

***Patient-reported outcome measures beim Genitourinary Syndrome
of Menopause (GSM):***

systematische Übersichtsarbeit,
Validierung des „*Day-to-Day Impact of Vaginal Aging (DIVA)*
Questionnaire“ und
Exploration der urologischen Komponente des GSM



Dissertation
zur Erlangung des Doktorgrades
der Humanwissenschaften
(Dr. sc. hum.)

der
Fakultät für Medizin
der Universität Regensburg

vorgelegt von
Michaela Gabes
aus
Regensburg

im Jahr
2020

***Patient-reported outcome measures beim Genitourinary Syndrome
of Menopause (GSM):***

systematische Übersichtsarbeit,
Validierung des „*Day-to-Day Impact of Vaginal Aging (DIVA)*
Questionnaire“ und
Exploration der urologischen Komponente des GSM



Dissertation
zur Erlangung des Doktorgrades
der Humanwissenschaften
(Dr. sc. hum.)

der
Fakultät für Medizin
der Universität Regensburg

vorgelegt von
Michaela Gabes
aus
Regensburg

im Jahr
2020

Dekan:

Prof. Dr. Dirk Hellwig

Betreuer:

Prof. Dr. sc. hum. Christian Apfelbacher PhD

Tag der mündlichen Prüfung:

30. April 2021

Inhaltsverzeichnis

Inhaltsverzeichnis.....	5
Tabellenverzeichnis.....	9
Abbildungsverzeichnis.....	11
Abkürzungsverzeichnis.....	13
1. Einleitung.....	15
1.1 <i>Patient-reported outcomes</i> (PROs) und <i>patient-reported outcome measures</i> (PROMs)...	15
1.2 Entwicklung und Validierung von PROMs: Qualität vor Quantität.....	16
1.2.1 Einführung in die Terminologie.....	17
1.2.2 Bedeutsamkeit von Validierung.....	18
1.3 PROMs am Anwendungsbeispiel des „ <i>Genitourinary Syndrome of Menopause</i> (GSM)“ bzw. der vulvovaginalen Atrophie (VVA).....	20
1.3.1 Physiologische Veränderungen, Symptome und Diagnostik.....	21
1.3.2 Auswirkungen des GSM bzw. der VVA auf die Lebensqualität und Sexualität.....	23
1.3.3 PROMs beim GSM bzw. der VVA.....	25
1.4 Ziele der Arbeit.....	26
2. Teil 1: Systematische Übersichtsarbeit.....	27
2.1 Material.....	27
2.1.1 Die COSMIN <i>Risk of Bias Checklist</i>	27
2.1.2 Kriterien guter Messeigenschaften.....	27
2.1.3 Der <i>GRADE approach</i>	29
2.2 Methode.....	29
2.2.1 Das Protokoll: Registrierung.....	29
2.2.2 Literaturrecherche.....	29
2.2.3 Einschlusskriterien und Studienauswahl.....	31
2.2.4 Datenextraktion.....	32
2.2.4.1 Beurteilung der Inhaltsvalidität.....	32
2.2.4.2 Beurteilung der übrigen Messeigenschaften.....	32
2.2.5 Abgabe einer Empfehlung.....	33
2.3 Ergebnisse.....	33
2.3.1 Literaturrecherche.....	33
2.3.2 Titel-, Abstract- und Volltext-Screening.....	34
2.3.3 Datenextraktion.....	35
2.3.3.1 Beurteilung der Inhaltsvalidität.....	36
2.3.3.2 Beurteilung der übrigen Messeigenschaften.....	38

2.3.3.3	Charakteristiken der eingeschlossenen PROMs und Studienpopulationen.....	38
2.3.3.4	Informationen zu Interpretierbarkeit und Machbarkeit.....	41
2.3.3.5	Zusammenfassung der Ergebnisse	41
2.3.4	Abgabe einer Empfehlung.....	43
2.4	Diskussion	43
3.	Teil 2a: Validierung des <i>Day-to-Day Impact of Vaginal Aging Questionnaire</i> in einer deutschsprachigen Population	46
3.1	Einleitung	46
3.2	Methode.....	47
3.2.1	Strukturvalidität.....	51
3.2.2	Interne Konsistenz.....	51
3.2.3	Test-Retest-Reliabilität.....	52
3.2.4	Konstruktvalidität: Hypothesen-Testen.....	52
3.2.4.1	Konvergente Validität	52
3.2.4.2	Diskriminante Validität.....	53
3.2.5	Änderungssensitivität.....	53
3.3	Ergebnisse	54
3.3.1	Strukturvalidität.....	58
3.3.2	Interne Konsistenz.....	58
3.3.3	Test-Retest-Reliabilität.....	58
3.3.4	Konstruktvalidität: Hypothesen-Testen.....	59
3.3.4.1	Konvergente Validität	59
3.3.4.2	Diskriminante Validität.....	62
3.3.5	Änderungssensitivität.....	63
3.4	Explorative Analysen	65
3.4.1	Einfluss von Hormonen auf die Auswirkung von vaginalen Beschwerden.....	65
3.4.2	Auswirkungen der vaginalen Beschwerden (im DIVA) über das Alter hinweg.....	66
3.4.3	„Verfeinerung“ des DIVA.....	67
3.4.3.1	Überprüfung der Angemessenheit der Antwortkategorien (<i>monotonicity</i>).....	67
3.4.3.2	Überprüfung der Eindimensionalität (<i>unidimensionality</i>).....	72
3.4.3.3	Überprüfung der lokalen Unabhängigkeit (<i>local independence</i>)	73
3.4.3.4	Überprüfung der Passung des Modells (<i>model fit</i>).....	74
3.4.3.5	Exploratorische Faktorenanalysen	75
3.4.3.6	Analyse der Item-Korrelationen mit den zugehörigen Skalen.....	75
3.4.3.7	Zusammenfassung.....	77
3.4.4	Validierung der „verfeinerten“ Version des DIVA.....	79

3.4.4.1 Strukturvalidität	79
3.4.4.2 Interne Konsistenz.....	79
3.4.4.3 Konstruktvalidität: Hypothesen-Testen	79
3.4.4.4 Zusammenfassung.....	80
3.5 Diskussion	81
4. Teil 2b: Exploration der urologischen Komponente des GSM mithilfe von Fokusgruppen (online und <i>face-to-face</i>)	86
4.1 Einleitung	86
4.2 Methode.....	87
4.2.1 Rekrutierung.....	87
4.2.2 Durchführung der Fokusgruppen	89
4.2.3 Datenanalyse	91
4.3 Ergebnisse	91
4.3.1 Alltag.....	92
4.3.2 Emotionales Wohlbefinden	93
4.3.3 Sexuelle Funktionsfähigkeit.....	93
4.3.4 Selbstkonzept und Körperbild.....	94
4.3.5 Zwischenmenschliche Beziehungen und Kommunikation	95
4.3.6 Unerfüllte Bedürfnisse im Gesundheitswesen	96
4.3.7 Umgang mit den Beschwerden	97
4.4 Diskussion	98
5. Allgemeine Diskussion	101
6. Zusammenfassung.....	105
Anhang A Suchstrategie für <i>MEDLINE</i>	106
Anhang B Suchstrategie für <i>EMBASE</i>	111
Anhang C Suchstrategie für <i>Web of Science</i>	115
Anhang D Deutsche Version des DIVA	120
Anhang E Flyer für die Rekrutierung.....	122
Anhang F Mplus Syntax zur Überprüfung der Strukturvalidität des DIVA	124
Anhang G SPSS Syntax für die Aufbereitung der Daten und die Validierung des DIVA	125
Anhang H Teilnehmerinformation für die Online-Fokusgruppe	133
Anhang I Teilnehmerinformation, Einverständniserklärung und kurzer Fragebogen für die Durchführung der <i>face-to-face</i> Fokusgruppen	134
Anhang J Leitfaden für die Durchführung der Fokusgruppen	138
Literaturverzeichnis.....	140

Danksagung.....

Lebenslauf.....

Selbstständigkeitserklärung.....

Tabellenverzeichnis

Tabelle 1	Die 10 Boxen der <i>COSMIN Risk of Bias Checklist</i>	27
Tabelle 2	Qualitätskriterien für gute Messeigenschaften gemäß COSMIN (52)	28
Tabelle 3	Ein- und Ausschlusskriterien.....	31
Tabelle 4	Studien, die zur Datenextraktion herangezogen wurden	36
Tabelle 5	COSMIN <i>Risk of Bias</i> Bewertung der Inhaltsvalidität	36
Tabelle 6	Bewertung der Inhaltsvalidität und Qualität der Evidenz der eingeschlossenen PROMs	37
Tabelle 7	COSMIN <i>Risk of Bias</i> Bewertung der übrigen Messeigenschaften	38
Tabelle 8	Charakteristiken der eingeschlossenen PROMs	39
Tabelle 9	Charakteristiken der eingeschlossenen Studien.....	40
Tabelle 10	<i>Summary of Findings</i>	42
Tabelle 11	Empfehlungen für künftige GSM-Studien.....	43
Tabelle 12	Erwartete Korrelationen des DIVA mit anderen Items und PROMs	53
Tabelle 13	Demographische Daten der Befragung (aufgeteilt nach Online- und Papierversion).....	56
Tabelle 14	Ergebnisse der Menopause Rating Scale II (MRS), n = 185.....	57
Tabelle 15	Fit-Maße für die Kurzversion (für alle Frauen, n = 152) und die Langversion (nur für sexuell aktive Frauen, n = 128) des DIVA zur Bestätigung der vier Skalen	58
Tabelle 16	Cronbach's alpha für die Skalen des DIVA	58
Tabelle 17	Anzahl an Teilnehmerinnen mit stabiler bzw. veränderter Schwere vaginaler Symptome bzw. Beeinträchtigungen anhand der fünf Ankerfragen	59
Tabelle 18	Deskriptive Gegenüberstellung der Mittelwerte (und Standardabweichungen) der Skalen des DIVA bei stabilen Teilnehmerinnen (Veränderung in der jeweiligen Ankerfrage = 0) bei <i>Baseline</i> und nach ca. einer Woche.....	60
Tabelle 19	Tatsächliche Korrelationen des DIVA mit anderen Items und PROMs.....	61
Tabelle 20	P-Werte der einzelnen Kruskal-Wallis-Tests zur Überprüfung der diskriminanten Validität	62
Tabelle 21	Anzahl an Teilnehmerinnen mit veränderter bzw. stabiler Schwere vaginaler Symptome bzw. Beeinträchtigungen anhand der fünf Ankerfragen	63

Tabelle 22	Deskriptive Gegenüberstellung der Mittelwerte (und Standardabweichungen) der Skalen des DIVA bei Teilnehmerinnen, die eine Verbesserung wahrnehmen (Veränderung in der jeweiligen Ankerfrage < 0), bei <i>Baseline</i> und nach ca. vier Wochen.....	64
Tabelle 23	P-Werte der einzelnen Mann-Whitney-U-Tests zur Überprüfung des Einflusses von Hormonen	65
Tabelle 24	Fit-Maße für die Kurzversion (für alle Frauen, n = 152) und die Langversion (nur für sexuell aktive Frauen, n = 128) des DIVA mit den adaptierten Antwortoptionen zur Bestätigung der vier Skalen.....	71
Tabelle 25	Fit-Maße für die Kurzversion (für alle Frauen, n = 152) und die Langversion (nur für sexuell aktive Frauen, n = 128) des DIVA zur Überprüfung der Eindimensionalität des DIVA.....	72
Tabelle 26	Fit-Maße aus den Faktorenanalysen pro Skala zur Überprüfung der Eindimensionalität jeder Skala des DIVA.....	73
Tabelle 27	Faktorladungen der EFAs für die Kurzversion und Langversion des DIVA ..	76
Tabelle 28	Korrelationen der einzelnen Items mit den Skalen des DIVA für die Kurzversion und die Langversion	77
Tabelle 29	Auffällige Items getrennt nach IRT bzw. Rasch-Analyse und KTT	78
Tabelle 30	Fit-Maße für die Kurzversion (für alle Frauen, n = 152) und die Langversion (nur für sexuell aktive Frauen, n = 128) der „verfeinerten“ Version des DIVA (ohne Item 5 und 17).....	79
Tabelle 31	Korrelationen der „verfeinerten“ Version des DIVA mit anderen PROMs	80

Abbildungsverzeichnis

Abbildung 1	Inhalte von <i>patient-reported outcomes</i> (PROs) (2)	16
Abbildung 2	Die COSMIN-Taxonomie der Messeigenschaften, verfügbar unter https://www.cosmin.nl/tools/cosmin-taxonomy-measurement-properties/	18
Abbildung 3	Schematische Darstellung des <i>Genitourinary Syndrome of Menopause</i> (GSM)	21
Abbildung 4	Checkliste, um die VVA bzw. das GSM in der klinischen Praxis zu diagnostizieren, übersetzt und zitiert nach Nappi <i>et al.</i> (30).....	23
Abbildung 5	Adaptiertes 2009 PRISMA Flussdiagramm (64). Weitere Informationen unter www.prisma-statement.org	34
Abbildung 6	Cluster der PROMs nach dem ersten Titel- und Abstract-Screening	35
Abbildung 7	Zusammensetzung des DIVA	46
Abbildung 8	Verwendete Datenquellen für die Validierung des DIVA.....	47
Abbildung 9	Aufbau der Online-Befragung	49
Abbildung 10	Zusammensetzung der Stichprobe der Online-Befragung.....	54
Abbildung 11	Mittelwerte der DIVA-Skalen über verschiedene Altersgruppen hinweg.....	66
Abbildung 12	Itemfunktionen (<i>item characteristic curves</i>) der Items 1 – 5 für die originalen Antwortkategorien des DIVA (a) und für Antwortkategorien nach der Adaption (b).....	68
Abbildung 13	Itemfunktionen (<i>item characteristic curves</i>) der Items 6 – 9 für die originalen Antwortkategorien des DIVA (a) und für Antwortkategorien nach der Adaption (b).....	69
Abbildung 14	Itemfunktionen (<i>item characteristic curves</i>) der Items 10, 11 und 16 – 18 (Kurzversion) bzw. der Items 10 – 18 (Langversion) für die originalen Antwortkategorien des DIVA (a) und für Antwortkategorien nach der Adaption (b).....	70
Abbildung 15	Itemfunktionen (<i>item characteristic curves</i>) der Items 19 – 23 für die originalen Antwortkategorien des DIVA (a) und für Antwortkategorien nach der Adaption (b).....	71
Abbildung 16	Scree-Plots für die Kurzversion (links) und die Langversion (rechts) des DIVA	75
Abbildung 17	Zusammensetzung der „verfeinerten“ Version des deutschen DIVA	78

Abbildung 18 Häufigkeit berichteter vulvovaginaler (links, n = 40) und urologischer Beschwerden (rechts, n = 45)	88
Abbildung 19 Ablauf der Rekrutierung und Zusammensetzung der Stichprobe.....	88
Abbildung 20 Konzeptuelles Modell (in Anlehnung an die Ergebnisse der Fokusgruppen, die zur Entwicklung des DIVA (65) durchgeführt wurden) mit den zwei zusätzlichen Themenblöcken „Erwartungshaltung“ und „Aufklärung von ärztlicher Seite“	90
Abbildung 21 Übersicht über die eingeschlossene Stichprobe.....	91
Abbildung 22 Themen-Cluster	92

Abkürzungsverzeichnis

AUC	<i>Area under the curve</i>
CFA	<i>Confirmatory factor analysis</i>
CFI	<i>Comparative fit index</i>
CI	Konfidenzintervall (<i>confidence interval</i>)
COS	<i>Core outcome set</i>
CTT	<i>Classical test theory</i>
CLOSER	<i>CLarifying Vaginal Atrophy's Impact On SEx and Relationships</i>
COSMIN	<i>COnsensus-based Standards for the selection of health Measurement INstruments</i>
df	Freiheitsgrade (<i>degrees of freedom</i>)
DIF	<i>Differential item functioning</i>
DIVA	<i>Day-to-Day Impact of Vaginal Aging Questionnaire</i>
EFA	Exploratorische Faktorenanalyse
FDA	<i>US Food and Drug Administration</i>
FSFI	<i>Female Sexual Function Index</i>
GCS	<i>Greene Climacteric Scale</i>
GRADE	<i>Grading of Recommendations Assessment, Development and Evaluation</i>
GSM	<i>Genitourinary Syndrome of Menopause</i>
ICC	<i>Intraclass correlation coefficient</i>
ICIQ-OAB	<i>International Consultation on Incontinence Questionnaire Overactive Bladder</i>
IRT	<i>Item response theory</i>
ISSWSH	<i>International Society for the Study of Women's Sexual Health</i>
KFA	Konfirmatorische Faktorenanalyse
KTT	Klassische Testtheorie
LoA	<i>Limits of agreement</i>
MCS	Psychische Summenskala (<i>mental component score</i>)
MIC/MID	<i>Minimal important change/difference</i>
MRS	<i>Menopause Rating Scale II</i>
NAMS	<i>North American Menopause Society</i>
QoL	<i>Quality of life</i>
PCA	<i>Principal component analysis</i>

PCS	Körperliche Summenskala (<i>physical component score</i>)
PGI	<i>Patients Global Impression of Improvement</i>
PHQ	<i>Patient Health Questionnaire</i>
PRISMA-P	<i>Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols</i>
PRO	<i>Patient-reported outcome</i>
PROM	<i>Patient-reported outcome measure</i>
PROSPERO	<i>International Prospective Register of Systematic Reviews</i>
RMSEA	<i>Root mean square error of approximation</i>
SD	Standardabweichung (<i>standard deviation</i>)
SDC	<i>Smallest detectable change</i>
SEM	<i>Standard error of measurement</i>
SF-12	<i>Short-form health survey</i>
SOEP	Sozio-oekonomisches Panel
SRMR	<i>Standardized root mean residuals</i>
TLI	<i>Tucker-Lewis Index</i>
UGAQoL	<i>Urogenital Atrophy Quality of Life Questionnaire</i>
VHI	<i>Vaginal Health Index</i>
VIVA	<i>Vaginal Health: Insights, Views & Attitudes</i>
VMI	<i>Vaginal Maturation Index</i>
VSQ	<i>Vulvovaginal Symptoms Questionnaire</i>
VVA	Vulvovaginale Atrophie
VVAQ	<i>Vulvovaginal Atrophy Questionnaire</i>
WHAP	<i>Women's Healthy Ageing Project</i>

1. Einleitung

“It is to say that the ultimate measure by which to judge the quality of a medical effort is whether it helps patients (and their families) as they see it. Anything done in health care that does not help a patient or family is, by definition, waste, whether or not the professions and their associations traditionally hallow it (1).”¹

Berwick, 1997, S. 1565

Mit dieser Aussage aus dem Jahr 1997 stellt Donald M. Berwick bereits klar dar, dass die Sicht eines Patienten bzw. einer Patientin (und seiner bzw. ihrer Familie) von entscheidender Bedeutung für den Nutzen und Erfolg einer Behandlung ist. Ohne die Patientenperspektive kann sich eine Behandlung nicht an den Betroffenen orientieren und deren Bedürfnisse, Werte und Präferenzen miteinbinden (2). Oder wie Valeria Billingham so schön sagte: *„Nothing about me without me.”*²

1.1 Patient-reported outcomes (PROs) und patient-reported outcome measures (PROMs)

Patient-reported outcomes (PROs) bezeichnen jegliche Auskunft eines Patienten oder einer Patientin über seinen oder ihren Gesundheitszustand, ohne dass diese durch einen Arzt bzw. eine Ärztin oder irgendjemanden sonst interpretiert wird (3). PRO ist ein Oberbegriff, der sich auf drei Inhalte von Patientenberichten beziehen kann (vgl. Abbildung 1). PatientInnen können von verschiedenen Domänen, wie beispielsweise von ihren Symptomen oder Nebenwirkungen, von ihrer Funktionsfähigkeit auf verschiedenen Ebenen oder von ihrer generellen Zufriedenheit berichten (2). PROs haben in den letzten zwei Jahrzehnten deutlich an Bedeutung zugenommen (4). Sie werden immer häufiger als Endpunkte bzw. *Outcomes* in klinischen Studien erfasst, da sie einen einzigartigen Einblick in die Auswirkung einer Erkrankung und deren Behandlung aus der Sicht eines Betroffenen bzw. einer Betroffenen geben (2). Diese Information kann durch biomedizinische Endpunkte, wie z. B. ein Blutbild, allein nicht gewonnen werden. Zudem sind bestimmte Bereiche, wie z. B. Schmerz oder Müdigkeit, objektiv schwer zu beobachten und auch der Grad der Beeinträchtigung ist sehr subjektiv und demzufolge am besten durch einen Patientenbericht beschreibbar. PROs können ÄrztInnen und künftigen PatientInnen auch

¹ Deutsche Übersetzung: „Es ist festzuhalten, dass die Qualität einer medizinischen Leistung letztlich daran gemessen wird, ob sie den PatientInnen (und ihren Familien) so hilft, wie sie sie sehen. Alles, was in der Gesundheitsversorgung getan wird und einem Patienten oder seiner Familie nicht hilft, ist per Definition Verschwendung, unabhängig davon, ob die Berufe und ihre Verbände es traditionellerweise heiligen oder nicht.“

² Deutsche Übersetzung: „Nichts an mir ohne mich.“ Aus Session 356, *Through the Patient's Eyes*, 1998, Salzburg Global Seminar, verfügbar unter <https://www.healthnewsreview.org/2012/03/nothing-about-me-without-me-nejm-perspective-pieces/> [zuletzt aufgerufen am 22.10.2020]

helfen, die beste Behandlung auszuwählen, da sie ein klares Bild über Kosten und Nutzen einer Behandlung liefern können (5).

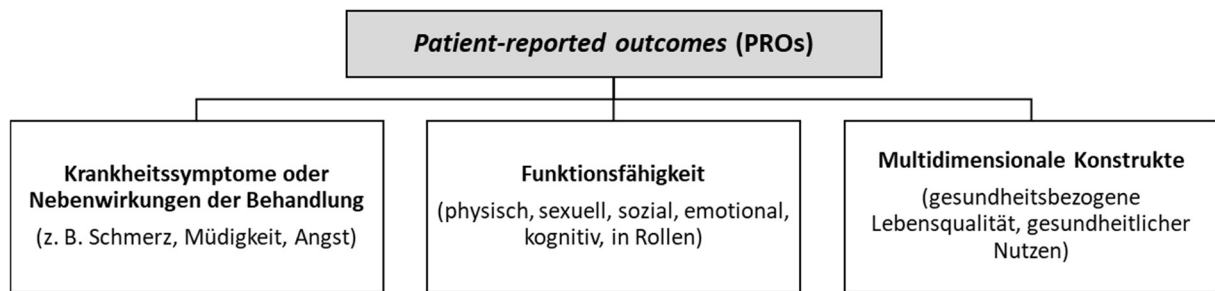


Abbildung 1 Inhalte von *patient-reported outcomes* (PROs) (2)

Patient-reported outcome measures (PROMs) werden eingesetzt, um PROs systematisch zu erfassen. PROMs sind somit Messinstrumente, die die Wahrnehmung der PatientInnen, sei es deren Gesundheitszustand oder die Effektivität einer Intervention, mittels selbst berichteten Fragebögen widerspiegeln (6). PROMs geben also Auskunft darüber, wie etwas gemessen wird, wohingegen PROs die Frage beantworten, was gemessen wird. In der klinischen Forschung werden PROMs z. B. zum Screening oder zur Beurteilung eines Behandlungsergebnisses eingesetzt. In der klinischen Praxis können PROMs beispielsweise für die Diagnosestellung oder zur Beobachtung des Behandlungsfortschritts genutzt werden (7) und so die Grundlage für eine gute Indikations- und Ergebnisqualität bilden. Dadurch, dass die PatientInnen miteinbezogen werden und z. B. mithilfe von PROMs Veränderungen ihrer Symptome wöchentlich dokumentieren können, kann das Selbstmanagement der PatientInnen gefördert werden und eine erhöhte Selbstwirksamkeit zur Folge haben (4, 8). Neben der klinischen Forschung und Praxis verwendet man PROMs auch im Bereich der Gesundheitsorganisation, um die Qualität von Leistungen zu verbessern und Leistungen miteinander zu vergleichen, und im Gesundheitssystem, um Leistungen zu bewerten und Kosten und Nutzen einer Leistung einschätzen zu können (7). PROMs werden typischerweise mit dem Input von PatientInnen, ÄrztInnen und PRO-ExpertInnen entwickelt, um sicherzustellen, dass die jeweilige PROM auch klinisch relevante Aspekte beinhaltet, die für die jeweilige Zielgruppe bedeutsam sind (2). PROMs sollen somit helfen, die Perspektive der PatientInnen zu verstehen. Mittlerweile zählt es zum Standard in der klinischen Forschung, Patientenberichte mittels PROMs zu erfassen, um die Wirksamkeit und Sicherheit einer Therapie beurteilen zu können (4).

1.2 Entwicklung und Validierung von PROMs: Qualität vor Quantität

In der klinischen Forschung werden viele PROMs eingesetzt, bei denen es Evidenz für deren schlechte Qualität gibt oder deren Qualität gar nicht bekannt ist. Wie bereits erwähnt, ist eine

PROM ein Messinstrument und die Verwendung eines Messinstruments, das nur schlecht funktioniert oder von dem gar nicht bekannt ist, wie gut oder schlecht es ist, kann zu missverständlichen Ergebnissen führen, die nicht richtig interpretiert werden können, und somit wertvolle Forschungsressourcen verschwenden (9).

In den letzten Jahren hat sich jedoch nicht nur die Anzahl an verfügbaren PROMs erhöht, sondern auch die Verfügbarkeit von qualitativ hochwertigen PROMs hat sich deutlich verbessert. Um überhaupt eine gute PROM zu bekommen, die das interessierende Konstrukt auch wirklich misst, ist eine intensive Entwicklung, bei der auch die Zielgruppe miteinbezogen wird, notwendig. Zudem sollte eine fertige PROM zuverlässige und genaue Ergebnisse liefern (Reliabilität) und auch tatsächlich das messen, was sie zu messen beabsichtigt (Validität). Sie sollte außerdem in der Lage sein, Veränderungen über die Zeit abzubilden (Änderungssensitivität). Diese Eigenschaften werden als „Messeigenschaften“ bezeichnet. Mit ihnen kann die Güte eines Messinstruments beurteilt werden (10).

Für die Auswahl der besten PROM, die auch wirklich den *Outcome* misst, der interessiert, sind qualitativ hochwertige Studien nötig, die die Evaluierung verschiedener Messeigenschaften dokumentieren. Außerdem werden hochwertige systematische Übersichtsarbeiten empfohlen, die alle Studien zu Messeigenschaften vereinen und bewerten, um abschließend eine Empfehlung für die beste(n) verfügbare(n) PROM(s) abzugeben (10).

1.2.1 Einführung in die Terminologie

In der Literatur gibt es eine große Variation in den Begrifflichkeiten und Definitionen für Messeigenschaften. Für ein und dieselbe Messeigenschaft werden häufig viele Synonyme benutzt. Beispielsweise wird im Englischen die Messeigenschaft ‚*reliability*‘ auch als ‚*reproducibility*‘, ‚*stability*‘, ‚*repeatability*‘ usw. bezeichnet (11). Die internationale *Consensus-based Standards for the selection of health Measurement Instruments* (COSMIN)-Initiative (<https://www.cosmin.nl/>) hat sich zum Ziel gemacht, die Auswahl von PROMs in Forschung und Praxis zu verbessern, indem sie eine Methodik sowie praktische Hilfsmittel, wie z. B. Checklisten und Handbücher (10, 12), entwickelt hat, um die am besten geeignete PROM für den jeweiligen Kontext auswählen zu können. Diese Initiative setzt sich aus einem multidisziplinären Team mit Expertise in der Entwicklung und Validierung von PROMs zusammen. Bevor diese Methodik mit ihren Werkzeugen entwickelt werden konnte, war eine einheitliche Terminologie vonnöten. Mithilfe der COSMIN Delphi-Studie (13) wurde Konsensus bezüglich Taxonomie, Terminologie sowie Definitionen für alle Messeigenschaften erreicht, die für die Evaluierung von PROMs relevant sind.

In der COSMIN-Taxonomie werden drei Qualitätsbereiche unterschieden: Reliabilität (*reliability*), Validität (*validity*) und Änderungssensitivität (*responsiveness*). Jeder Bereich enthält eine oder mehrere Messeigenschaften (vgl. Abbildung 2). Zum Bereich der Reliabilität gehört die interne Konsistenz (*internal consistency*), die Reliabilität (*reliability*) und der Messfehler (*measurement error*). Zum Bereich der Validität zählt die Inhaltsvalidität (*content validity*), zu der die Augenscheinvalidität (*face validity*) gehört, die Konstruktvalidität (*construct validity*), die wiederum die Strukturvalidität (*structural validity*), das Hypothesentesten für die Konstruktvalidität und die interkulturelle Validität (*cross-cultural validity*) umfasst, und die Kriteriumsvalidität (*criterion validity*). Die Änderungssensitivität steht als Messeigenschaft für sich. Die Interpretierbarkeit (*interpretability*) einer PROM wird nicht als Messeigenschaft bezeichnet. Sie ist aber trotzdem eine wichtige Eigenschaft einer PROM und taucht deshalb in der Taxonomie mit auf (10). Denn ohne eine reliable, valide und änderungs-sensitive PROM können erhobene Daten nicht aussagekräftig interpretiert werden (14).

