C_i \text{ und } y \in C_j$$

definiert wird. Daher wird diese Methode auch die Methode des nächsten Nachbarn (nearest-neighbour method) genannt. Die Single-Linkage-Methode wurde zuerst von Florek et al. (118) und Sneath (119) beschrieben.

In der Anwendung führt das Single-Linkage-Verfahren oft zu einer sogenannten „Verkettung“ (engl. „chaining“). Insbesondere in großen Datensätzen, wie sie in der vorliegenden Arbeit untersucht werden, bedeutet dies, dass dieses Verfahren dazu tendiert, sukzessive einen einzigen großen Cluster zu bilden.

4.3.3.2 Complete-Linkage-Methode

Die Complete-Linkage-Methode geht auf Sørensen (120) zurück. Im Gegensatz zur Single-Linkage-Methode wird der Abstand zweier Cluster C_i und C_j bei der Complete-Linkage-Methode nicht durch den kleinsten Abstand, sondern durch den größten Abstand zwischen einem Element des ersten Clusters und einem Element des zweiten Clusters definiert:

$$d(C_i, C_j) = \max \{d(x, y)\} \text{ mit } x \in C_i \text{ und } y \in C_j$$

Die Complete-Linkage-Methode wird auch als die Methode des fernsten Nachbarn (furthest-neighbour method) bezeichnet. Die Complete-Linkage-Methode tendiert dazu, kompakte Cluster mit ähnlichem Durchmesser zu finden. Laut Bacher (116, Seite 146) neigt die Complete-Linkage-Methode jedoch dazu, viele kleine Cluster zu bilden.

4.3.3.3 Average-Linkage-Methode

Bei der Average-Linkage-Methode wird der arithmetische Mittelwert aller Abstände zwischen je einem Element des ersten Clusters und je einem Element des zweiten Clusters als Maß für den Abstand zweier Cluster verwendet:

$$d(C_i, C_j) = \frac{\sum d(x, y)}{|C_i| \cdot |C_j|} \text{ mit } x \in C_i \text{ und } y \in C_j$$

wobei $|M|$ die Mächtigkeit der Menge M bezeichnet. Die Average-Linkage-Methode wurde zuerst von Sokal und Michener (121) beschrieben, sie stellt einen Kompromiss zwischen der Methode des nächsten und der Methode des fernsten Nachbarn dar. Die Average-Linkage-Methode gilt als relativ robust (115, Seite 62).

4.3.3.4 Methode von Ward

Die Methode von Ward (122) basiert auf dem Ansatz der kleinsten Quadrate. Es werden in jedem Schritt die beiden Cluster zusammengefasst, die den Anstieg der innerhalb-der-Cluster Quadratsumme (within cluster sum of squares)

$$E = \sum_{m=1}^g E_m$$

minimiert. Die einzelne der innerhalb-der-Cluster Quadratsumme ist durch

$$E_m = \sum_{l=1}^{n_m} \sum_{k=1}^p (x_{m,l,k} - \bar{x}_{m,k})^2$$

definiert, wobei

$$\bar{x}_{m,k} = \frac{\sum_{l=1}^{n_m} x_{m,l,k}}{n_m}$$

den arithmetischen Mittelwert des k -ten Merkmals des m -ten Clusters bezeichnet. Hierbei steht $x_{m,l,k}$ für den Wert des k -ten Merkmals des m -ten Clusters für die l -te Untersuchungseinheit. Die Anzahl der Cluster wird in den Formeln mit p , die Anzahl der Elemente im m -ten Cluster wird mit n_m bezeichnet.

Die Methode von Ward tendiert dazu, Cluster von gleicher Größe zu finden. Sie gilt jedoch als anfällig für einzelne Ausreißer (115, Seite 62). Dafür ist die Methode von Ward relativ robust gegenüber der Wahl eines falschen Abstandsmaßes (116, Seite 150), was für die empirische Untersuchung der Blutungskalender von nicht unerheblicher Bedeutung ist, da, wie später ausgeführt werden wird, es kein kanonisches Abstandsmaß für Blutungskalender gibt.

Die wesentlichen Methoden für das agglomerativ hierarchische Vorgehen sind die Single-Linkage-Methode, die Complete-Linkage-Methode, die Average-Linkage-Methode, sowie die Minimum-Varianz-Methode von Ward.

4.3.4 Dendrogramm

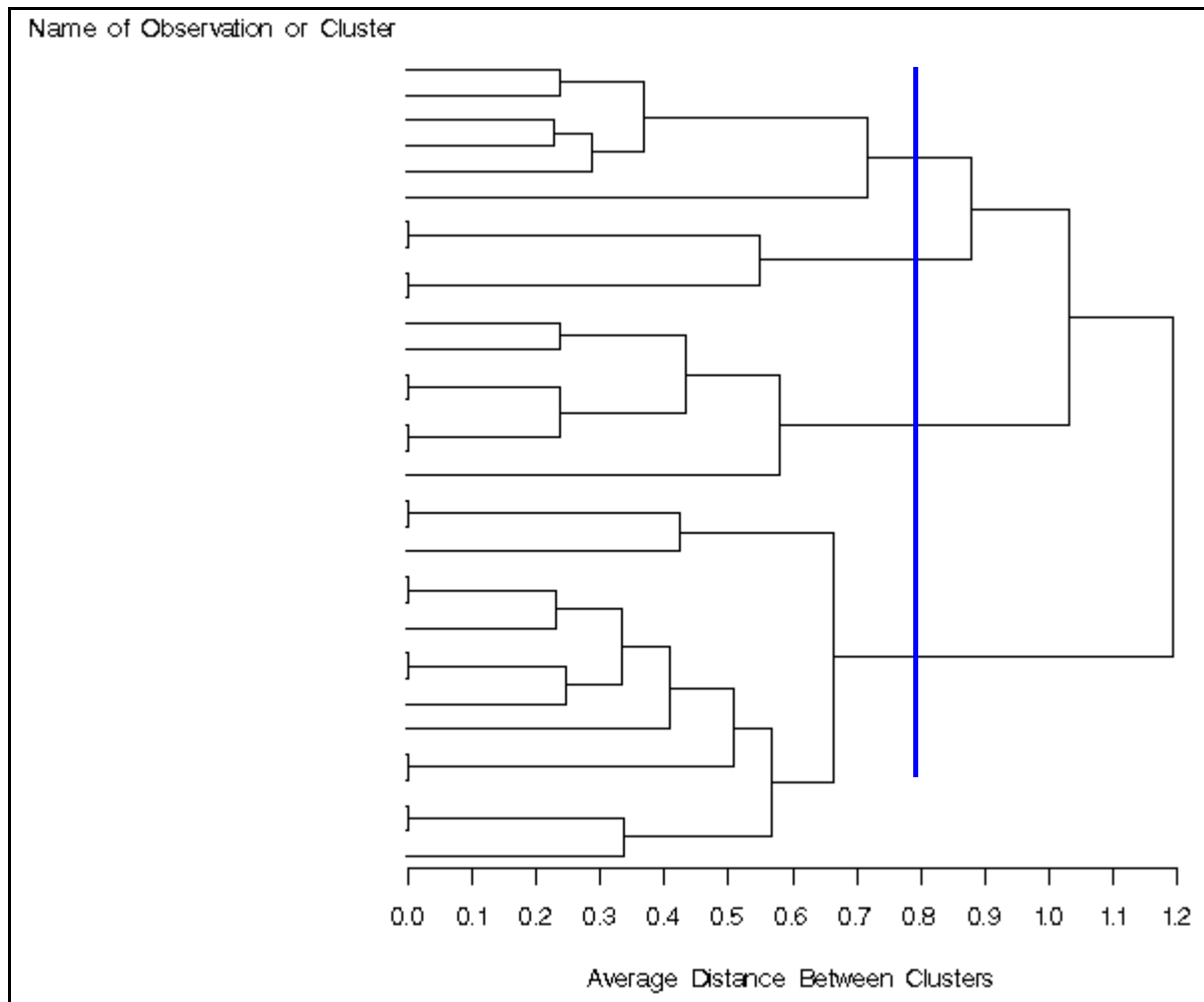
Die Ergebnisse einer hierarchischen Clusteranalyse lassen sich für hochdimensionale Daten am besten in einem sogenannten Dendrogramm beschreiben. Dabei wird die hierarchische Baumstruktur der vom Clusterverfahren erzeugten Partitionen der Grundgesamtheit graphisch dargestellt (vgl. 113 Seite 168, 115 Seite 70ff).

Die Objekte in der Grundgesamtheit stellen die Blätter des Baumgraphen dar. Die Knoten stellen die jeweils gebildeten Cluster dar und die Wurzel des Baumgraphen besteht aus dem obersten Cluster, der alle Objekte der Grundgesamtheit umfasst.

Der Abstand zwischen den auf dem jeweiligen Niveau fusionierten Clustern wird als Länge der Kante dargestellt.

Ein anschauliches Beispiel eines Dendrogramms ist in Abbildung 5 dargestellt.

Abbildung 5: Beispiel für ein Dendrogramm



adaptiert aus (123, Seite 3552) mit Schnitt bei 4 Clustern (senkrechte Linie)

Aus dem Dendrogramm lässt sich direkt entnehmen, in welcher Reihenfolge die Beobachtungen in der in Abbildung 5 dargestellten Clusteranalyse zu Clustern zusammengefasst werden.

Das Dendrogramm kann auch zur heuristischen Bestimmung der Anzahl der in den Daten vorhandenen Cluster verwendet werden (115 Seite 77ff und Seite 103). Typischerweise wird ein Sprung in den sukzessiven Abständen, zu denen die einzelnen Cluster fusioniert werden, als Indikator für den besten Schnitt (engl. „best cut“) eines Dendrogramms verwendet. Im Beispiel des Dendrogramms in Abbildung 5 könnte der beste Schnitt bei vier Clustern gezogen werden. Die Wahl der Anzahl der Cluster sollte aber zusätzlich auch von substanzwissenschaftlichen Überlegungen begleitet werden, da es für die Analyse eines Dendrogramms keine uniform optimale Methode gibt.

Abschließend sei noch darauf hingewiesen, dass die Darstellung der Ergebnisse einer hierarchischen Clusteranalyse in einem Dendrogramm nicht eindeutig ist. Durch Permutationen der einzelnen Objekte lassen sich für eine Darstellung mit n Objekten insgesamt 2^{n-1} verschiedene Dendrogramme erzeugen. Üblicherweise wird jedoch eine überschneidungsfreie Darstellung gewählt.

4.3.5 Wahl der Anzahl der Cluster

Wenn, wie in der vorliegenden Arbeit, die Anzahl der in den untersuchten Daten vorhandenen Cluster a priori unbekannt ist, lassen sich verschiedene statistische Kenngrößen zur Bestimmung der optimalen Anzahl der Cluster heranziehen. Im einzelnen sind dies das semi-partielle Bestimmtheitsmaß, das kubische Cluster-Kriterium, die pseudo-F- und die pseudo- t^2 -Statistik.

Für jeden Fusionsschritt gibt das semi-partielle Bestimmtheitsmaß $sp-R^2$ an, welcher Anteil an der durch die Clustereinteilung erklärten Varianz durch die anstehende Fusion der beiden Cluster verloren geht. Wenn in einem Schritt die Cluster a und b zum Cluster c fusioniert werden, so berechnet sich das semi-partielle Bestimmtheitsmaß $sp-R^2$ durch

$$sp - R^2 = \frac{\sum_{l=1}^{n_c} \sum_{k=1}^p (x_{c,l,k} - \bar{x}_{c,k})^2 - \sum_{l=1}^{n_a} \sum_{k=1}^p (x_{a,l,k} - \bar{x}_{a,k})^2 - \sum_{l=1}^{n_b} \sum_{k=1}^p (x_{b,l,k} - \bar{x}_{b,k})^2}{\sum_{l=1}^n \sum_{k=1}^p (x_{l,k} - \bar{x}_k)^2}$$

Das kubische Cluster-Kriterium wurde von Sarle (124) auf der Basis von heuristischen Simulationsstudien eingeführt. Werte des kubischen Cluster-Kriteriums von über 3 deuten meist auf eine gute Anzahl von Clustern hin. Da sich das kubische Cluster-Kriterium für das in dieser Arbeit untersuchte Problem als ungeeignet erwies, wird bezüglich seiner genauen Definition auf (124, Seite 8) verwiesen.

Die pseudo-F-Statistik beschreibt den Quotienten der Abweichungsquadratsumme zwischen den Clustern geteilt durch die Abweichungsquadrate innerhalb der Cluster. Für eine Einteilung der Grundgesamtheit in G Cluster berechnet sich die pseudo-F-Statistik durch

$$pseudo\ F = \frac{\frac{\sum_{l=1}^n \sum_{k=1}^p (x_{l,k} - \bar{x}_k)^2 - \sum_{g \in G} \sum_{l=1}^{n_g} \sum_{k=1}^p (x_{g,l,k} - \bar{x}_{g,k})^2}{G-1}}{\frac{\sum_{g \in G} \sum_{l=1}^{n_g} \sum_{k=1}^p (x_{g,l,k} - \bar{x}_{g,k})^2}{n-G}}$$

Es muss jedoch betont werden, dass die pseudo-F-Statistik keinesfalls einer F-Verteilung folgt. Ein großer Wert der pseudo-F-Statistik deutet oft auf eine gute Anzahl von Clustern hin.

Die pseudo- t^2 -Statistik gibt für die Fusion zweier Cluster a und b zum Cluster c die Zunahme der Abweichungsquadratsumme durch die Fusion dividiert durch die Abweichungsquadratsumme in den Clustern a und b an.

$$pseudo\,t^2 = \frac{\sum_{l=1}^{n_c} \sum_{k=1}^p (x_{c,l,k} - \bar{x}_{c,k})^2 - \sum_{l=1}^{n_a} \sum_{k=1}^p (x_{a,l,k} - \bar{x}_{a,k})^2 - \sum_{l=1}^{n_b} \sum_{k=1}^p (x_{b,l,k} - \bar{x}_{b,k})^2}{\frac{\sum_{l=1}^{n_a} \sum_{k=1}^p (x_{a,l,k} - \bar{x}_{a,k})^2 + \sum_{l=1}^{n_b} \sum_{k=1}^p (x_{b,l,k} - \bar{x}_{b,k})^2}{n_a + n_b - 2}}$$

Wie im Falle der pseudo-F-Statistik muss explizit darauf hingewiesen werden, dass die pseudo-t²-Statistik nicht t² verteilt ist. Eine gute Anzahl von Clustern wird durch die Anzahl der Cluster unmittelbar vor einem deutlichen Anstieg der pseudo-t²-Statistik gefunden (vgl. 123 Seite 839f).

4.3.6 Software

Die Clusteranalysen wurden mit dem statistischen Softwarepaket SAS[®], das von der Firma SAS Institute Inc. aus Cary, North Carolina, USA, (www.sas.com) herausgegeben wird, durchgeführt.

Es wurde die Version 8.e der SAS Module SAS/Base (125, 126, 127, 128), SAS/GRAPH (129) und SAS/STAT (123) verwendet. Die eigentliche Clusteranalyse wurde mit der SAS Prozedur PROC CLUSTER berechnet. PROC CLUSTER ist in Kapitel 23 von (123) beschrieben. Die Dendrogramme wurden mit der SAS Prozedur PROC TREE berechnet und graphisch dargestellt. PROC TREE ist in Kapitel 66 von (123) beschrieben.

Das SAS Programmpaket lief unter dem Betriebssystem Windows XP, Version 5.1 (130) auf einem Omnibook 6100 des Herstellers Hewlett-Packard Company (www.hp.com) aus Wilmington, Delaware, USA.

4.4 Anwendung auf Blutungskalender

Für die empirische Untersuchung der Blutungskalenderdaten wurden als Untersuchungsobjekte die einzelnen Blutungskalender definiert. Dies bedeutet, dass die einzelnen Frauen anhand ihres Blutungsverhaltens in die einzelnen Cluster

eingeordnet werden. Die statistische Untersuchungseinheit ist also die Frau und nicht der einzelne Zyklus. Die Merkmale der Untersuchungsobjekte, anhand derer die Clusteranalyse erfolgt, sind die im Blutungskalender eingetragenen Blutungsstärken der einzelnen Blutungskalendertage.

Alle der in der verwendeten SAS Software implementierten Algorithmen zur Clusteranalyse gehen von vollständigen Daten, d.h. von Beobachtungen ohne fehlende Werte, aus. Daher wurden einzelne fehlende Werte in den Blutungskalendern nach dem im Abschnitt 4.1.1 beschriebenen Verfahren in den Blutungskalendern ergänzt. Da die Blutungskalender in den untersuchten Datenbanken über unterschiedliche lange Zeiträume zwischen einem und 448 Tagen geführt wurden, war es erforderlich, die Länge der untersuchten Blutungskalender zu vereinheitlichen. Eine Ergänzung vieler fehlender Werte nach dem letzten Eintrag in einem kurzen Blutungskalender ist inhaltlich wenig sinnvoll. Daher wurden nur Blutungskalender, die über einen gewissen Mindestzeitraum geführt wurden, in die Auswertung mit einbezogen. Dabei sollte einerseits der Mindestzeitraum lang genug sein, um substanzwissenschaftlich sinnvolle Aussagen zu ermöglichen. Der Mindestzeitraum sollte andererseits aber auch nicht zu lang gewählt werden, damit möglichst viele der vorliegenden Blutungskalender in die Auswertung mit einbezogen werden können.

Für die vorliegende Arbeit wurde ein Mindestzeitraum von 90 Tagen festgelegt. Der Mindestzeitraum entspricht damit der Länge einer Referenzperiode gemäß den Definitionen der Weltgesundheitsorganisation WHO (10). Es kann daher davon ausgegangen werden, dass der Mindestzeitraum lang genug ist, um aussagefähige Ergebnisse zu erhalten. Der gewählte Mindestzeitraum von 90 Tagen ist jedoch nicht zu lang gewählt worden. Die Daten der überwiegenden Mehrheit von 86% der Frauen, für die Blutungskalender in den ausgewerteten Datenbanken vorlagen, konnten in die empirische Analyse mit einbezogen werden.

4.4.1 Abstandsdefinition zweier Blutungskalender

Die Anwendung der Clusteranalyse zur automatischen Klassifikation der Blutungsmuster setzt die Definition eines geeigneten Abstandes zwischen zwei Blutungskalendern voraus. Ein kanonisches Abstandsmaß, wie den euklidischen Abstand zweier Punkte im n -dimensionalen Raum, gibt es für Blutungskalender a priori nicht. Es muss daher ein geeignetes Abstandsmaß für Blutungskalender gefunden werden, mittels dessen sich substanzwissenschaftlich sinnvolle Ergebnisse ergeben.

Bei den einzelnen Merkmalen des Blutungskalenders, den an den jeweiligen Tagen beobachteten Blutungsstärken, handelt es sich genaugenommen um ordinal skalierte Größen. Die Blutungsstärke wurde in den drei Kategorien „keine Blutung“, „Spotting“ und „Blutung“ eingeteilt. Die inhärente Ordnung dieser Kategorien der gemessenen Blutungsstärke wird durch die Kodierung in die Scores „0 = keine Blutung“, „1 = Spotting“ und „2 = Blutung“ abgebildet.

Hierbei ist zu beachten, dass die Scores a priori zwar eine Reihenfolge, aber nicht notwendigerweise auch einen Abstand beschreiben. Um die Scores 0 bis 2 als Abstand verwenden zu können, müsste unterstellt werden, dass der Unterschied zwischen „keine Blutung“ und „Spotting“ genau so groß wäre wie der Unterschied zwischen „Spotting“ und „Blutung“. Eine Folge dieser Annahme wäre, dass die beiden Unterschiede zwischen „keine Blutung“ und „Spotting“ bzw. zwischen „Spotting“ und „Blutung“ genau halb so groß wären wie der Unterschied zwischen „keine Blutung“ und „Blutung“.

Man könnte zwar durch eine andere Wahl der Scores, wie zum Beispiel „0 = keine Blutung“, „2/3 = Spotting“ und „1 = Blutung“, die notwendige Forderung der Gleichabständigkeit der einzelnen Blutungsstärkekategorien umgehen, müsste aber wiederum unterstellen, dass die gewählten Scores den Abstand der Blutungsstärken tatsächlich beschreiben.

Das Problem des Abstandes der einzelnen Blutungsstärkekategorien ließe sich jedoch umgehen, indem man die Blutungsstärken „Spotting“ und „Blutung“ zu einer Kategorie zusammenfasst. Es werden dann nur noch die Kategorien „0 = keine Blutung“ und „1 = Spotting und/oder Blutung“ unterschieden. Der numerische Unterschied der Scores spielt jetzt keine Rolle mehr, da die Blutungsstärke jetzt de facto als dichotom angesehen wird. Die Dichotomisierung der Blutungsstärke ist jedoch mit einem Informationsverlust verbunden, da „Spotting“ und „Blutung“ nicht mehr unterschieden werden können. Dem entgegen steht die Tatsache, dass so gut wie alle in der Literaturrecherche gefundenen Methoden zur Analyse des Blutungsmusters sehr wohl eine Unterscheidung von „Spotting“ und „Blutung“ vornehmen. Daraus lässt sich schließen, dass eine Dichotomisierung der Blutungsstärke substanzwissenschaftlich gesehen eine unzulässige Vereinfachung des Problems wäre. Daher wurde dieser Ansatz nicht weiter verfolgt.

Die in Abschnitt 4.3.2 beschriebenen Abstandsdefinitionen haben alle den Nachteil, dass es für diese Abstandsmaße keinen Unterschied macht, an welchen Tagen ein Unterschied zwischen zwei Blutungskalendern auftritt. So wird jede Abweichung an einem einzelnen Tag mit dem Abstand 1 bewertet, unabhängig davon, ob die Abweichung eine neue Blutungsepisode erzeugt (vgl. Kalender 2 in Abbildung 6) oder nur eine bestehende Blutungsepisode verlängert (vgl. Kalender 3 in Abbildung 7).

	1	2	3	4	5	6	7	8	9	10
Kalender 1										
Kalender 2										

Abbildung 6: Beispiel neue Blutungsepisode

	1	2	3	4	5	6	7	8	9	10
Kalender 1										
Kalender 3										

Abbildung 7: Beispiel verlängerte Blutungsepisode

Es liegt jedoch nahe, die in Abbildung 6 und Abbildung 7 exemplarisch dargestellten Blutungsmuster nicht mit dem gleichen Abstand zu bewerten. Die Abweichung zwischen Kalender 1 und Kalender 3 in Abbildung 7, die eine neue Blutungsepisode darstellt, wäre demzufolge geringer zu gewichten ist als die Abweichung zwischen Kalender 1 und Kalender 2 in Abbildung 6, die lediglich die Verlängerung einer bereits bestehenden Blutungsepisode bedeutet.

Eine Gewichtung der Abweichungen könnte zum Beispiel dadurch geschehen, dass Abweichungen, die eine bestehende Blutungsepisode verlängern, nur mit einem Faktor von $1/2$ bewertet werden, während Abweichungen, die verschiedene Blutungsepisoden darstellen, mit einem Faktor von 1 bewertet werden.

Diese intuitive Definition erfüllt jedoch die Dreiecksungleichung nicht, wie anhand des Beispiels in Abbildung 8 gezeigt wird. Dabei ist es unerheblich, ob wie in der Beispielsrechnung der euklidische Abstand mit den oben definierten Faktoren gewichtet wird, oder ob eine gewichtete City Block Metrik verwendet wird.

	1	2	3	4	5	6	7	8	9	10		
Kalender 1												
Kalender 2												
Kalender 3												

Abbildung 8: Beispiel verletzte Dreiecksungleichung bei Gewichtung

Der gewichtete Abstand zwischen Kalender 1 und Kalender 3 in Abbildung 8 beträgt

$\sqrt{6} \approx 2,449$, während der gewichtete Abstand zwischen Kalender 1 und Kalender 2

$\sqrt{2/2} = 1$ und der gewichtete Abstand zwischen Kalender 2 und Kalender 3

$\sqrt{4/2} \approx 1,414$ beträgt. Der direkte gewichtete Abstand zwischen Kalender 1 und

Kalender 3 $d(\text{Kalender 1}, \text{Kalender 2})$ wäre also größer als die Summe der

gewichteten Abstände zwischen Kalender 1 und Kalender 2 sowie zwischen

Kalender 2 und Kalender 3 $d(\text{Kalender 1}, \text{Kalender 3}) + d(\text{Kalender 3}, \text{Kalender 2})$ was

einen Widerspruch zur Dreiecksungleichung bedeutet. Aus diesem Grunde wurde auf

eine wie immer geartete Gewichtung der Abstände der Blutungskalender in der

vorliegenden Arbeit verzichtet.

Da die Methode von Ward relativ robust gegenüber der Wahl eines Abstandsmaßes

ist (116, Seite 150), ist es für die empirische Untersuchung der Blutungskalender

auch nicht von erheblicher Bedeutung, welches Abstandsmaß verwendet wird.

4.4.2 Ergebnisse der Clusteranalyse

Insgesamt konnten die Blutungskalender von 4612 Frauen in die Auswertung mit einbezogen werden. Einziges Kriterium für die Aufnahme eines Blutungskalenders in die Auswertung war, dass die Einträge der Blutungsstärke im Blutungskalender in der ersten 90-tägigen Referenzperiode vollständig sein mussten. Hierbei wurden einzelne fehlende Werte nach dem im Abschnitt 4.1.1 beschriebenen Verfahren ersetzt.

Bei der explorativen Datenanalyse des Datensatzes der Blutungskalender fällt ein besonderer Cluster auf. Etwa ein Viertel der Frauen (n=1288) im untersuchten Datensatz besitzt das gleiche Blutungsmuster (vgl. Tabelle 9). Dagegen haben etwa drei Viertel der Frauen (n=3172) ein Blutungsmuster, das sich vom Blutungsmuster aller anderen Frauen in der vorliegenden Stichprobe unterscheidet.

Tabelle 9: Anzahl Blutungskalender mit identischem Blutungsmuster

Anzahl Blutungskalender mit identischem Blutungsmuster	Anzahl der verschiedenen Blutungsmuster
1	3172
2	39
3	11
4	5
5	3
6	1
1288	1

Bei dem besonderen Cluster der 1288 Frauen mit dem gleichen Blutungsmuster handelt es sich um das Blutungsmuster der Amenorrhö, d.h. es trat keine einzige Blutung im gesamten Untersuchungszeitraum von 90 Tagen auf. Obwohl der Begriff „Amenorrhö“ auch im Sinne einer Krankheitsbezeichnung (ICD-10 Klassifikation „N91“) verwendet wird, ist Amenorrhö hier nicht im pathologischen Sinne zu verstehen. Gut die Hälfte der verfügbaren Blutungskalender stammt aus klinischen Studien zur Hormonersatztherapie bei peri- und postmenopausalen Frauen (vgl. Tabelle 5). Das Ausbleiben von vaginalen Blutungsereignissen ist jedoch in dieser Indikation, insbesondere für die Untergruppe der postmenopausalen Frauen, keine Krankheit. Ganz im Gegensatz dazu ist die weitest mögliche Blutungsfreiheit eines der wesentlichen Zielkriterien für die Entwicklung von kontinuierlich kombinierten Präparaten zur Hormonersatztherapie.

Das nach der vollständigen Amenorrhö nächst häufige Blutungsmuster wurde hingegen nur genau sechs mal beobachtet. Es bestand aus vier zusammenhängenden Blutungstagen an den Tagen 1 bis 4 gefolgt von 86 blutungsfreien Tagen. Die weiteren Blutungsmuster traten noch weniger gehäuft auf. Wie bereits erwähnt, hatten 3172 (68,78 %) der Frauen ein Blutungsmuster, das in der untersuchten Datenbasis einmalig war.

Durch die große Anzahl von 1288 Frauen mit exakt dem gleichen Blutungsmuster tritt bei der Anwendung der Clusteralgorithmen das Problem der Bindungen auf. Eine Bindung tritt immer dann auf, wenn es in einem Schritt des Clusteralgorithmus mehr als zwei Cluster gibt, die das Minimum aller Abstände zwischen zwei Clustern annehmen. Die Wahl, welche beiden Cluster aus der Menge aller möglichen Paare von Clustern mit dem minimalen Abstand untereinander zum einem neuen Cluster zusammengefasst werden, ist im Falle des Auftretens von Bindungen willkürlich. Der Clusteralgorithmus kann folglich nicht mehr auf jedem Schritt eindeutige Ergebnisse liefern.

Für die praktische Anwendung der Clusteranalyse ist das Auftreten von Bindungen jedoch solange unerheblich, wie Bindungen nur in frühen Stadien der hierarchischen

Bildung von Clustern auftreten. Dies ist einleuchtend, da es für die in dieser Arbeit untersuchte Fragestellung völlig belanglos ist, in welcher der $1288! \approx 1,42697 \cdot 10^{3448}$ Möglichkeiten die 1288 gleichen Blutungskalender zu einem „Amenorrhö“ Cluster zusammengefasst werden.

Für die numerische Berechnung der Cluster beim Auftreten von Bindungen wurde das in der verwendeten Software SAS implementierte Verfahren (vgl. 123, Seite 865f) verwendet. Im Folgenden werden die Ergebnisse der einzelnen verwendeten Algorithmen zur Clusteranalyse berichtet.

4.4.2.1 Ergebnisse der Methode von Ward

Die Clusteranalyse wurde zunächst mit der Methode von Ward durchgeführt. Anschließend wurden zusätzlich die Single-Linkage-Methode, die Average-Linkage-Methode sowie die Complete-Linkage-Methode angewandt, um das mit der Methode von Ward erhaltene Ergebnis auf Robustheit hinsichtlich der Wahl des Verfahrens zur Clusteranalyse zu untersuchen.

Das Verfahren von Ward liefert eine Einteilung der Daten in klar unterschiedene Cluster, wie sich anhand des Dendrogramms der Clusteranalyse nach Ward (vgl. Abbildung 9) zeigt.

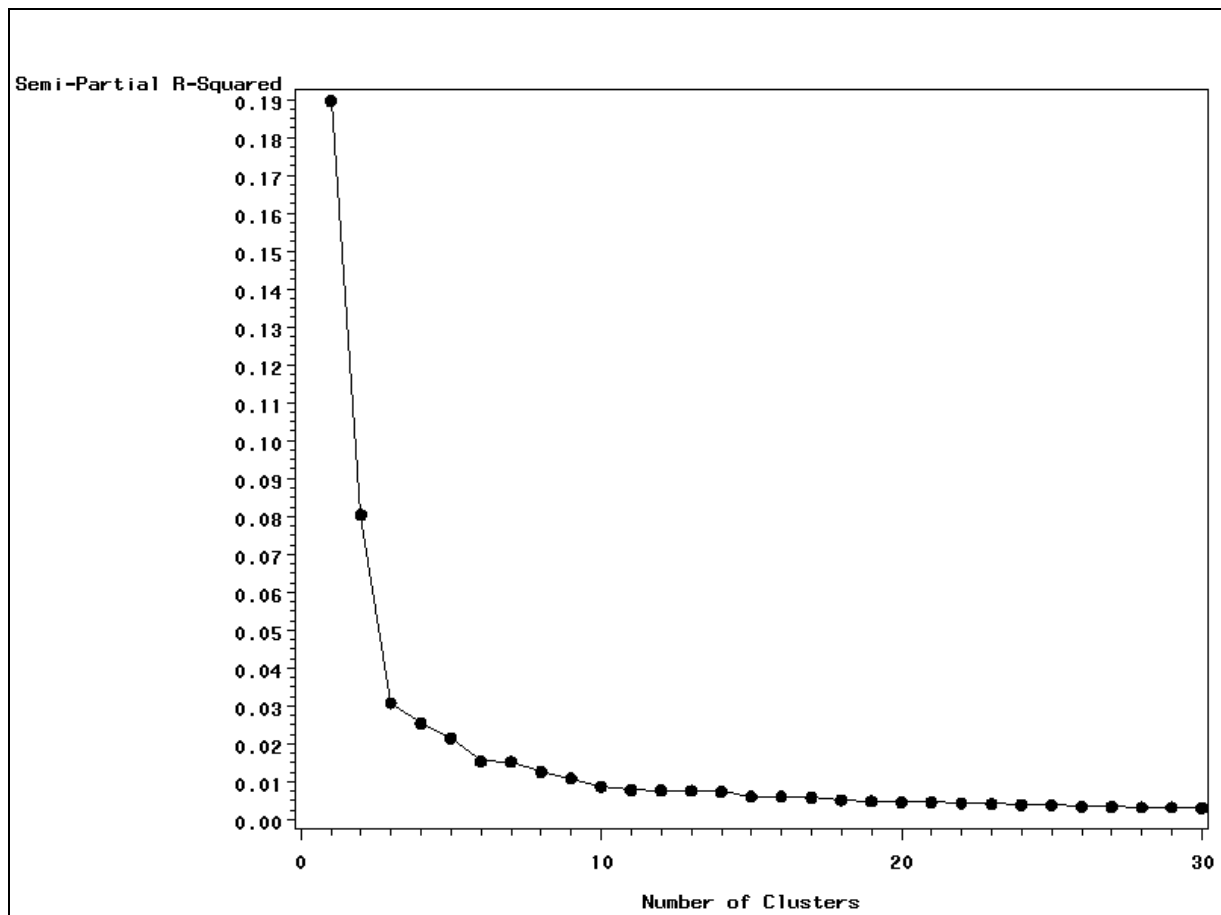


Nota bene: Die scheinbare Lücke im Dendrogramm stellt die 1288 identischen Blutungskalender dar (vgl. Tabelle 9).

Abbildung 9: Dendrogramm der Clusteranalyse nach Ward

4.4.2.1.1 Anzahl der Cluster

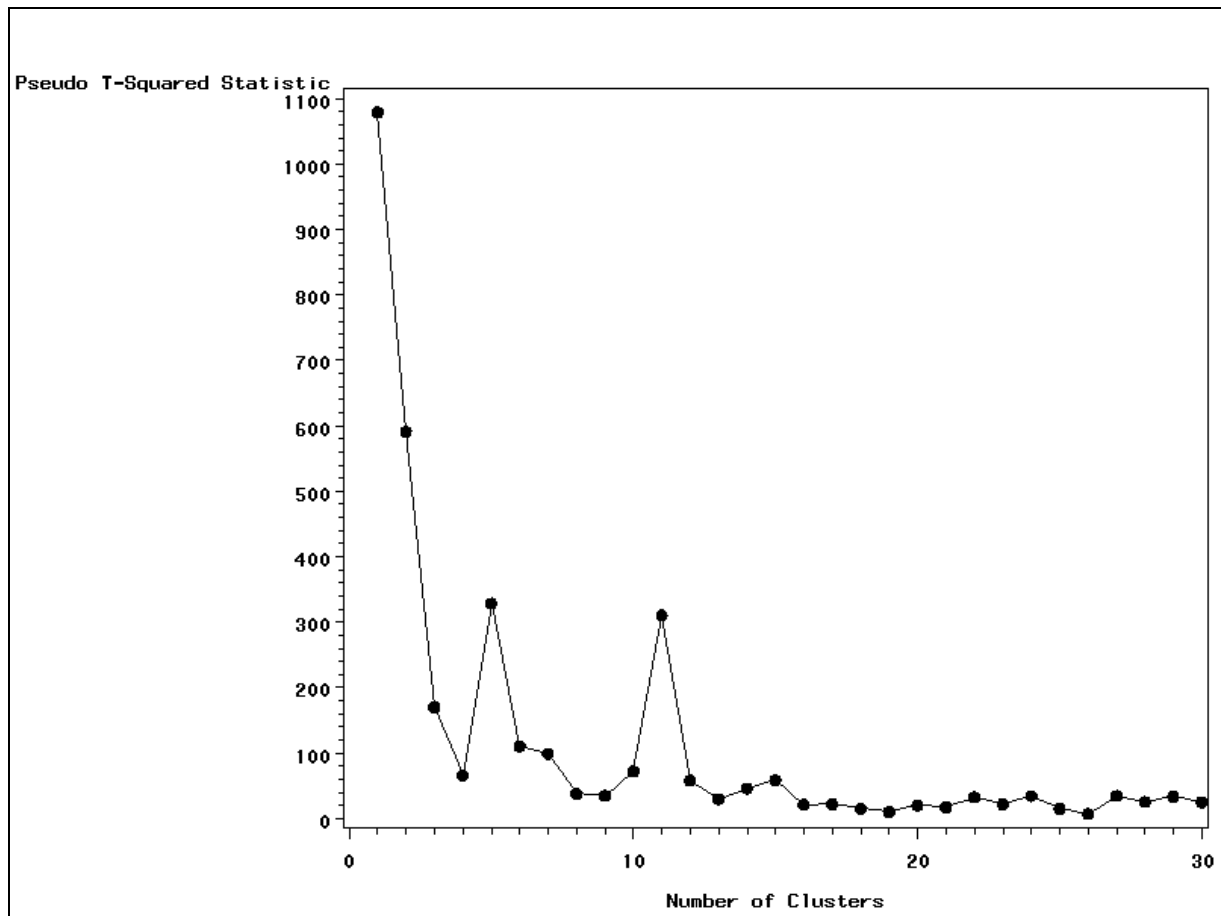
Zur Beurteilung der Frage, in wie viele Cluster der vorliegende Datensatz eingeteilt werden sollte, wurden das semi-partielle Bestimmtheitsmaß, das kubische Cluster Kriterium, die pseudo F- und die pseudo t-Statistik herangezogen (vgl. Abschnitt 4.3.5).



Nota: Die Anzahl der Cluster wurde nur bis 30 dargestellt.

Abbildung 10: Semipartielles Bestimmtheitsmaß über der Anzahl der Cluster

Die Graphik des semipartiellen Bestimmtheitsmaßes über die Anzahl der Cluster (Abbildung 10) zeigt einen streng monotonen Anstieg des semipartiellen Bestimmtheitsmaßes mit abnehmender Anzahl von Clustern. Der starke Anstieg des semipartiellen Bestimmtheitsmaßes bei weniger als drei Clustern, suggeriert eine Lösung mit drei Clustern. Die Lösung mit drei Clustern wird auch durch das Dendrogramm (Abbildung 9) nahe gelegt. Darüber hinaus ergibt sich jedoch kein Hinweis auf eine weitere spezielle Anzahl von Clustern durch das semipartielle Bestimmtheitsmaß.

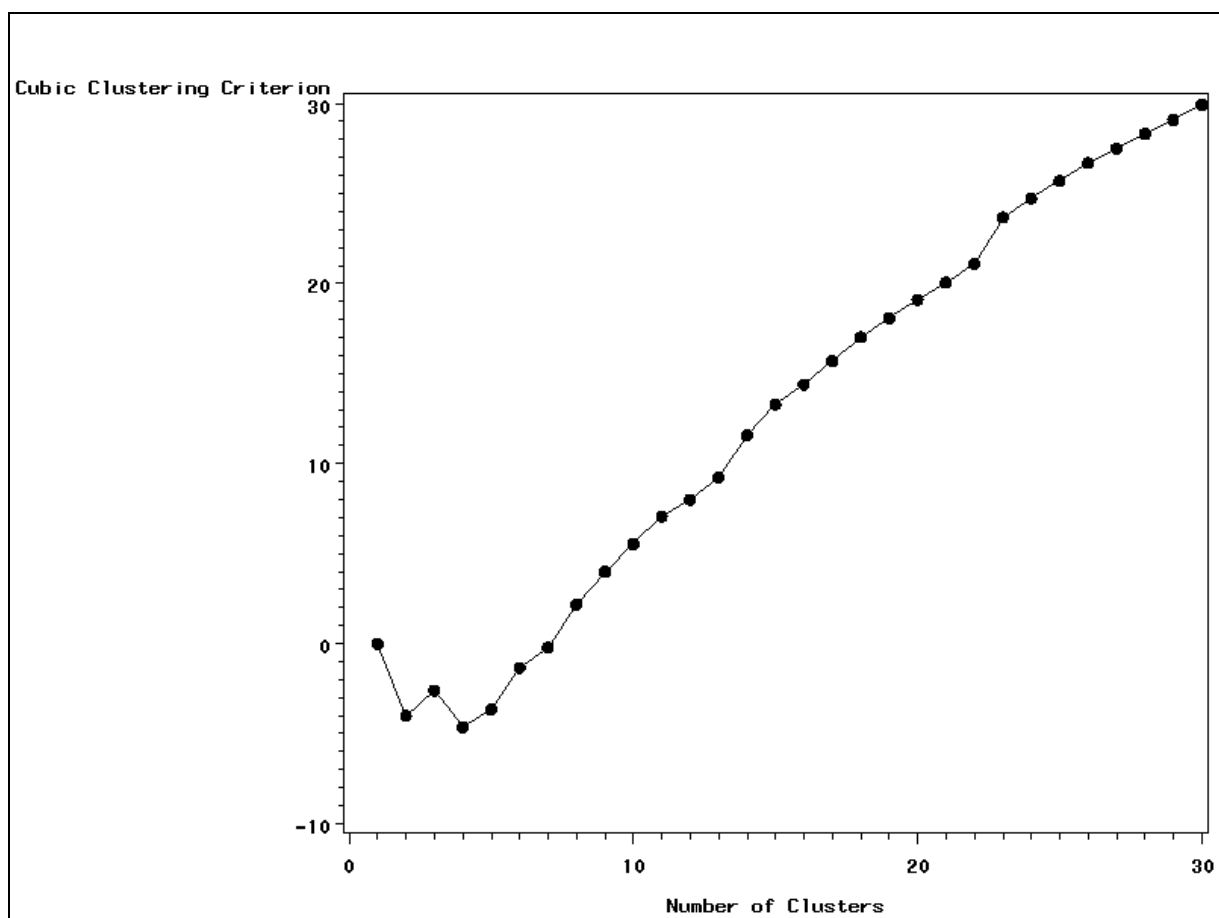


Nota: Die Anzahl der Cluster wurde nur bis 30 dargestellt.

Abbildung 11: Pseudo-t-Statistik über der Anzahl der Cluster

Die Graphik der pseudo t^2 -Statistik über die Anzahl der Cluster (Abbildung 11) zeigt deutliche lokale Maxima bei einem, fünf und elf Clustern. Da für die gesuchte Anzahl der Cluster der Wert unmittelbar vor dem starken Anstieg der pseudo- t^2 -Statistik maßgeblich ist (vgl. 4.3.5), deutet das Kriterium der pseudo- t^2 -Statistik auf Lösungen mit drei oder vier Clustern, sechs Clustern sowie mit zwölf Clustern hin.

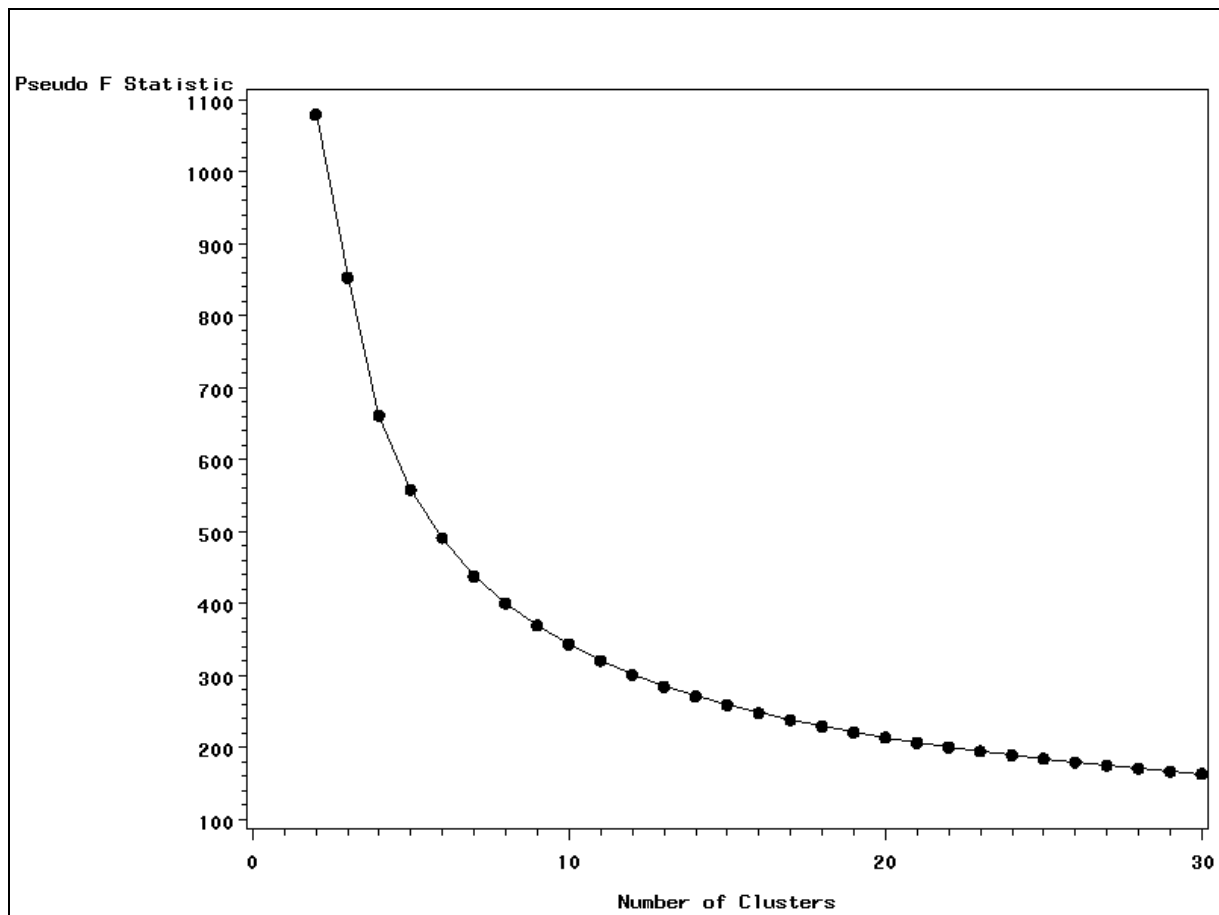
Der Vollständigkeit halber muss noch erwähnt werden, dass die pseudo- t^2 -Statistik für hohe Anzahlen von Clustern weitere lokale Maxima besitzt. Die Anzahl der Cluster ist jedoch so hoch, dass sie für die substanzwissenschaftliche, d.h. klinische, Betrachtung als wenig sinnvoll erscheint. Ziel dieser Arbeit ist es ja gerade, die immens große Anzahl der möglichen Blutungsmuster, wenn möglich, auf eine inhaltlich interpretierbare Anzahl von Clustern empirisch zusammenzufassen.



Nota: Die Anzahl der Cluster wurde nur bis 30 dargestellt.

Abbildung 12: Kubisches Cluster-Kriterium über der Anzahl der Cluster

Das kubische Cluster-Kriterium in Abbildung 12 gibt keinen klaren Hinweis auf die Anzahl der in den Daten enthaltenen Cluster. Zwischen einem und acht Clustern ist das kubische Cluster-Kriterium kleiner als zwei. Für größere Cluster-Anzahlen steigt es streng monoton an. Die Inkremente sind annähernd gleich, nur zwischen sieben und acht Clustern, zwischen 13 und 14 Clustern sowie zwischen 22 und 23 Clustern sind die Inkremente im Vergleich leicht höher.



Nota: Die Anzahl der Cluster wurde nur bis 30 dargestellt.

Abbildung 13: Pseudo-F-Statistik über der Anzahl der Cluster

Die Graphik der pseudo-F-Statistik über der Anzahl der Cluster (Abbildung 13) ergibt wie schon das kubische Cluster Kriterium keinen klaren Hinweis auf eine bestimmte

Anzahl von Clustern in den untersuchten Blutungsmustern. Die pseudo F-Statistik fällt anscheinend exponentiell mit steigender Anzahl der Cluster.

Zusammenfassend ist festzustellen, dass sich die Frage nach der Anzahl der in den untersuchten Blutungskalendern enthaltenen „wahren“ Cluster mit den zur Verfügung stehenden statistischen Kriterien nicht abschließend beantworten lässt. Vielmehr ist es erforderlich, diese Frage einer substanzwissenschaftlichen Betrachtung zu unterziehen.

Die mit den oben erläuterten statistischen Kriterien gefundenen möglichen Anzahlen von Clustern dienen hierbei als Ausgangspunkt für die weitergehende substanzwissenschaftliche Betrachtung. Im einzelnen sind dies die Lösungen mit drei, vier, sechs und mit zwölf Clustern, die im Folgenden näher untersucht werden sollen.

4.4.2.1.2 Lösung mit drei Clustern

Bei der Lösung mit drei Clustern ergeben sich Cluster, die 53,6, 35,1 und 11,3 Prozent der Beobachtungen umfassen (vgl. Tabelle 10). In zwei Clustern überwiegen die HRT-Fälle mit jeweils über 86 Prozent der Beobachtungen im jeweiligen Cluster. Der dritte Cluster enthält mit 97,5 Prozent seiner Beobachtungen fast ausschließlich Blutungsmuster aus OC Studien (vgl. Tabelle 11).

Tabelle 10: Anzahl der Elemente pro Cluster - Lösung mit drei Clustern

Clusternummer	Anzahl der Elemente	Prozent der Beobachtungen
1	2470	53,56
2	521	11,30
3	1621	35,15

Tabelle 11: Anzahl der Elemente pro Cluster nach Indikation - Lösung mit drei Clustern

Clusternummer	Indikation					
	Endometriose		HRT		OC	
	N	%	N	%	N	%
1	182	7,37	2129	86,19	159	6,44
2	20	3,84	452	86,76	49	9,40
3	24	1,48	16	0,99	1581	97,53

Die inhaltliche Interpretation der Cluster folgt im Abschnitt 4.5.1.

4.4.2.1.3 Lösung mit vier Clustern

Bei der Lösung mit vier Clustern entstehen Cluster mit 53,6, 26,8, 11,3 und 8,4 Prozent der Beobachtungen (vgl. Tabelle 12). Zwei Cluster enthalten jeweils zu über 86 Prozent ihrer Beobachtungen Blutungsmuster aus HRT Studien. Ein Cluster enthält mit 99,0 Prozent seiner Beobachtungen fast ausschließlich OC Blutungsmuster. Der vierte Cluster enthält mit 92,7 Prozent seiner Beobachtungen im wesentlichen OC Blutungsmuster (vgl. Tabelle 13).

Tabelle 12: Anzahl der Elemente pro Cluster - Lösung mit vier Clustern

Clusternummer	Anzahl der Elemente	Prozent der Beobachtungen
1	2470	53,56
2	521	11,30
3	1235	26,78
4	386	8,37

Tabelle 13: Anzahl der Elemente pro Cluster nach Indikation - Lösung mit vier Clustern

Clusternummer	Indikation					
	Endometriose		HRT		OC	
	N	%	N	%	N	%
1	182	7,37	2129	86,19	159	6,44
2	20	3,84	452	86,76	49	9,40
3	5	0,40	7	0,57	1223	99,03
4	19	4,92	9	2,33	358	92,75

Die inhaltliche Interpretation der Cluster folgt im Abschnitt 4.5.2.

4.4.2.1.4 Lösung mit sechs Clustern

Bei einer Lösung mit sechs Clustern reichen die Größen der einzelnen Cluster von 1,5 bis 40,8 Prozent der Beobachtungen (vgl. Tabelle 14). Drei der sechs Cluster bestehen überwiegend aus Blutungskalendern aus HRT Studien mit Anteilen von 84,2, 97,5 und 100 Prozent der Beobachtungen im jeweiligen Cluster. Zwei der sechs Cluster enthalten überwiegend Blutungskalender aus OC Studien. Ihre Anteile an der Gesamtzahl der Elemente im jeweiligen Cluster waren 92,7 und 99,0 Prozent. Ein Cluster (Nummer 2) lässt sich keiner Indikation zuordnen, er enthält zu 29,7 Prozent Blutungsmuster aus Studien der Endometriose, zu 50,2 Prozent Blutungsmuster aus HRT Studien und zu 20,2 Prozent Blutungsmuster aus OC Studien (vgl. Tabelle 15).

Tabelle 14: Anzahl der Elemente pro Cluster - Lösung mit sechs Clustern

Clusternummer	Anzahl der Elemente	Prozent der Beobachtungen
1	1880	40,76
2	590	12,79
3	450	9,76
4	1235	26,78
5	71	1,54
6	386	8,37

Tabelle 15: Anzahl der Elemente pro Cluster nach Indikation - Lösung mit sechs Clustern

Clusternummer	Indikation					
	Endometriose		HRT		OC	
	N	%	N	%	N	%
1	7	0,37	1833	97,50	40	2,13
2	175	29,66	296	50,17	119	20,17
3	20	4,44	381	84,67	49	10,89
4	5	0,40	7	0,57	1223	99,03
5	0	-	71	100,00	0	-
6	19	4,92	9	2,33	358	92,75

Die inhaltliche Interpretation der Cluster folgt im Abschnitt 4.5.3.

4.4.2.1.5 Lösung mit zwölf Clustern

Die Lösung mit zwölf Clustern führt im Vergleich zur Lösung mit sechs Clustern nicht zur Bildung von Clustern mit sehr geringem Umfang. Die Anzahl der Elemente pro Cluster reicht von 1,5 bis 36,7 Prozent der Beobachtungen (vgl. Tabelle 16). Der größte Cluster mit 1701 Elementen, d.h. 36,9 Prozent der Beobachtungen, umfasst dabei die 1288 identischen Blutungskalender (vgl. Tabelle 9).

Bei der Betrachtung der Cluster nach Indikationen in Tabelle 17 zeigt sich, dass elf der zwölf Cluster überwiegend Blutungskalender aus einer der drei untersuchten

Indikationen enthalten. Sieben Cluster (Nummern 1, 2, 4, 5, und 8 bis 10) enthalten hauptsächlich Blutungskalender von Frauen, die Hormonersatzpräparate eingenommen haben. Ihre Anteile am Umfang dieser Cluster reichen von 75,1 Prozent der Beobachtungen in Cluster 5 bis 100 Prozent der Beobachtungen in Cluster 10. Bis auf Cluster Nummer 5 enthalten alle HRT-dominierten Cluster zu mehr als 86 Prozent Beobachtungen aus Studien von Hormonersatzpräparaten.

Vier Cluster (Nummern 6, 7, 11 und 12) beschreiben fast ausschließlich Blutungsmuster von Frauen, die orale Kontrazeptiva einnehmen. Die Anteile am Umfang dieser Cluster reichen von 92,7 Prozent der Beobachtungen in Cluster 12 bis 99,2 Prozent der Beobachtungen in Cluster 6.

Cluster Nummer 3 enthält zu 48,4 Prozent Blutungskalender aus Studien zur Endometriose, zu 33,9 Prozent Blutungskalender aus Studien von oralen Kontrazeptiva und zu 17,7 Prozent Blutungskalender aus Studien von Hormonersatzpräparaten.

Tabelle 16: Anzahl der Elemente pro Cluster - Lösung mit zwölf Clustern

Clusternummer	Anzahl der Elemente	Prozent der Beobachtungen
1	1701	36,88
2	280	6,07
3	310	6,72
4	179	3,88
5	241	5,23
6	526	11,41
7	511	11,08
8	116	2,52
9	93	2,02
10	71	1,54
11	198	4,29
12	386	8,37

Tabelle 17: Anzahl der Elemente pro Cluster nach Indikation - Lösung mit zwölf Clustern

Clusternummer	Indikation					
	Endometriose		HRT		OC	
	N	%	N	%	N	%
1	6	0,35	1679	98,71	16	0,94
2	25	8,93	241	86,07	14	5,00
3	150	48,39	55	17,74	105	33,87
4	1	0,56	154	86,03	24	13,41
5	17	7,05	181	75,10	43	17,84
6	0	-	5	0,95	521	99,05
7	2	0,39	2	0,39	507	99,22
8	1	0,86	113	97,41	2	1,72
9	2	2,15	87	93,55	4	4,30
10	0	-	71	100,00	0	-
11	3	1,52	0	-	195	98,48
12	19	4,92	9	2,33	358	92,75

Die inhaltliche Interpretation der Cluster folgt im Abschnitt 4.5.4.

4.4.2.2 Ergebnisse der anderen Cluster Methoden

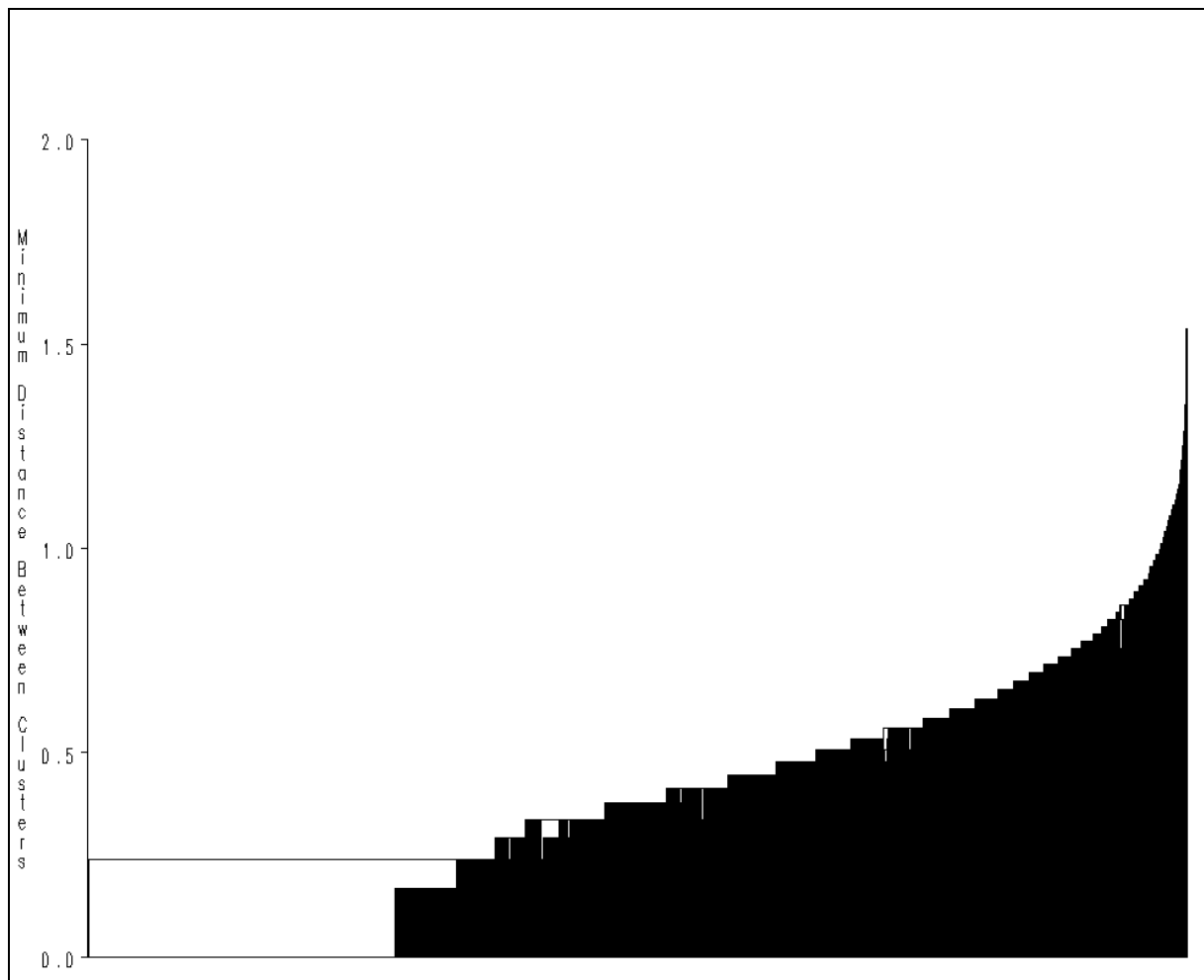
Die drei anderen verwandten Methoden führten alle zu keiner Einteilung der vorliegenden Blutungskalender in substanzwissenschaftlich interpretierbare Cluster.

4.4.2.2.1 Single Linkage

Der Single-Linkage-Algorithmus fasst wie erwartet zunächst die 1288 identischen Blutungsmuster zum ersten Cluster zusammen. Danach kommt es jedoch zu dem als „Chaining“ bekannten Effekt, dass beim single linkage Verfahren ein einziger großer Cluster schrittweise die weiteren Datenpunkte absorbiert, ohne dass es zur Ausbildung von mehreren größeren Clustern kommt (vgl. Abbildung 14).

Dies hat beim untersuchten Datensatz zur Folge, dass beispielsweise die letzten 25 Cluster aus einem Cluster mit 4588 Elementen und 24 Clustern mit je einem einzigen Element bestehen. Die letzten zehn Cluster bestehen aus einem Cluster mit 4602 Elementen, einem Clustern mit zwei Elementen sowie acht Clustern mit je einem Element.

Da der Single-Linkage-Algorithmus den untersuchten Datensatz im Wesentlichen in nur einen einzigen Cluster unterteilt, sind die Resultate dieses Verfahrens ungeeignet, zur Fragestellung dieser Arbeit beizutragen. Auf eine Tabelle der Cluster wird daher verzichtet.