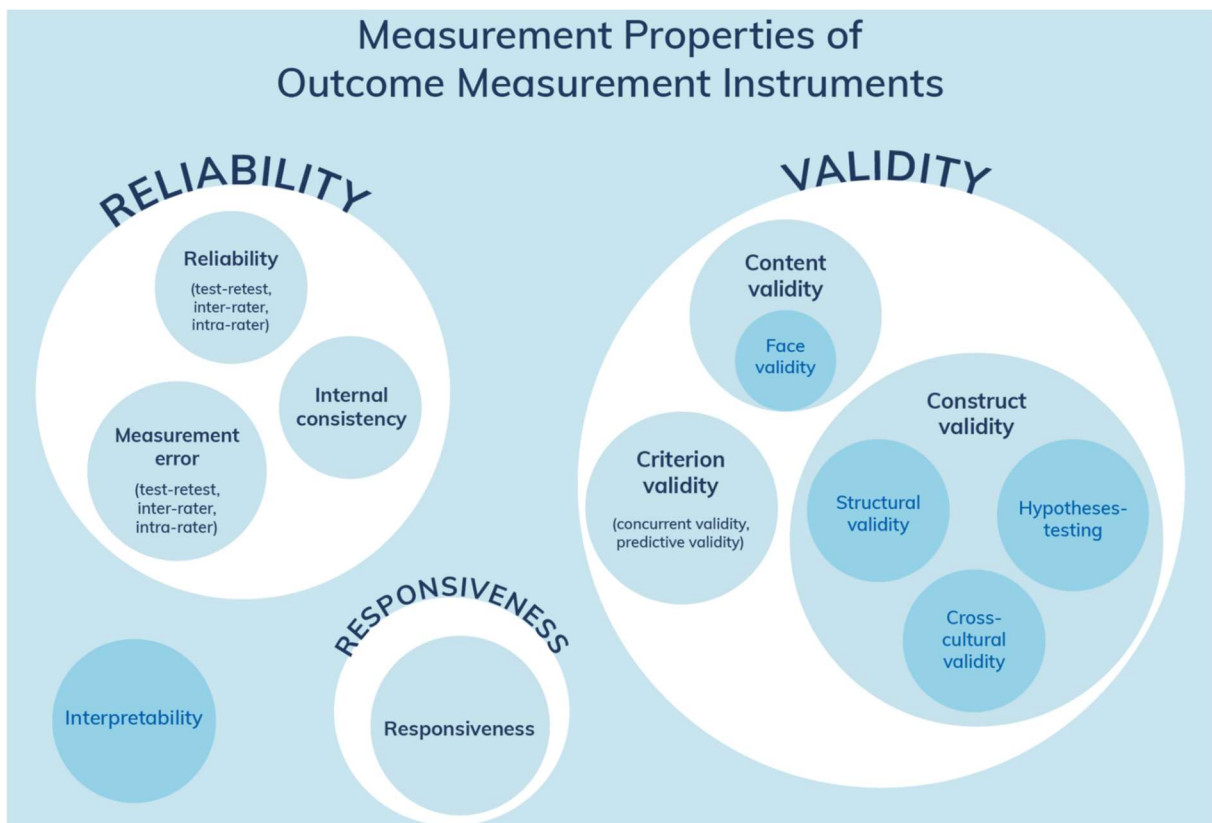


Abbildung 2 Die COSMIN-Taxonomie der Messeigenschaften, verfügbar unter <https://www.cosmin.nl/tools/cosmin-taxonomy-measurement-properties/>

1.2.2 Bedeutsamkeit von Validierung

Messeigenschaften sind nichts anderes als Qualitätsaspekte eines Messinstruments, also auch einer PROM. Auch klassische Messinstrumente, wie z. B. eine Waage, sollten diese

Eigenschaften erfüllen. Wiegt man sich beispielsweise direkt zweimal hintereinander, ohne in der Zwischenzeit etwas verändert zu haben, würde man das gleiche Gewicht wie beim ersten Wiegen erwarten. Erhält man tatsächlich bei zweimaligem Wiegen das gleiche Gewicht, würde man die Test-Retest-Reliabilität als gegeben annehmen. Hat sich z. B. nach vier Wochen der Bauchumfang um fünf Zentimeter verkleinert, weil eine Diät gemacht wurde, würde man ein geringeres Gewicht als das Ausgangsgewicht erwarten. Zeigt die Waage nach vier Wochen tatsächlich eine Gewichtsreduktion an, kann man davon ausgehen, dass diese Waage änderungssensitiv ist, d. h. sie registriert Veränderungen, wenn tatsächlich welche vorhanden sind (wie in diesem Beispiel der geringere Bauchumfang). Genau diese Eigenschaften lassen sich auch auf abstraktere Messinstrumente, wie eben PROMs, übertragen. Zeigt eine PROM nach einer gewissen Zeit keine Veränderung an, obwohl z. B. die Erkrankung komplett verschwunden ist, müsste man hier davon ausgehen, dass diese PROM die Messeigenschaft der Änderungssensitivität nicht erfüllt. Aus diesem Grund müssen PROMs validiert werden, um all ihre Messeigenschaften zu bestimmen. Erst wenn man weiß, dass eine PROM tatsächlich das misst, was sie vorgibt zu messen, dass sie reproduzierbare Ergebnisse erzielt, intern konsistent ist, und in der Lage ist, klinisch relevante Veränderungen über die Zeit abzubilden, kann man davon ausgehen, dass es sich um eine qualitativ hochwertige PROM handelt, deren Ergebnisse am Ende auch interpretiert werden können. Denn nur gute Messinstrumente liefern aussagekräftige und stabile Ergebnisse, auch über verschiedenen Studien hinweg. Ändert sich der Kontext (z. B. die Erkrankung), muss in dem neuen Kontext auch wieder eine Validierung stattfinden. Würde man beim Wiegebeispiel den Menschen durch einen Apfel ersetzen, müsste man auch erst überprüfen, ob sich die Personenwaage überhaupt auf Äpfel anwenden lässt. Ein halber Apfel sollte weniger wiegen als ein ganzer, aber ob eine Personenwaage in der Lage ist, diese Veränderung zu registrieren, müsste man zuerst ausprobieren, um eine Aussage über die Änderungssensitivität in diesem Kontext machen zu können.

Mittlerweile legen auch Lebensmittelüberwachungs- und Arzneimittelbehörden, wie z. B. die *US Food and Drug Administration* (FDA), großen Wert auf eine ‚patientenzentrierte‘ Arzneimittelentwicklung und den Einsatz von PROMs in klinischen Studien (4). Dabei achtet die FDA auch auf die Messeigenschaften von PROMs, die in den klinischen Studien verwendet werden, auf den Einbezug der Patientenpopulation in der PROM-Entwicklung sowie auf das konzeptuelle Modell, das einer PROM zugrunde liegt. Eine PROM, die keine Evidenz für ihre Eignung in der Zielpopulation aufweist, stuft die FDA als nicht glaubwürdig ein (3).

1.3 PROMs am Anwendungsbeispiel des „*Genitourinary Syndrome of Menopause (GSM)*“ bzw. der vulvovaginalen Atrophie (VVA)

Der Begriff *Genitourinary Syndrome of Menopause (GSM)* existiert seit dem Jahr 2014 und umfasst urogenitale Beschwerden, die im Zusammenhang mit der Menopause stehen (15). Die Menopause beschreibt den irreversiblen Verlust der physiologischen Fähigkeit, Nachwuchs zu produzieren (16, 17). Sie bezeichnet somit das Ende der reproduktiven Phase einer Frau. Bei einem natürlichen Verlauf lässt die Funktion der Eierstöcke nach und die Menstruation bleibt permanent aus. Retrospektiv spricht man vom Beginn der natürlichen Menopause, wenn länger als zwölf Monate keine spontane Regelblutung mehr aufgetreten ist und auch keine andere pathologische oder physiologische Ursache vorliegt (18, 19). Bei den meisten Frauen tritt die natürliche Menopause im Alter von 40 bis 60 Jahren ein (16), im Median mit 51,4 Jahren (18). Da die deutsche Bevölkerung immer älter wird und Frauen weltweit länger leben als Männer (20), verbringen viele Frauen mindestens das letzte Drittel ihres Lebens, wenn nicht sogar bis zu 50 % ihres Lebens in der Menopause. Die Menopause darf nicht als einzelner Zeitpunkt gesehen werden, sondern als eine Umstellung des Körpers über mehrere Jahre hinweg (19), die urogenitale Beschwerden mit sich ziehen kann.

Auf einer Konsensus-Konferenz einigten sich die *International Society for the Study of Women's Sexual Health (ISSWSH)* und die *North American Menopause Society (NAMS)* auf die Terminologie *Genitourinary Syndrome of Menopause*, da dieser Begriff sowohl Veränderungen im Genital- als auch im Harntrakt umfasst. GSM soll somit ein angemessenerer Begriff als vorherige Bezeichnungen wie „vulvovaginale Atrophie (VVA)“, „vaginale Atrophie“ oder „atrophische Vaginitis“ sein (21, 22). Dieser neue Begriff beschreibt umfassender als die vorherigen alle Anzeichen und Symptome, die mit einem Rückgang des Östrogenspiegels und anderer Sexualhormone im Rahmen der Menopause assoziiert sind. Des Weiteren wird eine negative Konnotation, die mit dem Wort „Atrophie“ einhergeht, vermieden. Auch der Begriff „Vagina“ ist in der heutigen Gesellschaft kein akzeptierter Begriff, der in der öffentlichen Diskussion oder in den Medien verwendet wird (15, 21). Für die neue Bezeichnung spricht zudem, dass urologische Anzeichen und Symptome häufig übersehen oder vernachlässigt werden, da Frauen sich oft dafür schämen und diese somit nicht bei ihrem Arzt bzw. ihrer Ärztin ansprechen. Weltweit sind bis zu 50 % aller postmenopausalen Frauen von urogenitalen Beschwerden betroffen (23). Betroffenen Frauen ist häufig der chronische und progressive Charakter dieser Beschwerden nicht bewusst (24, 25).

1.3.1 Physiologische Veränderungen, Symptome und Diagnostik

Das GSM bzw. die VVA geht mit einem verminderten Östrogenlevel einher. Östrogen-Rezeptoren sind nicht nur in der Vulva und Vagina stark vertreten, sondern auch in der Muskulatur und dem Bindegewebe des Beckenbodens, im Blasendreieck (= Schleimhautfeld am Harnblasenboden) und in der Harnröhre (24, 26). Dadurch, dass in der Menopause weniger Östrogen zirkuliert, verändert sich das urogenitale Gewebe einer Frau physiologisch. Der Kollagengehalt und das Elastin reduzieren sich, es kommt zu einer Hyalinisation (= Form der Gewebedegeneration), das Epithel wird dünner, das Bindegewebe weist eine höhere Dichte auf und weniger Blutgefäße sind vorhanden (15). Somit zirkuliert weniger Blut, die vaginale Feuchtigkeit reduziert sich, die Flexibilität und Elastizität des Vaginalgewölbes lassen nach und der pH-Wert steigt an (15, 26). All diese Faktoren begünstigen eine leichtere Verletzbarkeit des vaginalen Gewebes (21), beispielsweise bei sexueller Aktivität, und können bestimmte Symptome hervorrufen (vgl. Abbildung 3).

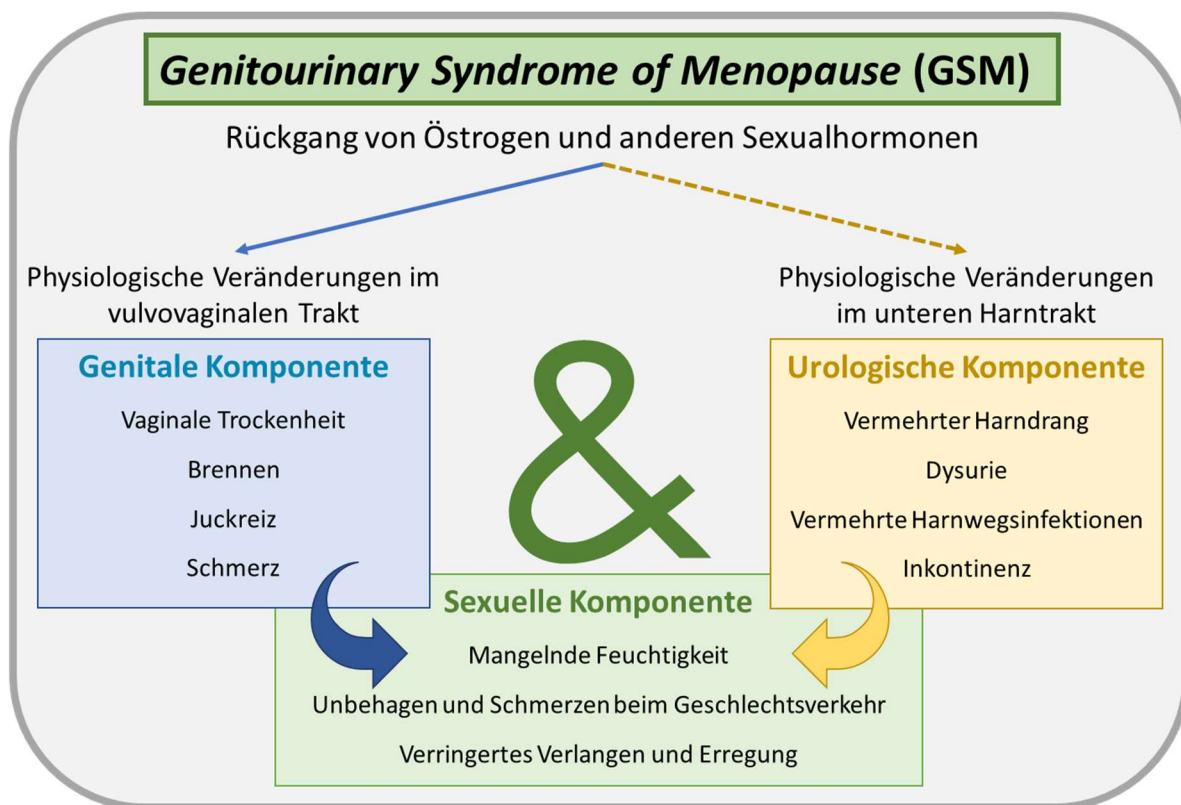


Abbildung 3 Schematische Darstellung des *Genitourinary Syndrome of Menopause* (GSM)

Genitale Symptome sind die prominentesten Symptome des GSM. Darunter fallen neben der vaginalen Trockenheit, die das Kardinalsymptom des GSM darstellt und von rund 75 % der symptomatischen, postmenopausalen Frauen berichtet wird, auch vaginales Brennen, vaginaler Juckreiz oder vaginale Schmerzen (15, 21, 25, 27). In der italienischen AGATA-Studie

(n = 722) berichteten sogar 100 % der Frauen mit dem GSM von vaginaler Trockenheit, 77,6 % von Schmerzen beim Geschlechtsverkehr (= Dyspareunie), 56,9 % von Brennen und 56,6 % von Juckreiz im Intimbereich (28).

Da auch der Harntrakt von den physiologischen Veränderungen betroffen ist, treten häufig auch urologische Symptome, wie vermehrter oder verstärkter Harndrang, Dysurie (= erschwerte und/oder schmerzhafte Blasenentleerung), Nykturie (= mehrfaches Wasserlassen während der Nacht), und Inkontinenz auf. Auch wiederkehrende Harnwegsinfektionen werden von rund 20 % der postmenopausalen Frauen berichtet, da die nachlassende Funktionsfähigkeit des Urogenitaltraktes das Auftreten von Infektionen begünstigen kann (21, 24, 25). Dysurie wurde in der italienischen AGATA-Studie von 36,1 % der Frauen mit dem GSM berichtet (28).

Sexuelle Symptome ergeben sich zum Teil als Konsequenz aus den beiden vorherigen Symptomkomplexen. Aufgrund physiologischer Veränderungen mangelt es beim Geschlechtsverkehr an Feuchtigkeit, was wiederum Dyspareunie, postkoitale Blutungen oder Unbehagen bei den Frauen auslösen kann. Genitale und urologische Symptome tragen dazu bei, dass das Verlangen und die Erregung der Frauen abnehmen und weniger Orgasmen erlebt werden (15, 24, 25). Der sinkende Östrogenspiegel steht dabei auch im Zusammenhang mit einer verringerten Libido.

Zur Diagnostik wird häufig der *Vaginal Health Index* (VHI) erhoben (29), der sich aus fünf verschiedenen Parametern zusammensetzt: der vaginalen Elastizität, dem Scheidenausfluss (Fluor), dem pH-Wert, der Beschaffenheit des Epithels bzw. der Schleimhaut und der vaginalen Feuchtigkeit (30). Ein niedriger VHI Score korreliert mit einer stärker ausgeprägten VVA (31). Wie bereits erwähnt, ist die vaginale Trockenheit als Kardinalsymptom des GSM bzw. der VVA aufzuführen, die auch bei älteren Damen sehr präsent ist (27). Viele Frauen haben Hemmungen, ihre Symptome bei ihrem Gynäkologen bzw. ihrer Gynäkologin anzusprechen, weshalb dem Arzt bzw. der Ärztin eine entscheidende Rolle bei der Befragung der Patientinnen und dem Deuten der Symptome zukommt. Einige äußere genitale und urologische Anzeichen können bei der Untersuchung festgestellt werden. Das GSM wird häufig diagnostiziert, wenn eine Patientin von Schmerzen beim Geschlechtsverkehr berichtet, die auf vaginale Trockenheit zurückzuführen sind (25). Eine objektive Einschätzung von ärztlicher Seite ist dennoch notwendig, da selbst berichtete Symptome nur schwach mit den objektiven Untersuchungsergebnissen korrelieren. Der vaginale pH-Wert stellt beispielsweise einen reliablen Indikator für die vaginale Gesundheit und das Ausmaß der Östrogen-Deprivation dar (32). Eine einfache Checkliste, um die VVA bzw. das GSM zu diagnostizieren, ist in Abbildung 4 dargestellt (30).

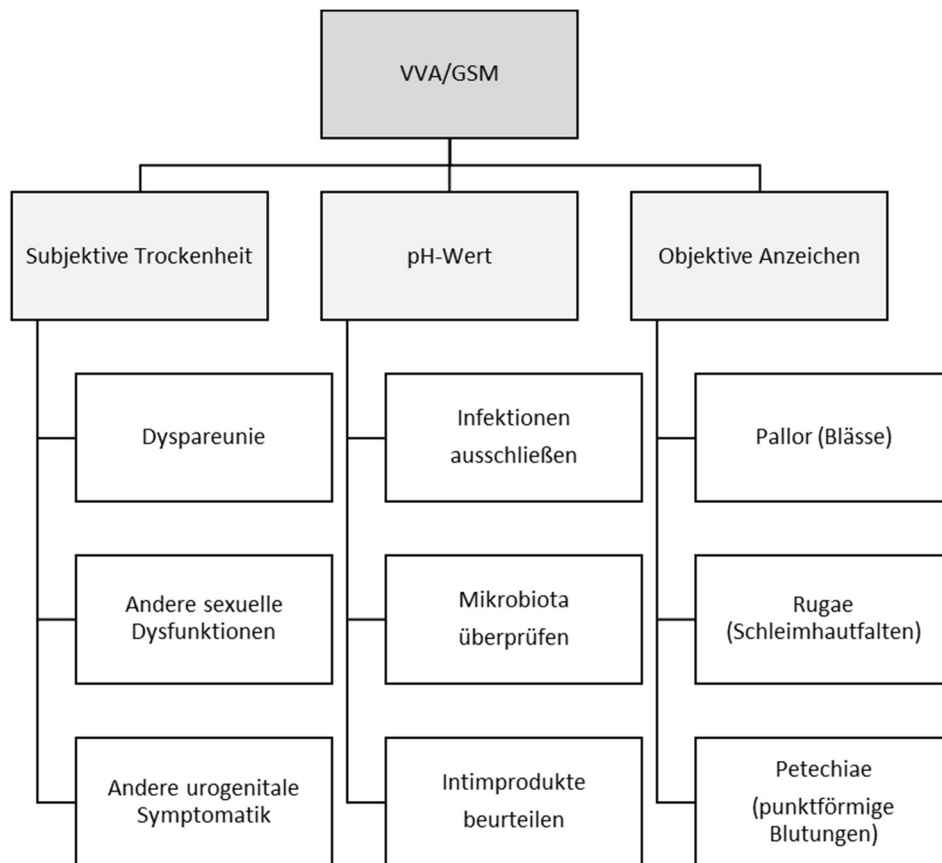


Abbildung 4 Checkliste, um die VVA bzw. das GSM in der klinischen Praxis zu diagnostizieren, übersetzt und zitiert nach Nappi *et al.* (30)

Abkürzungen. VVA = vulvovaginale Atrophie, GSM = Genitourinary Syndrome of Menopause.

Wie man sieht, enthält diese Checkliste neben objektiven Anzeichen und dem pH-Wert auch eine subjektive Komponente, in die die Patientin miteinbezogen wird.

1.3.2 Auswirkungen des GSM bzw. der VVA auf die Lebensqualität und Sexualität

In der letzten Dekade wurde häufig untersucht, wie sich das GSM bzw. die VVA auf die Sexualität und somit auch auf die Lebensqualität auswirkt. Laut internationalen Umfragen werden durch die Beschwerden verschiedene sexuelle Bereiche, wie die Befriedigung, die Intimität, die Spontaneität, eine liebevolle Beziehung mit dem Partner bzw. der Partnerin und die sexuelle Aktivität eingeschränkt. Man geht jedoch auch davon aus, dass unabhängig vom Alter das sexuelle Interesse bei vielen Frauen bestehen bleibt. Der wichtigste Faktor, der dazu beiträgt, dass eine Frau im Alter sexuell aktiv bleibt, ist eine romantische, partnerschaftliche Beziehung. In einer Umfrage in den USA wiesen verheiratete oder in Partnerschaft lebende Frauen eine ungefähr achtmal höhere sexuelle Aktivität auf, als diejenigen, die in keiner Partnerschaft lebten. Über 30 % der über 65-Jährigen berichteten von einer mindestens einmal wöchentlich stattfindenden sexuellen Aktivität (33). Obwohl Sex an Wichtigkeit beibehält,

nimmt die sexuelle Funktionsfähigkeit und somit auch zwangsläufig die sexuelle Aktivität der Frauen mit dem Alter ab (34). In einer Kohorte des *Women's Healthy Ageing Project* (WHAP) zeigte sich beispielsweise eine signifikante Abnahme der sexuellen Aktivität um 18 % zwischen Frauen der frühen Postmenopause und Frauen der späten Postmenopause (35). Durch die hormonellen und physiologischen Veränderungen lässt die vaginale Elastizität im Alter nach, das Vaginalsekret wird nur noch eingeschränkt produziert und häufig resultiert eine VVA, die Schmerzen beim Geschlechtsverkehr nach sich ziehen kann und betroffene Frauen in ihrer sexuellen Aktivität einschränken kann (34). In Bezug auf die Lebensqualität beeinträchtigt das GSM bzw. die VVA auch den Schlaf der Betroffenen, sportliche, berufliche oder soziale Aktivitäten, die weibliche Rolle, den „Jugendsinn“ und das allgemeine Vergnügen (30).

Die VIVA (*Vaginal Health: Insights, Views & Attitudes*) Umfrage in Europa und den USA im Jahre 2013 legte offen, dass 80 % aller postmenopausalen Frauen der vaginalen Atrophie negative Auswirkungen auf das Leben zuschrieben, vor allem in Bezug auf Intimität und Liebesbeziehungen. Mehr als die Hälfte aller Frauen stimmte zu, dass ihre Beschwerden ihre Freiheit, das zu tun, was sie wollen, einschränkten, einschließlich des Auslebens ihrer Sexualität. Viele Frauen fühlten sich alt und von ihrem Körper im Stich gelassen, da ihre vaginalen Beschwerden ihrer früheren Spontaneität im Wege standen (36).

Die CLOSER (*CLarifying Vaginal Atrophy's Impact On SEx and Relationships*) Umfrage in Europa und Nordamerika ein Jahr später (2014) hatte zum Ziel, Auswirkungen von vaginalen Beschwerden und lokaler Östrogen Therapie auf die Intimität, Beziehungen und das Selbstwertgefühl der Frauen zu eruieren. Frauen mit vulvovaginalen Symptomen und deren Partner füllten den Fragebogen aus. Es zeigte sich, dass vaginale Beschwerden bei rund 60 % der Frauen für eine Vermeidung von Intimität, für einen Verlust der Libido und Schmerzen beim Geschlechtsverkehr sorgten. Allein die Befürchtung, dass der Geschlechtsverkehr schmerzhaft sein könnte, führte zur Vermeidung von Intimitäten. Auch ein Großteil der Männer glaubte, dass die vaginalen Beschwerden ihrer Partnerinnen für den Verzicht auf Intimität, den Libidoverlust und die Schmerzen beim Geschlechtsverkehr verantwortlich waren. Ungefähr 30 % der Frauen und Männer berichteten, dass vaginale Beschwerden der Grund dafür waren, dass kein Geschlechtsverkehr mehr stattfand. Frauen mit einer lokalen Östrogen Therapie berichteten weniger häufig von derartigen Einschränkungen. Vaginale Beschwerden zeigten auch negative Auswirkungen auf den Selbstwert von betroffenen Frauen, die sich dadurch nicht mehr sexuell anziehend fanden und ihr Selbstvertrauen in sich als Sexualpartnerin verloren hatten. Aus dieser Studie lässt sich schließen, dass nicht nur die Frauen selbst mit den

Einschränkungen ihrer Beschwerden zu kämpfen haben, sondern auch deren Männer häufig auf eine gesunde partnerschaftliche Sexualität verzichten müssen (37).

Prinzipiell wirken sich alle drei Symptombereiche des GSM negativ auf die Lebensqualität von betroffenen Frauen aus. Je schwerer vor allem vulvovaginale Beschwerden sind, desto stärker ist die Lebensqualität von betroffenen Frauen beeinträchtigt (38). Tritt gleichzeitig zu vaginalen Symptomen beispielsweise Inkontinenz auf, reduziert sich die Lebensqualität der Frauen noch weiter, da sich die verschiedenen Symptomkomplexe kumulieren. Die Behandlung der Inkontinenz sollte somit bei der Behandlung der vaginalen Symptome berücksichtigt werden (39). Viele Frauen sprechen jedoch nicht mit ihren ÄrztInnen über ihre Beschwerden, weil sie entweder gar nicht wissen, dass eine entsprechende Behandlung verfügbar ist, oder weil eine gewisse Hemmschwelle besteht, die Symptome bei ihrem Arzt bzw. ihrer Ärztin anzusprechen. Diese Hemmschwelle zeigt sich vor allem, wenn es sich um einen jungen, männlichen Arzt handelt, wenn bereits eine schlechte Erfahrung mit einem Arzt bzw. einer Ärztin gemacht wurde oder wenn die Patientin ihre Symptome als Teil des ‚normalen Alterns‘ sieht, mit denen sie sich abfinden muss. Viele Frauen warten eher darauf, dass ihr Arzt bzw. ihre Ärztin mit entsprechenden Fragen auf sie zukommt. Demzufolge leiden viele Frauen im Stillen und viele Beschwerden werden nicht diagnostiziert und bleiben somit auch unbehandelt (24, 40).

1.3.3 PROMs beim GSM bzw. der VVA

Das GSM bzw. die VVA eignet sich sehr gut als Anwendungsbeispiel, da in diesem Bereich eine sehr hohe Forschungsaktivität stattfindet. Aufgrund der hohen Prävalenz des GSM, die davon ausgeht, dass jede zweite Frau im Laufe der Menopause urogenitale Beschwerden entwickelt, die von alleine nicht wieder verschwinden und in ihrer Intensität eher zunehmen (24), hat vor allem die Pharmaindustrie ein sehr großes Interesse daran, entsprechende Präparate und Therapien zu entwickeln und deren Wirksamkeit zu testen.