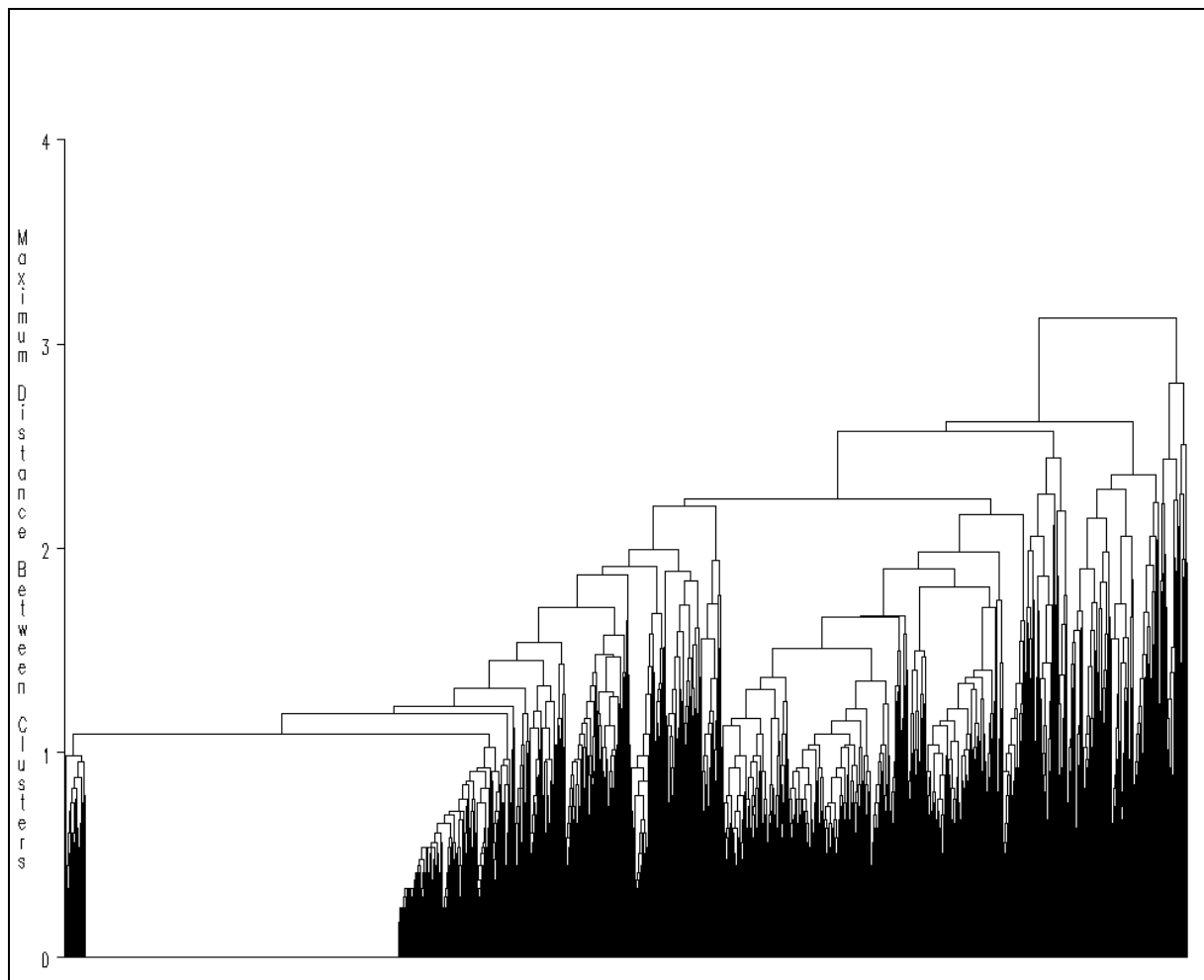


Nota bene: Die scheinbare Lücke im Dendrogramm stellt die 1288 identischen Blutungskalender dar (vgl. Tabelle 9).

Abbildung 14: Dendrogramm der Clusteranalyse - Single-Linkage-Algorithmus

4.4.2.2.2 Complete Linkage

Der Complete-Linkage-Algorithmus fasst wie erwartet zunächst die gleichen Beobachtungen zu Clustern zusammen. Ähnlich wie beim Single-Linkage-Algorithmus wird der Datensatz jedoch im Wesentlichen zu einem einzigen Cluster zusammengefasst (vgl. Abbildung 15 und Tabelle 18).



Nota bene: Die scheinbare Lücke im Dendrogramm stellt die 1288 identischen Blutungskalender dar (vgl. Tabelle 9).

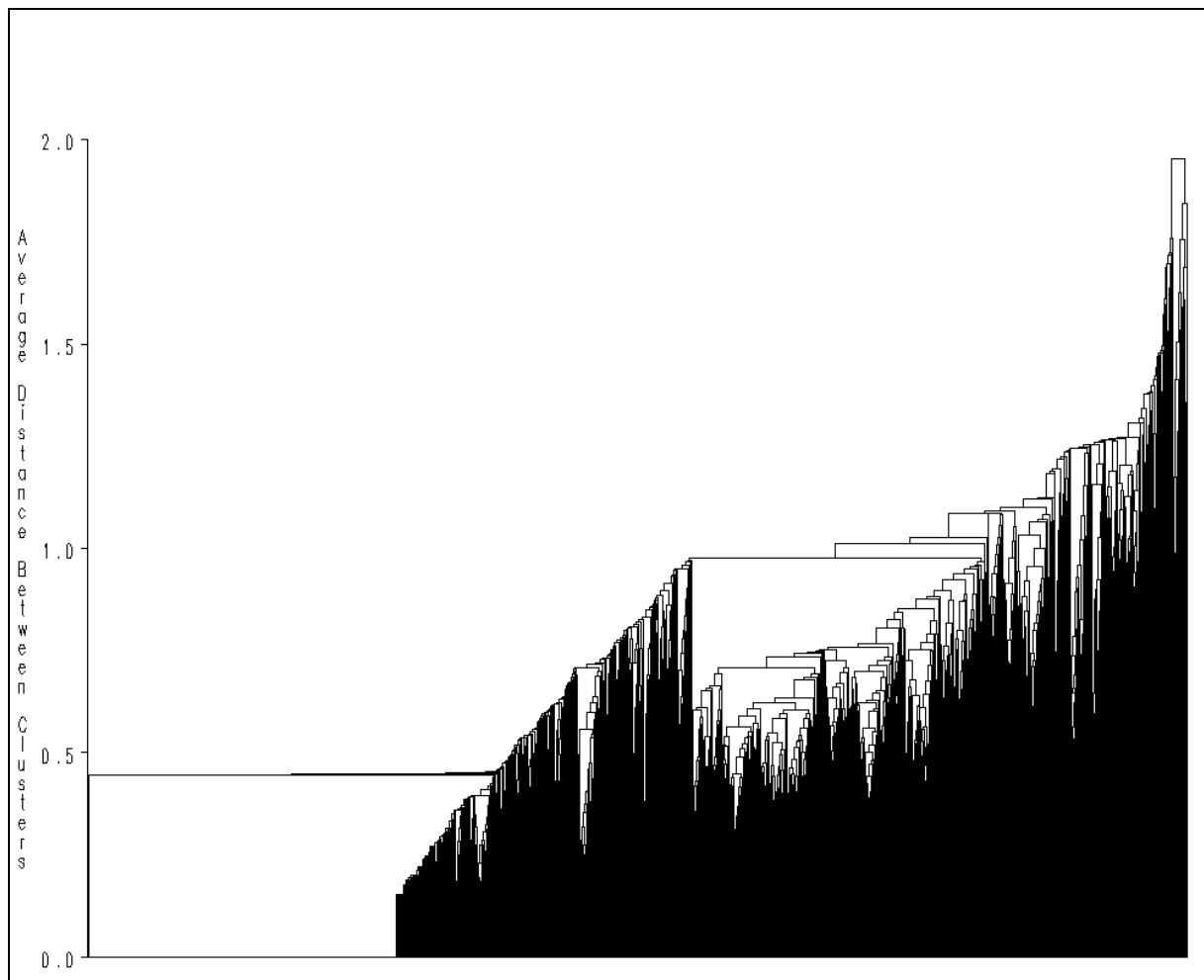
Abbildung 15 Dendrogramm der Clusteranalyse - Complete-Linkage-Algorithmus

Tabelle 18: Anzahl der Elemente pro Cluster - Complete-Linkage-Algorithmus

Clusternummer	Anzahl der Elemente	Prozent der Beobachtungen
1	3944	85.52
2	181	3.92
3	105	2.28
4	125	2.71
5	93	2.02
6	32	0.69
7	47	1.02
8	46	1.00
9	30	0.65
10	9	0.20

4.4.2.2.3 Average Linkage

Auch der Average-Linkage-Algorithmus erzeugt nur einen einzigen großen Cluster und war somit zur Unterteilung der verschiedenen Blutungsmuster ebenfalls ungeeignet (vgl. Abbildung 16 und Tabelle 19).



Nota bene: Die scheinbare Lücke im Dendrogramm stellt die 1288 identischen Blutungskalender dar (vgl. Tabelle 9).

Abbildung 16: Dendrogramm der Clusteranalyse - Average-Linkage-Algorithmus

Tabelle 19: Anzahl der Elemente pro Cluster - Average-Linkage-Algorithmus

Clusternummer	Anzahl der Elemente	Prozent der Beobachtungen
1	4525	98.11
2	40	0.87
3	18	0.39
4	18	0.39
5	6	0.13
6	1	0.02
7	1	0.02
8	1	0.02
9	1	0.02
10	1	0.02

4.5 Inhaltliche Interpretation der Ergebnisse

Von den vier untersuchten Methoden zur Clusteranalyse lassen sich drei Methoden schon anhand ihrer Dendrogramme als nicht substanzwissenschaftlich interpretierbar aussortieren.

Dies sind die Single-Linkage-Methode (vgl. Abbildung 14), die Complete-Linkage-Methode (vgl. Abbildung 15 und Tabelle 18) sowie die Average-Linkage-Methode (vgl. Abbildung 16 und Tabelle 19). Alle drei genannten Verfahren führen zur Bildung eines einzigen großen Clusters, das fast alle Blutungsmuster enthält. Es ist

offensichtlich, dass sich, wie zum Beispiel bei der Average-Linkage-Methode, mit einer Klassifizierung der Blutungsmuster in ein Cluster mit mehr als 98 Prozent der Beobachtungen und in weitere Cluster mit unter 1 Prozent der Beobachtungen, wenig zur Fragestellung dieser Arbeit beitragen lässt. Daher wird im Folgenden nur noch auf die Ergebnisse der Methode von Ward eingegangen.

Die Methode von Ward unterteilt die untersuchten Blutungskalender in mehrere Cluster ähnlicher Größe, wobei bei dieser Aussage der größte Cluster ausgenommen werden muss. Dies liegt daran, dass es in dem vorliegenden Datensatz 1288 identische Blutungskalender gibt (vgl. Tabelle 9), in denen keine einzige Blutung vorkommt. Diese Blutungsmuster alleine repräsentiert bereits 27,9 Prozent der 4612 in die Untersuchung einbezogenen Blutungskalender, so dass unabhängig von der Anzahl der betrachteten Cluster ein Cluster auf jeden Fall mindestens 27,9 Prozent der Beobachtungen umfassen muss.

4.5.1 Lösung mit drei Clustern

Die Lösung mit drei Clustern zeigt eine sehr deutlich Trennung der Blutungskalender (vgl. Die Länge des ersten Zyklus in Cluster 3 ist deutlich kürzer als die der darauffolgenden Zyklen. Dies liegt jedoch daran, dass in den vorliegenden Studien von oralen Kontrazeptiva der erste Tag der Pilleneinnahme gemäß Prüfplan am ersten Tag einer Menstruationsblutung erfolgen sollte. Der Hormonentzug durch die Pilleneinnahmepause, der im Allgemeinen die Abbruchblutung auslöst, ist im ersten Zyklus daher bereits nach 21 Tagen, was zu einer verkürzten Dauer des ersten Pillenzyklus führt.

) in zwei HRT und einen OC dominierten Cluster. Cluster Nummer 1 und Nummer 2 enthalten jeweils zu mehr als 86 Prozent der Beobachtungen Blutungskalender aus Studien eines HRT-Präparates, während Cluster 3 zu über 97 Prozent der Beobachtungen Blutungskalender von Frauen, die ein orales Kontrazeptivum einnehmen, enthält (vgl. Tabelle 11).

Cluster Nummer 1 beschreibt ein Blutungsmuster, in dem fast nicht geblutet wird und bei dem keine Periodizität erkennbar ist. Dieses Blutungsmuster entspricht dem typischen erwünschten Blutungsmuster unter der Einnahme von kontinuierlichen Präparaten zur Hormonersatztherapie. Bei diesen Präparaten soll es ja wie bei unbehandelten Frauen nach der Menopause zu keinen Blutungen kommen.

Der Cluster Nummer 2 zeigt dagegen ein unbefriedigendes Blutungsmuster von Frauen, die überwiegend ein HRT-Präparat einnehmen. Bei diesem Blutungsmuster kommt es weder zur Amenorrhö noch zu einem zyklischen Blutungsverlauf.

Der letzte der drei Cluster zeigt ein zyklisches Blutungsmuster mit einer Periodizität von 28 Tagen ab dem zweiten Zyklus. Die 28 Tage entsprechen genau dem Einnahmeschema von klassischen oralen Kontrazeptiva, bei denen 21 Tage lang täglich Hormone genommen werden, denen anschließend eine 7-tägige Einnahmepause (bzw. eine Placeboeinnahme) folgt.

Die Länge des ersten Zyklus in Cluster 3 ist deutlich kürzer als die der darauffolgenden Zyklen. Dies liegt jedoch daran, dass in den vorliegenden Studien von oralen Kontrazeptiva der erste Tag der Pilleneinnahme gemäß Prüfplan am ersten Tag einer Menstruationsblutung erfolgen sollte. Der Hormonentzug durch die Pilleneinnahmepause, der im Allgemeinen die Abbruchblutung auslöst, ist im ersten Zyklus daher bereits nach 21 Tagen, was zu einer verkürzten Dauer des ersten Pillenzyklus führt.

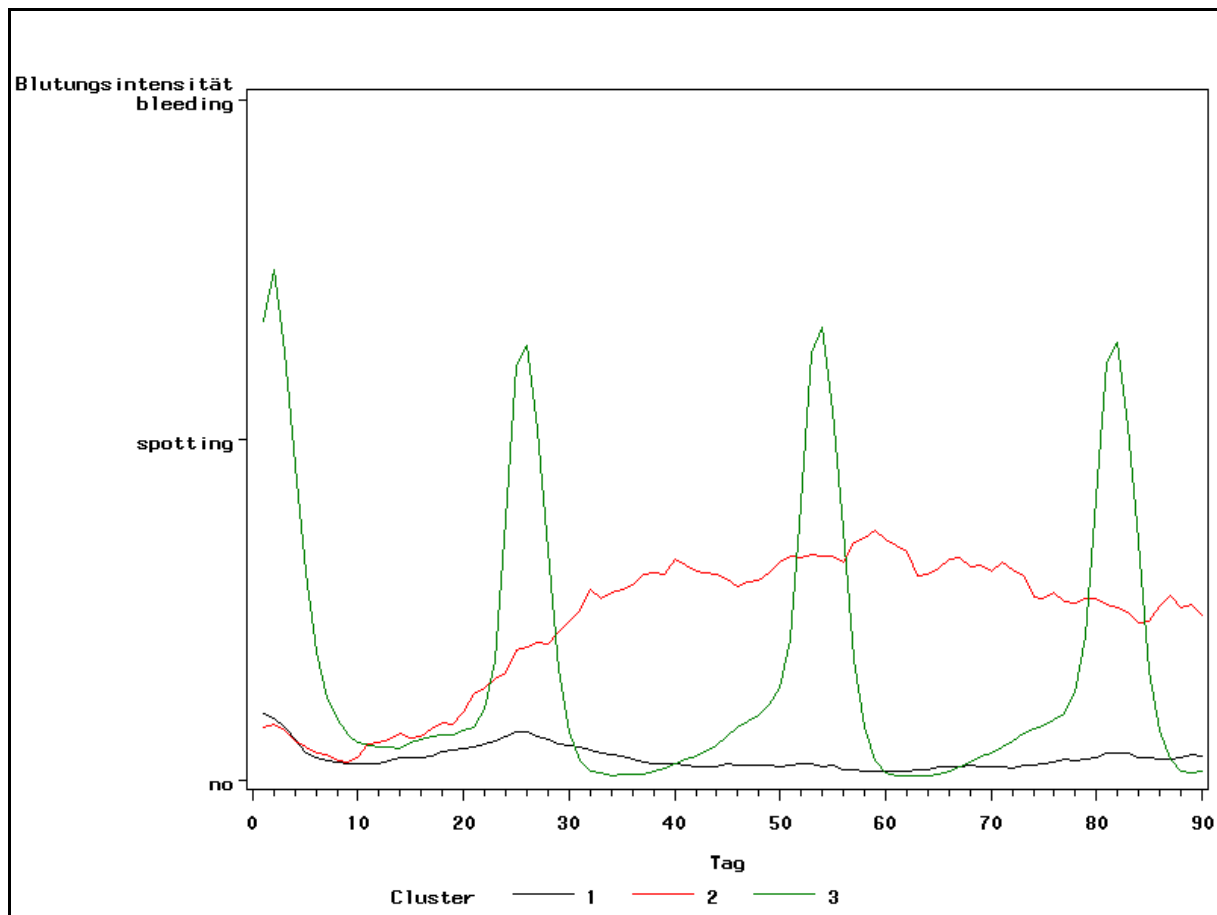


Abbildung 17: Mittlere Blutungsintensität über die Zeit - drei Cluster

Zusammenfassend lässt sich sagen, dass die Lösung mit drei Clustern substanzwissenschaftlich gesehen viel zu grob ist. Es werden lediglich zwei Blutungsmuster, die vorwiegend unter der Einnahme von Präparaten zur Hormonersatztherapie auftreten, und ein Blutungsmuster, das fast ausschließlich unter der Einnahme von oralen Kontrazeptiva vorkommt, unterschieden. Eine derartige Unterscheidung ist aber für die Beurteilung von Präparaten, die das Blutungsverhalten beeinflussen, wenig hilfreich, da hier im Wesentlichen nur die verschiedenen Einnahmeregime, kontinuierlich bei HRT und sequenziell bei OCs, unterschieden werden.

4.5.2 Lösung mit vier Clustern

Die Lösung mit vier Clustern (vgl. Abbildung 18) unterscheidet sich von der Lösung mit drei Clustern dadurch, dass nicht nur ein Cluster mit überwiegend OC Blutungskalendern gebildet wird. Es entstehen bei dieser Lösung zwei Cluster, die mit jeweils über 86 Prozent der Beobachtungen überwiegend Blutungskalender aus Studien zur Hormonersatztherapie enthalten, und zwei Cluster, die mit jeweils über 92 Prozent der Beobachtungen überwiegend Blutungskalender aus Studien zur hormonellen Empfängnisverhütung enthalten (vgl. Tabelle 13).

Wie im Falle der Lösung mit drei Clustern beschreibt Cluster Nummer 1 ein Blutungsmuster, in dem fast nicht geblutet wird und bei dem keine Periodizität erkennbar ist. Dieses Blutungsmuster entspricht dem typischen erwünschten Blutungsmuster unter der Einnahme von kontinuierlichen Präparaten zur Hormonersatztherapie. Der Cluster Nummer 2 zeigt dagegen ein unbefriedigendes Blutungsmuster von Frauen, die überwiegend ein HRT-Präparat einnehmen. Bei diesem Blutungsmuster kommt es weder zur Amenorrhö noch zu einem zyklischen Blutungsverlauf.

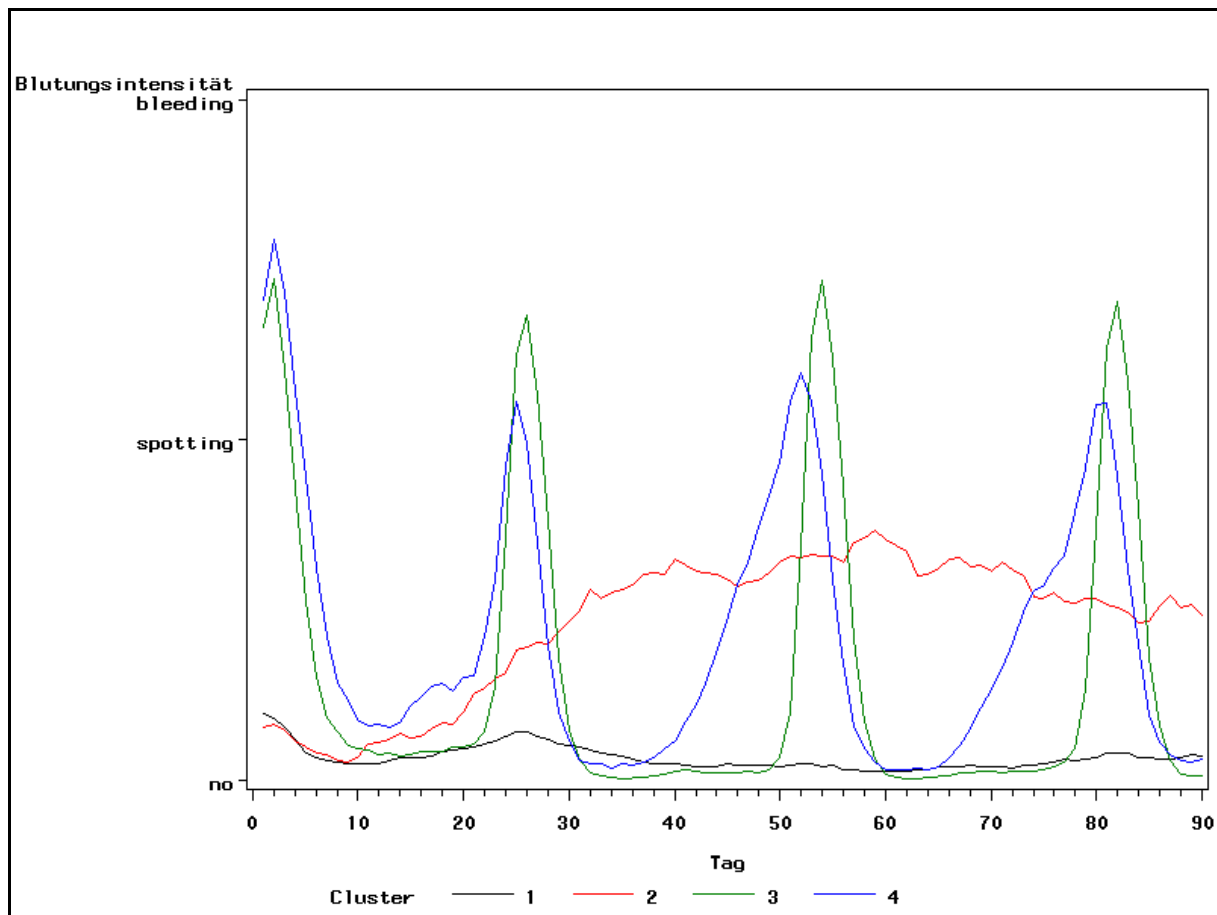


Abbildung 18: Mittlere Blutungsintensität über die Zeit - vier Cluster

Cluster Nummer 3 zeigt ein sehr deutlich zyklisches Blutungsmuster mit ausgeprägten Spitzen in den Pilleneinnahmefreien Zeiten. In diesem Cluster wurden die Blutungskalender zusammengefasst, die ein für sequenzielle Hormonpräparate erwünschtes Blutungsmuster zeigen. Insbesondere ist die mittlere Blutungsintensität während der Intervalle mit Pilleneinnahme sehr gering.

Im Gegensatz dazu zeigt der Cluster Nummer 4 ein eher unerwünschtes Blutungsmuster für ein sequenziell eingenommenes Hormonpräparat. Es kommt zwar wie in Cluster Nummer 3 zu einem ausgeprägten zyklischen Blutungsverhalten, jedoch steigt die mittlere Blutungsintensität schon bald nach Beginn eines

Pillenzyklus steil an. Dies bedeutet, dass ein nicht unerheblicher Teil der Frauen in diesem Cluster unerwünschte Zwischenblutungen hat.

Die Lösung der Clusteranalyse mit vier Clustern ermöglicht eine erste substanzwissenschaftlich akzeptable Einteilung der Blutungsmuster. Es werden zyklische (Cluster Nummer 3 und Cluster Nummer 4) und nichtzyklische (Cluster Nummer 1 und Cluster Nummer 2) Blutungsmuster klar unterschieden. Zusätzlich werden innerhalb der beiden zyklischen Blutungsmuster das erwünschte Blutungsmuster ohne Zwischenblutungen (Cluster Nummer 3) vom unerwünschten Blutungsmuster mit Zwischenblutungen (Cluster Nummer 4) aufgezeigt und deutlich voneinander getrennt.

Für die Fragestellung dieser Arbeit ist eine Einteilung der Blutungsmuster in die vier Kategorien „zyklisch, gute Blutungskontrolle“, „zyklisch, schlechte Blutungskontrolle“, „kontinuierlich, gute Blutungskontrolle“ und „kontinuierlich, schlechte Blutungskontrolle“ natürlich zu grob, um damit das Blutungsverhalten unter der Einnahme einzelner Präparate sinnvoll vergleichen zu können. Dies liegt daran, dass kleine, aber für die Anwenderin dennoch bedeutsame, Unterschiede zwischen Präparaten durch eine zu grobe Einteilung der Blutungsmusterkategorien nicht richtig erkannt werden können.

4.5.3 Lösung mit sechs Clustern

Die Lösung mit sechs Clustern (vgl. Abbildung 19) enthält mit Cluster Nummer 4 und Cluster Nummer 6 zwei Cluster mit einem zyklischen Blutungsmuster, die den Clustern Nummer 3 und 4 aus der Lösung mit vier Clustern entsprechen.

Die beiden Cluster mit einem sequenziellen Blutungsmuster aus der Lösung mit vier Clustern (Cluster Nummer 1 und Cluster Nummer 2 in Abbildung 18) erscheinen bei der Lösung mit sechs Clustern nicht mehr. Statt dessen ergeben sich drei Cluster, die sich überwiegend aus Blutungskalendern aus Studien von Präparaten zur Hormonersatztherapie zusammensetzen (Cluster Nummer 1, Cluster Nummer 3 und

Cluster Nummer 5) und ein Cluster (Cluster Nummer 2), das sich zu 50,2 Prozent der Beobachtungen aus Blutungskalendern unter der Einnahme von HRT-Präparaten, 29,7 Prozent aus Blutungskalendern unter der Einnahme von Präparaten zur Behandlung von Endometriose und zu 20,2 Prozent aus Blutungskalendern unter der Einnahme von oralen Kontrazeptiva zusammensetzt (vgl. Tabelle 15). Da das Einnahmeschema für Präparate zur sequenziellen Hormonersatztherapie und für Präparate zur Behandlung der Endometriose gleich ist, kann Cluster Nummer 2 jedoch durchaus zusammen mit den drei Clustern, die sich überwiegend aus Blutungskalendern unter der Einnahme von HRT-Präparaten zusammensetzen, behandelt werden.

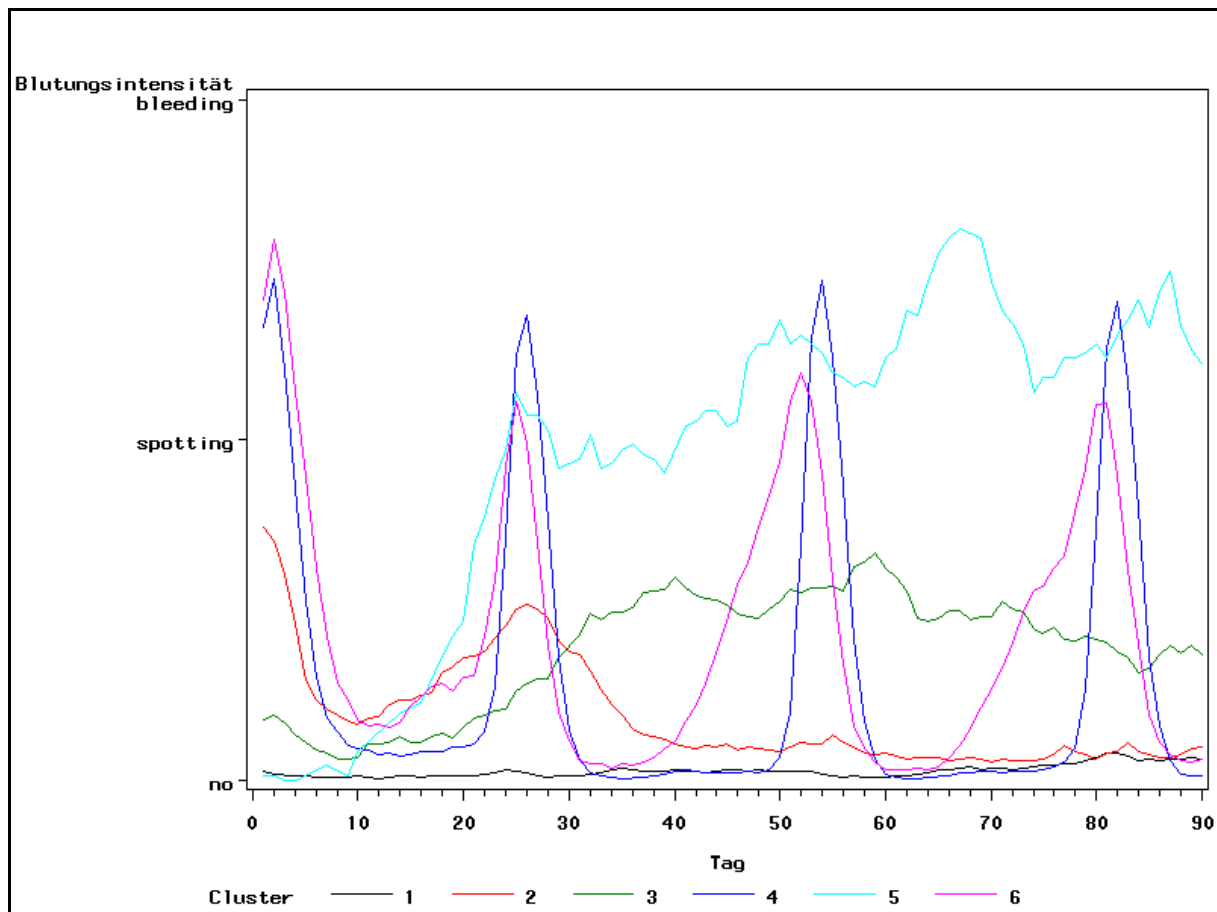


Abbildung 19: Mittlere Blutungsintensität über die Zeit - sechs Cluster

Die vier Cluster lassen sich in einen Cluster, der ein erwünschtes Blutungsmuster für kontinuierliche Präparate (Cluster Nummer 1) und drei Cluster (Cluster Nummer 2, Cluster Nummer 3 und Cluster Nummer 5), die ein für ein solches Präparat unerwünschte Blutungsmuster zeigen, einteilen. Hierbei zeigt Cluster Nummer 3 ein besseres Blutungsverhalten als Cluster Nummer 5, aber beiden Clustern ist gemein, dass die mittlere Blutungsintensität nach den ersten drei Wochen sehr deutlich ansteigt und sich dann auf hohem Niveau einpendelt.

Dem hingegen zeigt Cluster Nummer 2 ein deutlich anderes Blutungsmuster. Die mittlere Blutungsintensität ist zu Beginn der Behandlung erhöht, sinkt dann rasch ab, um nach etwa vierwöchiger Behandlung einen deutlichen Peak zu haben. Danach klingt die mittlere Blutungsintensität ab und erreicht zum Ende des Beobachtungszeitraumes von 90 Tagen das Niveau des Clusters Nummer 1. Ein Grund für ein derartiges Blutungsmuster liegt sicherlich im Anteil von 29,7 Prozent Blutungskalendern unter der Einnahme von Präparaten zur Behandlung von Endometriose. Die Patientinnen mit Endometriose sind noch im geschlechtsreifen Alter und ihr natürliches zyklisches Blutungsmuster wird durch die kontinuierliche Einnahme von Hormonpräparaten unterdrückt. Es ist daher nicht weiter verwunderlich, dass es bei diese Patientinnen zu dem beobachteten Peak nach etwa einem Monat nach dem Beginn der Behandlung kommt.

Ein derartiges Blutungsmuster ist zwar im Prinzip unerwünscht, es ist jedoch deutlich besser als die Blutungsmuster in Cluster Nummer 3 und Cluster Nummer 5, da die unerwünschten Blutungen unter der Einnahme eines sequenziellen Präparates nur vorübergehender Natur sind.

Für die Fragestellung dieser Arbeit sind die vier verschiedenen Blutungsmuster unter der Einnahme von sequenziellen Präparate n bereits ausreichend. Für die zwei Blutungsmuster unter der Einnahme von zyklischen Präparaten zur Geburtenkontrolle kann dies jedoch nicht unbedingt gesagt werden.

4.5.4 Lösung mit zwölf Clustern

Als weitere Lösung der Clusteranalyse wurde in Abschnitt 4.4.2.1 die Lösung mit zwölf Clustern erkannt. Diese Lösung ist in Abbildung 20 dargestellt und zeigt deutlich zyklische und nicht zyklische Blutungsmuster.

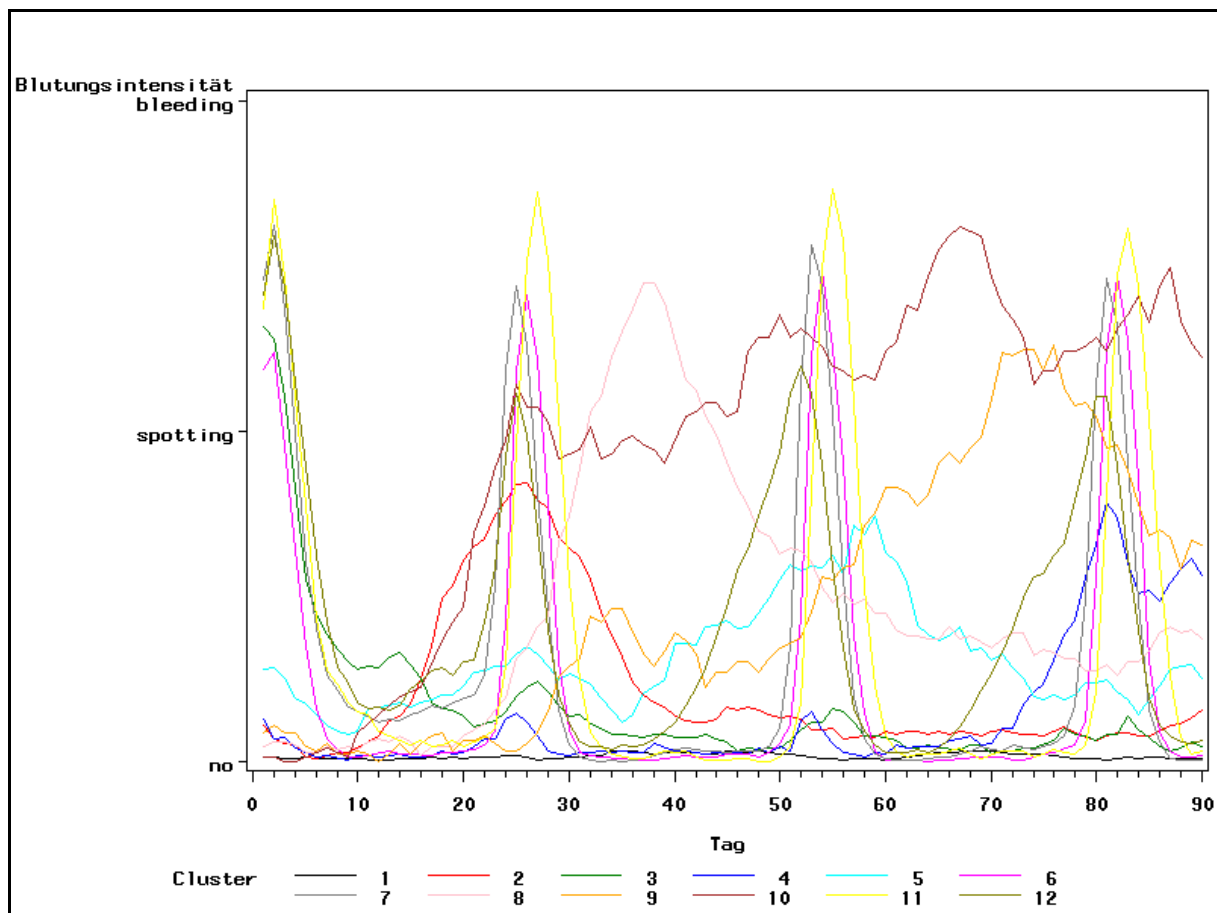


Abbildung 20: Mittlere Blutungsintensität über die Zeit - zwölf Cluster

Sieben dieser zwölf Cluster (Cluster Nummer 1, Cluster Nummer 2, Cluster Nummer 4, Cluster Nummer 5, Cluster Nummer 8, Cluster Nummer 9 und Cluster Nummer 10) enthalten, bis auf eine Ausnahme mit 75,1 Prozent der Beobachtungen, zu mehr als 86 Prozent Beobachtungen aus Studien von Hormonersatzpräparaten (vgl. Tabelle 17). Cluster Nummer 10 enthält sogar ausschließlich Beobachtungen von Frauen, die

ein Hormonersatzpräparat eingenommen haben. Diese sieben Cluster sind in Abbildung 21 der Übersichtlichkeit halber nochmals einzeln dargestellt.

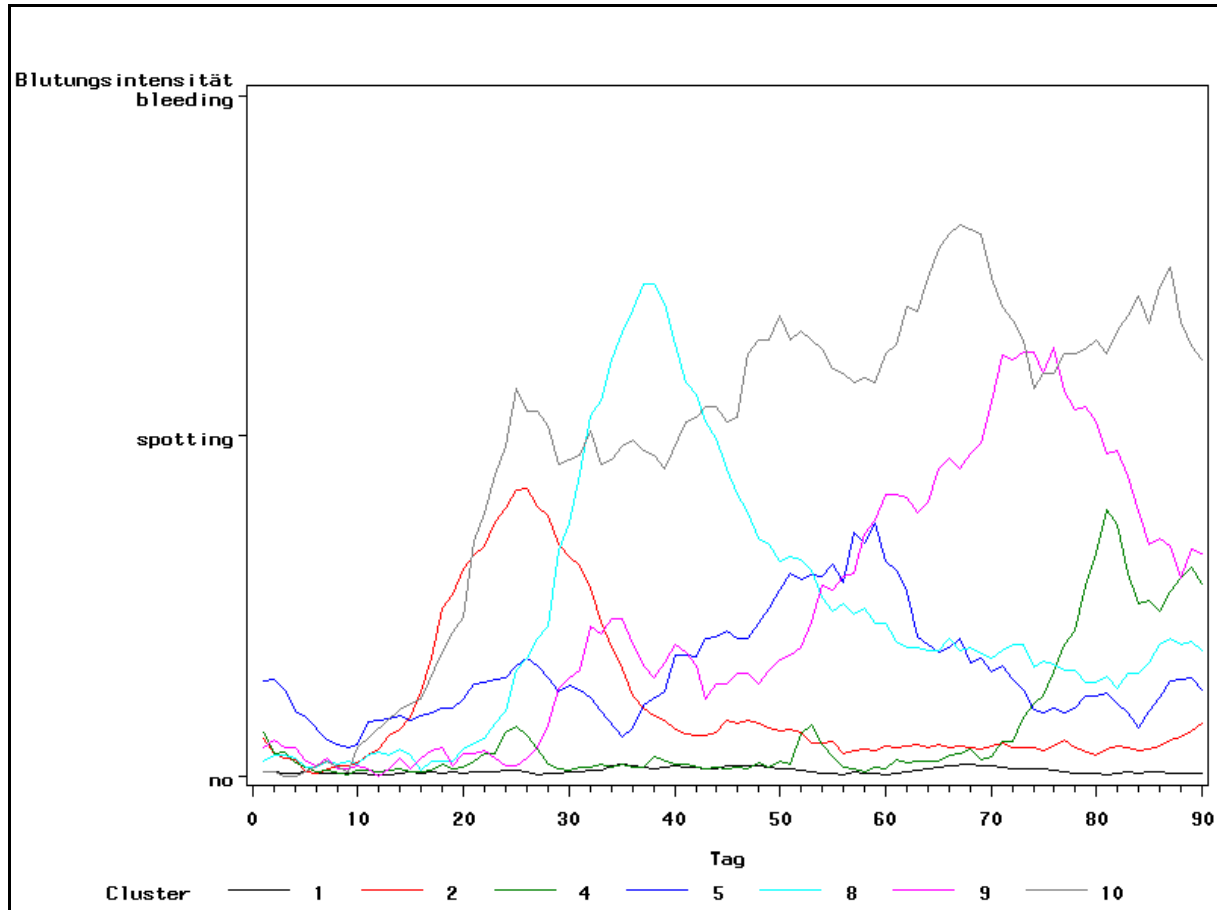


Abbildung 21: Mittlere Blutungsintensität über die Zeit - zwölf Cluster - 7 HRT dominierte Cluster

Von diesen sieben Clustern zeigt Cluster Nummer 1 das Blutungsmuster, das für Präparate zur kontinuierlichen Hormonersatztherapie erwünscht ist, während die anderen Cluster alle unerwünschte Blutungsmuster in verschiedenen Ausprägungen zeigen.

Von den verbleibenden fünf Clustern zeigen vier Cluster (Cluster Nummer 6, Cluster Nummer 7, Cluster Nummer 11 und Cluster Nummer 12) fast ausschließlich

Blutungsmuster von Frauen, die orale Kontrazeptiva einnehmen (vgl. Tabelle 17). Die Anteile der OC Nutzerinnen am Umfang dieser Cluster reichen von 92,7 bis 99,2 Prozent der Beobachtungen. Diese vier Cluster sind in Abbildung 22 der Übersichtlichkeit halber nochmals einzeln dargestellt.

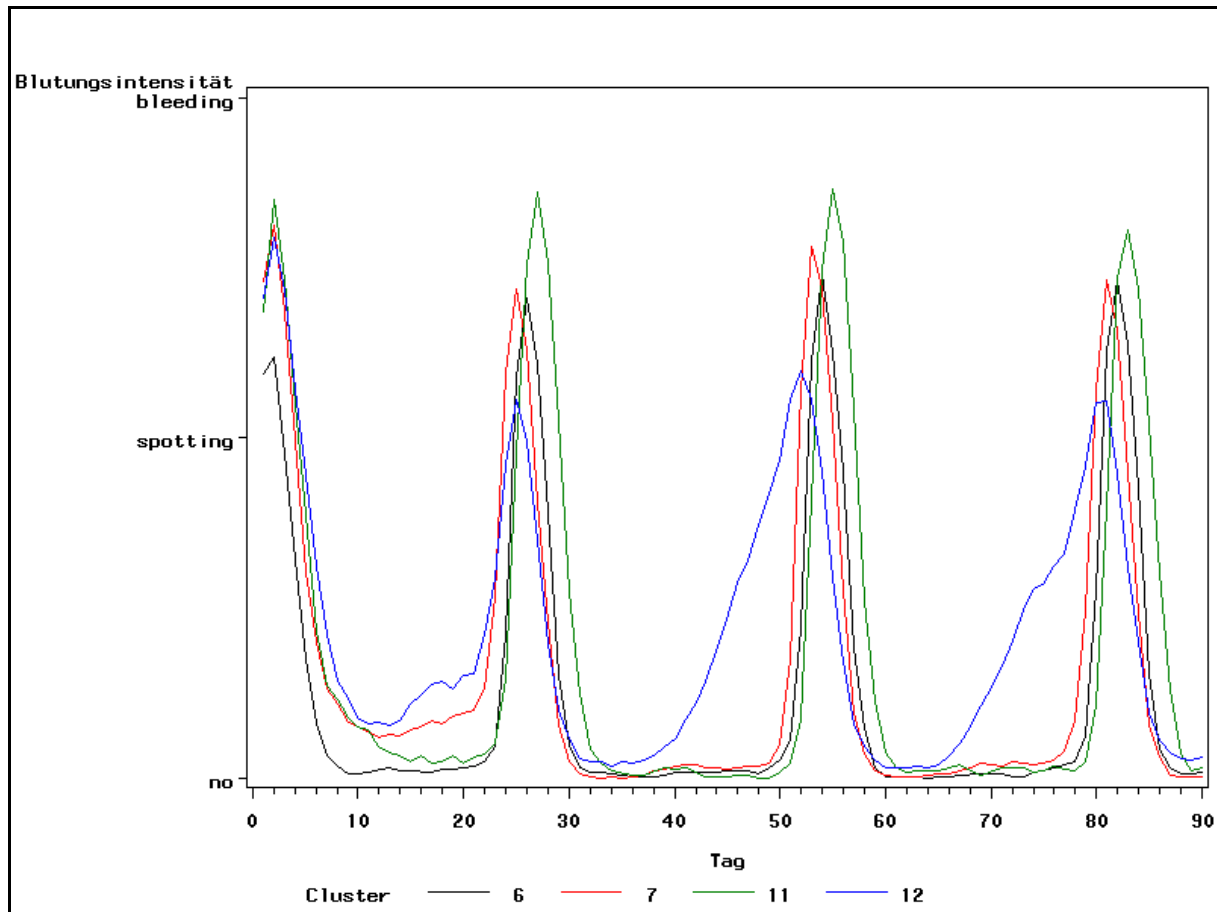


Abbildung 22: Mittlere Blutungsintensität über die Zeit - zwölf Cluster- 4 OC dominierte Cluster

Von den vier Clustern, die Blutungsmuster unter der Einnahme von oralen Kontrazeptiva zeigen, sind drei Cluster (Cluster Nummer 6, Cluster Nummer 7 und Cluster Nummer 11) sich relativ ähnlich und zeigen ein erwünschtes Blutungsmuster. Der einzige nennenswerte Unterschied zwischen diesen drei Clustern zeigt sich

zwischen Tag 8 und Tag 21 im ersten Pilleneinnahmezyklus. Cluster Nummer 12 zeigt das aus der Lösung mit sechs Clustern bereits bekannte unerwünschte Blutungsmuster.

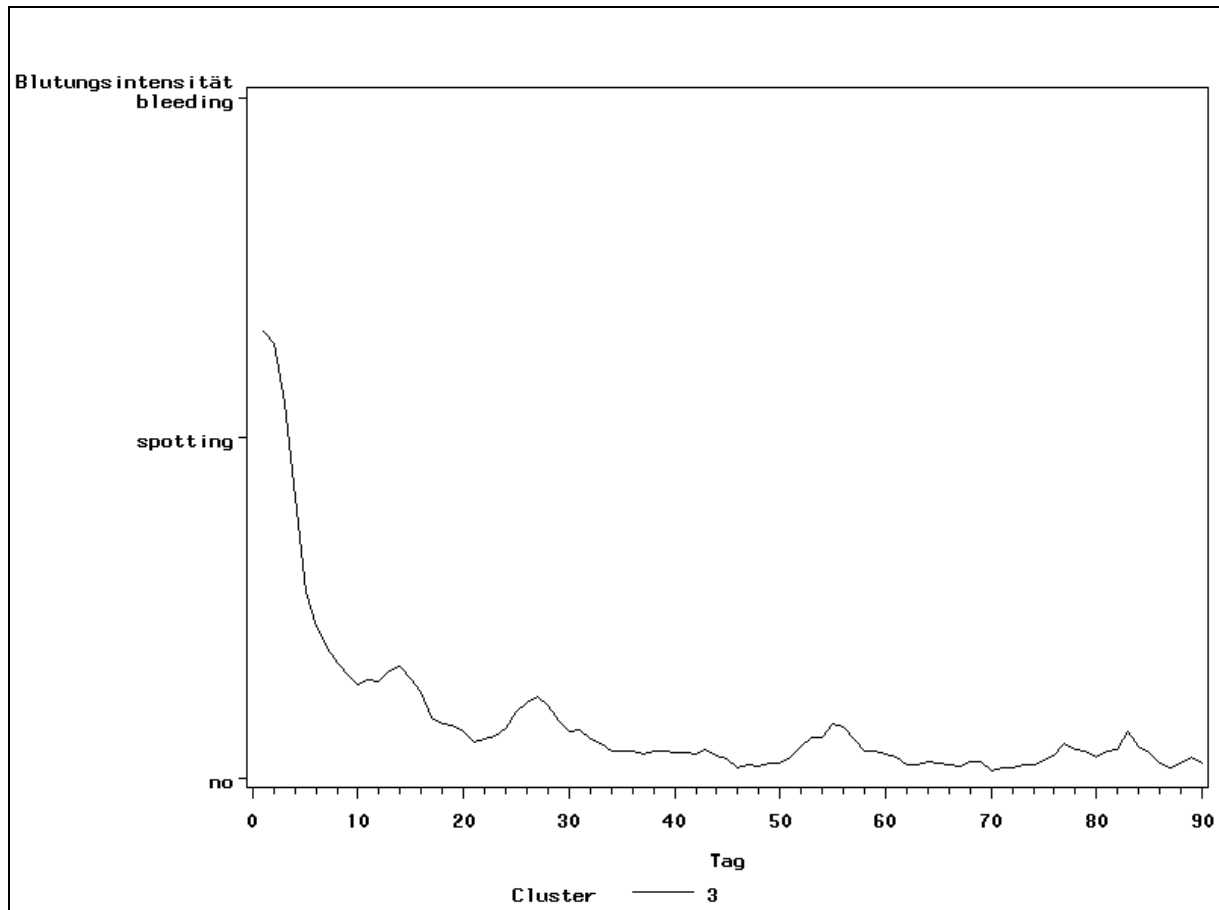


Abbildung 23: Mittlere Blutungsintensität über die Zeit - zwölf Cluster- 1 nicht dominierter Cluster

Der letzte Cluster (Cluster Nummer 3) enthält zu 48,4 Prozent Blutungskalender aus Studien zur Endometriose, zu 33,9 Prozent Blutungskalender aus Studien von oralen Kontrazeptiva und zu 17,7 Prozent Blutungskalender aus Studien von Hormonersatzpräparaten (vgl. Tabelle 17) und ist in Abbildung 23 nochmals dargestellt. Das Blutungsmuster dieses Clusters zeigt die abklingende mittlere

Blutungsintensität unter der Einnahme eines kontinuierlichen Hormonpräparates auch bei geschlechtsreifen Frauen.

4.5.5 Vergleich der Lösungen

Mit Hilfe empirischer Methoden wurden die Lösungen der Clusteranalyse mit drei, vier, sechs und zwölf Clustern als möglicherweise substanzwissenschaftlich interpretierbare Blutungsmuster erkannt.

Es zeigte sich schnell, dass die Lösungen mit drei und vier Clustern zu grob sind, um eine für den Vergleich von Hormonpräparaten hinsichtlich des von ihnen induzierten Blutungsverhaltens sinnvolle Unterteilung von Blutungsmustern zu ermöglichen. Die Lösung mit drei Clustern fasst fast alle Blutungskalender aus Studien von kontinuierlich verabreichten oralen Kontrazeptiva zu einem einzigen Cluster zusammen. Eine Unterscheidung von Blutungsmustern unter der Gabe oraler Kontrazeptiva war somit nicht möglich. Bei der Lösung mit vier Clustern konnten zwar bereits zwei Blutungsmuster unter der Gabe oraler Kontrazeptiva unterschieden werden, eine Unterscheidung bezüglich von Präparaten zur Behandlung der Endometriose war jedoch bei nur vier Clustern noch nicht möglich.

Die erste Lösung, die dies für die kontinuierlich verabreichten Präparate zur Hormonersatztherapie und zur Behandlung der Endometriose leistet, ist die Lösung mit sechs Clustern, von denen vier Blutungsmuster von vorwiegend Frauen, die kontinuierlich verabreichte Präparate einnehmen, darstellen. Allerdings werden die Blutungskalender der Frauen, die zyklisch verabreichte Präparate zur hormonellen Geburtenkontrolle einnehmen, nur in zwei Blutungsmuster zerlegt, was auf den ersten Blick als eventuell zu grob erscheint.

Die Lösung mit zwölf Clustern bringt jedoch keine substanzwissenschaftlich bessere Interpretierbarkeit im Vergleich zur Lösung mit sechs Clustern. Für die kontinuierlich

verabreichten Präparate zur Therapie der Endometriose und zur HRT werden bei der Lösung mit zwölf Clustern sieben anstelle der vier Cluster der Lösung mit sechs Clustern gebildet. Mit Ausnahme des Clusters, der das erwünschte Blutungsmuster zeigt, lassen sich die Cluster jedoch deutlich schlechter inhaltlich interpretieren, als die entsprechenden Cluster der Lösung mit sechs Clustern.

Für die zyklisch verabreichten Präparate zur oralen Kontrazeption führt die Lösung mit zwölf Clustern ebenfalls zu keiner wesentlichen Verbesserung der substanzwissenschaftlichen Interpretierbarkeit. Drei der vier Cluster, die Blutungsmuster überwiegend dieser Indikation beschreiben, sind nahezu gleich. Sie unterscheiden sich lediglich geringfügig im Blutungsverhalten während des ersten Zyklus.

Daraus ergibt sich, dass die sinnvollste Darstellung die Version mit sechs Clustern ist.

4.6 Zusammenfassung der empirischen Untersuchung

Von den insgesamt 5602 zur Verfügung stehenden Blutungskalendern konnten 4612 (82,3%) in die empirische Untersuchung aufgenommen werden. Es wurden alle Blutungskalender ausgewertet, die mindestens 90 Tage lang geführt wurden und die keine unüberbrückbaren Lücken (vgl. Abschnitt 4.1.1) hatten.

Es wurden Clusteranalysen mit den Methoden single linkage, complete linkage, average linkage sowie mit dem Verfahren von Ward durchgeführt (vgl. Abschnitt 4.3.3). Hierbei zeigte es sich, dass von diesen Verfahren zur Clusteranalyse nur die Methode von Ward für die empirische Analyse der Blutungskalender geeignet war (vgl. Abschnitt 4.4.2). Als Abstandsmaß wurde der Euklidische Abstand verwendet (vgl. Abschnitt 4.4.1), da das Verfahren Ward robust gegenüber der Wahl des Abstandsmaßes ist.

Zur Bestimmung der optimalen Anzahl der Cluster wurden das semi-partielle Bestimmtheitsmaß, das kubische Cluster-Kriterium, die pseudo-F- und die pseudo-t-

Statistik herangezogen (vgl. Abschnitt 4.3.5). Als mögliche Lösungen der Clusteranalyse wurden die Lösungen mit drei, vier, sechs und zwölf Clustern als möglicherweise substanzwissenschaftlich interpretierbare Blutungsmuster erkannt (vgl. Abschnitt 4.4.2.1.1).

Beim Vergleich der verschiedenen Lösungen der Clusteranalyse (vgl. Abschnitt 4.5.5) stellte sich heraus, dass die Lösung mit sechs Clustern die substanzwissenschaftlich am besten interpretierbare Lösung war. Bei dieser Lösung wurden vier Blutungsmuster für kontinuierlich verabreichte Hormonpräparate und zwei Blutungsmuster für zyklisch verabreichte Hormonpräparate unterschieden.

Zur besseren Unterscheidung sind die sechs Blutungsmuster der Lösung mit sechs Clustern aus Abbildung 19 im folgenden getrennt für kontinuierlich und für sequenziell verabreichte Präparate dargestellt.

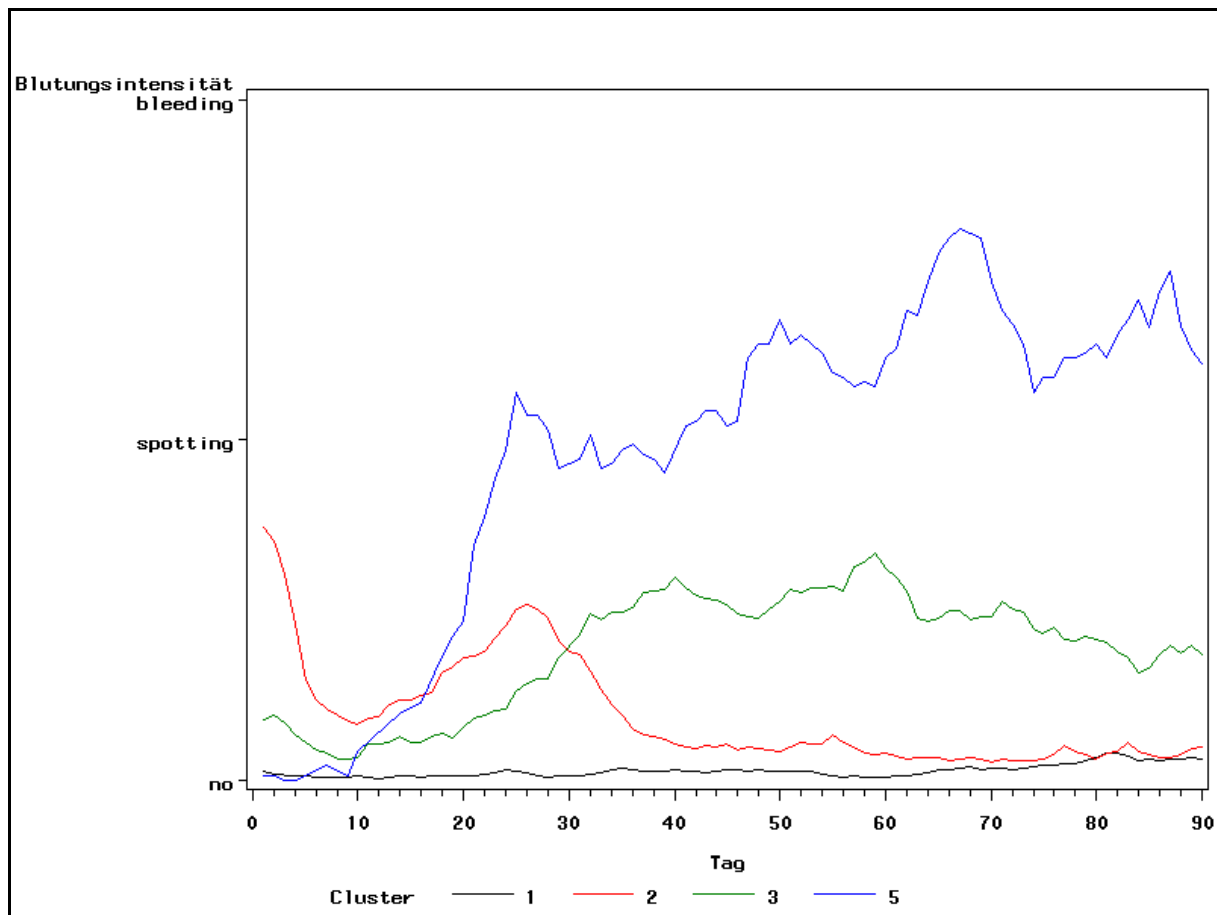


Abbildung 24: Mittlere Blutungsintensität über die Zeit - 4 kontinuierlich verabreichte Präparate

Das erwünschte Blutungsmuster wird in Abbildung 24 durch Cluster Nummer 1 dargestellt. Dieses Blutungsmuster entspricht dem natürlichen Blutungsmuster postmenopausaler Frauen, bei dem de facto eine Amenorrhö auftritt. Dies bedeutet, dass während der ganzen betrachteten 90 Tage keine vaginale Blutung stattfindet.

Das durch Cluster Nummer 2 dargestellte Blutungsmuster kann als akzeptabel für kontinuierliche Hormonpräparate bezeichnet werden. Die mittlere Blutungsintensität ist zu Beginn der Behandlung erhöht, sinkt dann rasch ab um nach etwa vierwöchiger Behandlung einen deutlichen Peak zu haben. Danach klingt die mittlere

Blutungsintensität ab und erreicht zum Ende des Beobachtungszeitraumes von 90 Tagen das Niveau des Clusters Nummer 1. Dieses Blutungsmuster beschreibt sehr gut die Umstellung des Blutungsverhaltens von Frauen, die vor Beginn der kontinuierlichen Behandlung mit Hormonen ein natürliches zyklisches Blutungsverhalten zeigen.

Die Cluster Nummer 3 und Cluster Nummer 5 zeigen unerwünschte Blutungsmuster, da es hier zu einem deutlichen Anstieg der Blutungen nach 20 bis 30 Tagen Behandlung kommt.

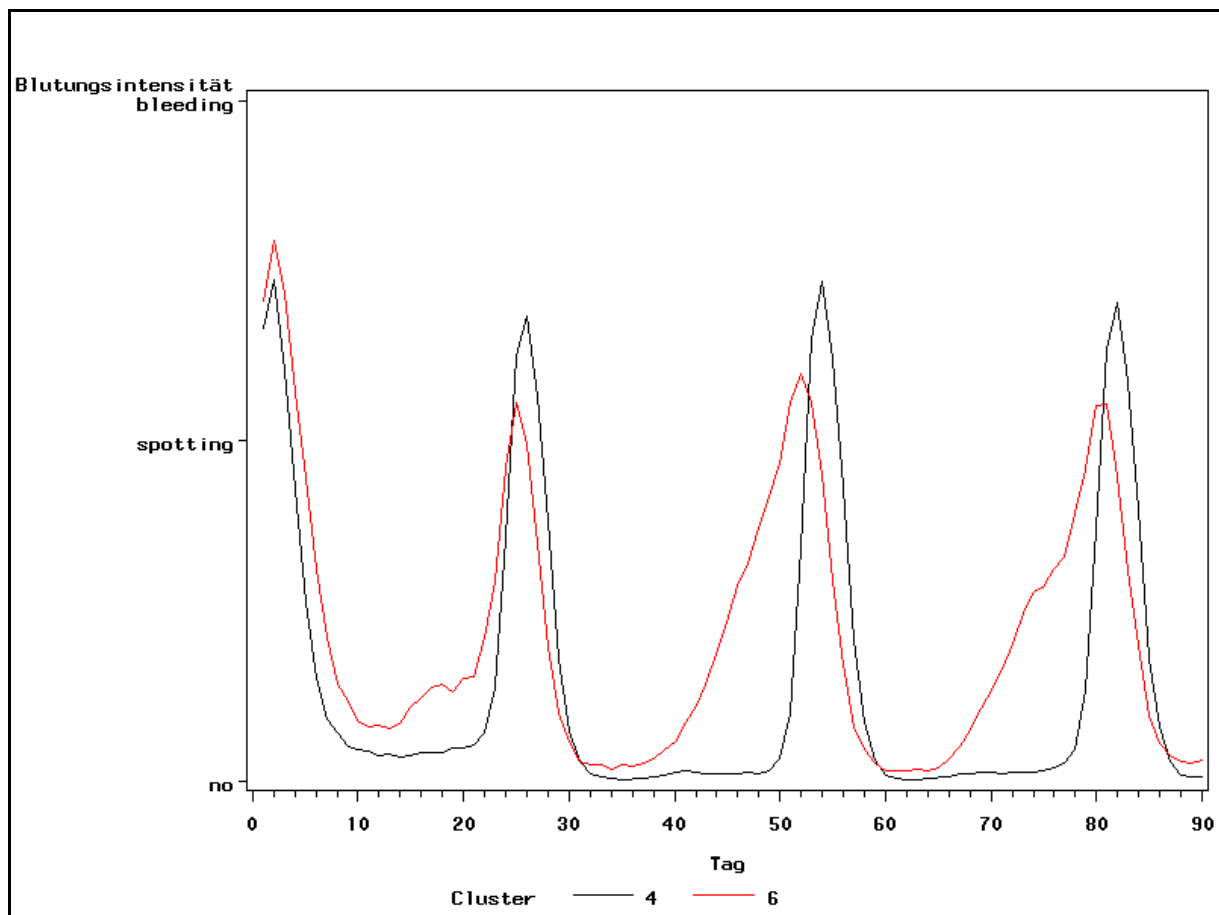


Abbildung 25: Mittlere Blutungsintensität über die Zeit - zyklisch verabreichte Präparate

Die beiden zyklischen Blutungsmuster in Abbildung 25 lassen sich in ein erwünschtes und ein unerwünschtes Blutungsmuster einteilen. Das durch Cluster Nummer 4 dargestellte Blutungsmuster zeigt das erwünschte Blutungsmuster mit deutlichen Abbruchblutungen nach dem Hormonentzug in jedem Zyklus ohne wesentliche Zwischenblutungen. Im Gegensatz dazu zeigt Cluster Nummer 6 ein unerwünschtes zyklisches Blutungsmuster mit vermehrten Zwischenblutungen.