Genau in solchen klinischen Studien, die die Wirksamkeit einer neuen Behandlung überprüfen, kommen PROMs zum Einsatz, um die Patientenperspektive abzubilden. Viele bereits publizierte Interventionsstudien (41-48) setzten mindestens eine PROM ein, um einen PRO zu erfassen, und auch bei den registrierten Studien auf *clinicaltrials.gov* ist bei fast allen Studien ein PRO entweder als primärer oder sekundärer Endpunkt gelistet. Jedoch ist der Einsatz von PROMs bei klinischen GSM-Studien sehr heterogen, beinahe willkürlich. Neben generischen PROMs (z. B. die *Patient Global Impression of Improvement* (PGI)) werden auch spezifische PROMs zur sexuellen Dysfunktion (z. B. der *Female Sexual Function Index* (FSFI)) oder zur Inkontinenz (z. B. der *International Consultation on Incontinence Questionnaire Overactive*

Bladder (ICIQ-OAB) eingesetzt. Wie vorher beschrieben, hat das GSM viele Auswirkungen auf die Lebensqualität einschließlich der Sexualität von betroffenen Frauen. Diese Auswirkungen sind nicht nur sehr subjektiv und können somit nur von den Betroffenen selbst berichtet werden, sie sind vor allem auch spezifisch für das GSM. Viele PROMs, die in klinischen Studien eingesetzt werden, wurden jedoch nie in der Zielpopulation, nämlich bei postmenopausalen Frauen mit urogenitalen Beschwerden validiert, d. h. es ist gar nicht klar, ob diese PROMs überhaupt in dieser Population funktionieren. Laut Vieira-Baptista *et al.* (49) werden dringend Instrumente und Fragebögen benötigt, um jede Komponente des GSM beurteilen zu können. Generell herrscht große Unklarheit darüber, welche GSM- bzw. VVA-spezifischen PROMs es überhaupt gibt, welche PROMs in der GSM- bzw. VVA-Population validiert wurden und wie diese Validierungsergebnisse aussehen.

1.4 Ziele der Arbeit

Bisher ist gänzlich unklar, welche und wie viele PROMs für das GSM bzw. die VVA existieren. Es wurde weder ein systematischer Vergleich noch eine Qualitätsbeurteilung von PROMs und deren Entwicklungs- und Validierungsstudien, die für dieses spezielle Beschwerdebild existieren, durchgeführt.

Erstes Ziel dieser Arbeit ist es somit, alle existierenden PROMs für das GSM bzw. die VVA zu identifizieren und deren methodische Qualität zu beurteilen, um am Ende eine Empfehlung abgeben zu können, welche PROM(s) in künftigen klinischen GSM-Studien eingesetzt werden sollte(n) (vgl. Kapitel 2).

In einem zweiten Schritt wird das Ziel verfolgt, das beste Kandidateninstrument in die deutsche Sprache zu übersetzen (sofern nicht in deutscher Sprache vorhanden) und in einer deutschsprachigen, peri- und postmenopausalen Population quantitativ zu validieren. Mit dieser Validierung soll weitere Evidenz zu verschiedenen Messeigenschaften, wie der Strukturvalidität, der internen Konsistenz, der Test-Retest-Reliabilität, der Konstruktvalidität und der Änderungssensitivität (*responsiveness*) generiert werden (vgl. Kapitel 3).

Als drittes und letztes Ziel soll der bisher unklare Stellenwert der urologischen Komponente mit einer qualitativen Methode exploriert werden, um daraus Handlungsimpulse für die künftige Erfassung der Auswirkung urologischer Beschwerden bei Frauen mit dem GSM ableiten zu können (vgl. Kapitel 4).

2. Teil 1: Systematische Übersichtsarbeit

Ziel dieser systematischen Übersichtsarbeit ist es, alle existierenden PROMs für das GSM bzw. die VVA zu identifizieren und deren methodische Qualität zu beurteilen, um am Ende die PROM(s) empfehlen zu können, die in künftigen GSM-Studien eingesetzt werden sollte(n). Die systematische Übersichtsarbeit wurde bereits im Journal *Menopause* veröffentlicht (50).

2.1 Material

2.1.1 Die COSMIN *Risk of Bias Checklist*

Die COSMIN *Risk of Bias Checklist* wurde speziell für systematische Übersichtsarbeiten entwickelt (51). Diese Checkliste hilft bei der Beurteilung der methodischen Qualität von Studien. Sie besteht aus zehn Boxen (vgl. Tabelle 1), wobei nur die Boxen für diejenigen Messeigenschaften ausgefüllt werden, die in einem Artikel berichtet werden. Die Messeigenschaften jeder Studie werden auf einer Likert-Skala mit vier Antwortmöglichkeiten bewertet („*inadequate*“, „*doubtful*“, „*adequate*“ und „*very good*“). Beim *Overall-Rating* einer Box gilt das Prinzip „*the worst score counts*“, d. h. die niedrigste Bewertung in einer Box bestimmt die Gesamtqualität der einzelnen Studie. Jede Studie zu einer Messeigenschaft wird dabei separat bewertet (10).

Tabelle 1 Die 10 Boxen der COSMIN *Risk of Bias Checklist*

Box 1.	<i>PROM development</i>	<i>Content validity</i>
Box 2.	<i>Content validity</i>	
Box 3.	<i>Structural validity</i>	<i>Internal structure</i>
Box 4.	<i>Internal consistency</i>	
Box 5.	<i>Cross-cultural validity/measurement invariance</i>	
Box 6.	<i>Reliability</i>	<i>Remaining measurement properties</i>
Box 7.	<i>Measurement error</i>	
Box 8.	<i>Criterion validity</i>	
Box 9.	<i>Hypotheses testing for construct validity</i>	
Box 10.	<i>Responsiveness</i>	

2.1.2 Kriterien guter Messeigenschaften

Die COSMIN-Gruppe formulierte Qualitätskriterien für gute Messeigenschaften (52), anhand derer die Ergebnisse jeder einzelnen Studie als entweder ausreichend (+), ungenügend (–) oder unbestimmt (?) beurteilt werden sollen (10). Diese Kriterien sind in Tabelle 2 dargestellt.

Tabelle 2 Qualitätskriterien für gute Messeigenschaften gemäß COSMIN (52)

Measurement property	Rating	Criteria
Structural validity	+	CTT: CFA: CFI or comparable measure > 0.95 OR RMSEA < 0.06 OR SRMR < 0.08 ^a IRT/Rasch: No violation of <u>unidimensionality</u> ^b : CFI or TLI or comparable measure > 0.95 OR RMSEA < 0.06 OR SMRM < 0.08 AND no violation of <u>local independence</u> : residual correlations among the items after controlling for the dominant factor < 0.20 OR Q3's < 0.37 AND no violation of <u>monotonicity</u> : adequate looking graphs OR item scalability > 0.30 AND adequate <u>model fit</u> IRT: $\chi^2 > 0.001$ Rasch: infit and outfit mean squares ≥ 0.5 and ≤ 1.5 OR Z-standardized values > -2 and < 2
	?	CTT: not all information for '+' reported; IRT/Rasch: model fit not reported
	-	Criteria for '+' not met
Internal consistency	+	At least low evidence ^c for sufficient structural validity ^d AND Cronbach's alpha(s) ≥ 0.70 for each unidimensional scale or subscale ^e
	?	Criteria for "At least low evidence ^c for sufficient structural validity ^d " not met
	-	At least low evidence ^c for sufficient structural validity ^d AND Cronbach's alpha(s) < 0.70 for each unidimensional scale or subscale ^e
Reliability	+	ICC or weighted Kappa ≥ 0.70
	?	ICC or weighted Kappa not reported
	-	ICC or weighted Kappa < 0.70
Measurement error	+	SDC or LoA < MIC ^d
	?	MIC not defined
	-	SDC or LoA > MIC
Hypotheses testing for construct validity	+	The result is in accordance with the hypothesis ^f
	?	No hypothesis defined (by the review team)
	-	The result is not in accordance with the hypothesis ^f
Cross-cultural validity/measurement invariance	+	No important differences found between group factors (such as age, gender, language) in multiple group factor analysis OR no important DIF for group factors (McFadden's $R^2 < 0.02$)
	?	No multiple group factor analysis OR DIF analysis performed
	-	Important differences between group factors OR DIF was found
Criterion validity	+	Correlation with gold standard ≥ 0.70 OR AUC ≥ 0.70
	?	Not all information for '+' reported
	-	Correlation with gold standard < 0.70 OR AUC < 0.70
Responsiveness	+	The result is in accordance with the hypothesis ^f OR AUC ≥ 0.70
	?	No hypothesis defined (by the review team)
	-	The result is not in accordance with the hypothesis ^f OR AUC < 0.70

Anmerkungen. Diese Kriterien basieren auf Terwee et al. (53) und Prinsen et al. (54)

Abkürzungen. AUC = area under the curve, CFA = confirmatory factor analysis, CFI = comparative fit index, CTT = classical test theory, DIF = differential item functioning, ICC = intraclass correlation coefficient, IRT = item response theory, LoA = limits of agreement, MIC = minimal important change, RMSEA = root mean square error of approximation, SEM = standard error of measurement, SDC = smallest detectable change, SRMR = standardized root mean residuals, TLI = Tucker-Lewis Index, "+" = hinreichend, "-" = ungenügend, "?" = unbestimmt.

^aDie Faktorenstruktur sollte über die Studien hinweg gleich sein, ^bEindimensionalität bezieht sich auf eine Faktorenanalyse pro Subskala, während sich Strukturvalidität auf eine Faktorenanalyse einer (mehrdimensionalen) *patient-reported outcome measure* bezieht, ^cEvidenz wird gemäß des *GRADE approach*s herabgestuft, ^dEvidenz kann von verschiedenen Studien kommen, ^eDas Kriterium 'Cronbach's alpha < 0,95' wurde gelöscht, weil es nur in der PROM-Entwicklungsphase relevant ist und nicht, wenn eine existierende PROM bewertet wird, ^fAlle Ergebnisse sollten zusammengefasst werden, dann sollte entschieden werden, ob 75 % der Ergebnisse mit den Hypothesen übereinstimmen.

2.1.3 Der *GRADE approach*

Der *GRADE* (= *Grading of Recommendations Assessment, Development and Evaluation approach*) hilft, zu entscheiden, ob ein zusammengefasstes Ergebnis zuverlässig ist oder nicht. Die Qualität der Evidenz, d. h. die Qualität aller Studien, die im zusammengefassten Ergebnis enthalten sind, steht hier im Fokus und soll gezielt betrachtet werden, um irreführende Empfehlungen zu vermeiden (55). Es wird bestimmt, ob den Schätzungen der Messeigenschaften vertraut werden kann. Die COSMIN-Gruppe empfiehlt, einen *GRADE approach* mit vier *GRADE* Faktoren zu nutzen: Verzerrungspotential (*risk of bias*), Inkonsistenz (*inconsistency*), Ungenauigkeit (*imprecision*) und Indirektheit (*indirectness*). Die Qualität der Evidenz kann in vier Stufen eingeteilt werden (hoch, moderat, niedrig oder sehr niedrig) und sinkt, je nachdem, wie stark dem jeweiligen Faktor bzw. wie vielen Faktoren nicht vertraut werden kann. Jede PROM wird hierbei einzeln eingestuft. Ist das *Overall-Rating* einer Messeigenschaft unbestimmt (?), kann die Qualität der PROM hinsichtlich dieser Messeigenschaft nicht beurteilt werden und somit auch die Qualität der Evidenz nicht eingestuft werden (52).

2.2 Methode

2.2.1 Das Protokoll: Registrierung

Die systematische Übersichtsarbeit entstand im Einklang mit den Empfehlungen der „*Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols* (PRISMA-P)“ Statements (56) und der „*COSMIN guideline and manual for systematic reviews of PROMs*“ (10, 52). Das Protokoll wurde im *International Prospective Register of Systematic Reviews* (PROSPERO, <https://www.crd.york.ac.uk/prospero/>) registriert: CRD42018092384.

2.2.2 Literaturrecherche

Die Suchstrategie für die systematische Übersichtsarbeit wurde unter Mithilfe des Regensburger Bibliothekars Dr. rer. nat. Helge Knüttel entwickelt. Die großen Datenbanken *MEDLINE* (via Ovid, 1946 – heute), *EMBASE* (via Ovid, 1974 – heute), *Science Citation Index Expanded* (*Web of Science*, 1965 – heute) und *Social Sciences Citation Index* (*Web of Science*, 1990 – heute) wurden durchsucht. Die Suchstrategien für alle durchsuchten bibliographischen Datenbanken sind im Anhang zu finden (Anhang A – C). Folgende Elemente waren in der Suchstrategie enthalten (10):

- A. Zielgruppe: Frauen mit dem GSM. Für maximale Sensitivität wurden viele verschiedene Terme verwendet. GSM ist ein sehr neuer und variabler Begriff und nicht immer wird

das Syndrom beim Namen genannt (15). Deshalb wurden zudem alle wichtigen Symptome des GSM in die Suchstrategie aufgenommen.

- B. Konstrukt: Alle PROMs interessierten, egal welches Konstrukt diesen zugrunde lag. Um optimale Sensitivität zu erhalten, wurde für dieses Element der Suchstrategie eine Kombination aus dem *PubMed*-Filter „*Quality of life (QoL)*“ von Vissers und de Vries (57), dem *PubMed*-Filter „*Patient-reported outcome measures (PROMs)*“ von Jansma und de Vries (58) und zusätzlichen Suchbegriffen aus dem „*PROM group construct & instrument type filter*“ von Mackintosh *et al.* (59) verwendet. *Patient-reported outcome measures* ist ein sehr weiter Begriff, der alle Instrumente, die auf Patienteneinschätzungen beruhen, umfasst. Ein Großteil aller PROMs misst Konstrukte wie Lebensqualität oder den Gesundheitszustand (6, 60).
- C. Messeigenschaften: Ein sensitiver Suchfilter, der von der COSMIN-Gruppe empfohlen wird (52) und speziell für die Suche nach Studien zu Messeigenschaften von Terwee *et al.* (61) entwickelt wurde, wurde verwendet.
- D. Machbarkeit von Instrumenten: Die Suchbegriffe für dieses Element wurden aus dem Protokoll von Heintz *et al.* (62) zu Messinstrumenten zum Atopischen Ekzem übernommen (die letzten fünf Suchbegriffe des Statements #1).
- E. Individuelle Instrumente: Bereits bekannte Messinstrumente wurden aufgenommen.
- F. Exklusionsfilter: Ein Exklusionsfilter von Terwee *et al.* (61), der irrelevante Publikationsformate und Tierstudien ausschloss, wurde verwendet.

Die einzelnen Elemente sollten alle Artikel zu Messeigenschaften oder Machbarkeit von PROMs für Frauen mit dem GSM oder Artikel, in denen GSM-spezifische Messinstrumente genannt werden, auffindig machen. Die Suchelemente wurden auf folgende Weise miteinander kombiniert: ((A UND B UND (C ODER D)) ODER (C UND E)) NICHT F, oder in Worten: ((Zielgruppe UND Konstrukt UND (Messeigenschaften ODER Machbarkeit)) ODER (individuelle Instrumente UND Messeigenschaften)) NICHT Exklusionsfilter.

Zusätzlich wurden spezifische Datenbanken nach Treffern für Messinstrumente, die für die Zielgruppe relevant sein könnten, durchsucht: PROQOLID (<https://eprovide.mapi-trust.org/about/about-proqolid>), die COSMIN *database of systematic reviews of outcome measurement instruments* (<http://www.cosmin.nl/database-of-systematic-reviews.html>), das Testarchiv des Leibniz-Zentrums für Psychologische Information und Dokumentation (<https://www.testarchiv.eu/>) und *PubPsych* (<https://pubpsych.zpid.de/pubpsych/>).

Im Anschluss wurden die Literaturlisten eingeschlossener Studien durchgesehen, um weitere Instrumente spezifisch für das GSM zu finden. Bezüglich des Publikationsdatums gab es keine Restriktionen. Nur Artikel in den Sprachen Englisch, Deutsch, Französisch und Italienisch wurden eingeschlossen. Alle Treffer wurden in *Covidence* (<https://www.covidence.org/>), einer kostenpflichtigen Software speziell für systematische Übersichtsarbeiten, die das Screenen und Verwalten der Literatur erleichtert, hochgeladen.

2.2.3 Einschlusskriterien und Studiena Auswahl

Die Einschlusskriterien wurden gemäß der „COSMIN *guideline for systematic reviews of patient-reported outcome measures*“ (52) festgelegt (vgl. Tabelle 3).

Tabelle 3 Ein- und Ausschlusskriterien

	Einschlusskriterien	Ausschlusskriterien
Population	Postmenopausale Frauen	Alle anderen
Studiendesign	PROM Entwicklungsstudie, Validierungsstudie	Alle anderen Studiendesigns
Outcome	Alle <i>patient-reported outcomes</i>	<i>Non-patient-reported outcomes</i> , wie Biomarker oder Physiologie der Haut
Typ des Mess-instruments	<i>Patient-reported outcome measures</i> spezifisch für mind. eine der beiden Hauptkomponenten (genital und urologisch) des GSM	Alle anderen
Publikationstyp	Artikel mit verfügbarem Volltext	Abstracts

Abkürzungen. PROM = *patient-reported outcome measure*.

Die relevante Population waren postmenopausale Frauen, da davon ausgegangen wurde, dass bis zu 50 % aller postmenopausalen Frauen vulvovaginale und/oder urologische Symptome aufweisen. Um mit diesem Populationskriterium keine PROMs einzuschließen, die nur auf die Menopause zugeschnitten waren (z. B. die *Menopause Rating Scale* oder der *MENQOL*), aber nicht speziell auf urogenitale Beschwerden, die mit der Menopause einhergingen, wurde ein zusätzliches Kriterium bezüglich der PROM formuliert. Somit wurden nur PROMs eingeschlossen, die spezifisch für mindestens eine der beiden Hauptkomponenten (genital und urologisch) des GSM waren. Die sexuelle Komponente wurde als Einschlusskriterium nicht aufgenommen, da PROMs, die nur diese Komponente bei postmenopausalen Frauen abfragten,

zu unspezifisch waren, um das Syndrom zu erfassen, und sich das Krankheitsbild z. B. in Richtung sexuelle Dysfunktion verschieben konnte. Zudem wurden Entwicklungs- und/oder Validierungsstudien von PROMs als relevant erachtet, d. h. Studien, in denen die Entwicklung der jeweiligen PROM beschrieben wurde und/oder Informationen zu Messeigenschaften, Interpretierbarkeit oder Machbarkeit enthalten waren. Nur Studien mit verfügbarem Volltext wurden eingeschlossen, da Abstracts oft nur wenig Auskunft über das Studiendesign geben.

Die gefundenen Titel und Abstracts wurden von zwei unabhängigen Reviewern beurteilt. Anschließend wurden für alle relevanten Studien Volltexte besorgt. Alle Volltexte waren entweder frei zugänglich oder wurden per Fernleihe bestellt. Unstimmigkeiten wurden im Forschungsteam diskutiert. Sobald eine Studie von nur einem Reviewer anhand des Abstracts als relevant erachtet wurde, ging die Studie ins Volltext-Screening über.

2.2.4 Datenextraktion

2.2.4.1 Beurteilung der Inhaltsvalidität

Die Inhaltsvalidität der eingeschlossenen Studien wurde mithilfe des COSMIN-Manuals für die Beurteilung der Inhaltsvalidität von PROMs evaluiert (12). In einem ersten Schritt wurden die ersten beiden Boxen der COSMIN *Risk of Bias Checklist*, *PROM development* und *Content validity*, ausgefüllt. Hierbei wurden die Qualität der PROM-Entwicklung und die Qualität von Inhaltsvalidierungsstudien beurteilt. Im Anschluss wurde jedes Ergebnis anhand von zehn Kriterien für gute Inhaltsvalidität bewertet (+/-/±/?) und auch die Reviewer gaben für jedes der Kriterien eine Bewertung ab (+/-/±). Alle Ergebnisse verfügbarer Studien wurden in einem nächsten Schritt qualitativ zusammengefasst und die Qualität der Evidenz wurde mithilfe des *GRADE approaches* eingestuft. War Evidenz von hoher Qualität gegeben, dass die Inhaltsvalidität einer PROM ungenügend ist, wurden keine weiteren Beurteilungen mehr durchgeführt und die PROM wurde sofort in die schlechteste Kategorie (Kategorie C, vgl. Absatz 2.2.5) eingestuft.

2.2.4.2 Beurteilung der übrigen Messeigenschaften

Hierbei erfolgte wieder der Einsatz der COSMIN *Risk of Bias Checklist*, allerdings wurden nur die Boxen mit den Messeigenschaften ausgefüllt, die im jeweiligen Artikel beurteilt wurden. Anschließend wurden in sogenannten Evidenztabelle die Ergebnisse aus den einzelnen Artikeln anhand von Qualitätskriterien guter Messeigenschaften bewertet und jede Messeigenschaft pro Artikel als ausreichend (+), unbestimmt (?) oder ungenügend (-) eingestuft. In einem nächsten Schritt wurden alle Ergebnisse zu jeder einzelnen Messeigenschaft einer PROM zusammengefasst und ein *Overall-Rating* anhand der

Qualitätskriterien vergeben. Zusätzlich wurden weitere Informationen zu Charakteristiken der Studienpopulationen und der eingeschlossenen PROMs und zur Interpretierbarkeit und Machbarkeit der PROMs aus den Artikeln extrahiert und in Tabellen festgehalten. Zum Schluss wurden *Summary of Findings* Tabellen erstellt, in die für jede Messeigenschaft jeder PROM das zusammengefasste Ergebnis und das *Overall-Rating* eingetragen wurde sowie in einer letzten Spalte die Qualität der Evidenz mithilfe des *GRADE approach* eingestuft wurde.

2.2.5 Abgabe einer Empfehlung

Um eine abschließende Empfehlung abgeben zu können, welche PROM(s) in Zukunft verwendet werden sollte(n), wurden die eingeschlossenen PROMs in drei Kategorien entsprechend ihrer Messeigenschaften eingeteilt:

- A. PROMs mit Evidenz für hinreichende Inhaltsvalidität (egal welches Level) UND mindestens Evidenz niedriger Qualität für hinreichende interne Konsistenz
- B. PROMs, die weder in A noch C eingeordnet werden können
- C. PROMs mit Evidenz hoher Qualität für eine ungenügende Messeigenschaft

PROMs, die Kategorie A zugeordnet werden, können verwendet werden und die Ergebnisse, die mit diesen PROMs erzielt werden, können als zuverlässig angesehen werden. Für PROMs aus Kategorie B bedarf es weiterer Validierung, bis eine Entscheidung ausgesprochen werden kann. PROMs in Kategorie C sollten nicht eingesetzt werden. Werden nur PROMs aus Kategorie B gefunden, kann vorläufig die PROM mit der besten Evidenz für Inhaltsvalidität empfohlen werden, bis zusätzliche Evidenz vorliegt.

2.3 Ergebnisse

2.3.1 Literaturrecherche

Die Literatursuche in den drei Datenbanken *MEDLINE*, *EMBASE* und *Web of Science* wurde am 23. April 2018 durchgeführt. Insgesamt wurden 9224 Treffer gefunden. Eine vorläufige De-Duplizierung wurde bereits in *EndNote* mittels einer Methode von Bramer *et al.* (63) vorgenommen und reduzierte die Treffer auf 5769. Anschließend erfolgte der Import in *Covidence*, wobei noch einmal 85 Duplikate erkannt wurden. Somit gingen 5684 Studien in das Screening ein. Das Durchsuchen der ausgewählten kleineren Datenbanken erbrachte keine zusätzlichen Studien. Eine zweite Literatursuche am 7. Dezember 2018 erbrachte in der finalen Phase der systematischen Übersichtsarbeit weitere 638 Treffer, von denen 393 Studien nach der De-Duplizierung ins Screening gingen.

2.3.2 Titel-, Abstract- und Volltext-Screening

Alle Titel und Abstracts der 6077 gefundenen Studien wurden von zwei unabhängigen Reviewern in *Covidence* beurteilt. Durch das Screening sollten irrelevante Artikel aussortiert werden. Hierbei wurde sehr liberal gescreent, indem alle Artikel zu PROMs behalten wurden, die im weitesten Sinne mit den relevanten Symptomen zu tun hatten. Konflikte wurden im Team besprochen und direkt in der Software gelöst. Insgesamt 142 Studien wurden als relevant erachtet und gingen ins Volltext-Screening ein (vgl. Abbildung 5).

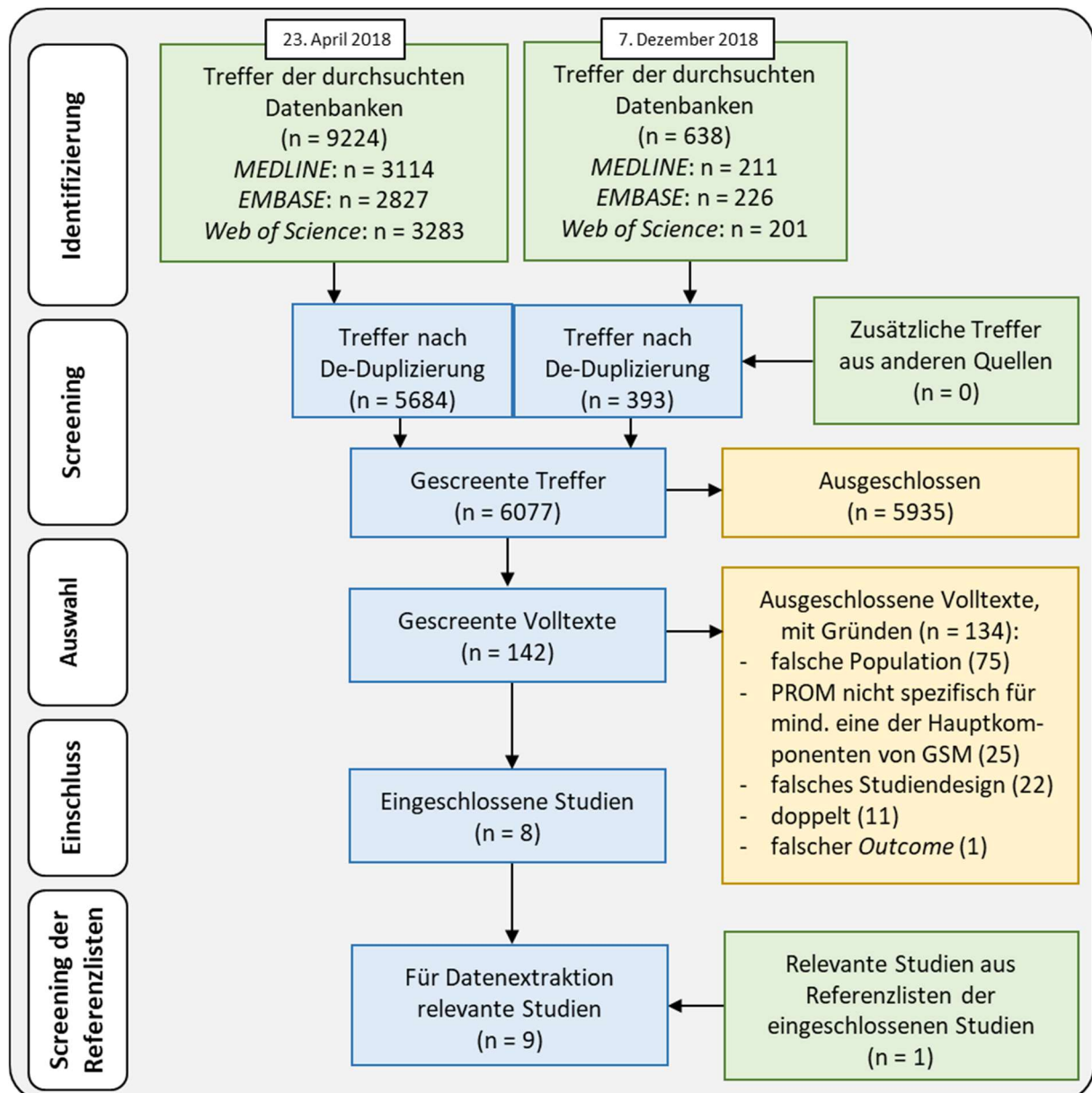


Abbildung 5 Adaptiertes 2009 PRISMA Flussdiagramm (64). Weitere Informationen unter www.prisma-statement.org

Anmerkungen. n = Anzahl an Studien.

Um einen ersten Überblick zu gewinnen, wurden die im Screening identifizierten PROMs geclustert. Da die Suchstrategie sehr umfangreich und detailliert war, wurden PROMs zum relevanten Bereich des GSM bzw. der VVA und den einzelnen Symptomkomponenten gefunden, aber auch PROMs zur Menopause im Allgemeinen und PROMs zu verschiedenen Erkrankungen des Beckenbodens, da urologische Symptome in der Suchstrategie enthalten waren (vgl. Abbildung 6).

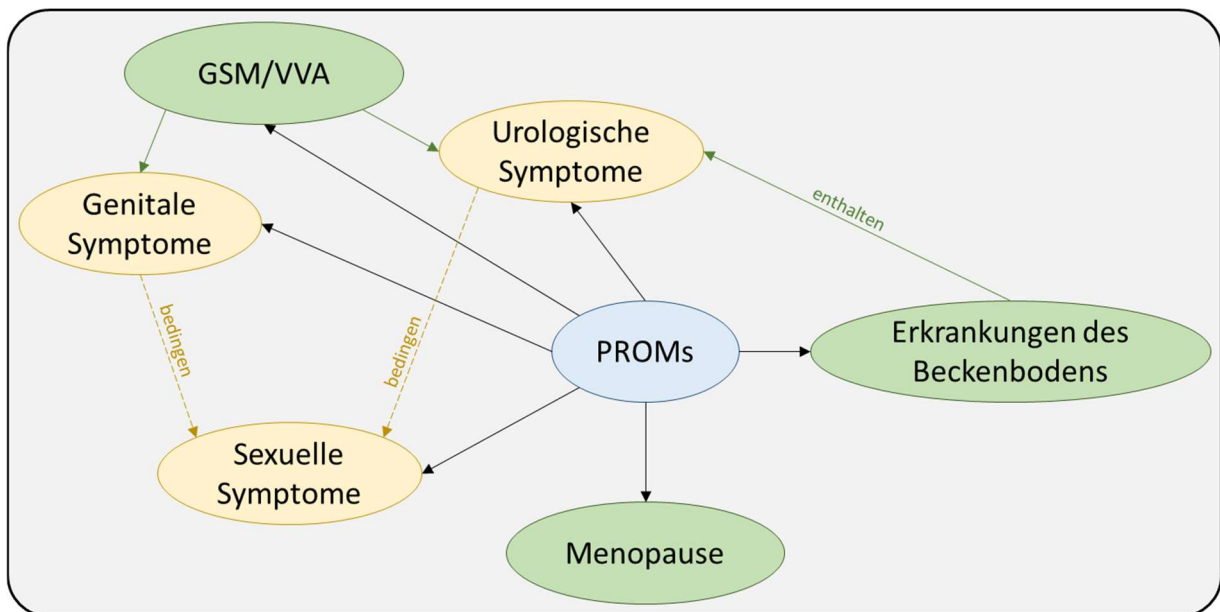


Abbildung 6 Cluster der PROMs nach dem ersten Titel- und Abstract-Screening

Im anschließenden Volltext-Screening wurden die definierten Ein- und Ausschlusskriterien von zwei unabhängigen Reviewern konsequent angewandt.

2.3.3 Datenextraktion

Acht Studien aus dem Volltext-Screening erfüllten die Einschlusskriterien und wurden für die Datenextraktion herangezogen. Bei diesen Studien wurden die Referenzlisten nach weiteren relevanten Artikeln durchsucht. Hierbei wurde ein weiterer Artikel gefunden, der als relevant erachtet wurde. Dieser beschrieb die Fokusgruppen, die zur Entwicklung des *Day-to-Day Impact of Vaginal Aging (DIVA) Questionnaire* durchgeführt wurden. Somit wurde er für die Beurteilung der Inhaltsvalidität des DIVA miteinbezogen (65). Insgesamt wurden aus neun Artikeln Daten zu vier PROMs von zwei Reviewern extrahiert (vgl. Tabelle 4).

Tabelle 4 Studien, die zur Datenextraktion herangezogen wurden

Studien	PROMs
Erekson <i>et al.</i> , 2013 (66)	
Erekson <i>et al.</i> , 2016 (67)	<i>Vulvovaginal Symptoms Questionnaire (VSQ)</i>
Fernandez-Alonso <i>et al.</i> , 2017 (68)	
Huang <i>et al.</i> , 2015 (69)	
Huang <i>et al.</i> , 2010 (65)*	<i>Day-to-Day Impact of Vaginal Aging (DIVA) Questionnaire</i>
Hunter <i>et al.</i> , 2016 (39)	
Nappi <i>et al.</i> , 2019 (38)	
McKenna <i>et al.</i> , 1999 (70)	<i>Urogenital Atrophy Quality of Life (UGAQoL) Questionnaire</i>
Chen <i>et al.</i> , 2010 (71)	<i>Urogenital symptom scale</i>

Anmerkungen. *nur für die Beurteilung der Inhaltsvalidität, Abkürzungen. PROMs = *patient-reported outcome measures*.

2.3.3.1 Beurteilung der Inhaltsvalidität

Die Entwicklungsstudie des DIVA wurde als inadäquat bewertet, da sie nicht in einer Stichprobe durchgeführt wurde, die exakt die Zielpopulation repräsentierte. Der DIVA wurde mithilfe von Frauen, die moderate bis schwere vulvovaginale Symptome hatten, entwickelt. Zielpopulation des DIVA sind jedoch Frauen mit vulvovaginalen Symptomen jeglichen Schweregrades. Die inadäquate Bewertung des *Urogenital Atrophy Quality of Life (UGAQoL) Questionnaire* ist auf die unzureichende Beschreibung des zugrunde liegenden Konstruktes zurückzuführen. Die Bedürfnisse des *needs-based model* wurden nicht weiter erläutert. Sowohl beim *Vulvovaginal Symptoms Questionnaire (VSQ)* als auch bei der *Urogenital symptom scale* wurde die Zielpopulation nicht in die Item-Generierung miteinbezogen, was zur inadäquaten Bewertung der Entwicklungsstudie führte. Die Inhaltsvalidität jeder PROM wurde mit *doubtful* bewertet, da zu wenige Informationen zum Ablauf der Studien in den Artikeln enthalten waren (vgl. Tabelle 5).

Tabelle 5 COSMIN *Risk of Bias* Bewertung der Inhaltsvalidität

	VSQ	DIVA	UGAQoL	<i>Urogenital symptom scale</i>
Box 1. PROM development	<i>inadequate</i>	<i>inadequate</i>	<i>inadequate</i>	<i>inadequate</i>
Box 2. Content validity	<i>doubtful</i>	<i>doubtful</i>	<i>doubtful</i>	<i>doubtful</i>

Abkürzungen. VSQ = *Vulvovaginal Symptoms Questionnaire*, DIVA = *Day-to-Day Impact of Vaginal Aging Questionnaire*, UGAQoL = *Urogenital Atrophy Quality of Life Questionnaire*.

Da beim VSQ und DIVA jeweils mindestens eine Studie von zweifelhafter Qualität zur Inhaltsvalidität vorlag, war die Qualität der Evidenz als moderat einzustufen. Eine Version des UGAQoL wurde von den Autoren leider nicht zur Verfügung gestellt, weshalb es für die Reviewer nicht möglich war, Relevanz, Vollständigkeit und Verständlichkeit des Fragebogens zu beurteilen. Nur die erste Seite des Fragebogens war frei zugänglich und ließ zumindest eine Beurteilung der Antwortkategorien zu. Aus Mangel an Informationen blieb den Reviewern nichts anderes übrig, als die Inhaltsvalidität als unbestimmt (?) zu bewerten. Die *Urogenital symptom scale* bestand nur aus drei Items und erhielt aus diesem Grund eine ungenügende Bewertung der Vollständigkeit und demzufolge eine inkonsistente Gesamtbewertung. Bei unbestimmten (?) oder inkonsistenten (±) Gesamtbewertungen wird die Qualität der Evidenz nicht bestimmt (vgl. Tabelle 6).

Tabelle 6 Bewertung der Inhaltsvalidität und Qualität der Evidenz der eingeschlossenen PROMs

		<i>relevance</i>	<i>comprehensiveness</i>	<i>comprehensibility</i>	<i>content validity rating</i>
VSQ	<i>overall rating</i>	+	+	+	<i>sufficient (+)</i>
	<i>quality of evidence</i>	<i>moderate (due to risk of bias)</i>			
DIVA	<i>overall rating</i>	+	+	+	<i>sufficient (+)</i>
	<i>quality of evidence</i>	<i>moderate (due to risk of bias)</i>			
UGAQoL	<i>overall rating</i>	?	?	+	<i>indeterminate (?)</i>
	<i>quality of evidence</i>	<i>no grading if overall rating is indeterminate</i>			
<i>Urogenital symptom scale</i>	<i>overall rating</i>	+	-	+	<i>inconsistent (±)</i>
	<i>quality of evidence</i>	<i>no grading if overall rating is inconsistent</i>			

Abkürzungen. VSQ = Vulvovaginal Symptoms Questionnaire, DIVA = Day-to-Day Impact of Vaginal Aging Questionnaire, UGAQoL = Urogenital Atrophy Quality of Life Questionnaire.

Bei keiner PROM gab es Evidenz von hoher Qualität für ungenügende Inhaltsvalidität. Somit wurden auch die restlichen Messeigenschaften aller PROMs beurteilt.

2.3.3.2 Beurteilung der übrigen Messeigenschaften

Insgesamt konnte die methodische Qualität von 24 Messeigenschaften bewertet werden. Davon waren 14 von sehr guter Qualität (58 %), jeweils drei (13 %) von adäquater oder zweifelhafter Qualität und vier (16 %) von inadäquater Qualität (vgl. Tabelle 7).

Tabelle 7 COSMIN *Risk of Bias* Bewertung der übrigen Messeigenschaften

	VSQ	DIVA	UGAQoL	<i>Urogenital symptom scale</i>
<i>Structural validity</i>	<i>very good^a</i>	<i>very good^d</i>		
	<i>inadequate^a</i>	<i>very good^d</i>		
	<i>very good^b</i>			
	<i>very good^b</i>			
<i>Internal consistency</i>	<i>very good^a</i>	<i>very good^d</i>	<i>doubtful^g</i>	<i>very good^h</i>
	<i>very good^c</i>			
<i>Cross-cultural validity/ measurement error</i>				
<i>Reliability</i>	<i>adequate^a</i>	<i>adequate^d</i>	<i>doubtful^g</i>	<i>very good^h</i>
<i>Measurement error</i>				
<i>Criterion validity</i>				
<i>Hypotheses testing for construct validity</i>	<i>very good^a</i>	<i>inadequate^d</i>	<i>doubtful^g</i>	
	<i>very good^c</i>	<i>adequate^d</i>		
		<i>very good^e</i>		
		<i>inadequate^f</i>		
		<i>very good^f</i>		
<i>Responsiveness</i>				<i>inadequate^h</i>

Anmerkungen. ^aErekson *et al.*, 2013; ^bErekson *et al.*, 2016; ^cFernandez-Alonso *et al.*, 2017; ^dHuang *et al.*, 2015; ^eHunter *et al.*, 2016; ^fNappi *et al.*, 2019; ^gMcKenna *et al.*, 1999; ^hChen *et al.*, 2010; Abkürzungen. VSQ = Vulvovaginal Symptoms Questionnaire, DIVA = Day-to-Day Impact of Vaginal Aging Questionnaire, UGAQoL = Urogenital Atrophy Quality of Life Questionnaire.

2.3.3.3 Charakteristiken der eingeschlossenen PROMs und Studienpopulationen

Verschiedene Charakteristiken der einzelnen PROMs, wie das zugrunde liegende Konstrukt, die Zielpopulation, die Anwendung, das Erinnerungsintervall, die Skalen und Anzahl der Items, das Antwortformat, der Score-Bereich und die verfügbaren Sprachen, sind in Tabelle 8 dargestellt. Tabelle 9 hingegen beschreibt relevante Charakteristiken der eingeschlossenen Studien, wie die Stichprobengröße, das Durchschnittsalter, das Setting, das Land, die Sprache und die Rücklaufquote.

Tabelle 8 Charakteristiken der eingeschlossenen PROMs

Charakteristiken	PROMs			
	VSQ	DIVA	UGAQoL	<i>Urogenital symptom scale</i>
Konstrukt	Symptome, Emotionen, Auswirkungen von vulvovaginalen Symptomen auf das (Sexual-)Leben	Aktivitäten des täglichen Lebens, emotionales Wohlbefinden, sexuelle Funktionsfähigkeit, Selbstkonzept und Körperbild	Lebensqualität (= Ausmaß, in dem Individuen in der Lage sind, ihre Bedürfnisse zu befriedigen)	Urogenitale Atrophie
Zielpopulation	Postmenopausale Frauen	Symptomatische postmenopausale Frauen	Frauen mit urogenitaler Atrophie	Frauen zwischen 40 und 60 Jahren
Anwendung	selbst berichtet	selbst berichtet	selbst berichtet	selbst berichtet
Recall-Periode	eine Woche	vier Wochen	in diesem Moment ^b	in diesem Moment (s. GCS ^c)
Skalen, Anzahl der Items	3 oder 4 Skalen, 17 Items (21 Items) ^a	4 Skalen, 19 Items (23 Items) ^a	20 Items	3 Items
Antwortformat	dichotom: ja/nein	5-Punkt Likert Skala (0 – 4)	dichotom: ja/nein (2007 ^b : <i>Yes, a lot; Yes, a little bit; No, not at all</i>)	keine Information vorhanden (GCS: 4-Punkt Likert Skala (0 – 3))
Score-Bereich	0 – 20	0 – 4 (Mittelwerte pro Skala werden gebildet)	0 – 20	keine Information vorhanden
Ausgangssprache	Englisch	Englisch	Englisch, Schwedisch	Chinesisch
Verfügbare Übersetzungen	Spanisch	Spanisch, Italienisch	-	-

Anmerkungen. ^afür sexuell aktive Frauen; ^bsiehe: <http://www.galen-research.com/measures-database/>.
 Abkürzungen. PROMs = *patient-reported outcome measures*, VSQ = *Vulvovaginal Symptoms Questionnaire*, DIVA = *Day-to-Day Impact of Vaginal Aging Questionnaire*, UGAQoL = *Urogenital Atrophy Quality of Life Questionnaire*, GCS = *Greene Climacteric Scale*.

Tabelle 9 Charakteristiken der eingeschlossenen Studien

PROM	Ref.	Population		Anwendung			Rücklaufquote
		n	Altersdurchschnitt (SD)	Setting	Land	Sprache	
VSQ	Erekson <i>et al.</i> , 2013	120	66,3 (10,9)	Gynäkologische Praxen	USA	Englisch	99 %
	Erekson <i>et al.</i> , 2016	358	62,0 (9,9); mit Symptomen an der Vulva: 60,9 (9,2); ohne Symptome an der Vulva: 63,1 (10,4)	Praxen für innere Medizin und Seniorenzentren	USA	Englisch	Seniorenzentren: 96 % Primärversorgungs- praxen: 85,7 %
	Fernandez-Alonso <i>et al.</i> , 2017	150	59,5 (4,9); mit vulvovaginalen Symptomen: 59,9 (5,1); ohne vulvovaginale Symptome: 58,2 (4,0)	Ambulanzen	Spanien	Spanisch	81,3 %
DIVA	Huang <i>et al.</i> , 2015	757	56,2 (8,5)	Parent RRISK3 Kohorte (Langzeit- Teilnehmer des <i>Kaiser</i> <i>Permanente Northern California</i> Gesundheitsversorgungssystems)	USA	Englisch	98,4 %
	Hunter <i>et al.</i> , 2016			vgl. Huang <i>et al.</i> , 2015			
	Nappi <i>et al.</i> , 2019	2160	58,9 (6,7)	46 ambulante Menopause- und Gynäkologie-Zentren	Spanien und Italien	Spanisch und Italienisch	90 %
UGAQoL	McKenna <i>et al.</i> , 1999	50	55,9	Postalische Umfrage	UK	Englisch	100 %
		54	67,7	Postalische Umfrage	Schweden	Schwedisch	100 %
Urogenital symptom scale	Chen <i>et al.</i> , 2010	611	Seminar: 49,7 (5,0) Postalisch: 48,2 (5,6)	Gesundheitsseminar und postalische Umfrage	China	Chinesisch	Seminar: 100 % Postalisch: 78,5 %

Anmerkungen. n = Stichprobengröße. Abkürzungen. PROM = *patient-reported outcome measure*, VSQ = *Vulvovaginal Symptoms Questionnaire*, DIVA = *Day-to-Day Impact of Vaginal Aging Questionnaire*, UGAQoL = *Urogenital Atrophy Quality of Life Questionnaire*, Ref. = Referenz, SD = Standardabweichung.

2.3.3.4 Informationen zu Interpretierbarkeit und Machbarkeit

Zur Interpretierbarkeit wurden bei einer Studie zum DIVA weniger als 1 % der Item-Antworten als fehlend berichtet (69). In einer Studie zum VSQ zeigte sich Evidenz für eine Normalverteilung der Daten (68). Keine weiteren Informationen zu Aspekten der Interpretierbarkeit, wie zu Boden- und Deckeneffekten, dem kleinsten klinisch bedeutsamen Unterschied (*minimal important change* (MIC), *minimal important difference* (MID)) oder *response shift*, konnten aus den eingeschlossenen Studien extrahiert werden.

Zu Aspekten der Machbarkeit kann berichtet werden, dass alle vier PROMs im Selbstbericht ausgefüllt werden und weder ein hohes kognitives Niveau noch bestimmte körperliche Fähigkeiten vorausgesetzt werden müssen. Die Gesamtscores von VSQ und UGAQoL werden durch einfaches Aufsummieren der einzelnen Items erreicht. Beim DIVA wird der Gesamtscore für jede Skala durch den Mittelwert abgebildet (Summe der Items/Anzahl der Items). Für die *Urogenital symptom scale* ist keine Information zur Berechnung des Gesamtscores verfügbar. Der DIVA unterliegt einem Copyright, jedoch werden weder Kosten noch eine schriftliche Einverständniserklärung von den Autoren für den Einsatz des DIVA verlangt. Für das Ausfüllen des UGAQoL wurde ein Zeitintervall von 4 – 15 Minuten angegeben, was durchaus in einem machbaren Rahmen liegt. Diese PROM ist jedoch nicht frei zugänglich, da die Autoren mit einer Verbreitung des UGAQoL nicht einverstanden waren.

2.3.3.5 Zusammenfassung der Ergebnisse

Die zusammengefassten Ergebnisse pro Messeigenschaft pro PROM sind in Tabelle 10 dargestellt. Beim VSQ zeigte sich ein inkonsistentes Gesamtergebnis bei der internen Konsistenz, da nicht alle Studien ein Cronbach's alpha $\geq 0,7$ berichteten. Das Gesamtergebnis wurde demzufolge aus der Mehrheit der Ergebnisse gebildet (fünf Studien (+) vs. drei Studien (-)) und wurde als hinreichend angenommen. Wegen der inkonsistenten Ergebnisse wurde die Qualität der Evidenz um ein Level herabgestuft. Beim DIVA wurde die Qualität der Evidenz von drei Messeigenschaften, die in der gleichen Studie untersucht wurden, jeweils um ein Level herabgestuft, da die Studie (69) nicht exakt in der Zielpopulation durchgeführt wurde (Grund: *indirectness*).

Tabelle 10 *Summary of Findings*

Strukturvalidität	Zusammengefasstes Ergebnis	Overall-Rating	Qualität der Evidenz
VSQ	CFI: 0,97 – 0,99	hinreichend	hoch
DIVA	CFI: 0,978 – 0,979	hinreichend	moderat (wegen <i>indirectness</i>)
Interne Konsistenz	Zusammengefasstes Ergebnis	Overall-Rating	Qualität der Evidenz
VSQ	0,623 – 0,87; n = 31 – 244	inkonsistent → <i>Overall-Rating</i> basiert auf Mehrheit (5+, 3-) → hinreichend	moderat (wegen <i>inconsistency</i>)
DIVA	0,82 – 0,94, konsistente Ergebnisse; n = 462 – 745	hinreichend	moderat (wegen <i>indirectness</i>)
UGAQoL	0,89 – 0,90, keine Evidenz für hinreichende Strukturvalidität; n = 104	unbestimmt	-
<i>Urogenital symptom scale</i>	0,43; n = 290	ungenügend	hoch
Reliabilität	Zusammengefasstes Ergebnis	Overall-Rating	Qualität der Evidenz
VSQ	0,55 – 0,75; n = 91	ungenügend	niedrig (wegen <i>risk of bias</i> und <i>imprecision</i>)
DIVA	0,47 – 0,72; n = 462 – 745	ungenügend	niedrig (wegen <i>risk of bias</i> und <i>indirectness</i>)
UGAQoL	0,85 – 0,92, ICC or Kappa not reported; n = 104	unbestimmt	-
<i>Urogenital symptom scale</i>	0,81; n = 52	hinreichend	moderat (wegen <i>imprecision</i>)
Hypothesen-Testen	Zusammengefasstes Ergebnis	Overall-Rating	Qualität der Evidenz
VSQ	3 von 3 Hypothesen bestätigt	hinreichend	hoch
DIVA	5 von 6 Hypothesen bestätigt, Erwartungen ggü. <i>Known-groups validity</i> bestätigt	hinreichend	hoch
UGAQoL	2 von 2 Hypothesen nicht bestätigt	ungenügend	niedrig (wegen <i>risk of bias</i>)
Änderungs-sensitivität	Zusammengefasstes Ergebnis	Overall-Rating	Qualität der Evidenz
<i>Urogenital symptom scale</i>	Effektstärke: 0,46; n = 19	unbestimmt	-

Anmerkungen. n = Stichprobengröße. Abkürzungen. ICC = *intraclass correlation coefficient*, VSQ = *Vulvovaginal Symptoms Questionnaire*, DIVA = *Day-to-Day Impact of Vaginal Aging Questionnaire*, UGAQoL = *Urogenital Atrophy Quality of Life Questionnaire*, CFI = *Comparative Fit Index*.

2.3.4 Abgabe einer Empfehlung

Anhand der *Summary of Findings* Tabellen wurden Empfehlungen für künftige GSM-Studien gemäß der COSMIN-Richtlinien (10) abgeleitet (vgl. Tabelle 11). Zwei PROMs, VSQ und DIVA, erfüllten beide Kriterien für Kategorie A, hinreichende Inhaltsvalidität und mindestens Evidenz niedriger Qualität für hinreichende interne Konsistenz, und konnten somit für künftige GSM-Studien empfohlen werden. Da die *Urogenital symptom scale* Evidenz hoher Qualität für eine ungenügende Messeigenschaft aufwies, wurde sie in Kategorie C platziert und konnte somit nicht empfohlen werden. Der UGAQoL erfüllte weder die Kriterien für Kategorie A noch für Kategorie C und wurde somit in Kategorie B eingeordnet. Weitere Validierungsstudien sind vonnöten.

Tabelle 11 Empfehlungen für künftige GSM-Studien

PROM	Kategorie A		Kategorie C	Empfehlung
	Hinreichende Inhaltsvalidität (jedes Level)	Mindestens Evidenz niedriger Qualität für hinreichende interne Konsistenz	Evidenz hoher Qualität für mindestens eine ungenügende Messeigenschaft	
VSQ	✓	✓	✗	A
DIVA	✓	✓	✗	A
UGAQoL	✗	✗	✗	B
<i>Urogenital symptom scale</i>	✗	✗	✓	C

Abkürzungen. PROM = *patient-reported outcome measure*, VSQ = *Vulvovaginal Symptoms Questionnaire*, DIVA = *Day-to-Day Impact of Vaginal Aging Questionnaire*, UGAQoL = *Urogenital Atrophy Quality of Life Questionnaire*.

2.4 Diskussion

In dieser systematischen Übersichtsarbeit wurden die Messeigenschaften von vier verschiedenen PROMs für postmenopausale Frauen mit urogenitalen Beschwerden untersucht. Zwei dieser PROMs, VSQ und DIVA, konnten für den künftigen Einsatz in klinischen Studien empfohlen werden und den Ergebnissen, die mit diesen PROMs erzielt werden, kann vertraut werden. Der DIVA sollte dem VSQ vorgezogen werden, da bei dessen Item-Generierung betroffene Frauen miteinbezogen wurden. Mithilfe von Fokusgruppen, einer anerkannten qualitativen Methode, wurde die Perspektive der Patientinnen abgebildet, um daraus entsprechende Items zu entwickeln. Der VSQ hingegen wurde mithilfe von ExpertInnen aus dem Skindex (72), einem Hautfragebogen, entwickelt. Ein zweiter Aspekt, der für die Präferenz des DIVA spricht, ist die deutlich höhere Stichprobengröße der eingeschlossenen Studien, die in die Analysen miteinging und die Glaubwürdigkeit der Ergebnisse stärkt. Nichtsdestotrotz sind weitere Validierungen beider PROMs wünschenswert. Vor allem die Test-Retest-

Reliabilität sollte weiter überprüft werden, da beide PROMs ein ungenügendes Gesamtergebnis erreichten, allerdings auf einem niedrigen Evidenzlevel. Auch der UGAQoL hat noch immer die Möglichkeit, weiterempfohlen zu werden, allerdings sprechen auch Gründe der Machbarkeit gegen die Aussprache einer Empfehlung, da diese PROM von den Autoren nicht herausgegeben wird und somit nicht frei zugänglich ist. Die *Urogenital symptom scale* sollte aus verschiedenen Gründen nicht weiterverwendet werden. Zum einen wurde Qualität hoher Evidenz für eine ungenügende Messeigenschaft, nämlich die interne Konsistenz, gefunden. Zum anderen wurde keine qualitative Methode für die Entwicklung der PROM verwendet. Eine Gruppe von ExpertInnen entwickelte die Skala und betroffene Frauen waren daran nicht beteiligt. Die *Urogenital symptom scale* entstand als Zusatzskala zur *Greene Climacteric Scale* (GCS), einer Symptomskala zur Menopause, jedoch verschlechterte diese Zusatzskala die Modellpassung der Standard-GCS. Sogar die Autoren unterstützten den Einschluss dieser Skala zur Standard-GCS nicht (71). Künftige Validierungsstudien sollten auch Aspekte der Interpretierbarkeit erheben, da für die eingeschlossenen PROMs hierzu bisher nur wenige Informationen verfügbar waren.

Alle vier PROMs wurden vor der Begriffsdefinition des GSM entwickelt, jedoch wurde die urologische Komponente bereits von zwei PROMs, dem UGAQoL und der *Urogenital symptom scale*, berücksichtigt, da sich diese auf den Begriff der urogenitalen Atrophie bezogen. Die präferierten PROMs, VSQ und DIVA, umfassten beide nur Aspekte der VVA, d. h. die genitale und sexuelle Komponente des GSM, und deckten daher nicht das gesamte GSM-Konstrukt ab.

Als Stärken dieser systematischen Übersichtsarbeit sind ein registriertes Protokoll, die Anwendung eines umfassenden und sensiblen Suchfilters, die Suche in den großen Datenbanken *MEDLINE*, *EMBASE* und *Web of Science*, mehreren kleineren Datenbanken und Referenzlisten eingeschlossener Studien, vordefinierte Ein- und Ausschlusskriterien und der Einsatz der *COSMIN Risk of Bias Checklist* für die Beurteilung der methodischen Qualität der eingeschlossenen Studien zu nennen. Die Tatsache, dass keine weiteren Treffer in den kleineren Datenbanken und nur eine relevante Studie in den Referenzlisten gefunden wurden, spricht für die Qualität der Suchstrategie, die mit einem universitären Bibliothekar entwickelt wurde. In jedem Schritt des Beurteilungsprozesses waren mindestens zwei unabhängige Reviewer beteiligt, wobei eine Wissenschaftlerin jeden Schritt ausführte, um eine gewisse Konsistenz zu wahren. Diskrepanzen wurden im Forschungsteam diskutiert und gelöst. Vor Abschluss der systematischen Übersichtsarbeit fand noch einmal ein Update der Literatursuche statt, um

potentielle Treffer, die während der Datenextraktion publiziert wurden, nicht zu verpassen. Als mögliche Limitation dieser Studie ist aufzuführen, dass nicht alle Referenzlisten der relevanten Volltexte nach weiteren Treffern durchsucht wurden. Als kritisch sollte auch die Tatsache betrachtet werden, dass die COSMIN-Kriterien klar definierte Schwellenwerte beinhalten und keine Konfidenzintervalle. Das würde z. B. bedeuten, dass eine Stichprobengröße von 99 schlechter bewertet würde als eine Stichprobengröße von 100, da ab einer Stichprobengröße von 100 die Qualität der Evidenz nicht herabgestuft werden müsste. Nichtsdestotrotz lassen diese klaren Schwellenwerte wenig Interpretationsspielraum, was wiederum zu einheitlicheren Bewertungen über unterschiedliche Reviewer hinweg führt.

Systematische Übersichtsarbeiten können helfen, die am besten geeignete PROM für einen *Outcome* zu identifizieren, der bereits Teil eines *core outcome sets* (COS) ist (52). Unter einem COS versteht man ein von ExpertInnen vereinbartes Set an *Outcomes*, das als Minimum in allen klinischen Studien gemessen und berichtet werden sollte (54, 73, 74). Ein weiterer Vorteil von systematischen Übersichtsarbeiten ist die Identifikation von Wissenslücken, da sehr übersichtlich dargestellt wird, welche Messeigenschaften von PROMs bereits untersucht wurden und welche noch ausstehen. Somit können anhand dieser umfassenden Übersicht neue Validierungsstudien geplant werden, die genau diese Wissenslücken füllen (52). Beispielsweise wurde bis dato weder beim VSQ noch beim DIVA die Änderungssensitivität (*responsiveness*) überprüft.

Zusammenfassend lässt sich sagen, dass diese qualitativ hochwertige systematische Übersichtsarbeit die evidenz-basierte Erkenntnis lieferte, dass derzeit zwei PROMs, VSQ und DIVA, für den Einsatz in klinischen Studien geeignet sind und somit empfohlen werden können. Beide PROMs decken die genitale und die sexuelle Komponente des GSM ab, jedoch nicht die urologische Komponente. Künftige Validierungsstudien sollten den Fokus auf diese beiden PROMs legen.