5 Vergleich der Ergebnisse der Literaturrecherche und der empirischen Analyse

Nach Abschluss der Literatursauswertung in Abschnitt 3.4 und der empirischen Analyse der Blutungsmuster in Abschnitt 4.6 stellt sich die Frage, inwieweit die Ergebnisse dieser beiden Methoden zur Erkenntnisgewinnung übereinstimmen und inwiefern sie sich ggf. unterscheiden.

5.1 Methodik des Vergleichs

Um die Ergebnisse der empirischen Analyse der Blutungsmuster mit denen der Literatursauswertung vergleichen zu können, wurden in der Literatur gefundene Methoden auf die selben Blutungskalender angewandt, auf denen auch die empirische Analyse beruhte. Somit lassen sich die Klassifizierungen der Literatur unmittelbar mit der empirisch gefundenen Clusterstruktur der Blutungskalender vergleichen.

Der Vergleich beschränkt sich auf die Erkenntnisse zum Blutungsmuster, da ein Vergleich der verschiedenen Definitionen der Blutungsstärke aus Mangel an empirischen Daten nicht möglich war. Da die verschiedenen Definitionen der Blutungsstärke in der Literatur ähnlich sind, ist dieser Mangel an Daten für die vorliegende Arbeit nur von unerheblicher Bedeutung.

5.2 Ergebnis des Vergleichs

5.2.1 WHO Klassifikation

Die Klassifikation der klinisch relevanten Blutungsmuster gemäß den Empfehlungen der Weltgesundheitsorganisation (vgl. Tabelle 4) wurde auf den in dieser Arbeit untersuchten Datensatz angewandt. Dabei ergab sich die folgende Verteilung der lt. WHO klinisch relevanten Blutungsmuster:

Tabelle 20: Klassifikation der Blutungskalender gemäß WHO

Klassifikation in klinisch relevante Blutungsmuster nach WHO	n	%
Amenorrhö	1288	27,93
häufige Blutungen	158	3,43
normale Blutungen	1034	22,42
seltene Blutungen	828	17,95
unregelmäßige Blutungen	742	16,09
verlängerte Blutungen	562	12,19
Summe	4612	100,00

Die theoretisch mögliche Doppelklassifikation eines Blutungsmusters als „verlängert“ und gleichzeitig als entweder „häufig“, „selten“ oder „unregelmäßig“ trat im untersuchten Datensatz nicht auf. Da die vorliegenden Blutungskalender aus klinischen Prüfungen von hormonhaltigen Präparaten aus verschiedenen Indikationen stammen, bietet sich eine Aufgliederung der Daten in Tabelle 20 nach Indikationen an.

Tabelle 21: Klassifikation der Blutungskalender gemäß WHO nach Indikationen

Klassifikation in klinisch relevante Blutungsmuster nach WHO	Endometriose		HRT		OC	
	n	%	n	%	n	%
Amenorrhö	2	0,88	1286	49,52	0	0,00
häufige Blutungen	9	3,98	38	1,46	111	6,20
normale Blutungen	4	1,77	52	2,00	978	54,67
seltene Blutungen	109	48,23	679	26,15	40	2,24
unregelmäßige Blutungen	42	18,58	194	7,47	506	28,28
verlängerte Blutungen	60	26,55	348	13,40	154	8,61
Summe	226	100,00	2597	100,00	1789	100,00

In Tabelle 21 zeigt sich sehr deutlich, dass die WHO Klassifikation sich am Blutungsverhalten menstruierender Frauen orientiert. Nur 2 % der postmenopausalen Frauen und 1,77% der Patientinnen mit Endometriose, die ein kontinuierlich verabreichtes Präparat zur Hormonersatztherapie einnehmen, haben ein „normales“ Blutungsmuster gemäß WHO. Dagegen hat die Mehrzahl der Frauen die ein hormonelles Kontrazeptivum einnehmen, ein „normales“ Blutungsmuster gemäß WHO.

Dieses Ergebnis lässt bereits vermuten, dass die Anwendung der WHO Klassifikation höchstens für zyklische Präparate sinnvoll sein könnte. Die nachfolgende Analyse gemäß den in Abschnitt 4.6 gefundenen Clustern bestätigt dies.

In den Clustern mit kontinuierlich verabreichten Präparaten befinden sich nur zu 1,9 bis 9,6 % Frauen mit einem normalen Blutungsmuster gemäß der Definition der Weltgesundheitsorganisation (vgl. Tabelle 22).

**Tabelle 22 Klassifikation der Blutungskalender gemäß WHO nach Clustern
- kontinuierlich verabreichte Präparate**

	Cluster 1		Cluster 2		Cluster 3		Cluster 5	
Klassifikation in klinisch relevante Blutungsmuster nach WHO	n	%	n	%	n	%	n	%
Amenorrhö	1288	68,51	0	0,00	0	0,00	0	0,00
häufige Blutungen	13	0,69	27	4,58	14	3,11	2	2,82
normale Blutungen	35	1,86	45	7,63	43	9,56	4	5,63
seltene Blutungen	439	23,35	294	49,83	89	19,78	0	0,00
unregelmäßige Blutungen	81	4,31	130	22,03	81	18,00	2	2,82
verlängerte Blutungen	24	1,28	94	15,93	223	49,56	63	88,73
Summe	1880	100,00	590	100,00	450	100,00	71	100,00

Im Cluster Nummer 1, dem Cluster mit dem für derartige Präparate erwünschten Blutungsmuster, haben jedoch 68,5 % der Frauen eine Amenorrhö und 23,4 5 % der Frauen seltene Blutungen. Das heißt, dass 91,9 % der Frauen in diesem Cluster weniger als normale Blutungen haben und auch keine unregelmäßigen oder verlängerten Blutungsepisoden haben.

Für den Cluster Nummer zwei, der ein akzeptables Blutungsmuster für kontinuierliche Präparate darstellt, trifft dies jedoch nur noch auf 49,8 % der Frauen zu. Bereits 15,9 % der Frauen in diesem Cluster haben verlängerte Blutungen. Für die Cluster Nummer 3 und Nummer 5 sinkt die Rate der Frauen mit seltenen Blutungen auf 19,8 bzw. 0 Prozent, während der Anteil der Frauen mit verlängerten Blutungen auf 49,6 und 88,7 % steigt. Hier zeigt sich also eine gute Übereinstimmung der WHO Definitionen mit der empirischen Analyse der Blutungskalender.

Ein anderes Bild ergibt sich für die Cluster, die hauptsächlich Blutungskalender unter der Einnahme von zyklischen Präparaten enthalten. Der Cluster Nummer 4 mit dem erwünschten Blutungsmuster für derartige Präparate enthält zu 65,4 Prozent Frauen mit einem normalen Blutungsmuster gemäß der Definition der WHO. Cluster Nummer 6 enthält dagegen zu 25,6 % Frauen mit einem normalen Blutungsmuster (vgl. Tabelle 23).

Hier zeigt sich sehr deutlich, dass die Klassifikation des Blutungsverhaltens der Weltgesundheitsorganisation in normale und nicht normale Blutungen nicht sehr gut mit den Ergebnissen der empirischen Analyse der Blutungskalender übereinstimmt. Der Grund dafür liegt in der Orientierung der WHO Klassifikation am normalen zyklischen Blutungsverhalten geschlechtsreifer Frauen.

**Tabelle 23: Klassifikation der Blutungskalender gemäß WHO nach Clustern
 - zyklisch verabreichte Präparate**

Klassifikation in klinisch relevante Blutungsmuster nach WHO	Cluster 4		Cluster 6	
	n	%	n	%
Amenorrhö	0	0,00	0	0,00
häufige Blutungen	64	5,18	38	9,84
normale Blutungen	808	65,43	99	25,65
seltene Blutungen	2	0,16	4	1,04
unregelmäßige Blutungen	299	24,21	149	38,60
verlängerte Blutungen	62	5,02	96	24,87
Summe	1235	100,00	386	100,00

Zusammenfassend zeigte sich im Vergleich der Ergebnisse der Literaturrecherche mit denen der empirischen Analyse, dass die gängige Definition des Blutungsmusters gemäß den Empfehlungen der Weltgesundheitsorganisation WHO für die Beurteilung von Blutungsmustern unter der Einnahme von hormonellen Präparaten nicht ausreicht.

5.3 Entwicklung der neuen Methode zur Analyse der Zykluskontrolle

Wie im Abschnitt 3.4 gezeigt, lassen sich die relativen Stärken und Schwächen der Zyklus-basierten Methoden und der Episoden-basierten Methoden wie folgt zusammenfassen. Die Zyklus-basierten Methoden eignen sich zur Analyse der Zykluskontrolle, da sie zwischen den erwünschten Entzugsblutungen und den unerwünschten Zwischenblutungen unterscheiden können. Sie haben aber den Nachteil, dass sie auf zusammenhängende Blutungsereignisse keine Rücksicht nehmen können. Die Episoden-basierten Methoden verwenden die Blutungsereignisse als Untersuchungsgröße, sie können aber per se nicht zwischen Entzugsblutungen und den Zwischenblutungen unterscheiden.

Die Idee, die beiden bisherigen Methoden zur Analyse der Blutungsmuster zu kombinieren, verfolgt einen zweistufigen Ansatz. Im ersten Schritt werden, wie bei den Episoden-basierten Methoden, die einzelnen Blutungsepisoden ermittelt. Danach werden im zweiten Schritt die einzelnen Blutungsepisoden in Entzugsblutungsepisoden und in Zwischenblutungsepisoden unterschieden. Dieser Ansatz vermeidet die Doppelzählung einer Blutungsepisode, die sich über mehr als einen Kalenderabschnitt erstreckt, wie dies bei den Zyklus-basierten Methoden der Fall ist. So wurden beispielsweise bei der Schering AG Blutungen an den Tagen 4 bis 21 als Zwischenblutungen und Blutungen an den anderen Tagen als Entzugsblutung gewertet (vgl. Abschnitt 3.3.3.1). Eine ununterbrochene Blutungsepisode, die an Tag 20 beginnt und an Tag 23 eines Zyklus endet wird demgemäss als zwei Tage Zwischenblutung (Tag 20 und Tag 21) und als zwei Tage Entzugsblutung (Tag 22 und Tag 23) gezählt.

5.3.1 Ausgangsmethode

Im Falle eines oralen Kontrazeptivums mit Standardregime, d.h. 21 Tage Wirkstoffeinnahme gefolgt von 7 Tagen ohne Einnahme bzw. von 7 Tagen mit

Einnahme von Placebopillen, kann man die Blutungsepisoden in Entzugsblutungsepisoden und in Zwischenblutungsepisoden unterscheiden. Hierbei werden die einzelnen Blutungsepisoden nach dem Tag ihres Beginns unterschieden.

Blutungsepisoden, die an einem Tag ohne Hormoneinnahme beginnen, werden als Entzugsblutungsepisoden klassifiziert, während Blutungsepisoden, die an einem Tag unter Hormoneinnahme beginnen, als Zwischenblutungsepisoden klassifiziert werden.

Diese Klassifizierung ist in der überwiegenden Anzahl der Fälle leicht anzuwenden. Es gibt jedoch einige Grenzfälle, die eine Adaption dieser simplen Klassifikation nötig machen. Auf diese Grenzfälle wird im folgenden eingegangen.

5.3.2 Erster Grenzfall der Ausgangsmethode

In der einwöchigen Pillenpause können theoretisch bis zu drei Blutungsepisoden auftreten. Ein Beispiel ist in Abbildung 26 dargestellt.

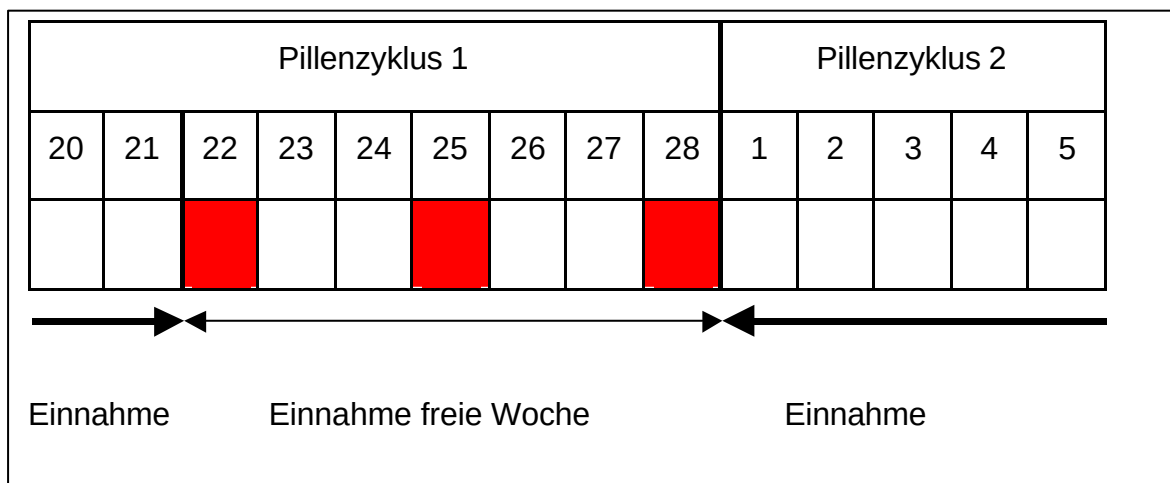


Abbildung 26: Beispiel mehrerer Blutungsepisoden im Einnahme freien Intervall

In dem in Abbildung 26 beschriebenen Blutungsmuster wird an den Tagen 22, 25 und 28 des Pilleneinnahmezyklus geblutet. Zwischen den einzelnen Tagen mit vaginaler Blutung liegen genau zwei Tage und weder an den letzten beiden Tagen der Pilleneinnahme, d.h. an den Tagen 20 und 21 des ersten Pilleneinnahmezyklus noch an den ersten beiden Tagen des zweiten Pilleneinnahmezyklus wird geblutet. Daher handelt es sich bei dem beschriebenen Blutungsmuster definitionsgemäß um drei singuläre Blutungsepisoden.

Alle drei Blutungsepisoden liegen vollständig in der Woche ohne Hormoneinnahme und wären gemäß der obigen Definition als Entzugsblutungen zu werten. Dies steht im Widerspruch zu der Tatsache, dass es in einem Pillenzyklus nur einen einzigen Hormonentzug, nämlich das Absetzen der täglichen Hormoneinnahme am Tag 22, gibt.

Dieser Grenzfall ist jedoch mehr theoretischer Natur. Bei der empirischen Analyse von mehr als 5000 Blutungskalendern (vgl. Abschnitt 4) trat ein derartiger Fall nicht auf.

5.3.3 Zweiter Grenzfall der Ausgangsmethode

Ein Grenzfall, der in der Praxis vorkommt, ist die vorfristige Entzugsepisode. Sie ist in Abbildung 27 schematisch dargestellt.

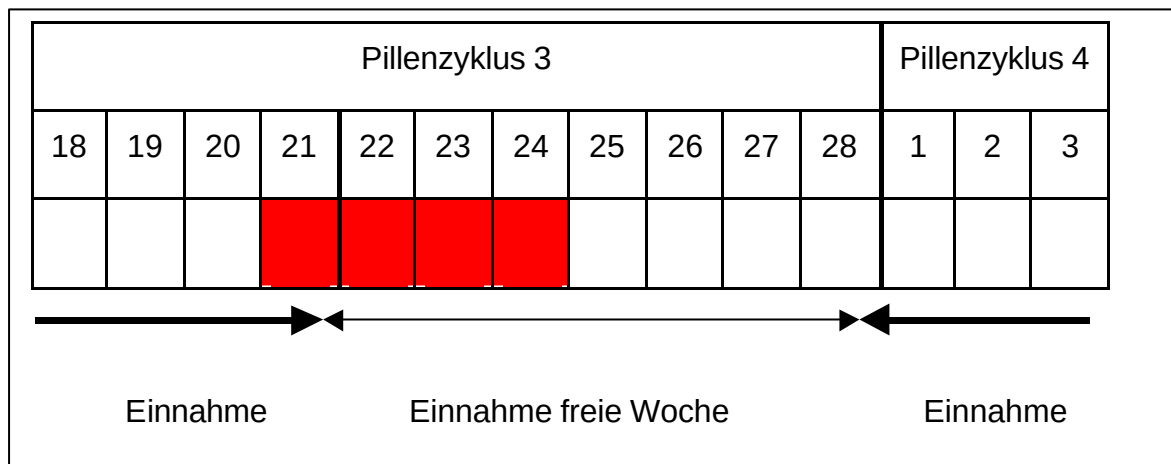


Abbildung 27: vorfristige Entzugsepisode

In diesem Fall findet vom 21. bis zum 24. Tag des Pillenzyklus eine Blutungsepisode statt. Nach der obigen Definition würde es sich bei dieser Blutungsepisode um eine Zwischenblutungsepisode handeln, da die Episode an Tag 21, dem letzten Tag der Hormoneinnahme, beginnt. Gleichzeitig würde für diesen Zyklus das Ausbleiben einer Entzugsblutung konstatiert werden, da im hormonfreien Intervall von Tag 22 bis Tag 28 keine Blutungsepisode beginnt. Aus medizinischer Sicht würde das in Abbildung 27 dargestellte Blutungsmuster jedoch eindeutig als die erwartete Monatsblutung bewertet werden.

Im Gegensatz zu dem eher theoretischen Beispiel des Auftretens mehrerer Entzugsblutungsepisoden in Abbildung 26 treten vorfristige Entzugsblutungen in der Praxis häufiger auf.

5.3.4 Dritter Grenzfall der Ausgangsmethode

Ein weiteres Problem tritt auf, wenn die Monatsblutung verspätet, d.h. erst nach der Wiederaufnahme der Hormoneinnahme im folgenden Pillenzyklus, eintritt. Dies ist

insbesondere dann relevant, wenn z.B. bei neuartigen Einnahmeschemata die Anzahl der hormonfreien Tage reduziert wird.

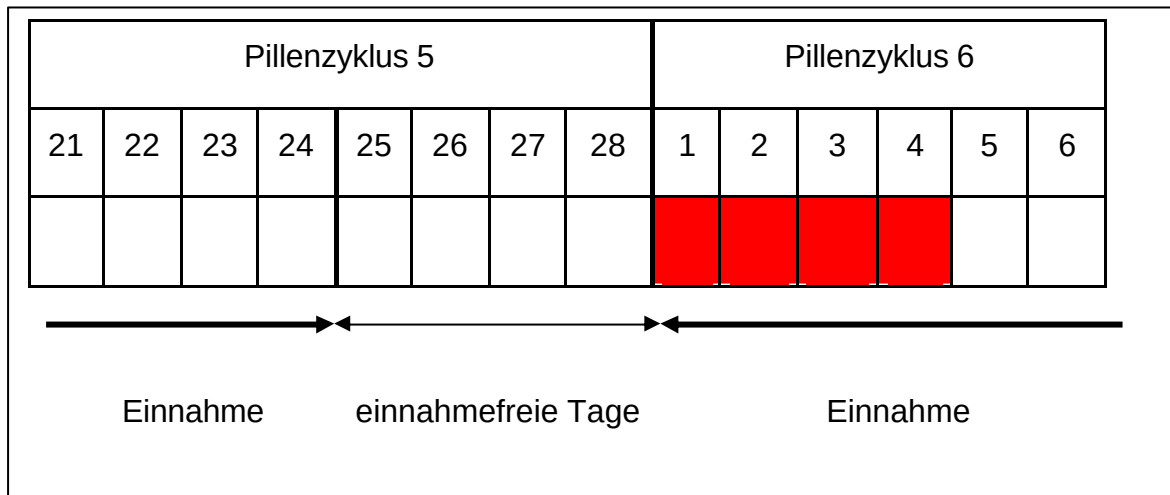


Abbildung 28: verspätete Entzugsepisode bei verkürzter einnahmefreier Zeit

Auch in diesem Fall würde aus medizinischer Sicht das in Abbildung 28 dargestellte Blutungsmuster als die erwartete Monatsblutung bewertet werden.

5.3.5 Neue Methode zur Analyse der Zykluskontrolle

Daraus ergibt sich, dass die zunächst vorgeschlagene Klassifikation der Blutungsepisoden in Entzugsepisoden und Zwischenblutungsepisoden allein anhand des Zyklustages, in dem die Blutungsepisoden beginnt, sich nicht als zufriedenstellend erwies. Um den erläuterten Spezialfällen (vgl. Abbildung 26, Abbildung 27 und Abbildung 28) Rechnung zu tragen, muss die Klassifikation der einzelnen Blutungsepisoden in Entzugsblutungsepisoden und Zwischenblutungsepisoden wie folgt adaptiert werden.

Die erste Adaption betrifft das Auftreten von mehreren Blutungsepisoden, die als Entzugsblutungsepisoden in Frage kämen (vgl. Abbildung 26). Dieses Problem kann ganz einfach dadurch gelöst werden, dass nur die zeitlich erste der in Frage kommenden Blutungsepisoden als Entzugsblutung klassifiziert wird. Die folgenden Blutungsepisoden werden alle als Zwischenblutungsepisoden klassifiziert.

Die zweite Adaption behandelt diejenigen Blutungsepisoden, die zum Zeitpunkt des Hormonentzugs andauern (vgl. Abbildung 27). Diese werden als die Entzugsblutungsepisode gewertet, vorausgesetzt, dass sie nicht zu früh beginnen. Als geeignete Zeitspanne hat sich empirisch eine Frist von 4 Tagen herausgestellt.

Die letzte Adaption betrifft das Problem der verzögerten Entzugsblutung (vgl. Abbildung 28). Dieses Problem kann dadurch gelöst werden, dass eine Blutungsepisode als Entzugsblutungsepisode gewertet werden kann, auch wenn sie nach dem Beginn des neuen Pilleneinnahmezyklus im eigentlichen Folgezyklus auftritt. Voraussetzung ist jedoch, dass diese Blutungsepisode nicht zu spät beginnt. Als geeignete Zeitspanne hat sich empirisch eine Frist von maximal 5 Tagen nach Wiederaufnahme der Hormoneinnahme herausgestellt.

Ein pharmazeutischer Hersteller verwendet gegenwärtig als Grenze den Zeitpunkt des erneuten Hormonentzugs im Folgezyklus. Dieses Vorgehen stellt sicherlich eine obere Grenze für den zulässigen Zeitpunkt des Beginns einer verzögerten Entzugsblutung dar. Besser wäre es, eine kürzere Frist zu wählen, da es bei dieser Definition bei sehr späten Entzugsblutungsepisoden zu einem Abgrenzungsproblem kommt. Nach dieser Definition wäre eine Blutungsepisode, die von Tag 20 bis 24 des zweiten Zyklus dauert, die Entzugsblutung des ersten Zyklus, falls seit Tag 22 des ersten Zyklus keine andere Blutungsepisode auftrat. Falls jedoch eine solche Blutungsepisode auftrat, wäre die betrachtete Blutungsepisode tatsächlich aber die Entzugsblutung des zweiten Zyklus. Eine kontextfreie Definition lässt sich durch eine stärkere Begrenzung der maximal zulässigen Verzögerung einer Entzugsblutungsepisode erreichen. Die maximal zulässige Verzögerung einer

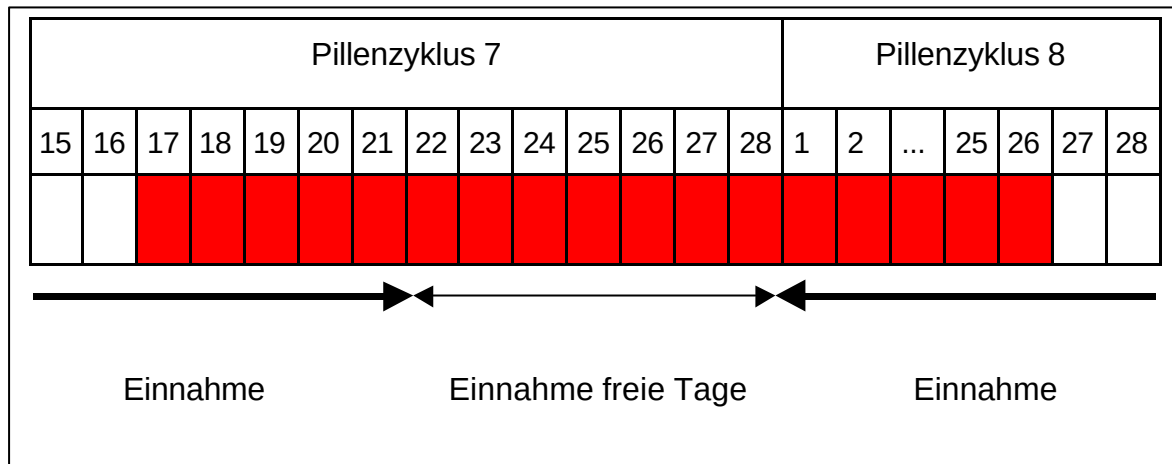
Entzugsblutungsepisode muss so kurz gewählt werden, dass sie sich nicht mit der Frist für eine vorzeitige Entzugsblutungsepisode überlappt.

Die genaue Klassifizierung der Blutungsepisoden in Entzugsblutungsepisoden und in Zwischenblutungsepisoden lässt sich wie folgt zusammenfassen:

- Ist am Tag des Hormonentzugs eine Blutungsepisode im Gange, so ist diese Blutungsepisode die Entzugsblutungsepisode dieses Zyklus, vorausgesetzt, dass diese Blutungsepisode nicht früher als 4 Tage vor dem Hormonentzug begonnen hat.
- Ist am Tag des Hormonentzugs keine Blutungsepisode im Gange, so ist die nächste Blutungsepisode die Entzugsblutungsepisode dieses Zyklus, vorausgesetzt, dass diese Blutungsepisode nicht später als 5 Tage vor dem Hormonentzug des darauffolgenden Zyklus begonnen hat.
- Alle anderen Blutungsepisoden werden als Zwischenblutungsepisoden kategorisiert. Sie werden dem Zyklus zugeordnet, in dem sie beginnen.

Durch die beschriebenen Verbesserungen der ursprünglichen Klassifikation der Blutungsepisoden wurde diese deutlich verbessert, da sie jetzt auch die erwähnten Grenzfälle mit abdeckt.

Lediglich ein Extremfall wird von der verbesserten Klassifikation nicht richtig erfasst. Dies ist genau dann der Fall, wenn eine Blutungsepisode auftritt, deren Länge nicht mit der Periodizität des Einnahmeschemas in Einklang steht. Ein Beispiel für dieses Problem ist in Abbildung 29 dargestellt. In diesem Falle liefern die obigen Definitionen ein verzerrtes Abbild des wahren Blutungsmusters.



als die Länge der gewählten Referenzperiode. Sollte dies nicht der Fall sein, müsste die Länge der Referenzperiode entsprechend erhöht werden.

5.4 Diskussion der neuen Methode

Die vorgeschlagene neue Methode bietet einige Vorzüge gegenüber den in der Literatur verwendeten Methoden. Sie verbindet die Vorzüge der Zyklus-basierten mit denen der Episoden-basierten Methoden.

Die neue Methode fasst zunächst aufeinander folgende Tage mit Blutungen zu einer Episode zusammen. Damit ist es ausgeschlossen, dass ein Blutungsereignis sowohl als Entzugs- als auch als Zwischenblutung gewertet wird, wie dies bei den Zyklus-basierten Methoden auftreten kann (vgl. Abschnitt 3.3.3.1).

Die Trennung der Blutungsepisoden in Entzugs- und Zwischenblutungsepisoden wird durch die neue Methode in eindeutiger Weise gewährleistet, während es bei den bisher gebräuchlichen Methoden durchaus zu Klassifizierungsproblemen kommen konnte (vgl. Abschnitt 5.3.4).

Ein noch offenes Problem ist die Behandlung von extrem langen Blutungsepisoden (vgl. Abbildung 29). Dieses Problem besteht jedoch auch bei den bisherigen Methoden und stellt daher keinen spezifischen Nachteil der vorgeschlagenen Methode dar.

Zusammenfassend lässt sich sagen, dass die neue Methode den bisherigen Methoden eindeutig überlegen ist. Eine Ausnahme bildet lediglich die Behandlung von extrem langen Blutungsepisoden, die weder von der neuen, noch von den bisherigen Methoden voll zufriedenstellend gelöst wird.

6 Klinische Relevanz und Beitrag zu Public Health

6.1 Klinische Relevanz

Das Blutungsmuster ist ein wichtiges Kriterium für die individuelle Auswahl eines hormonalen Empfängnisverhütungsmittels, eines Präparates zur Behandlung der Endometriose oder eines Präparates zur Hormonersatztherapie während und nach den Wechseljahren. Unregelmäßige und nicht vorhersehbare Blutungen können die Lebensqualität der Frau einschränken, da sie im täglichen Leben hinderlich sein können. Die regelmäßige Blutung während der Einnahme eines hormonalen Präparates zur Empfängnisverhütung gilt der Frau als Zeichen, nicht schwanger zu sein.

Blutungsprobleme sind einer der wichtigsten Gründe für die Entscheidung der Frau, ein Präparat zu wechseln, eine andere Form der Verhütung zu wählen oder eine Hormonersatztherapie ganz abubrechen. Im letzteren Falle führt der Verzicht auf die Osteoporoseprophylaxe zu einer erheblichen Erhöhung des Risikos, Knochenbrüche insbesondere der Wirbel und der Hüftknochen zu erleiden.

Ein weltweit einheitlich erhobenes und ausgewertetes Blutungsmuster erlaubt sowohl dem verschreibenden Arzt als auch der Konsumentin hormonhaltiger Präparate einen sinnvollen prospektiven Vergleich der angebotenen Produkte hinsichtlich ihrer Eigenschaften bezüglich der Blutungskontrolle. Sie werden damit unabhängiger von den einzelnen Herstellern, die bislang unterschiedliche und daher nur bedingt vergleichbare Angaben veröffentlichen.

Im wesentlichen trifft dies auf die folgenden Auswahlprobleme zu:

- **Präparate gleicher Indikation, gleicher Applikationsform und mit gleichen chemischen Inhaltstoffen**

Unter den genannten Voraussetzungen beruht das Blutungsverhalten lediglich auf der Dosierung der chemischen Inhaltsstoffe, wenn es sich um fixe

Kombinationen und nicht um Mehrstufenpräparate handelt. Der Arzt kann das individuelle Blutungsverhalten seiner Patientin mit dem in klinischen Studien ermittelten durchschnittlichen Blutungsprofil behandelter Frauen vergleichen. Bei unbefriedigender individueller Zykluskontrolle könnte die Dosis des Präparates verändert werden, um eine bessere Zykluskontrolle zu erreichen. Vergleichbare Angaben zum Blutungsmuster ermöglichen somit eine fundierte Entscheidung, ob im individuellen Fall eine Dosisreduktion oder -erhöhung in Frage kommt. Eine Dosistitration im Hinblick auf das Blutungsverhalten wird somit wissenschaftlich abgesichert.

- **Präparate gleicher Indikation, gleicher Applikationsform und mit verschiedenen chemischen Inhaltsstoffen**

Dies ist sicher die am häufigsten zu fällende klinische Entscheidung für den verschreibenden Arzt, da die Applikationsform häufig von der Anwenderin der hormonalen Kontrazeptiva bzw. der Hormonersatzpräparate vorgegeben wird. Hier wird das Blutungsverhalten nicht nur von der Dosierung der chemischen Inhaltsstoffe, sondern auch von deren Eigenschaften wie die individuelle Gestagenverträglichkeit bei gegebener Östrogendosis bestimmt. In diesem Fall ist eine wissenschaftlich fundierte Vergleichbarkeit der Auswirkungen der Präparate auf das Blutungsmuster besonders wichtig, da in der Regel Präparate von verschiedenen pharmazeutischen Herstellern miteinander verglichen werden müssen.

- **Präparate gleicher Indikation aber verschiedener Applikationsform**

Bei gleicher Indikation erlauben standardisiert erhobene und dargestellte Zyklusdaten den direkten Vergleich der Produkte in Bezug auf ihre Beeinflussung des Blutungsmusters. Im Falle der Kontrazeption wird es für Arzt und Anwenderin möglich, eine fundierte Wahl zwischen oralem Kontrazeptivum, hormonhaltiger Spirale (Mirena®), klassischer Spirale, Einmonats- oder Dreimonatsinjektat, Hormonpflaster oder Implantat bezüglich dem zu erwartenden Blutungsmuster zu treffen. Gleiches gilt für den Fall einer Hormonersatztherapie für die Auswahl

zwischen oralem sequentiellen, oralem kontinuierlichem, transdermalen sequenziellem, transdermalen kontinuierlichem oder Depotpräparat.

- **Präparate verschiedener Indikation**

Häufig werden hormonale Kontrazeptiva und Hormonersatzpräparate auch zur Behandlung von Blutungsunregelmäßigkeiten eingesetzt. In diesem Falle ist eine wissenschaftlich fundierte Vergleichbarkeit des von den jeweiligen Präparaten induzierten Blutungsmusters von besonderer Bedeutung. Dies gilt insbesondere in der Perimenopause beim Übergang von einem hormonalen Kontrazeptivum zu einem Hormonersatzpräparat.

Die klinisch relevanten Ergebnisse der vorliegenden Arbeit sind in der Leitlinie zur Verschreibung in Abschnitt 7.1 zusammengefasst.

6.2 Beitrag zu Public Health

Neben der Verbesserung der Möglichkeiten zur Berücksichtigung von Blutungsmustern bei der individuellen Verschreibung von Präparaten, die das Blutungsmuster verändern können, können die Ergebnisse dieser Arbeit auch zur Verbesserung der Entscheidung über neue Präparate im Rahmen eines arzneimittelrechtlichen Zulassungsverfahrens verwendet werden. Dies trifft auf eine Reihe von Entscheidungsproblemen zu, die im folgenden näher ausgeführt sind:

- **Alle Präparate und medizinische Hilfsmittel, die das Blutungsverhalten beeinflussen**

Durch eine weltweit einheitliche Erhebung und Auswertung der Daten zum Blutungsverhalten können die Gesundheitsbehörden erstmals das Blutungsverhalten von allen Präparaten und medizinischen Hilfsmitteln, die das Blutungsverhalten beeinflussen, wissenschaftlich fundiert vergleichen. Damit kann insbesondere gewährleistet werden, dass ein neu zuzulassendes Produkt keine wesentlich schlechtere Blutungskontrolle hat, als bereits zugelassene Präparate in der betreffenden Indikation. Dabei ist natürlich zu beachten, dass ein neues

Präparat mit einer wesentlich schlechteren Blutungskontrolle sehr wohl zulassungsfähig ist, wenn es über sonstige Vorteile (wie z.B. erhöhte Wirksamkeit oder höhere Sicherheit) verfügt.

- **Präparate mit gleicher Indikation**

Einheitlich erhobene und ausgewertete Daten zur Zykluskontrolle erlauben erstmalig einen genauen Vergleich der Blutungseigenschaften verschiedener Präparate, auch wenn die Daten aus verschiedenen Studien stammen und wenn die Studien von verschiedenen pharmazeutischen Herstellern oder von Universitäten durchgeführt wurden.

Zulassungsanträge von Präparaten mit deutlich schlechterer Zykluskontrolle als die bereits auf dem Markt befindlichen Präparate könnte so die Zulassung verweigert werden, sofern sie keinen sonstigen medizinischen Vorteil haben, der den Nachteil im Blutungsverhalten überwiegt. Vielleicht ist dies der Grund, warum die pharmazeutischen Hersteller sich bisher nicht auf eine einheitliche Darstellung verständigt haben.

- **Präparate verschiedener Indikationen**

Die beschriebene Methode erlaubt erstmals einen Vergleich von Präparaten verschiedener Indikationen. Präparate mit gleichzeitiger OC und HRT Indikation sind bereits in der klinischen Entwicklung. Vergleichbare Zykluskontrolldaten sind ein wichtiges Bewertungskriterium für diese neuartige Präparatklasse. Die neue Methode erlaubt daher auch zuverlässige Aussagen über verschiedene Altersgruppen.

- **Verschiedene Applikationsformen**

Eine einheitliche Erhebung und Auswertung von Blutungsdaten erlaubt den direkten Vergleich verschiedener Applikationsformen wie Tablette, Injektion, Depotinjektion, transdermale Applikation, Implantat oder intrauterine Applikation der Hormone. Claims der Hersteller, dass eine neuartige Applikationsform zu

besserer Blutungskontrolle führt, können so von den Zulassungsbehörden besser beurteilt werden.

- **Veränderte Einnahmeregimes**

Hormonpräparate mit veränderten Einnahmeregimes sind derzeit verstärkt in der klinischen Entwicklung und zum Teil auch bereits auf dem Markt, wie z.B. Seasonale® in den USA. Erst mit einer einheitlichen Methodologie können die Vor- und Nachteile dieser innovativen Einnahmeregimes beurteilt werden.

- **Vergleich mit dem physiologischen Blutungsverhalten**

Das beschriebene Verfahren kann auch eingesetzt werden, um das jeweilige „normale“ Blutungsmuster von Frauen in den verschiedenen Lebensphasen Reproduktionsphase, Prämenopause, Perimenopause und Postmenopause zu beschreiben. Dadurch könnten durch hormonhaltige Präparate verursachte Abweichungen vom „normalen“ Blutungsmuster quantifiziert werden. Auch interkulturelle Vergleiche würden möglich.

Die beschriebenen Entscheidungsprobleme im Rahmen der arzneimittelrechtlichen Zulassung neuer Präparate treten auch bei der Bewertung von Präparaten im Rahmen der Frage der Erstattungsfähigkeit durch gesetzliche wie private Krankenkassen und Health Maintenance Organisationen (HMOs) auf. Durch einen wissenschaftlich fundierten Vergleich des Blutungsmusters der Präparate können auch die im Hinblick auf die Wirtschaftlichkeit der einzelnen Präparate gestellten Fragen besser beantwortet werden.

7 Methodologische Leitlinien zur Verschreibung und Zulassung

7.1 Einleitung

Ziel dieser methodologischen Leitlinien ist es, durch eine Vereinheitlichung der Erhebung, der Darstellung und der Bewertung von vaginalen Blutungsmustern die Vergleichbarkeit von unterschiedlichen Präparaten in sämtlichen Indikationen sexualsteoridhaltiger Präparate wie zum Beispiel hormonelle Empfängnisverhütung, Blutungsstörungen, Endometriose oder Hormonersatztherapie zu ermöglichen.

Eine Transparenz hinsichtlich des Blutungsverhaltens ist sowohl im Rahmen des Zulassungsverfahrens als auch für die individuelle Therapieentscheidung erforderlich, um eine rationale Auswahl der Inhaltstoffe (insbesondere Wahl des Gestagens) und der Dosis (insbesondere Wahl der Ethinylestradioldosis) zu treffen.

Die bestehende Klassifikation des Blutungsmusters der WHO (10) ist für den Vergleich des Blutungsmusters von Präparaten zur hormonellen Empfängnisverhütung und zur Hormonersatztherapie wenig geeignet, da sie im wesentlichen für die Bewertung von Langzeitpräparaten wie zum Beispiel Depotinjektionen entwickelt wurde.

Hierbei ist zu betonen, dass das Blutungsmuster immer nur ein Teilaspekt einer Entscheidung bezüglich einer Verschreibung oder einer Arzneimittelzulassung darstellt. Das Blutungsmuster ist jedoch ein wichtiges Entscheidungskriterium, da sich die derzeit auf dem Markt befindlichen Präparate hinsichtlich der Wirksamkeit und der Sicherheit nur wenig unterscheiden. Dabei darf nicht unbeachtet bleiben, dass sich die individuelle Verträglichkeit einer chemischen Substanz oder einer speziellen Applikationsform sehr wohl deutlich unterscheiden kann, auch wenn sich bei der Betrachtung von Kohorten keine oder nur sehr kleine Unterschiede hinsichtlich der mittleren Verträglichkeit zeigen.

7.2 Erhebung des Blutungsmusters

Zur Bewertung des Blutungsmusters sind tägliche Angaben über das Auftreten von vaginalen Blutungen sowie gegebenenfalls Angaben über deren Intensität erforderlich. Um den Einfluss von Erinnerungslücken auf die Blutungsmustererhebung möglichst zu minimieren, sollte das Blutungsverhalten täglich mit Hilfe eines Blutungskalenders prospektiv erfasst werden.

Die Erhebung des Blutungsmusters kann in gleicher Weise sowohl für die individuelle Verschreibung als auch für die Zulassung von hormonhaltigen Präparaten durchgeführt werden.

Die Blutungsstärke sollte gemäß den WHO-Empfehlungen (10) als „Spotting“, „Blutung“, bzw. als „keine Blutung“ klassifiziert erhoben werden. Hierbei ist „Blutung“ als eine vaginale Absonderung definiert, welche die Anwendung mindestens einer Monatsbinde oder eines Tampons erforderlich macht. Im Gegensatz dazu wird eine vaginale Absonderung, die keine speziellen Schutzmaßnahmen erfordert, als „Spotting“ bezeichnet. Die Verwendung von dünnen Slipelinlagen, wie z.B. von Always Ultra für den ständigen Gebrauch, schließt die Kategorien „Spotting“ oder „keine Blutungen“ nicht aus. Welche Art von Binden bzw. Tampons tatsächlich verwendet wurden, ist bei der Bestimmung der Blutungsstärke unerheblich, wichtig ist nur, welche Art erforderlich gewesen wäre. Falls an einem Tag sowohl „Spotting“ als auch „Blutung“ auftreten, so wird die Blutungsstärke dieses Tags als Blutungstag gewertet, da die Blutung das schlechtere der beiden Fälle ist. Ein Tag, an dem keine vaginale Blutung auftrat, wird als „keine Blutung“ kategorisiert.

Genaugenommen müsste festgelegt werden, welche Periode von 24 Stunden als „Tag“ definiert wird. Die übliche Definition eines Kalendertages von 0 Uhr bis 24 Uhr kann in der Praxis kaum angewandt werden. Eine genaue kalendertägliche Erfassung würde eine Beobachtung um Mitternacht erfordern, um eine nächtliche Blutung dem richtigen Kalendertag zuordnen zu können. Blutungsereignisse erstrecken sich jedoch normalerweise über mehrere Tage, so dass es für die

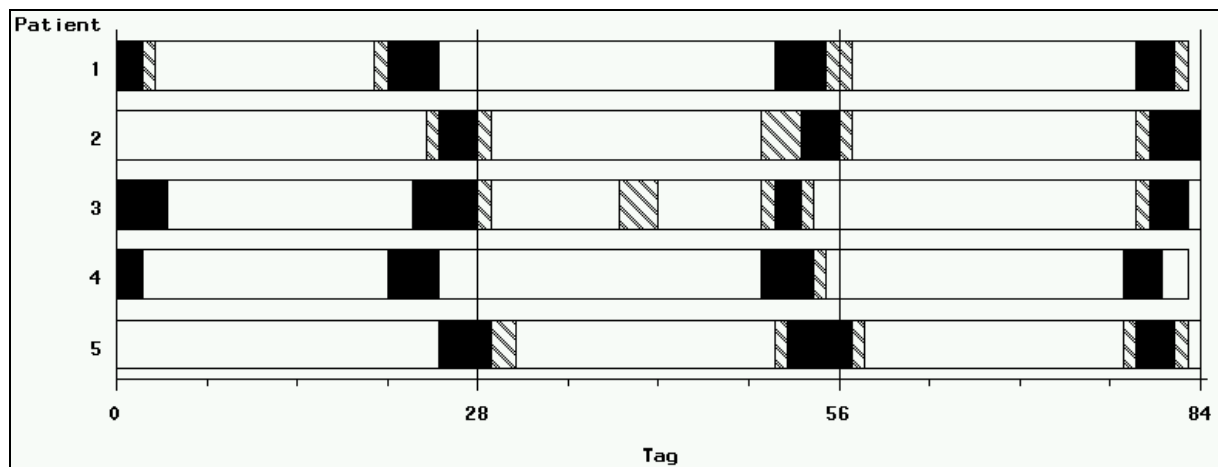
Vergleichbarkeit der erhobenen Daten von untergeordneter Bedeutung ist, wann die individuelle Frau ihren Beobachtungstag beginnt. Sie sollte nur dazu angehalten werden, ihre Definition über den gesamten Beobachtungszeitraum nach Möglichkeit beizubehalten.

7.3 Methodologische Leitlinie zur Verschreibung

7.3.1 Darstellung des Blutungsmusters für Einzelfälle

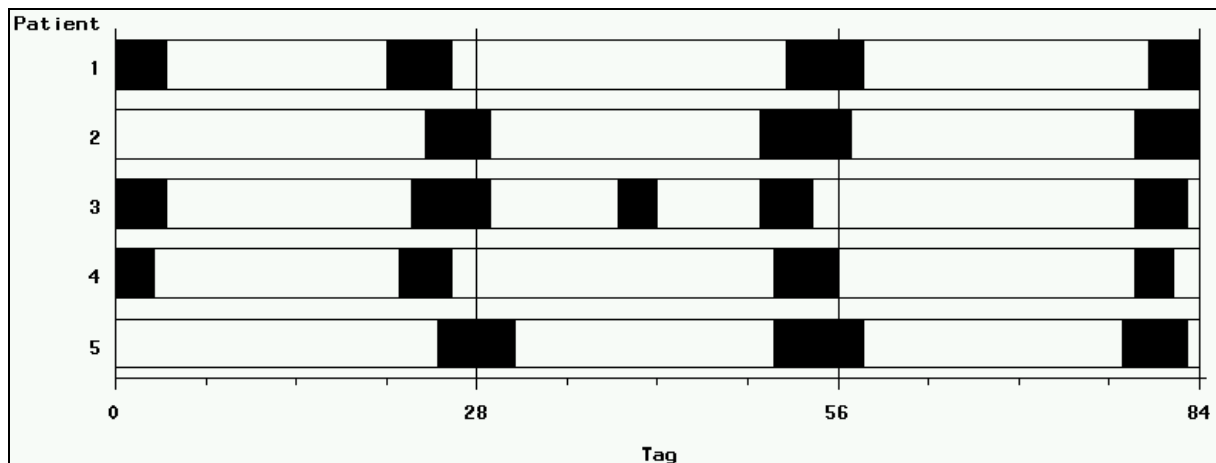
Bei der Darstellung von Blutungsmustern ist zwischen einer Einzelfallbetrachtung und einer Betrachtung von größeren Kohorten zu unterscheiden. Die Blutungsmuster einzelner Frauen lassen sich sehr anschaulich mit Hilfe von graphischen Methoden darstellen.

Für die Darstellung von einzelnen Blutungsmustern ist die graphische Darstellung geeignet. Aussagefähig ist eine spezielle Strichgraphik (vgl. Abbildung 30), bei der die Tage mit Blutung schwarz, die Tage mit Spotting gestreift und die Tage ohne Blutung weiß dargestellt sind. Die Zeitachse ist auf der Abszisse dargestellt. Die Periodizität des Blutungsmusters lässt sich auf der Graphik gut erkennen. Zur optischen Unterstützung wurden Referenzlinien alle vier Wochen eingezogen, die den Start eines neuen Pillenzyklus anzeigen.



**Abbildung 30: Graphische Darstellung des Blutungsmusters von Einzelfällen
 (Blutung schwarz und Spotting gestreift dargestellt)**

Diese Darstellung kann vereinfacht werden, in dem nicht zwischen den verschiedenen Blutungsstärken unterschieden wird (vgl. Abbildung 31). Bei dieser Form der Darstellung werden Tage mit Blutung und/oder Spotting schwarz und die Tage ohne Blutung weiß dargestellt.



**Abbildung 31: Graphische Darstellung des Blutungsmusters von Einzelfällen
 (Blutung und Spotting zusammengefasst)**

In beiden Darstellungen ist die Zwischenblutung der Patientin mit der Ordnungsnummer 3 im zweiten Zyklus gut zu erkennen. Für die Betrachtung des Einzelfalles ist die beschriebene graphische Darstellung völlig ausreichend. Die weitergehende Berechnung von Blutungsmusterindices (vgl. Tabelle 24 und Tabelle 25) führt nicht zu einem weiteren relevanten Erkenntnisgewinn für die Beurteilung von Einzelfällen.

7.3.2 Bewertung des Blutungsmusters für Einzelfälle

Bei der Bewertung des Blutungsmusters muss hinsichtlich der Indikation des zu untersuchenden Präparates unterschieden werden. So ist die regelmäßige Entzugsblutung bei einem Kontrazeptivum für die Anwenderin wichtig, da sie ein wesentliches Indiz dafür ist, dass die Anwenderin nicht schwanger ist. Bei einem sequenziellen Präparat zur hormonellen Substitution spielt dieser Aspekt hingegen keine Rolle. Ein Ausbleiben der regelmäßigen Entzugsblutung ist daher im Allgemeinen bei einem Präparat zur hormonellen Empfängnisverhütung als schwerwiegender zu beurteilen als bei einem sequenziellen Hormonersatzpräparat

zur Osteoporoseprophylaxe. Im Einzelfall kann die Wertigkeit bei HRT jedoch ebenso schwer wiegen.

Das Blutungsmuster kann meist schon anhand der graphischen Darstellung von Arzt und Anwenderin hinreichend beurteilt werden. Darüber hinaus können jedoch ohne weiteres die individuellen Blutungsmusterindices (vgl. Tabelle 24 und Tabelle 25) berechnet werden. Diese können dann mit den mittleren Ergebnissen aus den klinischen Prüfungen des betreffenden Präparates verglichen werden. Größere Abweichungen können auf eine bessere oder schlechtere individuelle Verträglichkeit hinweisen.

7.3.3 Empfehlung zur Präparatauswahl und zur Dosistitration anhand des Blutungsmusters

Das Blutungsmuster ist nur ein Kriterium unter vielen für die Auswahl eines geeigneten Präparates zur hormonellen Kontrazeption oder zur Hormonersatztherapie. Es ist jedoch für die Anwenderin ein Kriterium von zentraler Bedeutung, da das Blutungsmuster deutliche Auswirkungen auf die Lebensqualität der Frau haben kann. Ein schlechtes Blutungsmuster ist der häufigste Grund für den vorzeitigen Abbruch der Hormonersatztherapie bzw. den Wechsel der Verhütungsmethode oder des Präparates durch die Anwenderin.

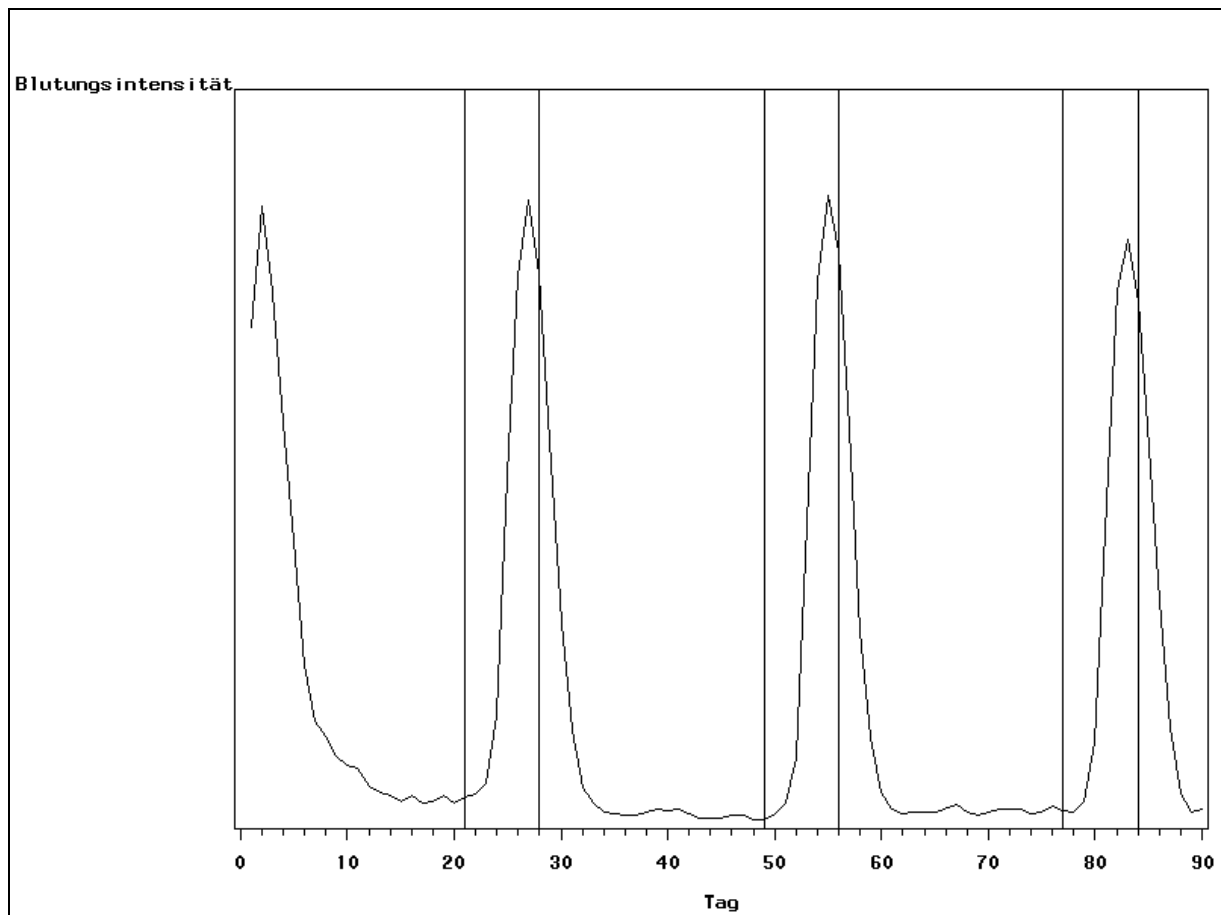
Ein Vergleich der individuellen Blutungsdaten mit den mittleren Ergebnissen aus den klinischen Prüfungen des betreffenden Präparates kann auf eine bessere oder schlechtere individuelle Verträglichkeit der Inhaltsstoffe hinweisen. Ein Wechsel des Präparates kann in Betracht gezogen werden. Unter der Voraussetzung, dass die Einnahmecompliance gut war, kann bei einer unbefriedigenden Blutungskontrolle die Östrogendosis gegebenenfalls erhöht werden. Bei einer zufriedenstellenden Blutungskontrolle kann die Östrogendosis eventuell weiter gesenkt werden, ohne die Gestagenkomponente zu wechseln.

7.4 Methodologische Leitlinie zur Zulassung

7.4.1 Darstellung des Blutungsmusters

Aussagen über das Blutungsverhalten einer größeren Anzahl von Frauen lassen sich nur mit den Methoden der beschreibenden Statistik treffen. Eine Darstellung der Einzelverläufe mit Hilfe von graphischen Methoden wie für das Blutungsmuster einzelner Frauen wäre zu unübersichtlich, um daraus Schlüsse ziehen zu können.

Die in Abschnitt 7.3.1 beschriebene graphische Darstellung von Einzelverläufen des Blutungsmusters ist zwar sehr anschaulich, sie findet jedoch rasch ihre Grenzen, wenn die Daten vieler Frauen betrachtet werden sollen. Als graphische Darstellung des Blutungsverhaltens über die Zeit hat sich hier die mittlere Blutungsintensität bewährt (vgl. Abbildung 32). In der Graphik sind der Beginn und das Ende des hormonfreien Intervalls durch senkrechte Linien dargestellt.



Nota: Die senkrechten Linien markieren Beginn und Ende des einnahmefreien Zeitraums. Im ersten Zyklus ist ein Peak am Anfang, da die Pilleneinnahme lt. Beipackzettel am ersten Tag einer Monatsblutung beginnen soll.

Abbildung 32: Graphische Darstellung der mittleren Blutungsintensität für eine Kohorte von OC Einnehmerinnen

Zum Vergleich der Blutungsmuster eines neuen Präparates zur Hormonersatztherapie mit einem herkömmlichen Präparat in dieser Indikation werden beispielsweise klinische Prüfungen an mehreren hundert von Frauen durchgeführt. Es ist daher zweckmäßig, das Blutungsmuster jeder einzelnen Frau durch geeignete Blutungsmusterindices zu beschreiben und die

Blutungsmusterindices anschließend mittels schließender und mittels beschreibender statistischer Verfahren zu analysieren.

In Anlehnung an die Definitionen der WHO (10) werden zunächst Blutungsmusterindices für alle Präparate definiert, die später für die statistische Analyse herangezogen werden. Hierbei werden jedoch längst nicht alle von der WHO beschrieben Blutungsmusterindices für die vergleichende Bewertung von Blutungsmustern benötigt.

Tabelle 24: Blutungsmusterindices für alle Hormon-Präparate

Blutungsmusterindices pro Subjekt und Referenzperiode
• Anzahl der Blutungstage mit Spotting und/oder Blutung • Anzahl der Blutungstage mit nur Spotting
• Anzahl arithmetisches Mittel, Maximum der Länge und Spannweite der Länge der Blutungsepisoden mit Spotting und/oder Blutung • Anzahl arithmetisches Mittel, Maximum der Länge und Spannweite der Länge der Blutungsepisoden mit nur Spotting

Die vorstehend genannten Blutungsmusterindices werden für jede der jeweils 90-tägigen Referenzperioden mit den statistischen Kenngrößen Anzahl der nicht fehlenden Beobachtungen, arithmetrischer Mittelwert, Stichprobenstandardabweichung, Minimum, erster Quartil, Median, drittes Quartil und Maximum beschrieben.

Zusätzlich werden für zyklische Einnahmeregime die einzelnen Blutungsepisoden in Entzugsblutungsepisoden und in Zwischenblutungsepisoden unterschieden. Die genaue Klassifizierung der Blutungsepisoden in Entzugsblutungsepisoden und in Zwischenblutungsepisoden lässt sich wie folgt zusammenfassen:

- Ist am Tag des Hormonentzugs eine Blutungsepisode im Gange, so ist diese Blutungsepisode die Entzugsblutungsepisode dieses Zyklus, vorausgesetzt, dass diese Blutungsepisode nicht früher als 4 Tage vor dem Hormonentzug begonnen hat.
- Ist am Tag des Hormonentzugs keine Blutungsepisode im Gange, so ist die nächste Blutungsepisode die Entzugsblutungsepisode dieses Zyklus, vorausgesetzt, dass diese Blutungsepisode nicht später als 5 Tage vor dem Hormonentzug des darauffolgenden Zyklus begonnen hat.
- Alle anderen Blutungsepisoden werden als Zwischenblutungsepisoden kategorisiert. Sie werden dem Zyklus zugeordnet, in dem sie beginnen.

Tabelle 25: Blutungsmusterindices für alle zyklisch verabreichte Präparate

Blutungsmusterindices pro Subjekt und Zyklus
• Entzugsblutung aufgetreten (ja/nein)
• Länge der Entzugsblutung in Tagen
• maximale Blutungsintensität der Entzugsblutung
• Beginn der Entzugsblutung relativ zum Hormonentzug
• Zwischenblutung aufgetreten (ja/nein)
• Anzahl der Zwischenblutungstage
• Anzahl und maximale Länge der Zwischenblutungsepisoden
• maximale Blutungsintensität der Zwischenblutungsepisode(n)

Die vorstehend in Tabelle 25 genannten Blutungsmusterindices werden für jeden Zyklus mit den statistischen Kenngrößen Anzahl der nicht fehlenden Beobachtungen, arithmetischer Mittelwert, Stichprobenstandartabweichung, Minimum, erster Quartil, Median, drittes Quartil und Maximum beschrieben.

Für konfirmatorische Studien zur Zykluskontrolle ist, je nach der Dauer der Studie, die untersuchte Hauptzielgröße in der Regel das Auftreten von einer oder mehrerer Zwischenblutungen in den Zyklen 2 bis 6, 2 bis 13 oder 2 bis 26. Der erste Zyklus wird nicht betrachtet, da es bei Frauen, die mit der hormonellen Verhütung beginnen, wegen der Umstellung von der natürlichen Menstruationsblutung zu einer durch exogene Hormone gesteuerten Entzugsblutung vorübergehend zu

Zwischenblutungen kommen kann, was jedoch für die Beurteilung von Präparaten zur langfristigen Einnahme von vermindertem Belang ist.

Je nach Studienziel wird eine der folgenden statistischen Hypothesen gewählt.

Tabelle 26 Hypothesen und Tests für Zykluskontrollstudien

Test	Nullhypothese	Alternativ-hypothese	statistischer Test
Test auf Unterschied	$H_0: \pi_T = \pi_R$	$H_1: \pi_T \neq \pi_R$	zweiseitiger exakter Test von Fisher
Test auf Überlegenheit	$H_0: \pi_T \geq \pi_R$	$H_1: \pi_T < \pi_R$	einseitiger exakter Test von Fisher
Test auf Nichtunterlegenheit (einseitige Äquivalenz)	$H_0: \pi_T \geq \pi_R + \delta$	$H_1: \pi_T < \pi_R + \delta$	einseitiger Test, siehe Fußnote ⁷ .

hierbei bezeichnet π die Wahrscheinlichkeit für eine Zwischenblutung, T das getestete Präparat, R das Referenzpräparat und δ die Nichtunterlegenheitsschranke.

⁷ H_0 wird abgelehnt wenn die obere einseitige 95% Konfidenzgrenze (Normalapproximation)

$$p_{0.95} = (\hat{p}_T - \hat{p}_R) + z_{0.95} \sqrt{\frac{\hat{p}_T(1-\hat{p}_T)}{n_T} + \frac{\hat{p}_R(1-\hat{p}_R)}{n_R}}$$

des Behandlungsunterschieds ($\pi_T - \pi_R$) kleiner als δ ist. Hierbei bezeichnet $z_{0.95}$ das 95% Quantil der Standardnormalverteilung und n_T bzw. n_R die Anzahl der Beobachtungen in den jeweiligen Behandlungsgruppen.

Die Nichtunterlegenheitsschranke δ wird gemäß den internationalen Empfehlungen (131) festgelegt.