3. Teil 2a: Validierung des *Day-to-Day Impact of Vaginal Aging Questionnaire* in einer deutschsprachigen Population

3.1 Einleitung

Der DIVA, zu Deutsch „Fragebogen zu den alltäglichen Auswirkungen der vaginalen Atrophie (des vaginalen Alterns)“, setzt sich aus vier Skalen zusammen (vgl. Abbildung 7, Anhang D). Skala 1 bezieht sich auf die Auswirkungen der Beschwerden auf die Aktivitäten des täglichen Lebens, Skala 2 auf das emotionale Wohlbefinden, Skala 3 überprüft den Einfluss der Beschwerden auf die sexuelle Funktionsfähigkeit, wobei es hier eine kurze Version für alle Frauen (einschließlich nicht sexuell aktiver Frauen) und eine längere Version für sexuell aktive Frauen gibt, und Skala 4 bildet die Auswirkungen auf das Selbstkonzept und Körperbild ab. Der Fragebogen besteht aus 23 Items, wobei 4 dieser Items nur von kürzlich sexuell aktiven Frauen beantwortet werden können. Jedes Item kann auf einer Skala von 0 bis 4 bewertet werden. Der Gesamtscore einer Skala wird über den Mittelwert der zugehörigen Items bestimmt (Summe der Items einer Skala/Anzahl der Items der Skala). Somit liegt der Gesamtscore für jede Skala zwischen 0 und 4. Je höher der Score, desto größer wird die Auswirkung vaginaler Symptome wahrgenommen. Der Erinnerungszeitraum umfasst die letzten vier Wochen (69). Neben Englisch ist der Fragebogen derzeit noch in Spanisch und Italienisch verfügbar (38, 75, 76).

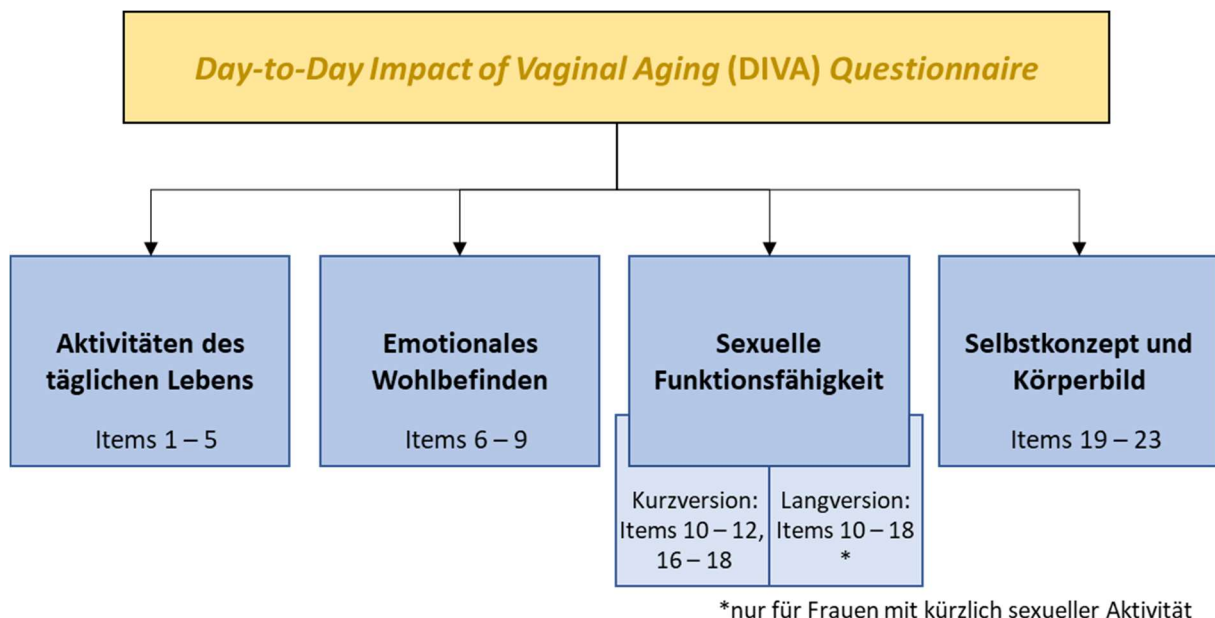


Abbildung 7 Zusammensetzung des DIVA

Mit dieser Studie wird das Ziel verfolgt, die hervorragenden Messeigenschaften des DIVA in einer deutschsprachigen Population zu replizieren sowie die Qualität der deutschen Version des DIVA zu überprüfen. Bei Bestehen dieser Prüfung kann der Einsatz der deutschen Version des DIVA in künftigen klinischen Studien empfohlen werden. Die deutsche Version des DIVA (vgl. Anhang D) wurde von einem Übersetzungsbüro (*TransPerfect*) erstellt und linguistisch mithilfe von Interviews mit der Zielpopulation, d. h. postmenopausalen Frauen mit vaginalen Beschwerden, validiert.

3.2 Methode

Die Daten wurden aus zwei verschiedenen Quellen bezogen (vgl. Abbildung 8). Zum einen fand eine quantitative Online-Befragung mithilfe eines eigens generierten Links statt. Zum anderen wurde die Papierversion der gleichen Befragung an die Teilnehmerinnen eines Studien-Screenings der Firma Dr. August Wolff GmbH & Co. KG Arzneimittel verteilt.

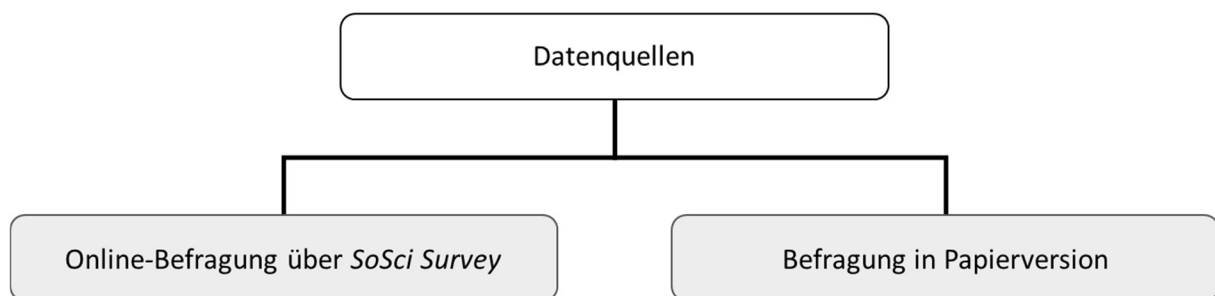


Abbildung 8 Verwendete Datenquellen für die Validierung des DIVA

Für die Online-Datenerhebung wurde in diversen Facebook-Gruppen für Frauen in den Wechseljahren rekrutiert, indem der entsprechende Link gepostet wurde. Außerdem wurden 5000 kleine Flyer mit dem Link zur Befragung an ca. 100 gynäkologische Praxen und Kliniken in Niederbayern, Oberbayern, Franken und der Oberpfalz verschickt sowie an der Universitätsklinik für Frauenheilkunde in Magdeburg und im Inselspital in Bern verteilt (vgl. Anhang E, (a)). Bevor der Link freigeschaltet wurde, wurde seine Funktionsfähigkeit von weiblichen Bekannten und Kolleginnen ($n = 7$) getestet und die sprachliche Richtigkeit der Befragung bestätigt. Ein positives Ethikvotum wurde durch die Ethikkommission der Universität Regensburg erteilt (Aktenzeichen: 19-1554-101).

Die Online-Datenerhebung fand vom 05.11.2019 bis zum 25.04.2020 statt. Über www.soscisurvey.de wurde ein Fragebogen erstellt, der folgendermaßen aufgebaut war (vgl. Abbildung 9): Zuerst wurde den TeilnehmerInnen eine Aufklärungsseite gezeigt. Eine

Einwilligung durch das Setzen eines Hakens war erforderlich, um zur nächsten Seite zu gelangen. Zu Beginn der Befragung wurde ein individueller Code vergeben, der sich aus dem ersten Buchstaben des Vornamens und dem sechsstelligen Geburtsdatum zusammensetzte, um eine Zuordnung der Datensätze bei wiederholtem Ausfüllen gewährleisten zu können. Soziodemographische Fragen zum Geschlecht, Alter, Heimatland und der Muttersprache folgten. Danach sollten der Menstruationszyklus und die Einnahme von Hormonen dokumentiert werden. Nicht-weibliche Teilnehmer oder diejenigen, die angaben, noch regelmäßige oder gelegentliche Blutungen zu haben, wurden sofort zur Schlussseite der Befragung geleitet. Nur Teilnehmerinnen, die Zykluspausen von über zwei Monaten angaben, seit über einem Jahr keine Periode mehr hatten oder angaben, keine Gebärmutter mehr zu haben, wurden zu den nächsten Fragen weitergeleitet. Anschließend sollte die *Menopause Rating Scale II* (MRS) ausgefüllt werden. Eine Filterfrage zu vaginalen Beschwerden wurde eingebaut, um nur Frauen, die auch an vaginalen Beschwerden litten, zu fünf Ankerfragen, dem DIVA, dem *Patient Health Questionnaire-4* (PHQ-4) (77) und der SOEP (= Sozio-oekonomisches Panel)-Version des *short-form health survey* (SF-12) (78) weiterzuleiten. Am Ende der Befragung wurden die Teilnehmerinnen gebeten, ihre E-Mail-Adresse zu hinterlassen, um nach einer bzw. vier Wochen erneut kontaktiert werden zu können. Das Ausfüllen der Befragung sollte 10 – 15 Minuten Zeit in Anspruch nehmen.

Zusätzlich zur Online-Datenerhebung wurde eine Papierversion der Befragung an 41 Frauen ausgegeben, die zum Zusatzscreening einer klinischen Studie der Firma Dr. August Wolff GmbH & Co. KG Arzneimittel kamen. Diese Frauen wurden vor der Ladung zum Zusatzscreening telefonisch kontaktiert und befragt und nur diejenigen, die vaginale Beschwerden wie Trockenheit, Brennen, Juckreiz oder vaginalen Schmerz angaben, und seit mindestens zwölf Monaten keine Periode mehr hatten (bzw. seit mindestens drei Monaten keine Gebärmutter mehr), wurden zum Zusatzscreening eingeladen und durften die Papierversion der Befragung ausfüllen. Die Daten wurden anschließend per Hand in eine Excel-Tabelle eingegeben und gingen in die Analyse mit ein.

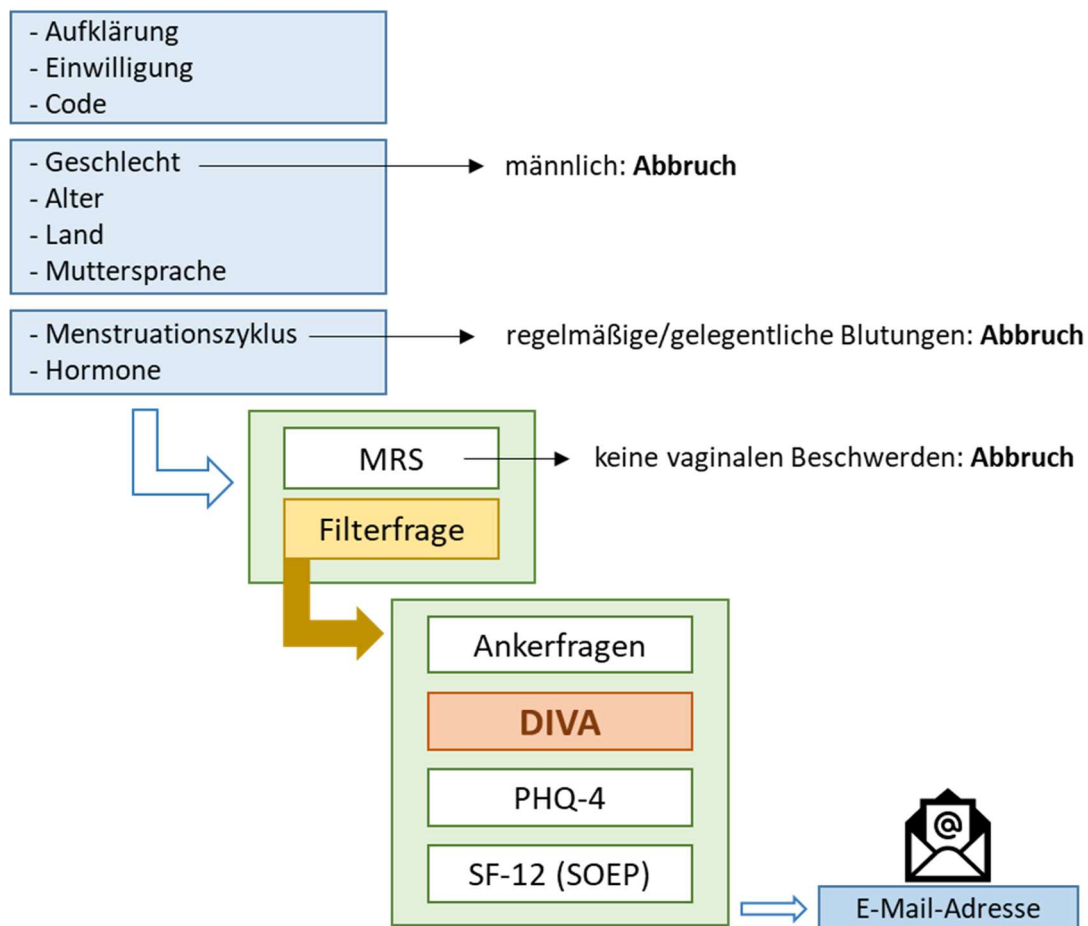


Abbildung 9 Aufbau der Online-Befragung

Abkürzungen. MRS = Menopause Rating Scale, DIVA = Day-to-Day Impact of Vaginal Aging Questionnaire, PHQ = Patient Health Questionnaire, SF-12 = short-form health survey, SOEP = Sozio-oekonomisches Panel.

Die MRS ist ein Fragebogen, der in den 90er Jahren entwickelt wurde, um die gesundheitsbezogene Lebensqualität oder den Schweregrad der Beschwerden alternder Frauen zu erfassen. Die Skala ist in über 20 Sprachen verfügbar und beinhaltet elf typische Beschwerden der Wechseljahre, unter anderem psychologische, somato-vegetative und urogenitale Symptome. Die Schwere jedes Items wird auf einer 5-Punkt-Likert-Skala eingeschätzt, somit kann der Gesamtscore maximal 44 Punkte erreichen (79-82).

Fünf Ankerfragen wurden in Anlehnung an Ofenloch *et al.* (83) kreiert und in die Befragung integriert. Die erste Ankerfrage wurde allgemein gehalten (Frage 1), alle anderen waren jeweils spezifisch für eine Skala des DIVA (Fragen 2 – 5):

1. Wie schwer stufen Sie die durch Ihre vaginalen Beschwerden hervorgerufenen Symptome ein? (hierzu zählen beispielsweise: vaginale Trockenheit, Brennen, Juckreiz, Schmerzen im Intimbereich, ...)

2. Wie schwer stufen Sie die funktionalen Beeinträchtigungen durch Ihre vaginalen Beschwerden ein? (dass Sie gewohnte Dinge schlechter oder gar nicht tun können)
3. Wie schwer stufen Sie die emotionalen Beeinträchtigungen durch Ihre vaginalen Beschwerden ein? (beispielsweise: ärgern, Niedergeschlagenheit, Zukunftsängste, genervt sein, schämen, ...)
4. Wie schwer stufen Sie die sexuellen Beeinträchtigungen durch Ihre vaginalen Beschwerden ein? (dass Sie Ihre Sexualität nicht mehr so wie früher ausleben können)
5. Wie schwer stufen Sie die Beeinträchtigung durch Ihre vaginalen Beschwerden auf Ihr Selbstwert- und Körpergefühl ein? (dass Sie sich in Ihrem eigenen Körper noch wohl fühlen)

Jede Ankerfrage konnte auf einer Skala von 1 bis 5 (keine, leichte, mäßige, starke oder sehr starke Beeinträchtigung durch Symptome) beantwortet werden.

Der PHQ-4 ist eine Ultrakurzform des *Patient Health Questionnaire* (PHQ) und beinhaltet zwei Fragen des Depressionsmoduls und zwei Fragen zur Erfassung der generalisierten Ängstlichkeit. Jede Frage kann auf einer 4-Punkt-Likert-Skala (von „überhaupt nicht“ bis „fast jeden Tag“) beantwortet werden und ein Summenwert von 0 bis 12 kann berechnet werden. Der PHQ und seine Subskalen dürfen kostenfrei verwendet werden (77). Das Depressionsmodul wurde in die Befragung aufgenommen, da depressive Symptome in der Menopause gehäuft auftreten. In einer longitudinalen Studie war beispielsweise die Diagnose einer Depression bei Frauen in den Wechseljahren 2,5-mal wahrscheinlicher als bei prämenopausalen Frauen (84).

Bei der SOEP-Version des SF-12 handelt es sich um eine Version des SF-12, die in einer der größten repräsentativen Erhebungen Deutschlands, genannt Sozio-oekonomisches Panel, zur Erfassung der gesundheitsbezogenen Lebensqualität eingesetzt wird (85). Die SOEP-Version des SF-12 ist an den originalen SF-12 (86) angelehnt, allerdings weichen Formulierungen, Antwortoptionen, die Reihenfolge der Fragen und das Layout von der Originalversion ab. Durch diese Unterschiede wurde ein eigener Berechnungsalgorithmus für die SOEP-Version des SF-12 erstellt. Der SF-12 besteht aus acht Subskalen und zwei übergeordneten Skalen, der „Körperlichen Summenskala (*physical component score* (PCS))“ und der „Psychischen Summenskala (*mental component score* (MCS))“, wobei je vier Subskalen einer übergeordneten Skala zugeordnet sind (78). Großer Vorteil der SOEP-Version ist, dass diese Version im Vergleich zum originalen SF-12 kostenfrei genutzt werden kann.

Bevor mit den Analysen begonnen wurde, wurde der online erhobene Datensatz aufbereitet und die über die Papierversion erhaltenen Daten eingegeben. Alle „Nicht-Frauen“, alle Teilnehmerinnen, die die Befragung vorzeitig (vor der MRS) abgebrochen haben sowie alle Frauen, die noch eine (fast) regelmäßige Periode angaben, wurden ausgeschlossen. Alle postmenopausalen Frauen gingen in die Auswertung der MRS mit ein. Alle postmenopausalen Frauen mit vaginalen Beschwerden gingen in die Analysen zur Validierung des DIVA ein. Fehlende Werte blieben „fehlend“ und wurden nicht durch Imputationen ersetzt. Explorativ wurde wie bei Nappi *et al.* (38) ein Score für die gesamte Skala des DIVA (mit 23 Items) berechnet sowie ein Gesamtscore für die Kurzversion des DIVA (mit 19 Items), ohne zu wissen, ob der gesamte DIVA überhaupt als eindimensionale Skala gesehen werden kann. Für die Validierung des DIVA wurden folgende Analysen geplant und in einem Analyseplan festgehalten:

3.2.1 Strukturvalidität

Strukturvalidität bezeichnet das Ausmaß, in dem die Scores einer PROM die Dimensionalität des zu messenden Konstruktes adäquat abbilden (10, 54). Zwei konfirmatorische Faktorenanalysen (KFAs) (eine für die Kurzversion, die alle Frauen mit vaginalen Beschwerden enthält, und eine für die Langversion, die nur kürzlich sexuell aktive Frauen enthält) wurden durchgeführt, um die vier-Faktoren-Struktur des DIVA zu bestätigen. Folgende Fit-Maße wurden für einen angemessenen (bzw. noch akzeptablen) Modell-Fit herangezogen: *comparative fit index* (CFI)/*Tucker-Lewis Index* (TLI) $\geq 0,95$ ($\geq 0,90$), *root mean square error of approximation* (RMSEA) $< 0,05$ ($< 0,08$), *standardized root mean square residual* (SRMR) $< 0,05$ ($< 0,08$) und Chi-Quadrat geteilt durch die Anzahl der Freiheitsgrade (χ^2/df) < 2 (< 3) (87-89). Gemäß der COSMIN-Initiative wird die Strukturvalidität einer PROM als hinreichend bewertet, wenn entweder der CFI oder TLI $> 0,95$ sind oder der RMSEA $< 0,06$ ist oder das SRMR $< 0,08$ ist (52). Die KFAs wurden mit Mplus 8.4 (Muthen & Muthen, Los Angeles, CA) gerechnet (vgl. Anhang F).

Alle weiteren Analysen wurden mit *IBM SPSS Statistics 25* gerechnet. Die zugehörige SPSS Syntax wurde mit einem Probedatensatz entwickelt und ist in Anhang G dargestellt.

3.2.2 Interne Konsistenz

Interne Konsistenz beschreibt die Inter-Korrelation der Items, also wie die einzelnen Items einer Skala miteinander zusammenhängen (54). Der Parameter Cronbach's alpha wurde für alle vier Skalen des DIVA sowie für die gesamte DIVA-Skala berechnet. Für die Skala 3 zur sexuellen Funktionsfähigkeit und die gesamte Skala wurde ein Cronbach's alpha sowohl für die

Kurzversion als auch für die Langversion berechnet. Ein Cronbach's alpha $\geq 0,7$ galt als angemessen (52).

3.2.3 Test-Retest-Reliabilität

Test-Retest-Reliabilität spiegelt die Übereinstimmung der Scores bei zwei verschiedenen Messzeitpunkten wider (54). Teilnehmerinnen, die ihre E-Mail-Adresse am Ende der Befragung hinterlassen hatten, wurden nach einer Woche erneut kontaktiert, um die einzelnen Fragen und Fragebögen über einen weiteren Link noch einmal auszufüllen. Nur Teilnehmerinnen, die als stabiles Merkmal keine Veränderung in der Ankerfrage zeigten, gingen in die Analysen mit ein. Die allgemein gehaltene Ankerfrage 1 sollte als Anker für alle Skalen dienen, Ankerfrage 2 für Skala 1, Ankerfrage 3 für Skala 2, Ankerfrage 4 für Skala 3 (kurz und lang) und Ankerfrage 5 für Skala 4. Zudem sollten alle Ankerfragen als Anker für die Gesamtskalen (kurz und lang) verwendet werden. Intraklassenkorrelationskoeffizienten (*intraclass correlation coefficients* (ICCs)) mit 95 % Konfidenzintervallen sollten zur Bestimmung der absoluten Übereinstimmung der DIVA-Skalen zwischen den zwei Messzeitpunkten berechnet werden. ICCs $\geq 0,7$ sollten als akzeptabel angenommen werden (52).

3.2.4 Konstruktvalidität: Hypothesen-Testen

Konstruktvalidität basiert auf der Annahme, dass eine PROM das zu messende Konstrukt zuverlässig misst (54). Dafür wurden *a priori* Hypothesen formuliert und anschließend getestet.

3.2.4.1 Konvergente Validität

Konvergente Validität ist ein Teil von Konstruktvalidität. Hypothesen über Spearman-Rangkorrelationen mit anderen PROMs und einzelnen Items wurden aufgestellt und überprüft (vgl. Tabelle 12). Generell wurden mittlere Korrelationen mit allen Ankerfragen (0,3 – 0,5) vermutet, mit den für jede Skala spezifischen Fragen sogar hohe Korrelationen ($> 0,5$). Bei den beiden Gesamtskalen wurden für alle Ankerfragen hohe Korrelationen erwartet, da diese alle Aspekte der einzelnen Fragen vereinen sollten. Bei den drei Items der MRS wurden mittlere Korrelationen angenommen, da diese nur auf Beschwerden abzielen und nicht auf deren Auswirkungen. Der PHQ-4 im Gesamten sowie seine beiden Module sollten zumindest moderat mit Skala 2, dem emotionalen Wohlbefinden, korrelieren. Bei der SOEP-Version des SF-12 wurden die Hypothesen postuliert, dass die physische Gesundheit zumindest moderat mit Skala 1, den Aktivitäten des täglichen Lebens, korrelieren sollte, und auch, dass die mentale Gesundheit ebenfalls eine mittlere Korrelation mit Skala 2, dem emotionalen Wohlbefinden, und Skala 4, dem Selbstkonzept und Körperbild, haben sollte. Eine geringe Korrelation ($< 0,3$) wurde bei keiner PROM bzw. bei keinem Item *a priori* postuliert.

Tabelle 12 Erwartete Korrelationen des DIVA mit anderen Items und PROMs

	MRS			Ankerfragen					PHQ			SF-12 SOEP*	
	Item 8 ¹	Item 9 ²	Item 10 ³	1	2	3	4	5	Depr.	Ängst.	Gesamt	PCS	MCS
Skala 1				Grün	Grün	Grün	Grün	Grün				Grün	
Skala 2				Grün	Grün	Grün	Grün	Grün	Gelb	Gelb	Gelb		Grün
Skala 3 (kurz)				Grün	Grün	Grün	Grün	Grün					
Skala 3 (lang)				Grün	Grün	Grün	Grün	Grün					
Skala 4				Grün	Grün	Grün	Grün	Grün					Grün
Gesamt (kurz)	Grün	Grün	Grün	Grün	Grün	Grün	Grün	Grün					
Gesamt (lang)	Grün	Grün	Grün	Grün	Grün	Grün	Grün	Grün					

Anmerkungen. Grün: hohe Korrelation erwartet (> 0,5), gelb: mittlere Korrelation erwartet (0,3 – 0,5), grau: keine Hypothese. Skala 1 = Aktivitäten des täglichen Lebens, Skala 2 = Emotionales Wohlbefinden, Skala 3 = Sexuelle Funktionsfähigkeit, Skala 4 = Selbstkonzept und Körperbild. Abkürzungen. PROMs = *patient-reported outcome measures*, MRS = *Menopause Rating Scale*, PHQ = *Patient Health Questionnaire*, Depr. = Depression, Ängst. = Ängstlichkeit, SF-12 = *short-form health survey*, SOEP = Sozio-oekonomisches Panel, PCS = *physical component score*, MCS = *mental component score*. ¹Sexualprobleme, ²Harnwegsbeschwerden, ³Trockenheit der Scheide. *negative Korrelationen erwartet (je höher die Lebensqualität, desto geringer die Auswirkung der Beschwerden).

3.2.4.2 Diskriminante Validität

Diskriminante Validität (im Englischen auch *known-groups validity*) ist ein weiterer Aspekt von Konstruktvalidität und überprüft Unterschiede im DIVA zwischen verschiedenen Gruppen (Altersgruppen, Heimatland, Schwere der Erkrankung). Kruskal-Wallis-Tests wurden gerechnet, um die Hypothese zu testen, ob die mittleren Rangsummen der Gruppen identisch sind. Zwischen den verschiedenen Altersgruppen (≤ 39 , 40 – 49, 50 – 59, 60 – 69, ≥ 70 Jahre) und den Ländern, aus denen die Teilnehmerinnen kamen (Deutschland, Österreich, Schweiz), wurden keine signifikanten Unterschiede erwartet. Bezüglich der Schwere der Erkrankung wurden die Teilnehmerinnen zum einen aufgrund von Frage 10 der MRS zu vaginaler Trockenheit in vier Gruppen (leichte, mittlere, starke, sehr starke Beschwerden) und zum anderen aufgrund von Ankerfrage 1 zur Schwere der Beschwerden in fünf Gruppen (keine, leichte, mäßige, starke, sehr starke Beeinträchtigung durch Symptome) eingeteilt. Signifikante Unterschiede in den DIVA-Skalen zwischen den Gruppen unterschiedlichen Schweregrads wurden erwartet.

3.2.5 Änderungssensitivität

Änderungssensitivität (im Englischen *responsiveness* oder *ability to detect change*) beschreibt die Fähigkeit eines Instruments, Veränderungen im zu interessierenden Konstrukt über die Zeit zu erfassen (54). Teilnehmerinnen, die ihre E-Mail-Adresse am Ende der Befragung hinterlassen hatten, wurden nach vier Wochen erneut kontaktiert, um die einzelnen Fragebögen

sowie die fünf Ankerfragen noch einmal auszufüllen. Um zu überprüfen, ob sich diese fünf Ankerfragen tatsächlich als Anker eignen, wurde die Veränderung in den Ankerfragen mit den Veränderungen in den einzelnen Skalen des DIVA korreliert. Eine Korrelation $> 0,3$ wurde als angemessen angenommen (90). Bei einer angemessenen Korrelation und einer ausreichend großen Stichprobe wären die Teilnehmerinnen, die eine Verbesserung in der allgemeinen Ankerfrage bzw. in der zur jeweiligen Skala zugehörigen Ankerfrage angaben, in die Analyse eingegangen. Die DIVA-Werte zu Beginn wären mit denen nach ca. vier Wochen mithilfe von non-parametrischen Wilcoxon-Tests verglichen worden. Zum Schluss wären noch die Effektstärken berechnet worden, um eine Aussage über die Größe der Effekte machen zu können ($r = |z/\sqrt{N}|$; mit $r < 0,3$: kleiner Effekt, $0,3 < r < 0,5$: mittlerer Effekt, $> 0,5$ großer Effekt).