Für Studien, in denen mehr als zwei Behandlungen miteinander verglichen werden, müssen die Hypothesen entsprechend angepasst werden. Hierbei ist zu beachten, dass der exakte Test von Fisher durch den χ^2 -Test ersetzt werden sollte. Die Fallzahl für Zykluskontrollstudien ist in der Regel so hoch, so dass der exakte Test von Fisher aufgrund seiner numerischen Komplexität nicht mehr in sinnvoller Zeit berechnet werden kann.

Für die Tests sollte gemäß den internationalen Gepflogenheiten (132) ein Signifikanzniveau α von 5% im Falle eines zweiseitigen Tests und ein Signifikanzniveau α von 2,5% im Falle eines einseitigen Tests verwendet werden.

7.5 Bewertung des Blutungsmusters im Rahmen der Arzneimittelzulassung

Bei der Bewertung des Blutungsmusters muss hinsichtlich der Indikation des zu untersuchenden Präparates unterschieden werden. So ist die regelmäßige Entzugsblutung bei einem Kontrazeptivum für die Anwenderin wichtig, da sie ein wesentliches Indiz dafür ist, dass die Anwenderin nicht schwanger ist. Bei einem Präparat zur hormonellen Substitution spielt dieser Aspekt hingegen keine so wesentliche Rolle. Ein Ausbleiben der regelmäßigen Entzugsblutung ist daher bei einem Präparat zur hormonellen Empfängnisverhütung schwerwiegender zu beurteilen als bei einem Hormonersatzpräparat zur Osteoporoseprophylaxe.

Als Zulassungsvoraussetzung für ein neues Präparat müsste verlangt werden, dass das neue Präparat kein wesentlich schlechteres Blutungsmuster verursacht als bereits auf dem Markt befindliche Präparate. Dies gilt selbstverständlich nur dann, wenn das neue Präparat keinen anderen medizinisch relevanten Vorteil bezüglich

Wirksamkeit oder Sicherheit hat, der die Nutzen-Risiko-Abwägung erheblich beeinflusst.

Für Präparate, die ein neuartiges Blutungsmuster erzeugen sollen, wie zum Beispiel verlängerte Einnahmeperioden für orale Kontrazeptiva zur Reduktion der Anzahl der Entzugsblutungen, kann die Bewertung im Vergleich zu den bereits zugelassenen Präparaten erfolgen.

8 Zusammenfassung und Ausblick

Das Blutungsmuster ist ein wichtiges Kriterium für die Beurteilung von hormonalen Präparaten wie zum Beispiel orale Kontrazeptiva oder Präparate zur Hormonersatztherapie, da die Einnahme eines solchen Präparates stets auch das Blutungsmuster beeinflusst (4). Die Einnahme eines sexualhormonhaltigen Präparates verändert das natürliche Blutungsmuster, da das Blutungsverhalten nicht mehr durch die endogenen zyklischen Hormonspiegelschwankungen, sondern durch die exogene Hormoneinnahme bestimmt wird. Unregelmäßige und nicht vorhersehbare Blutungen können die Lebensqualität der Frau erheblich einschränken, da sie im täglichen Leben hinderlich sind.

Für die arzneimittelrechtliche Zulassung neuer Präparate zur Empfängnisverhütung wird vom Committee for Proprietary Medicinal Products (CPMP) der europäischen Arzneimittelagentur (EMA) zusätzlich zum Nachweis der Sicherheit (Nebenwirkungsprofil) und der Wirksamkeit (Ovulationshemmung bzw. Schwangerschaftsverhütung) die Analyse des Blutungsmusters verlangt. Während die Anforderungen an die Nachweise für Sicherheit und Wirksamkeit (vgl. Gerlinger et al. 7) genau spezifiziert sind, sind die Anforderungen der europäischen Arzneimittelagentur an die Erhebung, die Auswertung und die Darstellung des Blutungsmusters sind jedoch recht vage. Obwohl sich das Blutungsmuster verschiedener steroidaler empfängnisverhütender Medikamente zum Teil deutlich unterscheidet, wird in der EMA-Leitlinie zu den Zulassungsvoraussetzungen (8) lediglich gefordert, dass das Blutungsmuster in einer klinischen Studie mit einem aktiven Vergleichspräparat erhoben werden muss.

Die Anforderungen der europäischen Arzneimittelagentur EMA für die Zulassung von hormonhaltigen Präparaten zur Behandlung von Klimakteriumsbeschwerden und zur Osteoporoseprophylaxe sind in Bezug auf das Blutungsmuster nicht wesentlich präziser (9).

Ziel dieser Arbeit war es, allgemeine methodische Leitlinien zur Beurteilung von Blutungsmustern zu entwickeln. Eine methodische Leitlinie soll den Vergleich zwischen den Auswertungen des Blutungsmusters verschiedenerer pharmazeutischer Hersteller im Rahmen der arzneimittelrechtlichen Zulassung neuer Präparate ermöglichen. Die andere methodische Leitlinie soll eine Hilfe bei der individuellen Verschreibung von hormonalen Kontrazeptiva und Hormonersatzpräparaten sein, da eine optimierte Blutungskontrolle entscheidend zur Akzeptanz eines Präparates bei der Anwenderin beiträgt.

Die methodischen Leitlinien zur Datenerhebung, Analyse und Bewertung von Blutungsmustern wurde im wesentlichen in zwei Schritten erstellt. Zum einen wurde eine systematische Literaturrecherche durchgeführt, um die derzeit verwendeten Methoden der pharmazeutischen Unternehmer und der von pharmazeutischen Herstellern unabhängigen Wissenschaftler zu erfassen und zu bewerten.

Zum anderen wurde eine empirische Untersuchung von insgesamt 4612 auswertbaren Blutungskalendern, die in klinischen Prüfungen von mehreren hormonhaltigen Präparaten in verschiedenen Indikationen geführt worden waren, durchgeführt. Ziel der empirischen Untersuchung war es, in den Blutungskalendern vorhandene Muster durch statistische Methoden zu erkennen und die so gefundenen Muster anschließend zu beschreiben.

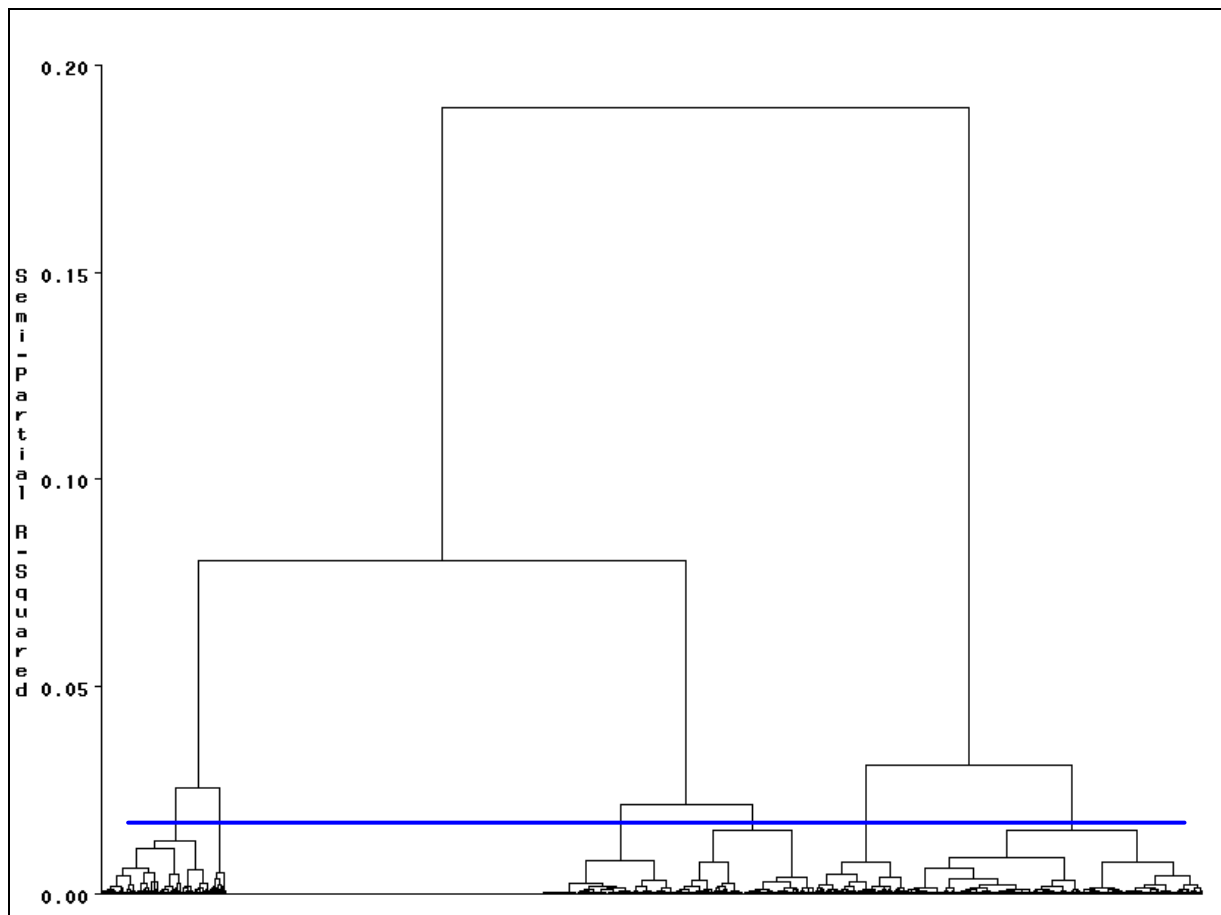
Hervorzuheben ist die Tatsache, dass die beiden in der vorliegenden Arbeit verwandten Erkenntnisquellen zur Formulierung der methodischen Leitlinien zur Beurteilung von Blutungsmustern an zwei diametral entgegengesetzten Ausgangspunkten beginnen. Bei der systematischen Literaturschau wird publiziertes Wissen möglichst vollständig erfasst, während die empirische Analyse ein Clusterverfahren zur automatischen Klassifikation ohne jegliches Vorwissen verwendet.

Die Literaturrecherche ergab zwei Klassen von Methoden, Zyklus-basierte und Episoden-basierte Methoden. Die Zyklus-basierten Blutungsmuster-Methoden haben

den Vorteil der einfachen Definitionen und der im Vergleich zu den Episoden-basierten Blutungsmuster-Methoden unkomplizierteren Datenerfassung. Der wesentliche Nachteil der Zyklus-basierten Blutungsmuster-Methoden liegt hingegen darin, dass nicht unterschieden wird, wie häufig unerwünschte Blutungsereignisse auftreten. Insbesondere im Falle von Durchbruchblutungen, d.h. von Zwischenblutungen, welche die Verwendung eines sanitären Produktes unbedingt erfordern, kommt es schon darauf an, ob es sich um ein Ereignis von dreitägiger Dauer handelt oder ob es sich um drei verschiedene Ereignisse handelt. Diese Unterscheidung wird nur bei den Episoden-basierten Methoden getroffen.

In die empirische Untersuchung konnten von den insgesamt 5602 zur Verfügung stehenden Blutungskalendern 4612 (82,3%) aufgenommen werden. Es wurden alle Blutungskalender ausgewertet, die mindestens 90 Tage lang geführt wurden und die keine unüberbrückbaren Lücken (vgl. Abschnitt 4.1.1) hatten.

Es wurden statistische Clusteranalysen mit den Methoden single linkage, complete linkage, average linkage sowie mit dem Verfahren von Ward durchgeführt. Hierbei zeigte es sich, dass von diesen Verfahren zur Clusteranalyse nur die Methode von Ward für die empirische Analyse der Blutungskalender geeignet war. Als mögliche Lösungen der Clusteranalyse wurden die Lösungen mit drei, vier, sechs und zwölf Clustern als möglicherweise substanzwissenschaftlich interpretierbare Blutungsmuster erkannt. Beim Vergleich der verschiedenen Lösungen der Clusteranalyse stellte sich heraus, dass die Lösung mit sechs Clustern die substanzwissenschaftlich am besten interpretierbare Lösung war. Diese Lösung ist in Abbildung 33 durch die waagrechte Linie dargestellt.



Nota bene: Die scheinbare Lücke im Dendrogramm stellt die 1288 identischen Blutungskalender dar (vgl. Tabelle 9).

Abbildung 33: Dendrogramm der Clusteranalyse nach Ward mit optimaler Anzahl der Cluster

Bei der empirisch ermittelten Lösung wurden vier Blutungsmuster für kontinuierlich verabreichte Hormonpräparate und zwei Blutungsmuster für zyklisch verabreichte Hormonpräparate unterschieden.

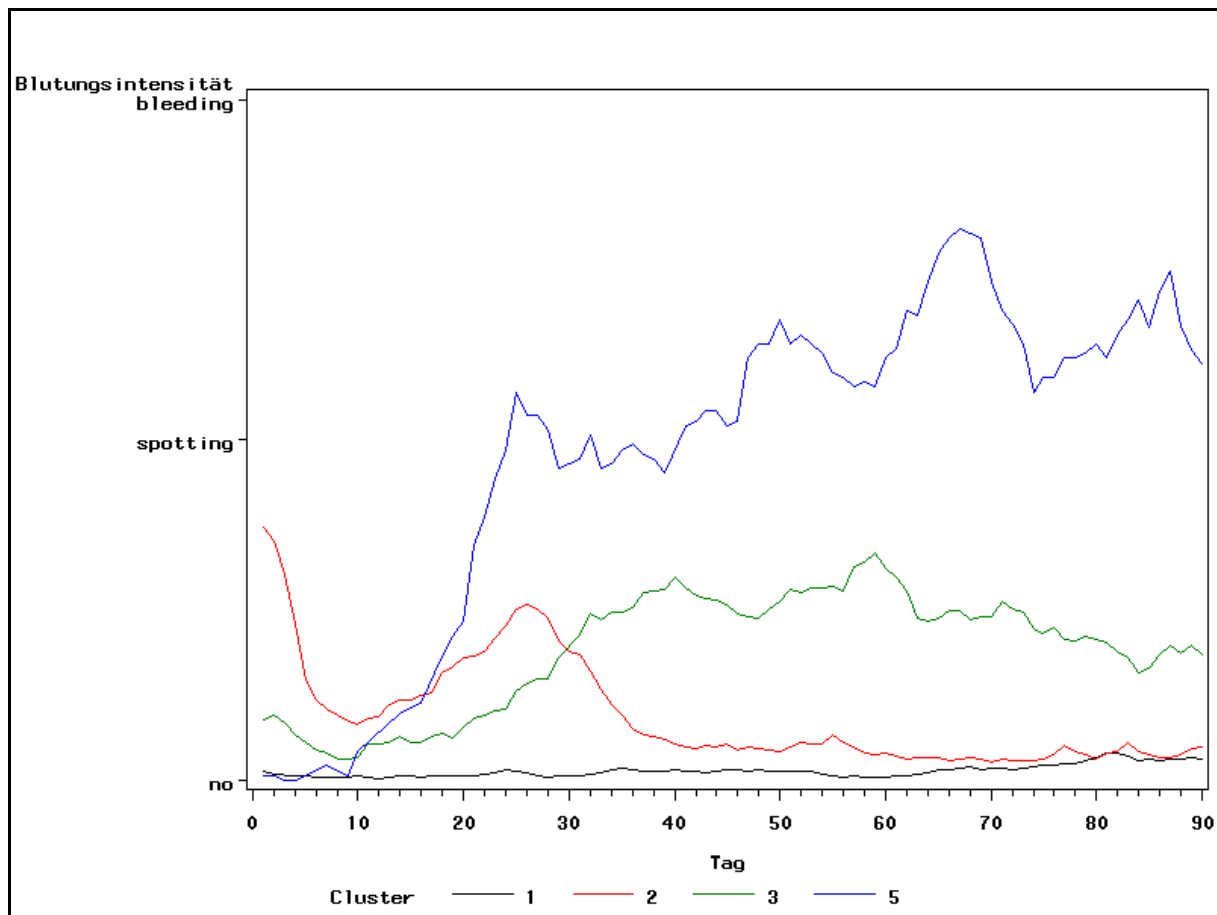


Abbildung 34: Mittlere Blutungsintensität über die Zeit - 4 kontinuierlich verabreichte Präparate

Das erwünschte Blutungsmuster für kontinuierlich verabreichte Präparate wird in Abbildung 34 durch Cluster Nummer 1 dargestellt. Dieses Blutungsmuster entspricht dem natürlichen Blutungsmuster postmenopausaler Frauen, bei dem keine vaginale Blutung mehr stattfindet.

Das durch Cluster Nummer 2 dargestellte Blutungsmuster kann als akzeptabel für kontinuierliche Hormonpräparate bezeichnet werden. Die mittlere Blutungsintensität ist zu Beginn der Behandlung erhöht, sinkt dann rasch ab um nach etwa vierwöchiger Behandlung einen deutlichen Peak zu haben. Danach klingt die mittlere Blutungsintensität ab und erreicht zum Ende des Beobachtungszeitraumes von 90 Tagen das Niveau des Clusters Nummer 1. Dieses Blutungsmuster beschreibt sehr

gut die Umstellung des Blutungsverhaltens von Frauen, die vor Beginn der kontinuierlichen Behandlung mit Hormonen ein natürliches zyklisches Blutungsverhalten zeigen.

Die Cluster Nummer 3 und Cluster Nummer 5 zeigen unerwünschte Blutungsmuster, da es hier zu einem deutlichen Anstieg der Blutungen nach 20 bis 30 Tagen Behandlung kommt.

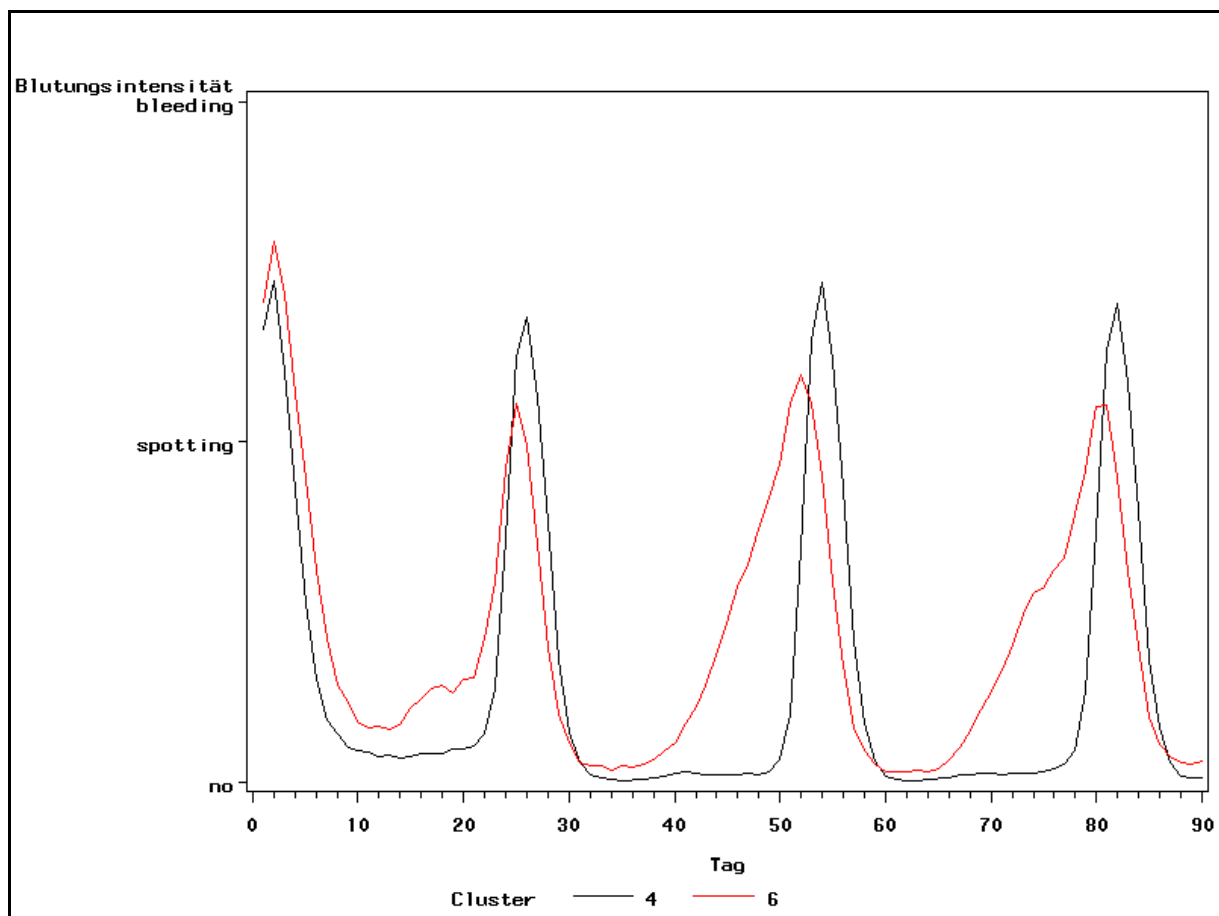


Abbildung 35: Mittlere Blutungsintensität über die Zeit - zyklisch verabreichte Präparate

Die beiden zyklischen Blutungsmuster in Abbildung 35 lassen sich in ein erwünschtes und ein unerwünschtes Blutungsmuster einteilen. Das durch Cluster Nummer 4 dargestellte Blutungsmuster zeigt das erwünschte Blutungsmuster mit deutlichen Abbruchblutungen nach dem Hormonentzug in jedem Zyklus ohne wesentliche Zwischenblutungen. Im Gegensatz dazu zeigt Cluster Nummer 6 ein unerwünschtes zyklisches Blutungsmuster mit vermehrten Zwischenblutungen. Eine weitere Unterteilung dieser beiden Cluster erwies sich als nicht sinnvoll (vgl. Abbildung 22).

Um die Ergebnisse der empirischen Analyse der Blutungsmuster mit denen der Literatúrauswertung vergleichen zu können, wurden in der Literatur gefundene Methoden auf die selben Blutungskalender angewandt, auf denen auch die empirische Analyse beruhte. Somit lassen sich die Klassifizierungen der Literatur unmittelbar mit der empirisch gefundenen Clusterstruktur der Blutungskalender vergleichen.

Der Vergleich beschränkte sich auf die Erkenntnisse zum Blutungsmuster, da ein Vergleich der verschiedenen Definitionen der Blutungsstärke aus Mangel an empirischen Daten nicht möglich war. Da die verschiedenen Definitionen der Blutungsstärke in der Literatur ähnlich sind, ist dieser Mangel an Daten für die vorliegende Arbeit nur von unerheblicher Bedeutung.

Aufbauend auf den Ergebnissen der Literaturanalyse und der empirischen Analyse wurde ein modifiziertes Verfahren zur Erhebung und Auswertung von Blutungsmusterdaten erstellt (vgl. Abschnitt 5.3):

Die Blutungsstärke sollte in Anlehnung an eine Definition der Weltgesundheitsorganisation täglich in den Kategorien

- „spotting“
- „bleeding“

sowie

- „no bleeding or spotting“

erhoben werden. Hierbei ist „spotting“ als ein blutiger vaginaler Ausfluss, der jedoch keine Verwendung von Schutzmaßnahmen, wie Monatsbinden oder Tampons, erforderlich macht, definiert. Die Kategorie „bleeding“ ist hingegen definiert als ein blutiger vaginaler Ausfluss, der eine Verwendung von Schutzmaßnahmen erforderlich macht. Tritt kein vaginaler Ausfluss auf, so wird dieses Ereignis als „no bleeding or spotting“ bezeichnet.

Im nächsten Schritt werden zusammenhängende Tage, an denen eine vaginale Blutung stattfand, zu einer Blutungsepisode (engl. „bleeding episode“) zusammengefasst. Hierbei ist zu beachten, dass einzelne Tage ohne Blutung eine Blutungsepisode nicht beenden. Eine Blutungsepisode ist folglich eine Folge von einem oder mehreren Tagen, an denen geblutet wurde, die auf beiden Seiten von jeweils mindestens zwei blutungsfreien Tagen begrenzt wird.

In einem dritten Schritt werden bei zyklisch verabreichten Präparaten die Blutungsepisoden in Entzugsblutungsepisoden und in Zwischenblutungsepisoden unterschieden. Die genaue Klassifizierung der Blutungsepisoden in Entzugsblutungsepisoden und in Zwischenblutungsepisoden lässt sich wie folgt zusammenfassen:

- Ist am Tag des Hormonentzugs eine Blutungsepisode im Gange, so ist diese Blutungsepisode die Entzugsblutungsepisode dieses Zyklus, vorausgesetzt, dass diese Blutungsepisode nicht früher als 4 Tage vor dem Hormonentzug begonnen hat.
- Ist am Tag des Hormonentzugs keine Blutungsepisode im Gange, so ist die nächste Blutungsepisode die Entzugsblutungsepisode dieses Zyklus, vorausgesetzt, dass diese Blutungsepisode nicht später als 5 Tage vor dem Hormonentzug des darauffolgenden Zyklus begonnen hat.

- Alle anderen Blutungsepisoden werden als Zwischenblutungsepisoden kategorisiert. Sie werden dem Zyklus zugeordnet, in dem sie beginnen.

Die vorgeschlagene neue Methode bietet einige Vorzüge gegenüber den in der Literatur verwendeten Methoden. Sie verbindet die Vorzüge der Zyklus-basierten mit denen der Episoden-basierten Methoden.

Die neue Methode fasst zunächst aufeinander folgende Tage mit Blutungen zu einer Episode zusammen. Damit ist es ausgeschlossen, dass ein Blutungsereignis sowohl als Entzugs- als auch als Zwischenblutung gewertet wird, wie dies bei den Zyklus-basierten Methoden auftreten kann (vgl. Abschnitt 3.3.3.1).

Die Trennung der Blutungsepisoden in Entzugs- und Zwischenblutungsepisoden wird durch die neue Methode in eindeutiger Weise gewährleistet, während es bei den bisher gebräuchlichen Methoden durchaus zu Klassifizierungsproblemen kommen konnte (vgl. Abschnitt 5.3.4).

Ein noch offenes Problem ist die Behandlung von extrem langen Blutungsepisoden (vgl. Abbildung 29). Dieses Problem besteht jedoch auch bei den bisherigen Methoden und stellt daher keinen spezifischen Nachteil der vorgeschlagenen Methode dar.

Zusammenfassend lässt sich sagen, dass die neue Methode den bisherigen Methoden eindeutig überlegen ist. Eine Ausnahme bildet lediglich die Behandlung von extrem langen Blutungsepisoden, die weder von der neuen, noch von den bisherigen Methoden voll zufriedenstellend gelöst wird.

Abschließend wurden die Ergebnisse der Arbeit in zwei methodologische Leitlinien zur Berücksichtigung des Blutungsmusters bei der arzneimittelrechtlichen Zulassung und bei der individuellen Verschreibung zusammengefasst. Während die Datenerhebung für die Analyse der Blutungsmuster für Verschreibung und arzneimittelrechtliche Zulassung in gleicher Weise durchgeführt werden kann, gibt es

Unterschiede in der Auswertung und der Bewertung der Blutungsmuster, je nachdem ob Einzelfälle oder Kohorten betrachtet werden sollen.

Es steht zu hoffen, dass die Ergebnisse dieser Arbeit Eingang in die Richtlinien der Medikamentenzulassung, beispielsweise in die Note for Guidance on Clinical Evaluation of Steroid Contraceptives in Women (8) und die Points to Consider on Hormone Replacement Therapy (9) der EMEA, finden. Dies hätte zur Folge, dass die hier vorgeschlagene Methode von allen pharmazeutischen Herstellern angewandt wird und dass die Ergebnisse entsprechend publiziert werden. Dann können die Auswirkungen der verschiedenen Präparate auf das Blutungsmuster fundiert miteinander verglichen werden, auch wenn die Daten aus verschiedenen Studien und von verschiedenen Herstellern stammen.

Darüber hinaus sollten die Ergebnisse dieser Arbeit auch einen Beitrag zu einer optimierten individuellen Verschreibung von hormonhaltigen Präparaten leisten, da das Blutungsmuster die individuelle Lebensqualität nachhaltig beeinflussen kann.

9 Abkürzungen

CPMP	Committee for Proprietary Medicinal Products der europäischen Arzneimittelagentur (EMA)
CRF	Case Record Form
EMA	europäische Arzneimittelagentur mit Sitz in London
engl.	englisch
FDA	Food and Drug Administration (USA)
GCP	Good Clinical Practices
H_0	Nullhypothese
H_1	Alternativhypothese
HMO	Health Maintenance Organisation
HRT	Hormonersatztherapie (Hormone Replacement Therapy)
IC	injizierbares Kontrazeptivum oder Implantatkontrazeptivum
ICH	International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use
ICD-10	International Classification of Diseases, 10te Version
IUD	Intrauterinpessar
LOCF	Last Observation Carried Forward (Ersetzungsalgorithmus für fehlende Werte)

NIH	National Institute of Health (USA)
NR	nicht relevant
NV	nicht verfügbar
OC	orales Kontrazeptivum
PubMed	Datenbank medizinischer Literatur der US-amerikanischen National Library of Medicine
SAG	Schering AG
TC	transdermales Kontrazeptivum
VR	Vaginalring zur Kontrazeption
WHO	World Health Organization der Vereinten Nationen mit Sitz in Genf

10 Literatur

- 1 Kösterling S. „Etwas Besseres als das Kondom“ Ludwig Haberlandt und die Idee der Pille. in: Staupe G, Vieht L (Hrsg.). Die Pille – Von der Lust und von der Liebe. Berlin, Rowohlt, 1996.
- 2 Sieg S. „Anovlar“ – die erste europäische Pille. in: Staupe G, Vieht L (Hrsg.). Die Pille – Von der Lust und von der Liebe. Berlin, Rowohlt, 1996.
- 3 Asbell B. Die Pille und wie sie die Welt veränderte. Frankfurt am Main, Fischer Taschenbuch Verlag, 1998. Seite 214.
- 4 Taubert HD, Kuhl H: Kontrazeption mit Hormonen, Ein Leitfaden für die Praxis; Stuttgart Georg Thieme Verlag, 1995, zweite überarbeitete und erweiterte Auflage, Seite 153ff.
- 5 Belsey EM, Pinol AP. Menstrual bleeding patterns in untreated women. Task Force on Long-Acting Systemic Agents for Fertility Regulation. Contraception 1997; 55(2):57-65.
- 6 Wallis J.: A survey of reproductive parameters in the free-ranging chimpanzees of Gombe National Park. J Reprod Fertil. 1997; 109(2):297-307.
- 7 Gerlinger C, Endrikat J, van der Meulen EA, Dieben TOM, Düsterberg B.: Recommendation for confidence interval and sample size calculation for the Pearl Index. Eur J Contracept Reprod Health Care 2003; 8(2):87-92.
- 8 The European Agency for the Evaluation of Medicinal Products, Committee for Proprietary Medicinal Products (CPMP): Note for Guidance on Clinical Evaluation of Steroid Contraceptives in Women. London 17 February 2000 CPMP/EWP/519/98. Abschnitt 4.3.
- 9 The European Agency for the Evaluation of Medicinal Products, Committee for Proprietary Medicinal Products (CPMP): Points to Consider on Hormone Replacement Therapy. London 17 November 1997 CPMP/EWP/021/97. Abschnitt 2.2.2.
- 10 World Health Organization Special Programme of Research, Development and Research Training in Human Reproduction: The analysis of vaginal bleeding patterns induced by fertility regulating methods; Contraception, 1986, Vol. 34, 253-260.