3.3 Ergebnisse

Bei der Online-Befragung wurden 284 Datensätze von peri- und postmenopausalen Frauen abgeschlossen. Unter diesen befanden sich 144 Datensätze von symptomatischen, peri- und postmenopausalen Frauen (vgl. Abbildung 10). Das durchschnittliche Alter der eingeschlossenen Stichprobe lag bei 53,7 Jahren (SD = 5,7; Median: 53; Range: 41 – 85). Die Anzahl der fehlenden Werte nahm augenscheinlich zum Ende der Befragung hin zu. Die durchschnittliche Dauer für das Ausfüllen der kompletten Online-Befragung (n = 122) lag bei 9,35 ($\pm 2,25$) Minuten.

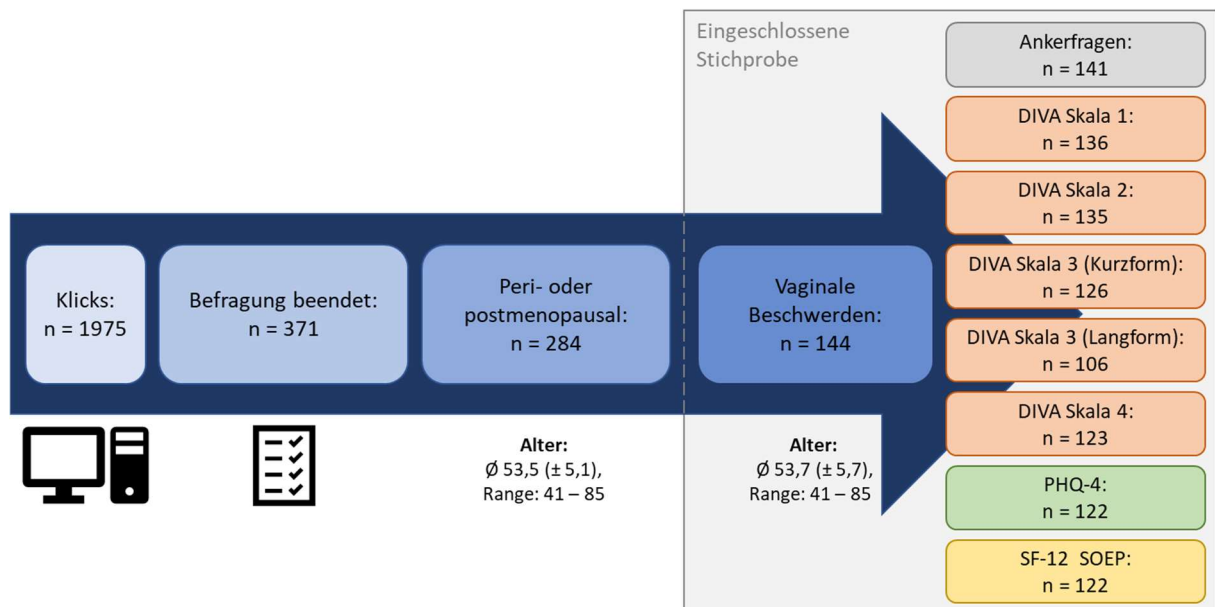


Abbildung 10 Zusammensetzung der Stichprobe der Online-Befragung

Anmerkungen. n = Stichprobengröße, Skala 1 = Aktivitäten des täglichen Lebens, Skala 2 = Emotionales Wohlbefinden, Skala 3 = Sexuelle Funktionsfähigkeit, Skala 4 = Selbstkonzept und Körperbild. *Abkürzungen.* DIVA = Day-to-Day Impact of Vaginal Aging Questionnaire, PHQ = Patient Health Questionnaire, SF-12 = short-form health survey, SOEP = Sozio-oekonomisches Panel.

Die Papierversion der Befragung wurde von 41 Frauen ausgefüllt. Diese waren im Durchschnitt 59,5 Jahre alt (SD = 6,1; Median: 58; Range: 50 – 71).

Bevor die beiden Stichproben (Online- und Papierversion der Befragung) zusammengefasst wurden, wurde untersucht, inwieweit sich die beiden voneinander unterscheiden und ob eine Zusammenfassung überhaupt zulässig wäre. Dazu wurde eine neue Variable zur Darbietungsform der Befragung (Online- vs. Papierversion) generiert. Für die Analysen wurden Frage 10 der MRS zur vaginalen Trockenheit und die fünf Ankerfragen zur Schwere der vaginalen Beschwerden herangezogen. Zum einen wurde mit Mann-Whitney-U-Tests überprüft, ob sich die beiden Stichproben hinsichtlich des Kardinalsymptoms des GSM, der vaginalen Trockenheit, voneinander unterscheiden und zum anderen, ob es signifikante Unterschiede in der Schwere vaginaler Beschwerden zwischen den beiden Stichproben gibt. Weder bei Frage 10 der MRS ($p = 0,702$) noch bei den fünf Ankerfragen ($p \geq 0,065$) zeigten sich signifikante Unterschiede in der Darbietungsform.

Beide Stichproben wurden somit zusammengefasst (\bar{x} Alter = 54,9 Jahre; SD = 6,2; Median: 54; Range: 41 – 85). Eine Übersicht zu den demographischen Variablen der beiden Stichproben (Alter, Heimatland, Muttersprache, Zyklus und Hormoneinnahme) ist in Tabelle 13 dargestellt. Die Teilnehmerinnen der Online-Befragung waren im Schnitt etwas jünger als die Teilnehmerinnen an der Papier-Befragung. Weniger als 15 % befanden sich noch im Übergang zur Menopause, dadurch dass noch unregelmäßige Blutungen auftraten. Der Großteil gab an, seit über einem Jahr keine Periode mehr zu haben. Mehr als die Hälfte der Teilnehmerinnen an der Online-Befragung berichteten, keine Hormone einzunehmen, knapp 35 % berichteten von einer Hormoneinnahme zur Behandlung von Wechseljahresbeschwerden.

Tabelle 13 Demographische Daten der Befragung (aufgeteilt nach Online- und Papierversion)

Variable	Online-Version n = 144	Papierversion n = 41
Durchschnittsalter	53,7 (\pm 5,7)	59,5 (\pm 6,1)
Heimatland		
- Deutschland	108 (75,0 %)	40 (97,6 %)
- Österreich	11 (7,6 %)	1 (2,4 %)
- Schweiz	22 (15,3 %)	-
- Anderes Land ¹	3 (2,1 %)	-
Muttersprache		
- Deutsch	143 (99,3 %)	39 (95,1 %)
- Sonstige ²	1 (0,7 %)	2 (4,9 %)
Zyklus		
- Pausen von mind. 2 Monaten	21 (14,6 %)	2 (4,8 %)
- Keine Periode mehr seit mind. einem Jahr	93 (64,4 %)	31 (75,6 %)
- Keine Gebärmutter mehr	30 (20,8 %)	7 (17,1 %)
Hormoneinnahme		
- Keine Hormoneinnahme	83 (57,7 %)	41 (100 %) ³
- Hormoneinnahme zur Schwangerschaftsverhütung	4 (2,8 %)	-
- Hormoneinnahme zur Behandlung von Wechseljahresbeschwerden	50 (34,7 %)	-
- Sonstiges	7 (4,9 %)	-

Anmerkungen. ¹Als andere Heimatländer wurden die USA, Italien und Spanien (nur im Winter) angegeben. ²Als sonstige Muttersprachen wurden Russisch, Polnisch und Französisch angegeben. ³Keine Hormoneinnahme war Voraussetzung für die Teilnahme am Zusatzscreening.

Die MRS wurde von 185 Teilnehmerinnen ausgefüllt. Fast drei Viertel der Teilnehmerinnen berichteten von mittlerer bis sehr starker vaginaler Trockenheit und mehr als 70 % der Teilnehmerinnen von mittleren bis sehr starken Sexualproblemen. Mittlere bis sehr starke Harnwegsbeschwerden wurden von ca. 40 % der Teilnehmerinnen angegeben (vgl. Tabelle 14).

Tabelle 14 Ergebnisse der Menopause Rating Scale II (MRS), n = 185

Aktuelle Beschwerden	keine	leicht	mittel	stark	sehr stark
Wallungen, Schwitzen (Aufsteigende Hitze, Schweißausbrüche)	23,2 %	28,6 %	29,2 %	11,4 %	7,6 %
Herzbeschwerden (Herzklopfen, Herzrasen, Herzstolpern, Herzbeklemmungen)	39,5 %	27,0 %	23,8 %	7,0 %	1,6 %
Schlafstörungen (Einschlafstörungen, Durchschlaf- störungen, zu frühes Aufwachen)	13,5 %	16,8 %	26,5 %	27,6 %	15,7 %
Depressive Verstimmung (Mutlosigkeit, Traurigkeit, Weinerlichkeit, Antriebslosigkeit, Stimmungsschwankungen)	27,0 %	24,9 %	22,7 %	17,3 %	8,1 %
Reizbarkeit (Nervosität, innere Anspannung, Aggressivität)	20,0 %	31,9 %	25,9 %	14,1 %	8,1 %
Ängstlichkeit (Innere Unruhe, Panik)	38,9 %	20,0 %	20,5 %	13,5 %	7,0 %
Körperliche und geistige Erschöpfung (Allgemeine Leistungsminderung, Gedächtnisminderung, Konzentrations- schwäche, Vergesslichkeit)	16,8 %	28,1 %	23,8 %	23,2 %	8,1 %
Sexualprobleme (Veränderung des sexuellen Verlangens, der sexuellen Betätigung und Befriedigung)	14,1 %	14,1 %	23,2 %	26,5 %	22,2 %
Harnwegsbeschwerden (Beschwerden beim Wasserlassen, häufiger Harndrang, unwillkürlicher Harnabgang)	35,1 %	24,3 %	21,6 %	11,9 %	7,0 %
Trockenheit der Scheide (Trockenheitsgefühl oder Brennen der Scheide, Beschwerden beim Geschlechtsverkehr)	6,5 %	20,0 %	29,7 %	23,8 %	20,0 %
Gelenk- und Muskelbeschwerden (Schmerzen im Bereich der Gelenke, rheuma-ähnliche Beschwerden)	21,6 %	16,8 %	20,0 %	24,3 %	17,3 %

3.3.1 Strukturvalidität

Die Ergebnisse der zwei KFAs sind in Tabelle 15 dargestellt.

Tabelle 15 Fit-Maße für die Kurzversion (für alle Frauen, n = 152) und die Langversion (nur für sexuell aktive Frauen, n = 128) des DIVA zur Bestätigung der vier Skalen

Fit-Maße	Angemessener (noch akzeptabler) Modell-Fit	DIVA Kurzversion	DIVA Langversion
CFI	≥ 0,95 (≥ 0,90)	0,942	0,932
TLI	≥ 0,95 (≥ 0,90)	0,932	0,923
RMSEA	< 0,05 (< 0,08)	0,070	0,072
		(90 % CI: 0,055 – 0,084)	(90 % CI: 0,058 – 0,084)
SRMR	< 0,05 (< 0,08)	0,053*	0,061*
χ^2/df	< 2 (< 3)	1,734	1,657

Abkürzungen. CFI = comparative fit index, TLI = Tucker-Lewis Index, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual, χ^2/df = Chi square/degrees of freedom, CI = Konfidenzintervall. *hinreichend gemäß den COSMIN-Kriterien (52).

3.3.2 Interne Konsistenz

Cronbach's alpha wurde für die vier Skalen des DIVA berechnet (vgl. Tabelle 16). Für die Skala zur sexuellen Funktionsfähigkeit wurde Cronbach's alpha sowohl für sexuell aktive Frauen (Langversion) als auch für nicht sexuell aktive Frauen (Kurzversion) berechnet.

Tabelle 16 Cronbach's alpha für die Skalen des DIVA

	Skala 1	Skala 2	Skala 3		Skala 4	Gesamte Skala	
			Kurzversion	Langversion		Kurzversion	Langversion
Cronbach's α	0,84	0,91	0,84	0,93	0,93	0,91	0,92
n	174	176	158	132	159	152	128

Anmerkungen. Skala 1 = Aktivitäten des täglichen Lebens, Skala 2 = Emotionales Wohlbefinden, Skala 3 = Sexuelle Funktionsfähigkeit, Skala 4 = Selbstkonzept und Körperbild, n = Stichprobengröße.

3.3.3 Test-Retest-Reliabilität

Vierundsechzig Teilnehmerinnen hinterließen ihre E-Mail-Adresse am Ende der Befragung und wurden nach einer Woche angeschrieben (und ggf. einmal erinnert). Von diesen 64 Frauen nahmen 31 erneut an der Befragung teil, wobei die Befragung nur von 26 Teilnehmerinnen vollständig abgeschlossen wurde. Die Befragung wurde im Schnitt nach 8.84 Tagen (SD = 2.4, Median: 8, Range: 7 – 17) ausgefüllt. Nach ca. einer Woche zeigten 15 – 18 Teilnehmerinnen keine Veränderung in einer der Ankerfragen (vgl. Tabelle 17).

Tabelle 17 Anzahl an Teilnehmerinnen mit stabiler bzw. veränderter Schwere vaginaler Symptome bzw. Beeinträchtigungen anhand der fünf Ankerfragen

Anker- frage	Grad der Veränderung nach ca. einer Woche					
	Starke Verbesserung	Mittlere Verbesserung	Leichte Verbesserung	Keine Veränderung	Leichte Verschlechterung	Mittlere Verschlechterung
	-3	-2	-1	0	1	2
1	1	1	7	17	5	-
2	-	2	7	17	3	2
3	-	3	10	15	2	1
4	-	1	6	18	6	-
5	-	-	9	18	4	-

Anmerkungen. Grau: Stabile Teilnehmerinnen.

Da die Stichprobengröße der stabilen Teilnehmerinnen nicht groß genug war, um valide ICCs berechnen zu können, wurden die Mittelwerte (und Standardabweichungen) zu beiden Erhebungszeitpunkten gegenübergestellt, um diese deskriptiv zu vergleichen (vgl. Tabelle 18). Es zeigten sich identische bis sehr ähnliche Mittelwerte zwischen den beiden Erhebungszeitpunkten, wodurch eine stabile Messung des DIVA angenommen werden konnte.

3.3.4 Konstruktvalidität: Hypothesen-Testen

3.3.4.1 Konvergente Validität

Von den 47 *a priori* postulierten Hypothesen konnten 28 bestätigt werden (vgl. Tabelle 19). Manche Korrelationen wurden leicht unterschätzt (z. B. mit dem PHQ), manche wurden leicht überschätzt (z. B. Skala 3 mit den Ankerfragen). Im Großen und Ganzen wurde keine Hypothese völlig verfehlt.

Tabelle 18 Deskriptive Gegenüberstellung der Mittelwerte (und Standardabweichungen) der Skalen des DIVA bei stabilen Teilnehmerinnen (Veränderung in der jeweiligen Ankerfrage = 0) bei *Baseline* und nach ca. einer Woche

		Ankerfragen				
		1	2	3	4	5
Skala 1	T0	0,647 (0,64) n = 17	0,918 (0,90) n = 17			
	T1	0,647 (0,83) n = 17	0,929 (1,03) n = 17			
Skala 2	T0	0,721 (0,93) n = 17		1,133 (1,24) n = 15		
	T1	0,721 (0,99) n = 17		1,150 (1,11) n = 15		
Skala 3 (kurz)	T0	2,075 (0,77) n = 16			2,000 (0,93) n = 18	
	T1	1,707 (0,89) n = 15			2,294 (1,05) n = 17	
Skala 3 (lang)	T0	2,082 (0,74) n = 15			2,153 (0,90) n = 16	
	T1	1,769 (0,83) n = 13			2,341 (1,06) n = 15	
Skala 4	T0	1,506 (0,94) n = 17				1,906 (1,12) n = 17
	T1	1,520 (1,15) n = 15				1,824 (1,18) n = 17
Gesamt (kurz)	T0	1,293 (0,70) n = 16	1,498 (0,77) n = 15	1,519 (0,92) n = 14	1,629 (0,70) n = 18	1,511 (0,84) n = 17
	T1	1,197 (0,90) n = 15	1,495 (0,93) n = 15	1,571 (1,03) n = 14	1,653 (0,83) n = 17	1,542 (0,93) n = 17
Gesamt (lang)	T0	1,435 (0,68) n = 15	1,600 (0,73) n = 15	1,606 (0,83) n = 14	1,799 (0,68) n = 16	1,681 (0,78) n = 15
	T1	1,284 (0,87) n = 13	1,682 (0,93) n = 13	1,716 (1,05) n = 13	1,849 (0,83) n = 15	1,733 (0,98) n = 14

Anmerkungen. Skala 1 = Aktivitäten des täglichen Lebens, Skala 2 = Emotionales Wohlbefinden, Skala 3 = Sexuelle Funktionsfähigkeit, Skala 4 = Selbstkonzept und Körperbild, T0 = *Baseline*, T1 = nach ca. 1 Woche, n = Stichprobengröße.

Tabelle 19 Tatsächliche Korrelationen des DIVA mit anderen Items und PROMs

	MRS			Ankerfragen					PHQ			SF-12 SOEP*	
	Item 8 ¹	Item 9 ²	Item 10 ³	1	2	3	4	5	Depr.	Ängst.	Gesamt	PCS	MCS
Skala 1				0,414	0,465	0,510	0,340	0,556				-0,452	
n				175	175	175	175	175				158	
Skala 2				0,365	0,537	0,698	0,487	0,727	0,533	0,533	0,579		-0,483
n				176	176	176	175	176	158	158	158		157
Skala 3 (kurz)				0,233	0,242	0,225	0,350	0,334					
n				158	158	158	157	158					
Skala 3 (lang)				0,372	0,275	0,286	0,417	0,376					
n				132	132	132	132	132					
Skala 4				0,447	0,595	0,603	0,673	0,685					-0,386
n				159	159	159	158	159					155
Gesamt (kurz)	0,402	0,377	0,423	0,535	0,643	0,710	0,681	0,799					
n	153	153	153	153	153	153	153	153					
Gesamt (lang)	0,430	0,271	0,455	0,587	0,630	0,669	0,689	0,764					
n	129	129	129	129	129	129	129	129					

Anmerkungen. Grün: hohe Korrelation erwartet (> 0,5), gelb: mittlere Korrelation erwartet (0,3 – 0,5), grau: keine Hypothese, **fett: Hypothese bestätigt**, Skala 1 = Aktivitäten des täglichen Lebens, Skala 2 = Emotionales Wohlbefinden, Skala 3 = Sexuelle Funktionsfähigkeit, Skala 4 = Selbstkonzept und Körperbild, n = Stichprobengröße. *Abkürzungen.* MRS = *Menopause Rating Scale*, PHQ = *Patient Health Questionnaire*, Depr. = Depression, Ängst. = Ängstlichkeit, SF-12 = *short-form health survey*, SOEP = Sozio-oekonomisches Panel, PCS = *physical component score*, MCS = *mental component score*. ¹Sexualprobleme, ²Harnwegsbeschwerden, ³Trockenheit der Scheide. *negative Korrelationen erwartet (je höher die Lebensqualität, desto geringer die Auswirkung der Beschwerden).

3.3.4.2 Diskriminante Validität

Wie vermutet wurden keine signifikanten Unterschiede zwischen den festgesetzten Alterskategorien (≤ 39 , 40 – 49, 50 – 59, 60 – 69, ≥ 70 Jahre) in den verschiedenen Skalen des DIVA gefunden. Wider Erwarten zeigten sich signifikante Unterschiede zwischen den Heimatländern der Teilnehmerinnen in den Skalen 1 und 2 und in den beiden Gesamtskalen (kurz und lang). Bei der Interpretation dieser Ergebnisse sollte jedoch berücksichtigt werden, dass der Großteil der Teilnehmerinnen in Deutschland wohnte ($n = 101 - 138$) und nur sehr wenige in Österreich ($n = 9 - 12$) bzw. der Schweiz ($n = 20 - 22$). Bezüglich der Schwere der Erkrankung wurde überprüft, ob die Skalen des DIVA in der Lage sind, zwischen verschiedenen Schweregruppen zu diskriminieren. Anhand der ersten Ankerfrage konnten alle Skalen des DIVA, wie erwartet, signifikant zwischen den verschiedenen Schweregraden unterscheiden, bei Frage 10 der MRS gelang dies auch allen Skalen bis auf Skala 3 zur sexuellen Funktionsfähigkeit (kurz und lang). Dies kann dadurch erklärt werden, dass Skala 3 nur gering mit Frage 10 der MRS korrelierte (Kurzversion: $r = 0,198$, Langversion: $r = 0,232$).

Insgesamt konnten 22 der 28 *a priori* postulierten Hypothesen bestätigt werden (vgl. Tabelle 20).

Tabelle 20 P-Werte der einzelnen Kruskal-Wallis-Tests zur Überprüfung der diskriminanten Validität

	p-Werte			
	Alter	Heimatland	Schwere der Erkrankung	
			Frage 10 (MRS)	Ankerfrage 1
Skala 1	0,215	< 0,001*	0,001*	< 0,001*
Skala 2	0,755	0,002*	0,010*	< 0,001*
Skala 3 (kurz)	0,294	0,143	0,107	0,028*
Skala 3 (lang)	0,462	0,102	0,096	0,001*
Skala 4	0,747	0,256	< 0,001*	< 0,001*
Gesamt (kurz)	0,620	0,035*	< 0,001*	< 0,001*
Gesamt (lang)	0,928	0,049*	< 0,001*	< 0,001*

Anmerkungen. Grau: keine signifikanten Gruppenunterschiede erwartet, grün: signifikante Gruppenunterschiede erwartet, **fett: Hypothese bestätigt**, Skala 1 = Aktivitäten des täglichen Lebens, Skala 2 = Emotionales Wohlbefinden, Skala 3 = Sexuelle Funktionsfähigkeit, Skala 4 = Selbstkonzept und Körperbild. *signifikant ($\alpha = 0,05$). *Abkürzungen.* MRS = *Menopause Rating Scale*.

3.3.5 Änderungssensitivität

Alle 64 Teilnehmerinnen, die ihre E-Mail-Adresse am Ende der Befragung hinterlassen hatten, wurden erneut nach vier Wochen mit einem weiteren Link zu den Fragen und Fragebögen kontaktiert (und ggf. einmal erinnert). Von diesen 64 Frauen nahmen 28 erneut an der Befragung teil, wobei die Befragung nur von 26 Teilnehmerinnen vollständig abgeschlossen wurde. Der letzte Teil der Befragung wurde im Schnitt nach 30,71 Tagen (SD = 5,0; Median: 29,5; Range: 28 – 50) ausgefüllt.

Die Spearman-Rangkorrelationen der Veränderung in den Ankerfragen mit den Veränderungen in den Skalen des DIVA lagen zwischen -0,032 und 0,529. Aufgrund der geringen Stichprobengröße konnte die Angemessenheit der Ankerfragen nur in 10 von 20 Fällen bestätigt werden. Die meisten Teilnehmerinnen gaben keine Veränderung in den Ankerfragen nach ca. vier Wochen an (vgl. Tabelle 21).

Tabelle 21 Anzahl an Teilnehmerinnen mit veränderter bzw. stabiler Schwere vaginaler Symptome bzw. Beeinträchtigungen anhand der fünf Ankerfragen

Anker- frage	Grad der Veränderung nach ca. vier Wochen					
	Starke Verbesserung	Mittlere Verbesserung	Leichte Verbesserung	Keine Veränderung	Leichte Verschlechterung	Mittlere Verschlechterung
	-3	-2	-1	0	1	2
1	1	1	11	14	1	-
2	1	5	5	15	2	-
3	-	4	5	17	2	-
4	-	3	7	15	3	-
5	-	1	10	12	5	-

Anmerkungen. Grau: Teilnehmerinnen, die eine Verbesserung angeben.

Alle Teilnehmerinnen, die eine Verbesserung in den Ankerfragen nach ca. vier Wochen angaben, wurden zusammengefasst und deren Mittelwerte (und Standardabweichungen) in den Skalen des DIVA an beiden Erhebungszeitpunkten gegenübergestellt. Deskriptiv ließ sich eine Verbesserung in allen Skalen des DIVA beobachten, wenn man die Mittelwerte von *Baseline* mit denen des letzten Erhebungszeitpunktes vergleicht (vgl. Tabelle 22).

Tabelle 22 Deskriptive Gegenüberstellung der Mittelwerte (und Standardabweichungen) der Skalen des DIVA bei Teilnehmerinnen, die eine Verbesserung wahrnehmen (Veränderung in der jeweiligen Ankerfrage < 0), bei *Baseline* und nach ca. vier Wochen

		Ankerfragen				
		1	2	3	4	5
Skala 1	T0	1,185 (0,90) n = 13	1,346 (1,01) n = 11			
	T2	0,917 (0,85) n = 12	1,146 (1,01) n = 11			
Skala 2	T0	0,981 (1,07) n = 13		1,694 (1,07) n = 9		
	T2	0,667 (0,76) n = 12		1,222 (0,94) n = 9		
Skala 3 (kurz)	T0	1,954 (0,70) n = 13			2,044 (0,73) n = 9	
	T2	1,667 (1,13) n = 12			1,467 (1,22) n = 9	
Skala 3 (lang)	T0	2,000 (0,75) n = 12			2,250 (0,90) n = 8	
	T2	1,354 (0,85) n = 11			1,476 (0,83) n = 7	
Skala 4	T0	1,985 (1,14) n = 13				1,891 (1,07) n = 11
	T2	1,850 (1,16) n = 12				1,740 (0,99) n = 10
Gesamt (kurz)	T0	1,555 (0,72) n = 13	1,900 (0,59) n = 10	2,105 (0,86) n = 9	1,374 (0,61) n = 9	1,563 (0,64) n = 10
	T2	1,307 (0,80) n = 12	1,440 (0,81) n = 11	1,608 (0,99) n = 9	1,146 (0,85) n = 9	1,211 (0,82) n = 10
Gesamt (lang)	T0	1,667 (0,66) n = 12	2,029 (0,53) n = 9	2,150 (0,85) n = 9	1,587 (0,71) n = 8	1,657 (0,68) n = 10
	T2	1,237 (0,71) n = 11	1,493 (0,68) n = 9	1,594 (0,94) n = 9	1,224 (0,77) n = 7	1,338 (0,74) n = 9

Anmerkungen. Skala 1 = Aktivitäten des täglichen Lebens, Skala 2 = Emotionales Wohlbefinden, Skala 3 = Sexuelle Funktions-fähigkeit, Skala 4 = Selbstkonzept und Körperbild, T0 = *Baseline*, T2 = nach ca. 4 Wochen, n = Stichprobengröße.

3.4 Explorative Analysen

Dieses Kapitel beinhaltet Analysen, die *post-hoc* durchgeführt wurden. Zur Erleichterung der Lektüre wurde dieses Kapitel nicht in „Methode“ und „Ergebnisse“ unterteilt.

3.4.1 Einfluss von Hormonen auf die Auswirkung von vaginalen Beschwerden

Als erstes wurde analysiert, ob sich Frauen, die Hormone zur Behandlung von Wechseljahresbeschwerden einnahmen, von Frauen, die keinerlei Hormone einnahmen, hinsichtlich der Schwere und Auswirkung ihrer Beschwerden unterschieden. Dazu wurde der non-parametrische Mann-Whitney-U-Test verwendet, um die beiden Gruppen miteinander zu vergleichen. Für keine Skala bzw. kein Item wurden signifikante Unterschiede zwischen den beiden Gruppen gefunden, d. h. ob Hormone eingenommen wurden oder nicht, hatte keinen Einfluss auf die Schwere bzw. auf die Auswirkungen der Beschwerden (vgl. Tabelle 23).

Tabelle 23 P-Werte der einzelnen Mann-Whitney-U-Tests zur Überprüfung des Einflusses von Hormonen

		n		p-Wert
		Gruppe A	Gruppe B	
DIVA	Skala 1	118	46	0,209
	Skala 2	119	46	0,304
	Skala 3 (kurz)	108	39	0,935
	Skala 3 (lang)	89	33	0,822
	Skala 4	110	39	0,096
	Gesamt (kurz)	105	38	0,312
	Gesamt (lang)	87	33	0,116
MRS	Item 8 ¹	124	50	0,929
	Item 9 ²	124	50	0,222
	Item 10 ³	124	50	0,640
PHQ	Depression	110	38	0,383
	Ängstlichkeit	110	38	0,899
	Gesamt	110	38	0,622
SF-12 SOEP	PCS	106	37	0,836
	MCS	106	37	0,433

Anmerkungen. Gruppe A: keine Hormoneinnahme, Gruppe B: Hormoneinnahme zur Behandlung von Wechseljahresbeschwerden, Skala 1 = Aktivitäten des täglichen Lebens, Skala 2 = Emotionales Wohlbefinden, Skala 3 = Sexuelle Funktionsfähigkeit, Skala 4 = Selbstkonzept und Körperbild, n = Stichprobengröße. ¹Sexualprobleme, ²Harnwegsbeschwerden, ³Trockenheit der Scheide. *Abkürzungen.* DIVA = *Day-to-Day Impact of Vaginal Aging Questionnaire*, MRS = *Menopause Rating Scale*, PHQ = *Patient Health Questionnaire*, SF-12 = *short-form health survey*, SOEP = *Sozio-oekonomisches Panel*, PCS = *physical component score*, MCS = *mental component score*.

3.4.2 Auswirkungen der vaginalen Beschwerden (im DIVA) über das Alter hinweg

Die Teilnehmerinnen wurden in Alterskategorien eingeteilt: 28 Frauen waren zwischen 40 und 49 Jahre alt, 121 Frauen zwischen 50 und 59, 26 Frauen zwischen 60 und 69 und 5 Frauen waren älter als 69 Jahre. Die Mittelwerte der Skalen des DIVA über die verschiedenen Alterskategorien hinweg sind in Abbildung 11 dargestellt. Die letzte Kategorie (> 69 Jahre) wurde nicht abgebildet, da die Größe der Stichprobe (n = 5) als zu gering und somit als nicht aussagekräftig angesehen wurde.

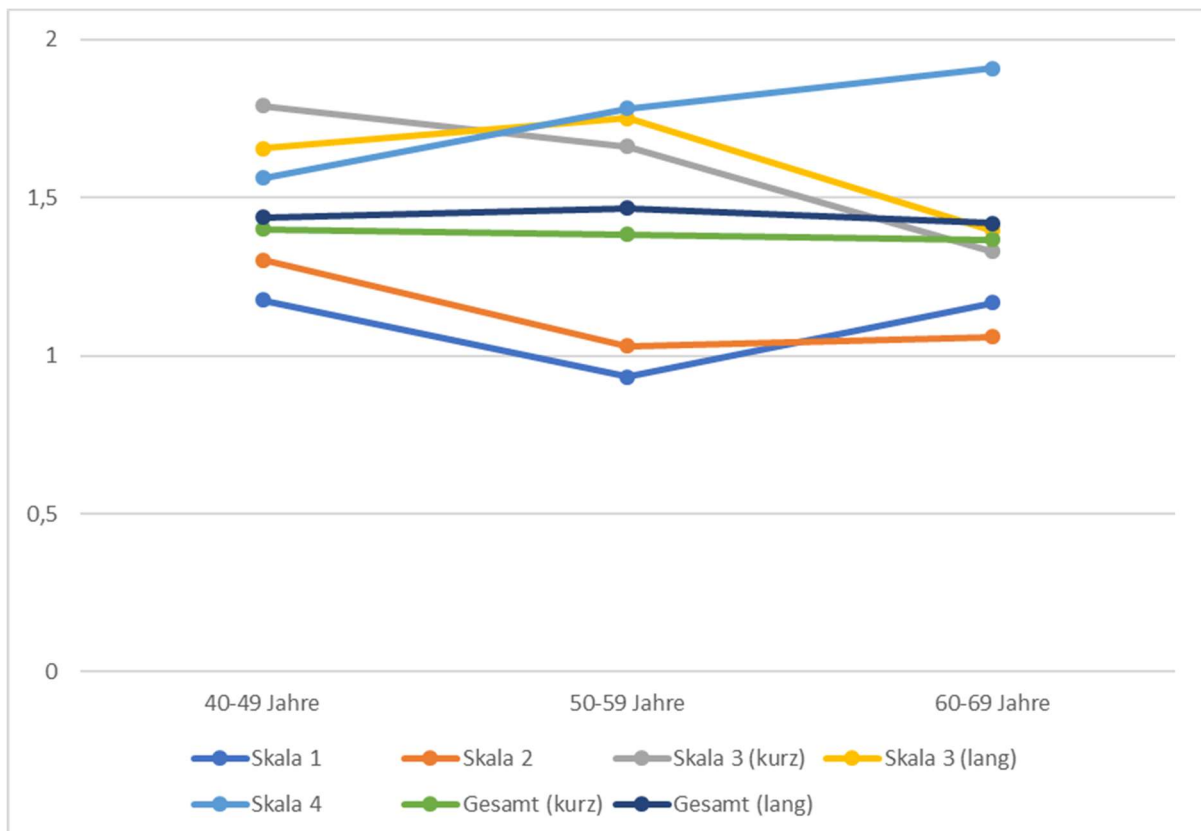


Abbildung 11 Mittelwerte der DIVA-Skalen über verschiedene Altersgruppen hinweg

Anmerkungen. Skala 1 = Aktivitäten des täglichen Lebens, Skala 2 = Emotionales Wohlbefinden, Skala 3 = Sexuelle Funktionsfähigkeit, Skala 4 = Selbstkonzept und Körperbild.

Die beiden Gesamtskalen verlaufen nahezu ohne Veränderung über das Alter hinweg. Bei Skala 4 zum Selbstkonzept und Körperbild ist eine deutlich größere Auswirkung der Beschwerden mit dem Alter zu beobachten, während die Auswirkungen bei allen anderen Skalen eher nachlassen. Nur bei Skala 1 zu den Aktivitäten des täglichen Lebens ist in höherem Alter wieder eine stärkere Beeinträchtigung zu beobachten, die auch mit der nachlassenden physischen Verfassung zusammenhängen könnte. Die Abbildung verdeutlicht noch einmal die Tatsache, dass keine signifikanten Unterschiede in den Skalen des DIVA zwischen den verschiedenen Altersgruppen gefunden wurden.

3.4.3 „Verfeinerung“ des DIVA

Für die „Verfeinerung“ des DIVA wurde dessen Strukturvalidität mithilfe der probabilistischen Testtheorie (*item response theory* (IRT) bzw. die Rasch-Analyse) und der klassischen Testtheorie (KTT) genauer untersucht. Die IRT und die Rasch-Analyse gehören beide zu den *Latent-trait*-Modellen, da sie den Zusammenhang zwischen einem latenten Merkmal und der Wahrscheinlichkeit, eine bestimmte Antwort zu wählen, die sich auf das latente Merkmal bezieht, beschreiben. Sie geben Auskunft darüber, wie gut die Items einer PROM funktionieren. Die Item- und Personenebene kann in einer Analyse betrachtet werden (91). Typischerweise beschreiben die IRT und KTT ein Datenset, wohingegen die Rasch-Analyse erreichen möchte, dass die Daten in das Modell passen. Die KTT beinhaltet das traditionelle psychometrische Vorgehen (92). Gemäß der COSMIN-Initiative wird die Strukturvalidität einer PROM bei Verwendung der IRT oder Rasch-Analyse als hinreichend bewertet, wenn die Annahmen zur Monotonie (*monotonicity*), zur Eindimensionalität (*unidimensionality*) und lokalen Unabhängigkeit (*local independence*) erfüllt sind und eine adäquate Passung des Modells (*model fit*) bestätigt werden konnte (52). Bei Verwendung der KTT zur Analyse der Strukturvalidität sollte eine exploratorische Faktorenanalyse (EFA) oder eine KFA durchgeführt werden (10, 93). Mithilfe einer EFA wurde in dieser Stichprobe überprüft, ob die Items einer Skala tatsächlich auf ihre zugehörige Skala laden. Bei allen Items sollte deren höchste Faktorladung ($> 0,4$) auf deren zugehörigem Faktor liegen und zudem sollten keine signifikanten Faktorladungen ($\leq 0,4$) auf einen anderen Faktor vorliegen (93, 94). Außerdem wurden die Spearman-Rangkorrelationen der einzelnen Items mit ihren zugehörigen Skalen überprüft. Hohe Korrelationen ($\geq 0,7$) wurden mit der zugehörigen Skala erwartet, niedrige Korrelationen ($< 0,7$) mit den restlichen Skalen (95). Zu guter Letzt wurden die Ergebnisse aus der IRT bzw. Rasch-Analyse und der KTT zusammengebracht und schlecht funktionierende Items aussortiert.

3.4.3.1 Überprüfung der Angemessenheit der Antwortkategorien (*monotonicity*)

Die Angemessenheit der Antwortkategorien kann mithilfe von Itemfunktionen (*item characteristic curves*) überprüft werden. Itemfunktionen sind ein wichtiger Bestandteil der IRT. Mit ihnen kann die passende Anzahl an Antwortoptionen basierend auf den Antworten der Testpersonen bestimmt werden (91). Die Itemfunktionen des DIVA wurden für jede Skala separat analysiert, um zu überprüfen, ob die Antwortkategorien gut funktionieren bzw. ob eine Anpassung der Kategorien des DIVA vonnöten ist. Bei optimal funktionierenden Antwortkategorien hat jede Kategorie ihren eigenen Höhepunkt. Die Wahrscheinlichkeit, eine

bestimmte Kategorie zu wählen (y-Achse), sollte mit dem latenten Merkmal, dem Ausmaß der Auswirkung der Beschwerden (x-Achse), monoton zunehmen. Hat eine Kategorie keinen eigenen Höhepunkt und somit auch keinen eigenen Bereich auf dem latenten Merkmal, ist diese Antwortkategorie niemals die wahrscheinlichste, die bei einer bestimmten Merkmalsausprägung gewählt wird (4, 96). Alle Itemfunktionen wurden mit SAS 9.4 berechnet.

Analyse der Antwortkategorien von Skala 1

Die erste Skala des DIVA zu den Aktivitäten des täglichen Lebens enthielt ursprünglich folgende Antwortkategorien: überhaupt nicht (0), ein wenig (1), mäßig (2), ziemlich (3) und extrem (4). Diese Kategorien funktionierten jedoch nicht optimal, da nicht jede Kategorie ihren eigenen Höhepunkt auf dem latenten Merkmal auf der x-Achse hatte. Durch eine Kombination benachbarter Kategorien (von 0-1-2-3-4 zu 0-1-1-2-2) konnte eine Ordnung der Itemfunktionen erreicht werden (vgl. Abbildung 12).

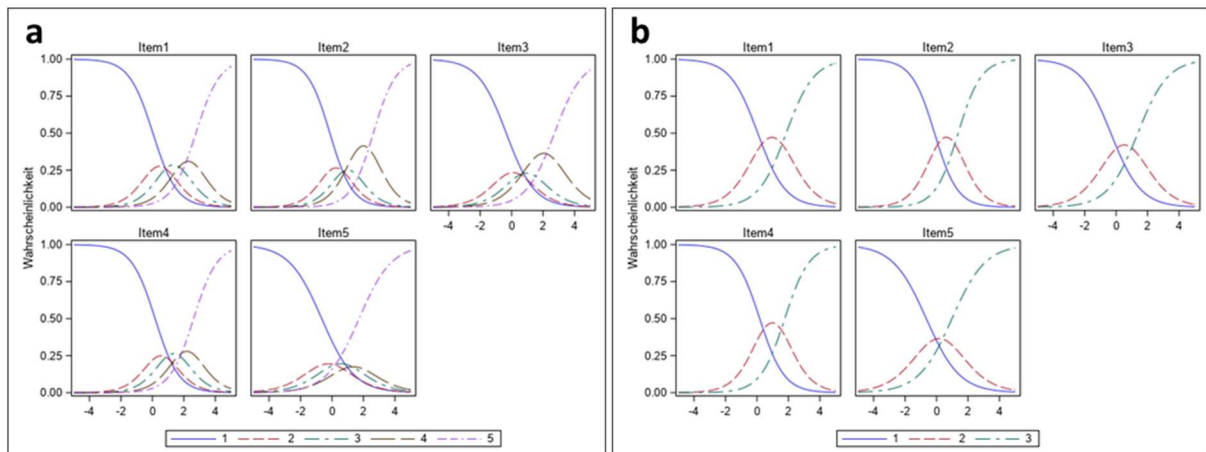


Abbildung 12 Itemfunktionen (*item characteristic curves*) der Items 1 – 5 für die originalen Antwortkategorien des DIVA (a) und für Antwortkategorien nach der Adaption (b)

Anmerkungen. 1 = Antwortkategorie 1 (0), 2 = Antwortkategorie 2 (1), 3 = Antwortkategorie 3 (2), 4 = Antwortkategorie 4 (3), 5 = Antwortkategorie 5 (4).

Analyse der Antwortkategorien von Skala 2

Die zweite Skala des DIVA zum emotionalen Wohlbefinden enthielt ursprünglich folgende Antwortkategorien: nie (0), selten (1), manchmal (2), ziemlich oft (3) und sehr oft (4). Bei den Items 6 und 7 konnten wieder vermischte Itemfunktionen beobachtet werden. Die Items 8 und 9 zeigten für jede Antwortkategorie einen eigenen, wenn auch kleinen Höhepunkt. Auch hier konnte wieder durch eine Kombination von Kategorie 1 und 2 und von Kategorie 3 und 4 eine deutlich klarere Trennung der Antwortkategorien erzielt werden (vgl. Abbildung 13).

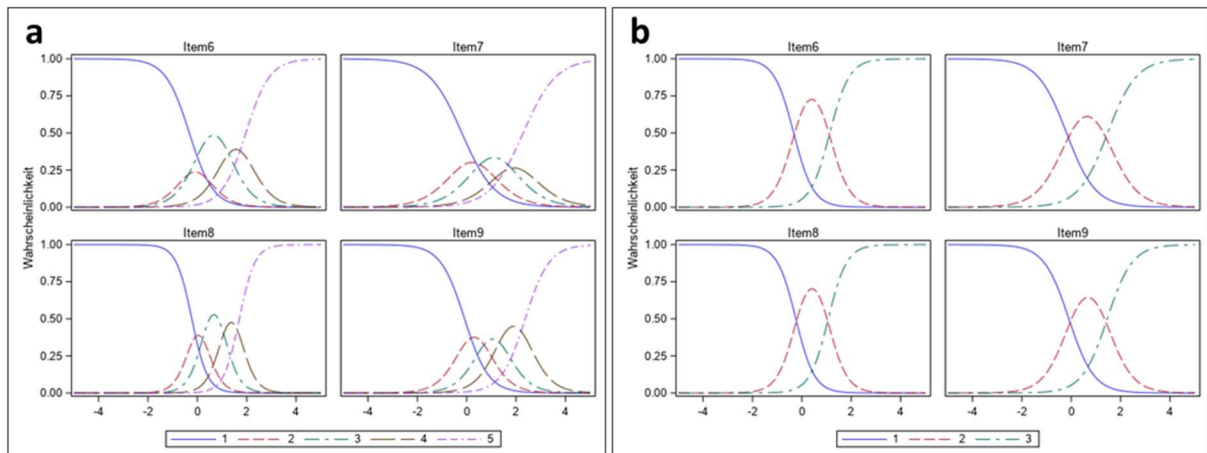


Abbildung 13 Itemfunktionen (*item characteristic curves*) der Items 6 – 9 für die originalen Antwortkategorien des DIVA (a) und für Antwortkategorien nach der Adaption (b)

Anmerkungen. 1 = Antwortkategorie 1 (0), 2 = Antwortkategorie 2 (1), 3 = Antwortkategorie 3 (2), 4 = Antwortkategorie 4 (3), 5 = Antwortkategorie 5 (4).

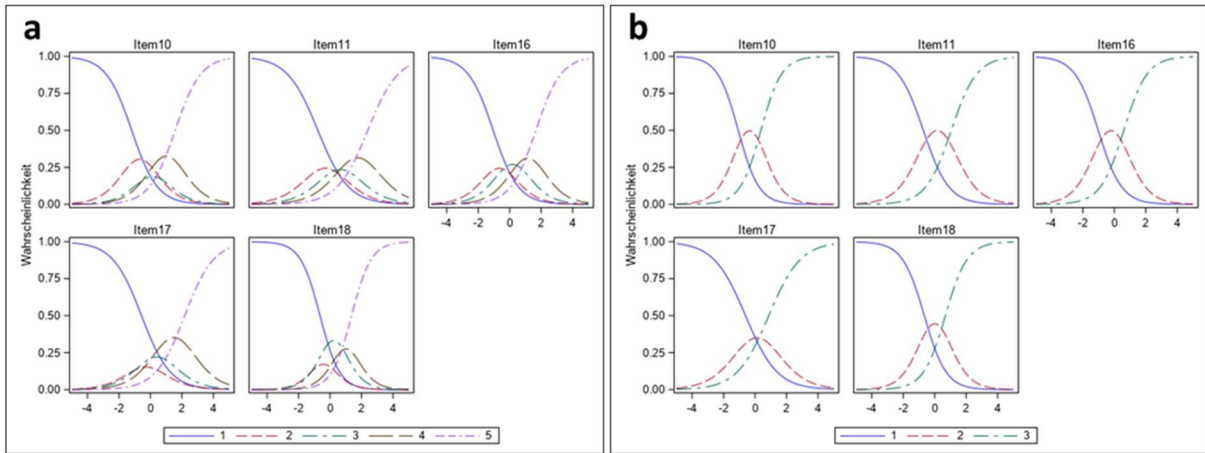
Analyse der Antwortkategorien von Skala 3 (kurz und lang)

Die dritte Skala des DIVA zur sexuellen Funktionsfähigkeit enthielt ursprünglich folgende Antwortkategorien: überhaupt nicht (0), ein wenig (1), mäßig (2), ziemlich (3) und extrem (4) bzw. ‚nicht zutreffend‘ für in letzter Zeit nicht sexuell aktive Frauen bei den Items 12 – 15. Auch hier konnten bei den fünf Antwortkategorien keine klaren Categorieschwellen beobachtet werden. Eine Kombination der Antwortkategorien 1 und 2 und der Antwortkategorien 3 und 4 konnte dieses Problem erneut sowohl für die Kurzversion als auch für die Langversion des DIVA beheben (vgl. Abbildung 14).

Analyse der Antwortkategorien von Skala 4

Die vierte Skala des DIVA zum Selbstkonzept und Körperbild enthielt ursprünglich folgende Antwortkategorien: trifft überhaupt nicht zu (0), trifft etwas zu (1), trifft zum Teil zu (2), trifft weitgehend zu (3) und trifft voll zu (4). Auch bei dieser Skala nahmen die Categorieschwellen nicht monoton mit der Kategorie zu. Bei Item 19 deutete sich dies zumindest an. Mit der gleichen Kombination von nebeneinander liegenden Kategorien wie schon bei den vorherigen Skalen (von 0-1-2-3-4 zu 0-1-1-2-2) wurden klare Kategorien auf dem latenten Merkmal ersichtlich (vgl. Abbildung 15).

Kurzversion



Langversion

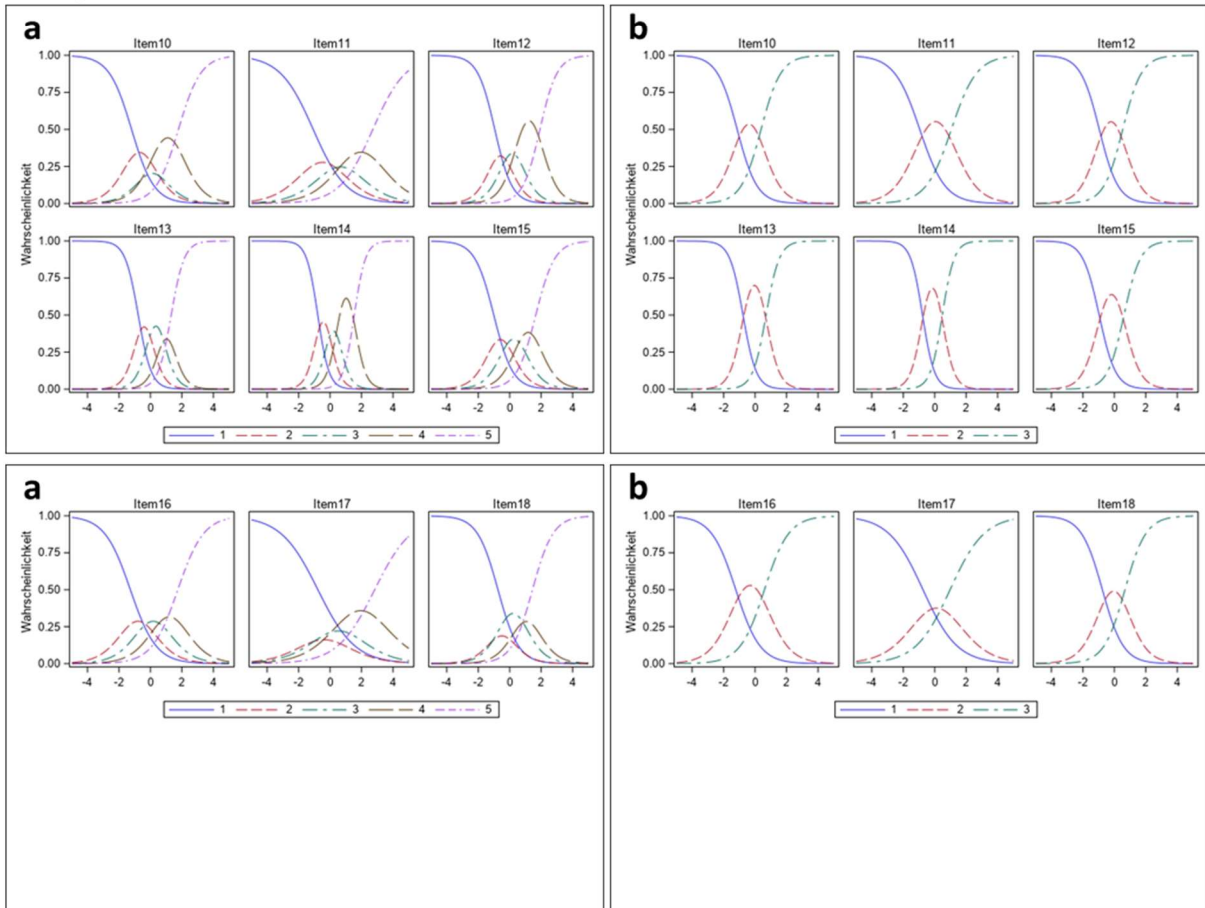


Abbildung 14 Itemfunktionen (*item characteristic curves*) der Items 10, 11 und 16 – 18 (Kurzversion) bzw. der Items 10 – 18 (Langversion) für die originalen Antwortkategorien des DIVA (a) und für Antwortkategorien nach der Adaption (b)

Anmerkungen. 1 = Antwortkategorie 1 (0), 2 = Antwortkategorie 2 (1), 3 = Antwortkategorie 3 (2), 4 = Antwortkategorie 4 (3), 5 = Antwortkategorie 5 (4).

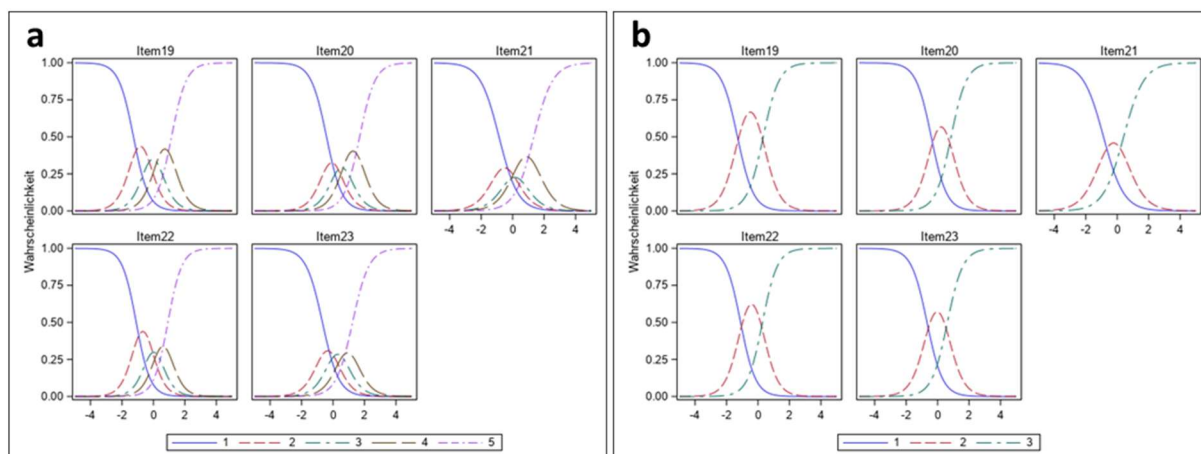


Abbildung 15 Itemfunktionen (*item characteristic curves*) der Items 19 – 23 für die originalen Antwortkategorien des DIVA (a) und für Antwortkategorien nach der Adaption (b)

Anmerkungen. 1 = Antwortkategorie 1 (0), 2 = Antwortkategorie 2 (1), 3 = Antwortkategorie 3 (2), 4 = Antwortkategorie 4 (3), 5 = Antwortkategorie 5 (4).

Zusammenfassung

Durch die Reduktion von fünf auf drei Antwortkategorien konnten die Itemfunktionen in allen Skalen des DIVA dadurch verbessert werden, dass nun jede Antwortkategorie ihren eigenen Höhepunkt und dementsprechend auch ihren eigenen Bereich auf dem latenten Merkmal, dem Ausmaß der Auswirkung vaginaler Beschwerden, hatte. Rechnete man erneut KFAs über die Kurz- und die Langversion des DIVA mit den adaptierten Antwortoptionen, verbesserte sich zudem der Modell-Fit des DIVA (vgl. Tabelle 24).

Tabelle 24 Fit-Maße für die Kurzversion (für alle Frauen, n = 152) und die Langversion (nur für sexuell aktive Frauen, n = 128) des DIVA mit den adaptierten Antwortoptionen zur Bestätigung der vier Skalen

Fit-Maße	Angemessener (noch akzeptabler) Modell-Fit	DIVA Kurzversion	DIVA Langversion
CFI	≥ 0,95 (≥ 0,90)	0,955	0,948
TLI	≥ 0,95 (≥ 0,90)	0,947	0,941
RMSEA	< 0,05 (< 0,08)	0,055 (90 % CI: 0,038 – 0,070)	0,057 (90 % CI: 0,042 – 0,071)
SRMR	< 0,05 (< 0,08)	0,055	0,062
χ^2/df	< 2 (< 3)	1,456	1,413

Abkürzungen. CFI = comparative fit index, TLI = Tucker-Lewis Index, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual, χ^2/df = Chi square/degrees of freedom, CI = Konfidenzintervall.

Durch die adaptierten Antwortoptionen konnte die Annahme der Monotonie bestätigt werden. Somit wurden alle weiteren Analysen mit den adaptierten Antwortoptionen durchgeführt.

3.4.3.2 Überprüfung der Eindimensionalität (*unidimensionality*)

Eindimensionalität bedeutet, dass alle Items einer Skala das gleiche latente Merkmal messen. Um diese Annahme zu überprüfen, rechnet man eine Hauptkomponentenanalyse (*principal component analysis* (PCA)) über die Residuen der Items. Die erste Komponente der PCA sollte nicht mehr als 30 % der Varianz in den Daten erklären und einen Eigenwert ≤ 3 haben (4). Laut COSMIN rechnet man zur Überprüfung von Eindimensionalität eine Faktorenanalyse pro Skala bzw. Subskala. Die Annahme der Eindimensionalität wird nicht verletzt, wenn entweder der CFI oder TLI $< 0,95$ ist oder der RMSEA $< 0,06$ ist oder das SRMR $< 0,08$ ist (52). Alle Analysen zur Eindimensionalität wurden mit Winsteps 4.5.5 und Mplus 8.4 gerechnet.

Bei der PCA wurde bei der ersten Komponente ein Eigenwert von 6,675 für die Kurzversion und ein Eigenwert von 6,878 für die Langversion gefunden. Diese erste Komponente erklärte 18,7 % der Varianz bei der Kurzversion und 18,6 % der Varianz bei der Langversion. Diese Verletzung der Eindimensionalität deutet auf mindestens eine weitere Dimension in beiden Versionen des DIVA hin. Zudem wurde eine KFA über alle Items des DIVA gerechnet, um zu sehen, ob der DIVA als eine Skala die COSMIN-Kriterien für Eindimensionalität erfüllt (vgl. Tabelle 25).

Tabelle 25 Fit-Maße für die Kurzversion (für alle Frauen, n = 152) und die Langversion (nur für sexuell aktive Frauen, n = 128) des DIVA zur Überprüfung der Eindimensionalität des DIVA

Fit-Maße	Angemessener Modell-Fit	DIVA Kurzversion	DIVA Langversion
CFI	> 0,95	0,621	0,467
TLI	> 0,95	0,574	0,414
RMSEA	< 0,06	0,156	0,179
SRMR	< 0,08	0,139	0,197

Abkürzungen. DIVA = Day-to-Day Impact of Vaginal Aging Questionnaire, CFI = comparative fit index, TLI = Tucker-Lewis Index, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual.

Eindimensionalität konnte weder für die Kurz- noch für die Langversion des DIVA bestätigt werden. Beim DIVA handelt es sich somit sehr wahrscheinlich um eine multidimensionale PROM. Aus diesem Grund wurde die Annahme der Eindimensionalität für jede Skala des DIVA separat überprüft.