- 11 URL: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>, (Zugriff am 7. Oktober 2002)
- 12 World Medical Association. Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2000; 284(23):3043-5
- 13 The European Agency for the Evaluation of Medicinal Products, Committee for Proprietary Medicinal Products (CPMP): Note for Guidance on Good Clinical Practice. London 1 May 1996 CPMP/ICH/135/958.
- 14 Bauer FL, Goos G: Informatik, Eine einführende Übersicht, Erster Teil; Berlin: Springer-Verlag, 1973, zweite Auflage, Seite 176.
- 15 Nozaki M, Tada M, Kobayashi H, Zhang CL, Sawamura Y, Abe H, Ishii N, Van Meir EG. Roles of the functional loss of p53 and other genes in astrocytoma tumorigenesis and progression. Neuro-oncol 1999; 1(2):124-37
- 16 van de Weijer PH, Scholten PC, van der Mooren MJ, Barentsen R, Kenemans P. Bleeding patterns and endometrial histology during administration of low-dose estradiol sequentially combined with dydrogesterone. Climacteric 1999; 2(2):101-9.
- 17 Graser T, Koytchev R, Muller A, Oettel M. Comparison of the efficacy and endometrial safety of two estradiol valerate/dienogest combinations and Kliogest for continuous combined hormone replacement therapy in postmenopausal women. Climacteric 2000; 3(2):109-18.
- 18 Wren BG, McFarland K, Edwards L, O'Shea P, Sufi S, Gross B, Eden JA. Effect of sequential transdermal progesterone cream on endometrium, bleeding pattern, and plasma progesterone and salivary progesterone levels in postmenopausal women. Climacteric 2000; 3(3):155-60.
- 19 Massai MR, Diaz S, Quinteros E, Reyes MV, Herreros C, Zepeda A, Croxatto HB, Moo-Young AJ. Contraceptive efficacy and clinical performance of Nestorone implants in postpartum women. Contraception 2001; 64(6):369-76.
- 20 Petta CA, Hays M, Brache V, Massai R, Hua Y, Alvarez-Sanchez F, Salvatierra A, d'Arcangues C, Cook LA, Bahamondes L. Delayed first injection of the once-a-month injectable contraceptive containing 25 mg medroxyprogesterone acetate and 5 mg estradiol-cypionate: effects on cervical mucus. Contraception 2001; 64(6):363-8.
- 21 al-Azzawi F, Wahab M, Thompson J, Pornel B, Hirvonen E, Ylikorkala O, van der Mooren MJ, Dillon J, Magaril C. Acceptability and patterns of endometrial bleeding in

estradiol-based HRT regimens: a comparative study of cyclical sequential combinations of trimegestone or norethisterone acetate. *Climacteric* 2001; 4(4):343-54.

22 Bruhat M, Rudolf K, Vaheer R, Kainulainen P, Timonen U, Viitanen A. Effective bleeding control and symptom relief by lower dose regimens of continuous combined hormone replacement therapy: a randomized comparative dose-ranging study. *Maturitas* 2001; 40(3):259-71.

23 Smallwood GH, Meador ML, Lenihan JP, Shangold GA, Fisher AC, Creasy GW. Efficacy and safety of a transdermal contraceptive system. *Obstet Gynecol* 2001; 98(5 Pt 1):799-805.

24 Endrikat J, Cronin M, Gerlinger C, Ruebig A, Schmidt W, Dusterberg B. Open, multicenter comparison of efficacy, cycle control, and tolerability of a 23-day oral contraceptive regimen with 20 microg ethinyl estradiol and 75 microg gestodene and a 21-day regimen with 20 microg ethinyl estradiol and 150 microg desogestrel. *Contraception* 2001; 64(3):201-7.

25 Zhou L, Xiao B. Emergency contraception with Multiload Cu-375 SL IUD: a multicenter clinical trial. *Contraception* 2001; 64(2):107-12.

26 Endrikat J, Cronin M, Gerlinger C, Ruebig A, Schmidt W, Dusterberg B. Double-blind, multicenter comparison of efficacy, cycle control, and tolerability of a 23-day versus a 21-day low-dose oral contraceptive regimen containing 20 microg ethinyl estradiol and 75 microg gestodene. *Contraception* 2001; 64(2):99-105.

27 Fruzzetti F, Genazzani AR, Ricci C, De Negri F, Bersi C, Carmassi F. A 12-month clinical investigation with a 24-day regimen containing 15 microg ethinylestradiol plus 60 microg gestodene with respect to hemostasis and cycle control. *Contraception* 2001; 63(6):303-7.

28 Hampton RM, Short M, Bieber E, Bouchard C, Ayotte N, Shangold G, Fisher AC, Creasy GW. Comparison of a novel norgestimate/ethinyl estradiol oral contraceptive (Ortho Tri-Cyclen Lo) with the oral contraceptive Loestrin Fe 1/20. *Contraception* 2001; 63(6):289-95.

29 Endrikat J, Hite R, Bannemerschult R, Gerlinger C, Schmidt W. Multicenter, comparative study of cycle control, efficacy and tolerability of two low-dose oral contraceptives containing 20 microg ethinylestradiol/100 microg levonorgestrel and

20 microg ethinylestradiol/500 microg norethisterone. *Contraception* 2001; 64(1):3-10.

30 Jaisamrarn U, Reinprayoon D, Virutamasen P. Clinical study of a monophasic pill containing 20 microg ethinylestradiol and 150 microg desogestrel in Thai women. *J Med Assoc Thai* 2001; 84 Suppl 1:S377-83.

31 Ryan N, Rosner A. Quality of life and costs associated with micronized progesterone and medroxyprogesterone acetate in hormone replacement therapy for nonhysterectomized, postmenopausal women. *Clin Ther* 2001; 23(7):1099-115.

32 Vree ML, Schmidt J. A large observational clinical evaluation of a desogestrel-containing combiphasic oral contraceptive in Germany. *Eur J Contracept Reprod Health Care* 2001; 6(2):108-14.

33 Audet MC, Moreau M, Koltun WD, Waldbaum AS, Shangold G, Fisher AC, Creasy GW. Evaluation of contraceptive efficacy and cycle control of a transdermal contraceptive patch vs an oral contraceptive: a randomized controlled trial. *JAMA* 2001; 285(18):2347-54.

34 Petta CA, Hays M, Brache V, Massai R, Hua Y, Alvarez-Sanchez F, Croxatto H, d'Arcangues C, Cook LA, Bahamondes L. Delayed first injection of the once-a-month injectable contraceptive containing 25 mg of medroxyprogesterone acetate and 5 mg of E(2)-cypionate: effects on ovarian function. *Fertil Steril* 2001; 75(4):744-8.

35 Garceau RJ, Wajszczuk CJ, Kaunitz AM. Bleeding patterns of women using Lunelle monthly contraceptive injections (medroxyprogesterone acetate and estradiol cypionate injectable suspension) compared with those of women using Ortho-Novum 7/7/7 (norethindrone/ethinyl estradiol triphasic) or other oral contraceptives. *Contraception* 2000; 62(6):289-95.

36 Bassol S, Alvarado A, Celis C, Cravioto MC, Peralta O, Montano R, Novelli J, Alborno H, Kesseru E, Soares A, Petracco A, Isaia B, Mendes J, Bahamondes L, de Melo NR, Reyes-Marquez R, Albrecht G. Latin american experience with two low-dose oral contraceptives containing 30 microg ethinylestradiol/75 microg gestodene and 20 microg ethinylestradiol/150 microg desogestrel. *Contraception* 2000; 62(3):131-5.

37 Boulay JL, Perruchoud AP, Reuter J, Bolliger C, Herrmann R, Rochlitz C. P21 gene expression as an indicator for the activity of adenovirus-p53 gene therapy in non-small cell lung cancer patients. *Cancer Gene Ther* 2000; 7(9):1215-9.

- 38 Zichella L, Sbrignadello C, Tomassini A, Di Lieto A, Montoneri C, Zarbo G, Mancone M, Pietro Battista P, Bertoli G, Perrone G. Comparative study on the acceptability of two modern monophasic oral contraceptive preparations: 30 microgram ethinyl estradiol combined with 150 microgram desogestrel or 75 microgram gestodene. *Adv Contracept* 1999; 15(3):191-200.
- 39 Foidart JM, Wuttke W, Bouw GM, Gerlinger C, Heithecker R. A comparative investigation of contraceptive reliability, cycle control and tolerance of two monophasic oral contraceptives containing either drospirenone or desogestrel. *Eur J Contracept Reprod Health Care* 2000; 5(2):124-34.
- 40 Bassol S, Cravioto MC, Durand M, Bailon R, Carranza S, Fugarolas J, Gaona R, Parada LM, Celis C, Santoyo S, Garza-Flores J, Vazquez L, Lopez C, Gurucharri C, Novelli J, Carneiro de Oliveira H, Mendez J, de Andrade ME, de Mello NR, de Melo K, Chada E, Yassle ME, Castaneda A, Gomez P, Arboleda C, Trujillo L, Bucheli R, Hidalgo I, Olavide R, Parejarios J, Succar J, Reyes-Marquez R, Albrecht G. Mesigyna once-a-month combined injectable contraceptive: experience in Latin America. *Contraception* 2000; 61(5):309-16.
- 41 Kaunitz AM. Efficacy, cycle control, and safety of two triphasic oral contraceptives: Cyclessa (desogestrel/ethinyl estradiol) and ortho-Novum 7/7/7 (norethindrone/ethinyl estradiol): a randomized clinical trial. *Contraception* 2000; 61(5):295-302.
- 42 Ferguson H, Vree ML, Wilpshaar J, Eskes TK. Multicenter study of the efficacy, cycle control and tolerability of a phasic desogestrel-containing oral contraceptive. *Eur J Contracept Reprod Health Care* 2000; 5(1):35-45.
- 43 Huber J, Foidart JM, Wuttke W, Merki-Feld GS, The HS, Gerlinger C, Schellschmidt I, Heithecker R. Efficacy and tolerability of a monophasic oral contraceptive containing ethinylestradiol and drospirenone. *Eur J Contracept Reprod Health Care* 2000; 5(1):25-34.
- 44 Bruni V, Croxatto H, De La Cruz J, Dhont M, Durlot F, Fernandes MT, Andrade RP, Weisberg E, Rhoads M. A comparison of cycle control and effect on well-being of monophasic gestodene-, triphasic gestodene- and monophasic desogestrel-containing oral contraceptives. Gestodene Study Group. *Gynecol Endocrinol* 2000; 14(2):90-8.

- 45 Brouard R, Bossmar T, Fournie-Lloret D, Chassard D, Akerlund M. Effect of SR49059, an orally active V1a vasopressin receptor antagonist, in the prevention of dysmenorrhoea. *BJOG* 2000; 107(5):614-9.
- 46 Letterie GS. Inhibition of gonadotropin surge by a brief mid-cycle regimen of ethinyl estradiol and norethindrone: possible role in in vitro fertilization. *Gynecol Endocrinol* 2000; 14(1):1-4.
- 47 Limpaphayom KK, Bunyavejchevin S. Clinical effects of 17 beta-estradiol and norethisterone acetate in postmenopausal Thai women. *J Med Assoc Thai* 2000; 83(4):407-16.
- 48 Heikkinen JE, Vaheri RT, Ahomaki SM, Kainulainen PM, Viitanen AT, Timonen UM. Optimizing continuous-combined hormone replacement therapy for postmenopausal women: a comparison of six different treatment regimens. *Am J Obstet Gynecol* 2000; 182(3):560-7.
- 49 Endrikat J, Dusterberg B, Ruebig A, Gerlinger C, Strowitzki T. Comparison of efficacy, cycle control, and tolerability of two low-dose oral contraceptives in a multicenter clinical study. *Contraception* 1999; 60(5):269-74.
- 50 Rosenberg MJ, Meyers A, Roy V. Efficacy, cycle control, and side effects of low- and lower-dose oral contraceptives: a randomized trial of 20 micrograms and 35 micrograms estrogen preparations. *Contraception* 1999; 60(6):321-9.
- 51 Burch D, Bieshuevel E, Smith S, Fox H. Can endometrial protection be inferred from the bleeding pattern on combined cyclical hormone replacement therapy. *Maturitas* 2000; 34(2):155-60.
- 52 Saure A, Planellas J, Poulsen HK, Jaszczak P. A double-blind, randomized, comparative study evaluating clinical effects of two sequential estradiol-progestogen combinations containing either desogestrel or medroxyprogesterone acetate in climacteric women. *Maturitas* 2000; 34(2):133-42.
- 53 Belaisch J, Nappi C, Affinito P, De Gezelle H, Botelho M, Oliveira HM, Martinez-Oliveira J, Gonzaga F. [Effect of combined conjugated estrogen-medrogestone replacement therapy on lipid profiles, climacteric symptoms and the endometrium] *J Gynecol Obstet Biol Reprod (Paris)* 2000; 29(1):29-40.
- 54 Kaunitz AM, Garceau RJ, Cromie MA. Comparative safety, efficacy, and cycle control of Lunelle monthly contraceptive injection (medroxyprogesterone acetate and

estradiol cypionate injectable suspension) and Ortho-Novum 7/7/7 oral contraceptive (norethindrone/ethinyl estradiol triphasic). Lunelle Study Group. *Contraception* 1999; 60(4):179-87.

55 Reisman H, Martin D, Gast MJ. A multicenter randomized comparison of cycle control and laboratory findings with oral contraceptive agents containing 100 microg levonorgestrel with 20 microg ethinyl estradiol or triphasic norethindrone with ethinyl estradiol. *Am J Obstet Gynecol* 1999; 181(5 Pt 2):45-52.

56 Archer DF, Maheux R, DelConte A, O'Brien FB. Efficacy and safety of a low-dose monophasic combination oral contraceptive containing 100 microg levonorgestrel and 20 microg ethinyl estradiol (Alesse). North american Levonorgestrel Study Group (NALSG). *Am J Obstet Gynecol* 1999; 181(5 Pt 2):39-44.

57 Zheng SR, Zheng HM, Qian SZ, Sang GW, Kaper RF. A randomized multicenter study comparing the efficacy and bleeding pattern of a single-rod (Implanon) and a six-capsule (Norplant) hormonal contraceptive implant. *Contraception* 1999; 60(1):1-8.

58 Bjarnason K, Cerin A, Lindgren R, Weber T. Adverse endometrial effects during long cycle hormone replacement therapy. Scandinavian Long Cycle Study Group. *Maturitas* 1999; 32(3):161-70.

59 Weisberg E, Fraser IS, Mishell DR Jr, Lacarra M, Darney P, Jackanicz TM. A comparative study of two contraceptive vaginal rings releasing norethindrone acetate and differing doses of ethinyl estradiol. *Contraception* 1999; 59(5):305-10.

60 Ronnerdag M, Odland V. Health effects of long-term use of the intrauterine levonorgestrel-releasing system. A follow-up study over 12 years of continuous use. *Acta Obstet Gynecol Scand* 1999; 78(8):716-21.

61 Zheng SR, Zheng HM, Qian SZ, Sang GW, Kaper RF. A long-term study of the efficacy and acceptability of a single-rod hormonal contraceptive implant (Implanon) in healthy women in China. *Eur J Contracept Reprod Health Care* 1999; 4(2):85-93.

62 Chavez A, DelConte A. A comparison of cycle control with monophasic levonorgestrel/ethinylestradiol 100 micrograms/20 micrograms versus triphasic norethindrone/ethinylestradiol 500-750-1000 micrograms/35 micrograms: a multicenter, randomized, open-label study. *Eur J Contracept Reprod Health Care* 1999; 4(2):75-83.

- 63 Yang TS, Tsan SH, Chen CR, Chang SP, Yuan CC. Evaluation of conjugated estrogen plus medroxyprogesterone acetate versus tibolone in early postmenopausal Chinese women. *Zhonghua Yi Xue Za Zhi (Taipei)* 1999; 62(5):308-15.
- 64 DelConte A, Loffer F, Grubb GS. Cycle control with oral contraceptives containing 20 micrograms of ethinyl estradiol. A multicenter, randomized comparison of levonorgestrel/ethinyl estradiol (100 micrograms/20 micrograms) and norethindrone/ethinyl estradiol (1000 micrograms/20 micrograms). *Contraception* 1999; 59(3):187-93.
- 65 Affandi B, Korver T, Geurts TB, Coelingh Bennink HJ. A pilot efficacy study with a single-rod contraceptive implant (Implanon) in 200 Indonesian women treated for < or = 4 years. *Contraception* 1999; 59(3):167-74.
- 66 Sulak P, Lippman J, Siu C, Massaro J, Godwin A. Clinical comparison of triphasic norgestimate/35 micrograms ethinyl estradiol and monophasic norethindrone acetate/20 micrograms ethinyl estradiol. Cycle control, lipid effects, and user satisfaction. *Contraception* 1999; 59(3):161-6.
- 67 da Silva MO, Costa MM. Reason, myths and fantasies: preliminary data and reflections about the Portuguese experience with LNG-IUS-induced hypomenorrhea. *Eur J Contracept Reprod Health Care* 1999; 4(1):21-5.
- 68 van de Weijer PH, Barentsen R, de Vries M, Kenemans P. Relationship of estradiol levels to breakthrough bleeding during continuous combined hormone replacement therapy. *Obstet Gynecol* 1999; 93(4):551-7.
- 69 Swahn ML, Bygdeman M, Chen JK, Gemzell-Danielsson K, Song S, Yang QY, Yang PJ, Qian ML, Chang WF. Once-a-month treatment with a combination of mifepristone and the prostaglandin analogue misoprostol. *Hum Reprod* 1999; 14(2):485-8.
- 70 Serfaty D, Vree ML. A comparison of the cycle control and tolerability of two ultra low-dose oral contraceptives containing 20 micrograms ethinylestradiol and either 150 micrograms desogestrel or 75 micrograms gestodene. *Eur J Contracept Reprod Health Care* 1998; 3(4):179-89.
- 71 [No authors listed] A double-blind study comparing the contraceptive efficacy, acceptability and safety of two progestogen-only pills containing desogestrel 75 micrograms/day or levonorgestrel 30 micrograms/day. Collaborative Study Group on

the Desogestrel-containing Progestogen-only Pill. *Eur J Contracept Reprod Health Care*. 1998; 3(4):169-78.

72 Rowan JP. A double-blind study comparing the contraceptive efficacy, acceptability and safety of two progestogen-only pills containing desogestrel 75 micrograms/day or levonorgestrel 30 micrograms/day. Collaborative Study Group on the Desogestrel-containing Progestogen-only Pill. *Am J Obstet Gynecol* 1999; 180(2 Pt 2):302-6.

73 Cano A, Tarin JJ, Duenas JL. Two-year prospective, randomized trial comparing an innovative twice-a-week progestin regimen with a continuous combined regimen as postmenopausal hormone therapy. *Fertil Steril* 1999; 71(1):129-36.

74 Pal L, Lapensee L, Toth TL, Isaacson KB. Comparison of office hysteroscopy, transvaginal ultrasonography and endometrial biopsy in evaluation of abnormal uterine bleeding. *JSLs* 1997; 1(2):125-30.

75 Stanworth SJ, Bolton MJ, Hay CR, Shiach CR. Increased bleeding in HIV-positive haemophiliacs treated with antiretroviral protease inhibitors. *Haemophilia* 1998; 4(2):109-14.

76 Halbe HW, de Melo NR, Bahamondes L, Petracco A, Lemgruber M, de Andrade RP, da Cunha DC, Guazelli CA, Baracat EC. Efficacy and acceptability of two monophasic oral contraceptives containing ethinylestradiol and either desogestrel or gestodene. *Eur J Contracept Reprod Health Care* 1998; 3(3):113-20.

77 Vercellini P, Perino A, Consonni R, Oldani S, Parazzini F, Crosignani PG. Does preoperative treatment with a gonadotropin-releasing hormone agonist improve the outcome of endometrial resection? *J Am Assoc Gynecol Laparosc* 1998; 5(4):357-60.

78 Martinez GH, Castaneda A, Correa JE. Vaginal bleeding patterns in users of Perlutal, a once-a-month injectable contraceptive consisting of 10 mg estradiol enanthate combined with 150 mg dihydroxyprogesterone acetophenide. A trial of 5462 woman-months. *Contraception* 1998; 58(1):21-7.

79 Chen JH, Wu SC, Shao WQ, Zou MH, Hu J, Cong L, Miao L, Wang C, Dong J, Gao J, Xiao BL. The comparative trial of TCU 380A IUD and progesterone-releasing vaginal ring used by lactating women. *Contraception* 1998; 57(6):371-9.

80 Carranza-Lira S, Martinez-Chequer JC, Santa Rita MT, Ortiz de la Pena A, Perez Y, Fernandez RL. Endometrial changes according to hormone replacement therapy schedule. *Menopause* 1998; 5(2):86-9.

- 81 Croxatto HB, Kovacs L, Massai R, Resch BA, Fuentealba B, Salvatierra AM, Croxatto HD, Zalanyi S, Viski S, Krenacs L. Effects of long-term low-dose mifepristone on reproductive function in women. *Hum Reprod* 1998; 13(4):793-8.
- 82 Comparato MR, Yabur JA, Bajares M. Contraceptive efficacy and acceptability of a monophasic oral contraceptive containing 30 microg ethinyl estradiol and 150 microg desogestrel in Latin-American women. *Adv Contracept* 1998; 14(1):15-26.
- 83 Nand SL, Webster MA, Baber R, O'Connor V. Bleeding pattern and endometrial changes during continuous combined hormone replacement therapy. The Ogen/Provera Study Group. *Obstet Gynecol* 1998; 91(5 Pt 1):678-84.
- 84 Schipper I, Hop WC, Fauser BC. The follicle-stimulating hormone (FSH) threshold/window concept examined by different interventions with exogenous FSH during the follicular phase of the normal menstrual cycle: duration, rather than magnitude, of FSH increase affects follicle development. *J Clin Endocrinol Metab* 1998; 83(4):1292-8.
- 85 Graser T, Rossner P, Schubert K, Muller A, Bonisch U, Oettel M. A comparative study of two levonorgestrel-containing hormone replacement therapy regimens of efficacy and tolerability variables. *Maturitas* 1997 ; 28(2):169-79.
- 86 Lubbert H, Nauert C. Continuous versus cyclical transdermal estrogen replacement therapy in postmenopausal women: influence on climacteric symptoms, body weight and bleeding pattern. *Maturitas* 1997; 28(2):117-25.
- 87 Bannemerschult R, Hanker JP, Wunsch C, Fox P, Albring M, Brill K. A multicenter, uncontrolled clinical investigation of the contraceptive efficacy, cycle control, and safety of a new low dose oral contraceptive containing 20 micrograms ethinyl estradiol and 100 micrograms levonorgestrel over six treatment cycles. *Contraception* 1997; 56(5):285-90.
- 88 Bergqvist A, Jacobson J, Harris S. A double-blind randomized study of the treatment of endometriosis with nafarelin or nafarelin plus norethisterone. *Gynecol Endocrinol* 1997; 11(3):187-94.
- 89 Endrikat J, Muller U, Dusterberg B. A twelve-month comparative clinical investigation of two low-dose oral contraceptives containing 20 micrograms ethinylestradiol/75 micrograms gestodene and 30 micrograms ethinylestradiol/75 micrograms gestodene, with respect to efficacy, cycle control, and tolerance. *Contraception* 1997; 55(3):131-7.

- 90 Csemiczky G, Dieben T, Coeling Bennink HJ, Landgren BM. The pharmacodynamic effects of an oral contraceptive containing 3 mg micronized 17 beta-estradiol and 0.150 mg desogestrel for 21 days, followed by 0.030 mg desogestrel only for 7 days. *Contraception* 1996; 54(6):333-8.
- 91 AinMelk Y. Comparison of two continuous combined estrogen progestogen regimens in postmenopausal women: a randomized trial. *Fertil Steril* 1996 ;66(6):962-8.
- 92 Kalogirou D, Antoniou G, Karakitsos P, Kalogirou O, Antoniou D, Giannikos L. A comparative study of the effects of an estradiol-releasing vaginal ring combined with an oral gestagen versus transdermal estrogen combined with a levonorgestrel-releasing IUD: clinical findings and endometrial response. *Int J Fertil Menopausal Stud* 1996; 41(6):522-7.
- 93 Boonkasemsanti W, Reinprayoon D, Pruksananonda K, Niruttisard S, Triratanachai S, Leepipatpaiboon S, Wannakrairot P. The effect of transdermal oestradiol on bleeding pattern, hormonal profiles and sex steroid receptor distribution in the endometrium of Norplant users. *Hum Reprod* 1996; 11 Suppl 2:115-23.
- 94 Pansini F, De Paoli D, Albertazzi P, Bonaccorsi G, Campobasso C, Zanotti L, Pisati R, Giulini NA. Sequential addition of low dose of medrogestone or medroxyprogesterone acetate to transdermal estradiol: a pilot study on their influence on the endometrium. *Eur J Obstet Gynecol Reprod Biol* 1996; 68(1-2):137-41.
- 95 Saure A, Hirvonen E, Milsom I, Christensen A, Damber MG. A randomized, double-blind, multicentre study comparing the clinical effects of two sequential estradiol-progestin combinations containing either desogestrel or norethisterone acetate in climacteric women with estrogen deficiency symptoms. *Maturitas* 1996; 24(1-2):111-8.
- 96 Boerrigter PJ, van de Weijer PH, Baak JP, Fox H, Haspels AA, Kenemans P. Endometrial response in estrogen replacement therapy quarterly combined with a progestogen. *Maturitas* 1996; 24(1-2):63-71.
- 97 Rice C, Killick S, Hickling D, Coelingh Bennink H. Ovarian activity and vaginal bleeding patterns with a desogestrel-only preparation at three different doses. *Hum Reprod* 1996; 11(4):737-40.

- 98 Mayagoitia SB, Hernandez-Morales C, Macias AM, Vega RR. Luteinizing hormone releasing hormone agonist for postpartum contraception. *Adv Contracept* 1996; 12(1):27-41.
- 99 Messinis IE, Lolis D, Zikopoulos K, Milingos S, Kollis G, Seferiadis K, Templeton AA. Effect of follicle stimulating hormone or human chorionic gonadotrophin treatment on the production of gonadotrophin surge attenuating factor (GnSAF) during the luteal phase of the human menstrual cycle. *Clin Endocrinol (Oxf)* 1996; 44(2):169-75.
- 100 Dusterberg B, Ellman H, Muller U, Rowe E, Muhe B. A three-year clinical investigation into efficacy, cycle control and tolerability of a new low-dose monophasic oral contraceptive containing gestodene. *Gynecol Endocrinol* 1996; 10(1):33-9.
- 101 Rosenberg MJ, Waugh MS, Stevens CM. Smoking and cycle control among oral contraceptive users. *Am J Obstet Gynecol* 1996; 174(2):628-32.
- 102 Kunz J, Jaeger A. [Hormonal contraception has no effect on the atherogenicity of blood lipids: results of a prospective open-label multicenter study in 1321 patients] *Schweiz Rundsch Med Prax* 1996; 85(1-2):3-8.
- 103 Higham JM, O'Brien PM, Shaw RW. Assessment of menstrual blood loss using a pictorial chart. *Br J Obstet Gynaecol*. 1990; 97(8):734-9.
- 104 Higham JM, Shaw RW. Clinical associations with objective menstrual blood volume *Eur J Obstet Gynecol Reprod Biol*. 1999; 82(1):73-6.
- 105 Schering AG Studie 305220.
- 106 http://www.kaisersbkk.de/deutsch/download/zyklus_kalender.pdf Zugriff am 24. März 2003
- 107 Yasmin® [package insert]. Wayne, NJ: Berlex Laboratories Inc.; 2001.
- 108 Rodriguez G., Faundes-Latham A., Atkinson L.E.: An Approach to the Analysis of Menstrual Patterns in the Critical Evaluation of Contraceptives; *Studies in Family Planning* 1976; 7:42-51.
- 109 SAS Institute Inc. 2004. SAS/ETS_ 9.1 User's Guide. Cary, NC: SAS Institute Inc.

- 110 <http://www.spss.com/pdfs/STR13SPClr.pdf> Zugriff am 28. Januar 2005.
- 111 S-PLUS 6 for Windows Guide to Statistics, Volume 2, Insightful Corporation, Seattle, WA. Seite 196f.
- 112 Anderson T.W., An Introduction to Multivariate Statistical Analysis. New York, NY, John Wiley & Sons, 2nd edition 1984.
- 113 Dillon W.R., Goldstein M.: Multivariate Analysis. New York, NY, John Wiley & Sons, 1984.
- 114 Winston P.H.: Artificial Intelligence. Reading, Mass. Addison-Wesley 2nd edition, 1984.
- 115 Everitt B., Landau S. Leese M.: Cluster Analysis. London UK, Arnold, 4th edition 2001.
- 116 Bacher J.: Clusteranalyse. München und Wien, Oldenbourg, zweite Auflage, 1996.
- 117 Hands S, Everitt BS. A Monte Carlo study of the recovery of cluster structure in binary data by hierarchical clustering techniques. Multivariate Behavioral Research 1987; 22:235-245.
- 118 Florek K, Lukaszewicz L, Perkal L, Steinhaus H, Zortchi S. Sur la liason et la division des points d'un ensemble fini. Colloquium Mathematicum 1951; 2:282-285.
- 119 Sneath PHA. The application of computers to taxonomy. Journal of General Microbiology 1957; 17:201-226.
- 120 Sørensen T. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its applications to analyses of the vegetation on Danish commons. Biologiske Skrifter 1948; 5:1-34.
- 121 Sokal RR, Michener CD. A statistical method for evaluating systematic relationships. University of Kansas Science Bulletin 1958; 38:1409-38.
- 122 Ward JH. Hierarchical grouping to optimize an objective function. Journal of the American statistical Association 1963; 58:236-44.
- 123 SAS Institute Inc., SAS/STAT® User's Guide, Version 8, Cary, NC: SAS Institute Inc., 1999.

- 124 Sarle, W.S. (1983), Cubic Clustering Criterion, SAS Technical Report A-108, Cary, NC: SAS Institute Inc.
- 125 SAS Institute Inc., SAS Language Reference: Concepts, Cary, NC: SAS Institute Inc., 1999.
- 126 SAS Institute Inc., SAS Language Reference, Version 8, Cary, NC: SAS Institute Inc., 1999.
- 127 SAS Institute Inc., SAS Macro Language: Reference, Version 8, Cary, NC: SAS Institute Inc., 1999.
- 128 SAS Institute Inc., SAS Procedures Guide, Version 8, Cary, NC: SAS Institute Inc., 1999.
- 129 SAS Institute Inc., SAS/GRAPH® Software: Reference, Version 8, Cary, NC: SAS Institute Inc., 1999.
- 130 SAS Institute Inc., SAS Companion for the Microsoft Windows Environment, Version 8, Cary, NC: SAS Institute Inc., 1999.
- 131 The European Agency for the Evaluation of Medicinal Products: ICH Topic E10 Choice of Control Group in Clinical Trials. London, 27 July 2000 CPMP/ICH/364/96 <http://www.emea.eu.int/pdfs/human/ich/036496en.pdf> (Zugriff am 11. November 2004)
- 132 The European Agency for the Evaluation of Medicinal Products: ICH Topic E9 Statistical Principles for Clinical Trials. London, 18 March 1998 CPMP/ICH/363/96 <http://www.emea.eu.int/pdfs/human/ich/036396en.pdf> (Zugriff am 11. November 2004)