Für jede Skala des DIVA wurde eine separate PCA über die Residuen der Items gerechnet. Die erste Komponente hatte bei jeder der vier Skalen sowohl für die Kurzversion als auch für die Langversion einen Eigenwert < 2 und erklärte ebenfalls bei jeder Skala weniger als 30 % der Varianz. Zusätzlich wurde eine KFA pro Skala gerechnet, um zu überprüfen, ob jede Skala an sich nur ein latentes Merkmal misst (vgl. Tabelle 26). Jede Skala erfüllte mindestens eines der COSMIN-Kriterien, wodurch Eindimensionalität für jede der vier Skalen des DIVA angenommen werden konnte.

Tabelle 26 Fit-Maße aus den Faktorenanalysen pro Skala zur Überprüfung der Eindimensionalität jeder Skala des DIVA

Fit-Maße	Angemessener Modell-Fit	Skala 1	Skala 2	Skala 3 (kurz)	Skala 3 (lang)	Skala 4
CFI	$> 0,95$	0,993*	0,984*	0,943	0,971*	0,948
TLI	$> 0,95$	0,986*	0,953*	0,887	0,961*	0,897
RMSEA	$< 0,06$	0,044*	0,134	0,144	0,078	0,175
SRMR	$< 0,08$	0,025*	0,019*	0,041*	0,036*	0,033*

Anmerkungen. Skala 1 = Aktivitäten des täglichen Lebens, Skala 2 = Emotionales Wohlbefinden, Skala 3 = Sexuelle Funktionsfähigkeit, Skala 4 = Selbstkonzept und Körperbild. *hinreichend gemäß den COSMIN-Kriterien (52).
Abkürzungen. CFI = comparative fit index, TLI = Tucker-Lewis Index, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual.

3.4.3.3 Überprüfung der lokalen Unabhängigkeit (*local independence*)

Unter lokaler Unabhängigkeit versteht man, dass die Beantwortung eines Items nicht von der Antwort auf ein anderes Item abhängen sollte. Diese Annahme ist erfüllt, wenn die Korrelationen der Item-Residuen $< 0,2$ sind, nachdem für den dominanten Faktor kontrolliert wurde (4, 97). Alle Analysen zur lokalen Unabhängigkeit wurden mit Winsteps 4.5.5 gerechnet.

Die Annahme der lokalen Unabhängigkeit wurde sowohl für die Kurzversion als auch für die Langversion in jeweils 13 Fällen verletzt (standardisierte Residual-Korrelationen bei der Kurzversion: 0,41 – 0,70, bei der Langversion: 0,46 – 0,61). Da der DIVA als eine Skala somit nicht die Annahme der lokalen Unabhängigkeit erfüllte, wurde diese Annahme für jede der vier Skalen separat überprüft.

Betrachtet man jede der vier Skalen des DIVA wieder separat, wurde weder für die Kurzversion noch für die Langversion eine standardisierte Residual-Korrelation $> 0,2$ gefunden. Die Annahme der lokalen Unabhängigkeit wurde somit von keiner der Skalen verletzt.

3.4.3.4 Überprüfung der Passung des Modells (*model fit*)

Von einer adäquaten Passung des Modells spricht man, wenn sich die mittleren Quadrate des Infits und Outfits (*infit/outfit mean squares*) zwischen 0,5 und 1,5 befinden. Diese Statistik gibt Auskunft darüber, ob ein Item beispielsweise redundant ist ($< 0,5$: *overfit*) oder zu viel Messfehler beinhaltet ($> 1,5$: *underfit*) (52, 98). Das Modell als Gesamtes wird als passend angenommen, wenn das mittlere Fit-Residuum (*mean fit residual*) bei 0 mit einer Standardabweichung von 1 liegt (4, 99). Die Passung des Modells wurde mit Winsteps 4.5.5 bestimmt.

Für die Kurzversion und die Langversion zeigten alle Items akzeptable mittlere Quadrate des Infits und Outfits (Kurzversion: 0,74 – 1,35 (Infit), 0,71 – 1,37 (Outfit); Langversion: 0,77 – 1,27 (Infit), 0,74 – 1,35 (Outfit)) bis auf Item 5. Bei der Kurzversion war nur der Outfit von Item 5 zu hoch (1,66), bei der Langversion waren sowohl Infit (1,54) als auch Outfit (1,88) über der festgesetzten Obergrenze. Das mittlere Fit-Residuum lag für beide Versionen bei 0 (mit einer Standardabweichung von 0,60 für die Kurzversion und von 0,67 für die Langversion). Somit wurde die Passung des Modells als akzeptabel angenommen.

Zusätzlich wurde die Passung des Modells für jede der vier Skalen des DIVA separat bestimmt. Die mittleren Quadrate des Infits und Outfits lagen für die erste, zweite und vierte Skala (Skala 1: 0,86 – 1,09 (Infit), 0,79 – 1,09 (Outfit); Skala 2: 0,80 – 1,13 (Infit), 0,60 – 1,20 (Outfit); Skala 4: 0,87 – 1,21 (Infit), 0,85 – 1,24 (Outfit)) im angemessenen Bereich. Auch für die kurze Skala 3 lagen die mittleren Quadrate des Infits und Outfits für alle Items im Rahmen (Infit: 0,85 – 1,36, Outfit: 0,83 – 1,33). Bei der langen Skala 3 zeigten ebenfalls alle Items einen akzeptablen Infit und Outfit (Infit: 0,88 – 1,05, Outfit: 0,69 – 1,01), nur Item 17 lag außerhalb des angesetzten Bereichs (Infit: 1,66, Outfit: 1,60). Das mittlere Fit-Residuum lag bei allen Skalen bei 0, die Standardabweichungen reichten von 0,42 – 0,97. Eine adäquate Passung des gesamten Modells wurde für alle Skalen angenommen.

3.4.3.5 Exploratorische Faktorenanalysen

Die EFAs (mit Varimax-Rotation) brachten sowohl für die Kurzversion als auch für die Langversion des DIVA vier Eigenwerte > 1 hervor (Kaiser-Guttman-Kriterium), was bedeutet, dass aus den Daten vier Faktoren extrahiert werden konnten. Im Scree-Plot, der graphischen Veranschaulichung der Eigenwerte, deutet sich der charakteristische Knick augenscheinlich schon nach dem dritten Faktor an (vgl. Abbildung 16).

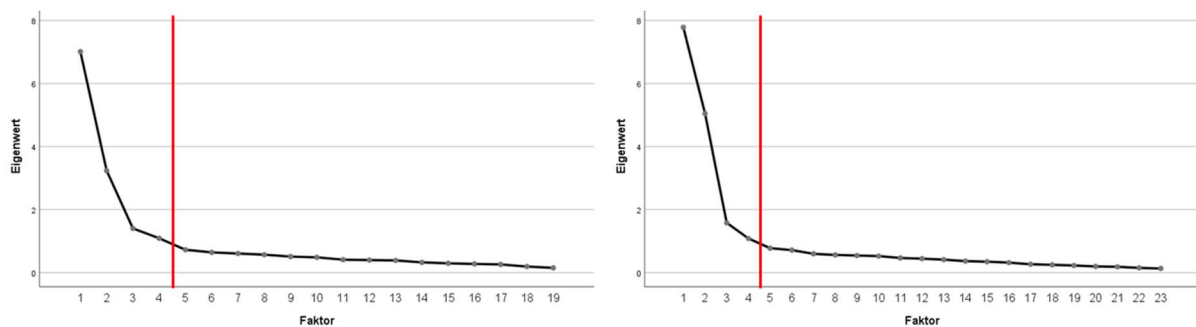


Abbildung 16 Scree-Plots für die Kurzversion (links) und die Langversion (rechts) des DIVA

Alle Items (außer Item 5) luden mit ihrer höchsten Faktorladung auf ihren zugehörigen Faktor (vgl. Tabelle 27). Item 5, Item 4 und Item 1 zeigten zudem signifikante Faktorladungen ($> 0,4$) auf Faktor 2.

3.4.3.6 Analyse der Item-Korrelationen mit den zugehörigen Skalen

Die Korrelationskoeffizienten für jedes Item mit jeder Skala (Kurz- und Langversion) sind in Tabelle 28 dargestellt. Item 3 und Item 11 blieben sowohl bei der Kurzversion als auch bei der Langversion knapp unter der Schwelle von 0,7. Item 17 zeigte nur bei der Langversion eine leicht niedrigere Korrelation als erwartet.

Die EFAs und die Spearman-Rangkorrelationen wurden mit IBM SPSS Statistics 25 gerechnet.

Tabelle 27 Faktorladungen der EFAs für die Kurzversion und Langversion des DIVA

	Kurzversion				Langversion			
	Faktor 1	Faktor 2	Faktor 3	Faktor 4	Faktor 1	Faktor 2	Faktor 3	Faktor 4
Item 1	0,560	0,425	< 0,4	< 0,4	0,610	0,402	< 0,4	< 0,4
Item 2	0,771	< 0,4	< 0,4	< 0,4	0,791	< 0,4	< 0,4	< 0,4
Item 3	0,810	< 0,4	< 0,4	< 0,4	0,822	< 0,4	< 0,4	< 0,4
Item 4	0,588	0,426	< 0,4	< 0,4	0,581	0,413	< 0,4	< 0,4
Item 5	0,438	0,619	< 0,4	< 0,4	0,384	0,639	< 0,4	< 0,4
Item 6	< 0,4	0,774	< 0,4	< 0,4	< 0,4	0,741	< 0,4	< 0,4
Item 7	< 0,4	0,725	< 0,4	< 0,4	< 0,4	0,713	< 0,4	< 0,4
Item 8	< 0,4	0,810	< 0,4	< 0,4	< 0,4	0,802	< 0,4	< 0,4
Item 9	< 0,4	0,798	< 0,4	< 0,4	< 0,4	0,802	< 0,4	< 0,4
Item 10	< 0,4	< 0,4	0,775	< 0,4	< 0,4	< 0,4	0,768	< 0,4
Item 11	< 0,4	< 0,4	0,754	< 0,4	< 0,4	< 0,4	0,755	< 0,4
Item 12	-	-	-	-	< 0,4	< 0,4	0,827	< 0,4
Item 13	-	-	-	-	< 0,4	< 0,4	0,835	< 0,4
Item 14	-	-	-	-	< 0,4	< 0,4	0,858	< 0,4
Item 15	-	-	-	-	< 0,4	< 0,4	0,784	< 0,4
Item 16	< 0,4	< 0,4	0,764	< 0,4	< 0,4	< 0,4	0,731	< 0,4
Item 17	< 0,4	< 0,4	0,706	< 0,4	< 0,4	< 0,4	0,655	< 0,4
Item 18	< 0,4	< 0,4	0,821	< 0,4	< 0,4	< 0,4	0,818	< 0,4
Item 19	< 0,4	< 0,4	< 0,4	0,783	< 0,4	< 0,4	< 0,4	0,775
Item 20	< 0,4	< 0,4	< 0,4	0,730	< 0,4	< 0,4	< 0,4	0,727
Item 21	< 0,4	< 0,4	< 0,4	0,804	< 0,4	< 0,4	< 0,4	0,787
Item 22	< 0,4	< 0,4	< 0,4	0,802	< 0,4	< 0,4	< 0,4	0,796
Item 23	< 0,4	< 0,4	< 0,4	0,744	< 0,4	< 0,4	< 0,4	0,740

Anmerkungen. Grau: signifikante Faktorladung (> 0,4) erwartet, **fett: Hypothese verletzt**, Faktor 1 = Skala 1: Aktivitäten des täglichen Lebens, Faktor 2 = Skala 2: Emotionales Wohlbefinden, Faktor 3 = Skala 3: Sexuelle Funktionsfähigkeit, Faktor 4 = Skala 4: Selbstkonzept und Körperbild.

Tabelle 28 Korrelationen der einzelnen Items mit den Skalen des DIVA für die Kurzversion und die Langversion

	Kurzversion				Langversion			
	Skala 1	Skala 2	Skala 3	Skala 4	Skala 1	Skala 2	Skala 3	Skala 4
Item 1	0,717	< 0,7	< 0,7	< 0,7	0,708	< 0,7	< 0,7	< 0,7
Item 2	0,731	< 0,7	< 0,7	< 0,7	0,717	< 0,7	< 0,7	< 0,7
Item 3	0,682	< 0,7	< 0,7	< 0,7	0,677	< 0,7	< 0,7	< 0,7
Item 4	0,741	< 0,7	< 0,7	< 0,7	0,753	< 0,7	< 0,7	< 0,7
Item 5	0,754	< 0,7	< 0,7	< 0,7	0,763	< 0,7	< 0,7	< 0,7
Item 6	< 0,7	0,880	< 0,7	< 0,7	< 0,7	0,876	< 0,7	< 0,7
Item 7	< 0,7	0,823	< 0,7	< 0,7	< 0,7	0,829	< 0,7	< 0,7
Item 8	< 0,7	0,895	< 0,7	< 0,7	< 0,7	0,898	< 0,7	< 0,7
Item 9	< 0,7	0,826	< 0,7	< 0,7	< 0,7	0,824	< 0,7	< 0,7
Item 10	< 0,7	< 0,7	0,791	< 0,7	< 0,7	< 0,7	0,797	< 0,7
Item 11	< 0,7	< 0,7	0,677	< 0,7	< 0,7	< 0,7	0,689	< 0,7
Item 12	-	-	-	-	< 0,7	< 0,7	0,795	< 0,7
Item 13	-	-	-	-	< 0,7	< 0,7	0,831	< 0,7
Item 14	-	-	-	-	< 0,7	< 0,7	0,849	< 0,7
Item 15	-	-	-	-	< 0,7	< 0,7	0,789	< 0,7
Item 16	< 0,7	< 0,7	0,735	< 0,7	< 0,7	< 0,7	0,730	< 0,7
Item 17	< 0,7	< 0,7	0,717	< 0,7	< 0,7	< 0,7	0,666	< 0,7
Item 18	< 0,7	< 0,7	0,776	< 0,7	< 0,7	< 0,7	0,770	< 0,7
Item 19	< 0,7	< 0,7	< 0,7	0,816	< 0,7	< 0,7	< 0,7	0,806
Item 20	< 0,7	< 0,7	< 0,7	0,830	< 0,7	< 0,7	< 0,7	0,819
Item 21	< 0,7	< 0,7	< 0,7	0,799	< 0,7	< 0,7	< 0,7	0,789
Item 22	< 0,7	< 0,7	< 0,7	0,838	< 0,7	< 0,7	< 0,7	0,851
Item 23	< 0,7	< 0,7	< 0,7	0,835	< 0,7	< 0,7	< 0,7	0,831

Anmerkungen. Grau: signifikante Korrelation (> 0,7) erwartet, **fett: Hypothese verletzt**, Faktor 1 = Skala 1: Aktivitäten des täglichen Lebens, Faktor 2 = Skala 2: Emotionales Wohlbefinden, Faktor 3 = Skala 3: Sexuelle Funktionsfähigkeit, Faktor 4 = Skala 4: Selbstkonzept und Körperbild.

3.4.3.7 Zusammenfassung

Nur Items, die bei beiden Ansätzen (IRT bzw. Rasch-Analyse und KTT) Auffälligkeiten zeigten, wurden nicht in die „verfeinerte“ Version des deutschen DIVA aufgenommen. Die jeweiligen Auffälligkeiten sind in Tabelle 29 zusammengefasst.

Tabelle 29 Auffällige Items getrennt nach IRT bzw. Rasch-Analyse und KTT

	Item	Auffälligkeit
IRT/Rasch-Analyse	5	Mittlere Quadrate des Infits und Outfits > 1,5 für beide Versionen
	17	Mittlere Quadrate des Outfits > 1,5 für die Langversion
KTT	1	EFA: sign. Faktorladungen (> 0,4) auf zwei Faktoren für beide Versionen
	4	EFA: sign. Faktorladungen (> 0,4) auf zwei Faktoren für beide Versionen
	3	Korrelation < 0,7 mit der zugehörigen Skala für beide Versionen
	5	EFA: sign. Faktorladungen (> 0,4) auf zwei Faktoren für die Kurzversion; keine sign. Faktorladung auf den zugehörigen Faktor (< 0,4), jedoch auf einen anderen Faktor für die Langversion
	11	Korrelation < 0,7 mit der zugehörigen Skala für beide Versionen
	17	Korrelation < 0,7 mit der zugehörigen Skala für die Langversion

Abkürzungen. IRT = *item response theory*, KTT = Klassische Testtheorie, EFA = exploratorische Faktorenanalyse.

Item 5 und 17 zeigten Auffälligkeiten in beiden Ansätzen und wurden demzufolge nicht in die „verfeinerten“ Versionen mitaufgenommen. Obwohl Item 17 nur in der Langversion des DIVA Auffälligkeiten zeigte, wurde das Item trotzdem in beiden Versionen eliminiert, um die Anwendung des DIVA nicht zusätzlich zu erschweren. Die „verfeinerte“ Version des DIVA besteht immer noch aus vier Skalen, hat jedoch nur noch drei Antwortoptionen und insgesamt 17 Items für die Kurzversion und 21 Items für die Langversion (vgl. Abbildung 17).

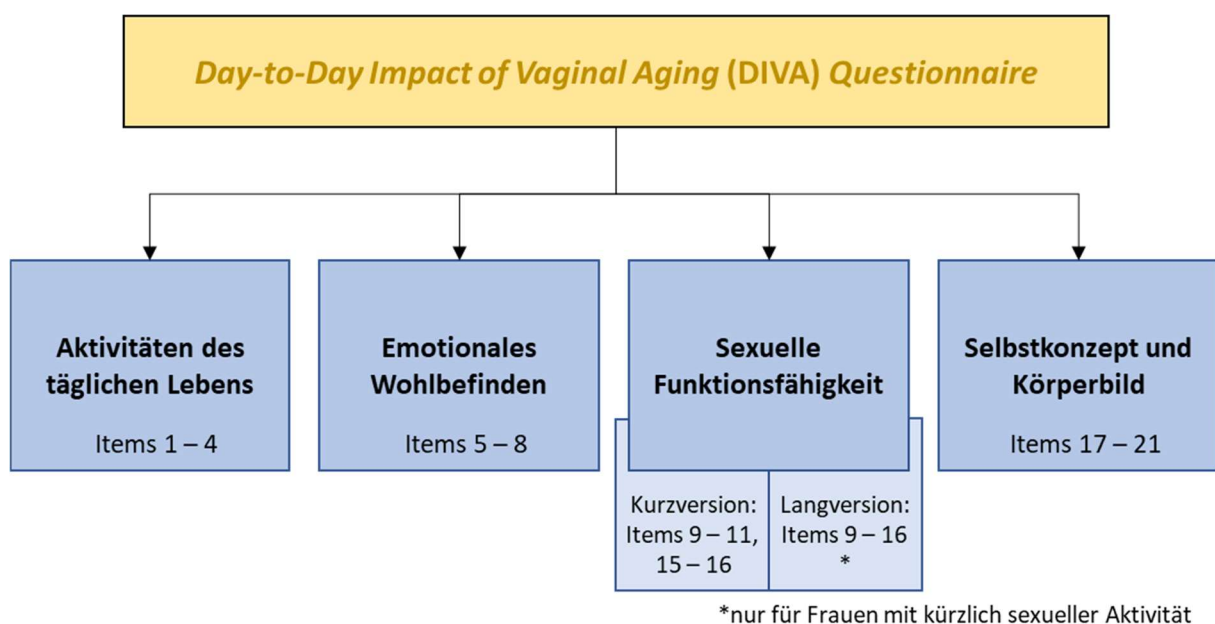


Abbildung 17 Zusammensetzung der „verfeinerten“ Version des deutschen DIVA

3.4.4 Validierung der „verfeinerten“ Version des DIVA

3.4.4.1 Strukturvalidität

Durch den Ausschluss von Item 5 und Item 17 erfüllten nun beide Versionen jedes Kriterium der COSMIN-Taxonomie für eine hinreichende Strukturvalidität (vgl. Tabelle 30).

Tabelle 30 Fit-Maße für die Kurzversion (für alle Frauen, n = 152) und die Langversion (nur für sexuell aktive Frauen, n = 128) der „verfeinerten“ Version des DIVA (ohne Item 5 und 17)

Fit-Maße	Angemessener (noch akzeptabler) Modell-Fit	DIVA Kurzversion	DIVA Langversion
CFI	≥ 0,95 (≥ 0,90)	0,971*	0,959*
TLI	≥ 0,95 (≥ 0,90)	0,966*	0,953
RMSEA	< 0,05 (< 0,08)	0,047* (90 % CI: 0,024 – 0,066)	0,054* (90 % CI: 0,036 – 0,069)
SRMR	< 0,05 (< 0,08)	0,050*	0,058*
χ^2/df	< 2 (< 3)	1,335	1,369

Abkürzungen. DIVA = Day-to-Day Impact of Vaginal Aging Questionnaire, CFI = comparative fit index, TLI = Tucker-Lewis Index, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual, χ^2/df = Chi square/degrees of freedom, CI = Konfidenzintervall. *hinreichend gemäß den COSMIN-Kriterien (52).

3.4.4.2 Interne Konsistenz

Für die erste Skala wurde ein Cronbach's alpha von 0,783 gefunden, für die zweite Skala ein Cronbach's alpha von 0,893, für die Kurzversion der dritten Skala ein Cronbach's alpha von 0,814 und von 0,926 für die Langversion und für die vierte Skala wurde ein Cronbach's alpha von 0,901 berechnet. Die interne Konsistenz aller Skalen war > 0,7 und somit als hinreichend zu bewerten.

3.4.4.3 Konstruktvalidität: Hypothesen-Testen

Bezüglich der konvergenten Validität konnten 14 der 31 *a priori* postulierten Hypothesen bestätigt werden (vgl. Tabelle 31).

Hinsichtlich der diskriminanten Validität wurden, wie erwartet, ebenfalls keine signifikanten Unterschiede zwischen den Alterskategorien in den vier Skalen der „verfeinerten“ Version des DIVA gefunden ($p \geq 0,108$). Auch hier zeigten sich wider Erwarten signifikante Unterschiede zwischen den Heimatländern der Teilnehmerinnen in den Skalen 1 und 2 ($p \leq 0,002$). Im Hinblick auf die Schwere der Erkrankung gelang es fast allen Skalen des DIVA, zwischen verschiedenen Schweregruppen zu unterscheiden. Anhand der ersten Ankerfrage gelang es Skala 3 der Kurzversion nicht, anhand von Frage 10 der MRS gelang es Skala 3 in beiden Versionen nicht. Dies war nicht verwunderlich, da Skala 3 nur eine geringe Korrelation mit der

ersten Ankerfrage in der Kurzversion (vgl. Tabelle 31) sowie ebenfalls eine geringe Korrelation mit Frage 10 der MRS in beiden Versionen ($r \leq 0,185$) aufwies.

Tabelle 31 Korrelationen der „verfeinerten“ Version des DIVA mit anderen PROMs

	Ankerfragen					PHQ			SF-12 SOEP*	
	1	2	3	4	5	Depr.	Ängst.	Gesamt	PCS	MCS
Skala 1	0,389	0,411	0,460	0,285	0,479				-0,413	
n	174	174	174	174	174				149	
Skala 2	0,365	0,537	0,698	0,487	0,727	0,533	0,533	0,579		-0,491
n	176	176	176	175	176	158	158	158		152
Skala 3 (kurz)	0,165	0,195	0,162	0,304	0,306					
n	160	160	160	159	160					
Skala 3 (lang)	0,310	0,236	0,226	0,374	0,337					
n	133	133	133	133	133					
Skala 4	0,447	0,595	0,603	0,673	0,685					-0,408
n	159	159	159	158	159					150

Anmerkungen. Grün: hohe Korrelation erwartet ($> 0,5$), gelb: mittlere Korrelation erwartet ($0,3 - 0,5$), grau: keine Hypothese, **fett: Hypothese bestätigt**, Skala 1 = Aktivitäten des täglichen Lebens, Skala 2 = Emotionales Wohlbefinden, Skala 3 = Sexuelle Funktionsfähigkeit, Skala 4 = Selbstkonzept und Körperbild, n = Stichprobengröße. *Abkürzungen.* MRS = *Menopause Rating Scale*, PHQ = *Patient Health Questionnaire*, Depr. = Depression, Ängst. = Ängstlichkeit, SOEP = Sozio-oekonomisches Panel, PCS = *physical component score*, MCS = *mental component score*.

3.4.4.4 Zusammenfassung

Die „verfeinerte“ Version des deutschen DIVA zeigte eine ausgezeichnete Strukturvalidität, da sogar alle vier COSMIN-Kriterien für eine hinreichende Bewertung dieser Messeigenschaft erfüllt werden konnten (52). Zudem konnte eine angemessene interne Konsistenz bestätigt werden. Die Ergebnisse zu den überprüften Hypothesen hinsichtlich der Konstruktvalidität fielen ähnlich aus wie die Ergebnisse zur Validierung der „originalen“ Version des DIVA. Es kristallisierte sich heraus, dass weder die Ankerfragen noch die Fragen 8, 9 und 10 der MRS hoch mit der dritten Skala des DIVA korrelierten. Das könnte dafür sprechen, dass sich diese Fragen eher weniger als Bezugspunkte für Skala 3 eignen, dass die eigens für Skala 3 kreierte Ankerfrage 4 das Konstrukt der sexuellen Funktionsfähigkeit nur oberflächlich abbildet bzw. hinter Skala 3 noch mehr steckt als nur „sexuelle Funktionsfähigkeit“.

3.5 Diskussion

Mehrere exzellente Messeigenschaften des DIVA konnten auch in seiner deutschen Version repliziert werden: In der Stichprobe der Online- bzw. Papier-Befragung wurde die Vier-Faktoren-Struktur des DIVA bestätigt. Konsistent mit zwei vorherigen Validierungsstudien (69, 100) zeigten alle Skalen des DIVA eine starke interne Konsistenz. Die Ergebnisse zur Konstruktvalidität stimmten größtenteils mit den vorher postulierten Hypothesen überein. Es wurde sogar eine noch stärkere Korrelation der Skala „Emotionales Wohlbefinden“ mit dem PHQ gefunden als in einer kürzlich veröffentlichten Studie zur konvergenten und divergenten Validität des DIVA ($r = 0,39$) (100). Die eigens generierten Ankerfragen korrelierten besser als erwartet mit der ersten, zweiten und vierten Skala des DIVA, jedoch eher schwach bis moderat ($r = 0,225 - 0,417$) mit der dritten Skala zur sexuellen Funktionsfähigkeit. Die Langversion der dritten Skala zeigte tendenziell etwas höhere Korrelationen mit den fünf Ankerfragen zur Symptomschwere als die Kurzversion der dritten Skala, was darauf schließen lässt, dass sexuelle Aktivität die negative Auswirkung vaginaler Symptome verstärkt. Hunter *et al.* (100) stellten bei Frauen mit kürzlich sexueller Aktivität jedoch eher eine schwächere Auswirkung der vaginalen Symptome auf die sexuelle Funktionsfähigkeit fest, die dadurch erklärt wurde, dass Frauen mit vaginalen Schmerzen sexuelle Aktivität eher vermeiden würden. Im Nachhinein wäre eine PROM zur sexuellen Funktionsfähigkeit als Referenzinstrument, wie z. B. der FSFI (101), noch hilfreich gewesen, um das Konstrukt der dritten Skala überprüfen zu können. Eine akzeptable Test-Retest-Reliabilität und Änderungssensitivität konnten aufgrund der geringen Stichprobengröße mit diesem Datensatz nicht bestimmt werden. Diese deuteten sich jedoch deskriptiv in den Daten an. In einer kürzlich publizierten, randomisierten, klinischen Studie, in der die Änderungssensitivität des DIVA überprüft wurde (102), zeigte sich nach zwölf Wochen für diejenigen, die auch eine Verbesserung in der Symptomschwere angaben, für jede Skala des DIVA eine mittlere Verbesserung von mindestens 0,3 bis maximal 1,6 (je nach Stärke der Verbesserung im Anker zur Symptomschwere). Diese Ergebnisse nach zwölf Wochen sind konsistent mit den oben beschriebenen Ergebnissen nach ca. vier Wochen. Im Mittel wurde für die Frauen der Online- bzw. Papier-Befragung, die eine Verbesserung in der allgemeinen Ankerfrage bzw. der zugehörigen Ankerfrage angaben, für die erste Skala eine Verbesserung von 0,317 bzw. 0,200 gefunden, für die zweite Skala eine Verbesserung von 0,396 bzw. 0,472, für die Kurzversion der dritten Skala eine Verbesserung von 0,333 bzw. 0,425, für die Langversion der dritten Skala eine Verbesserung von 0,687 bzw. 0,873 und für die vierte Skala wurde eine Verbesserung von 0,267 bzw. 0,300 beobachtet. Wie in der

randomisierten, klinischen Studie (102) wirkte sich auch hier eine Verbesserung der Symptomschwere vor allem positiv auf die sexuelle Funktionsfähigkeit aus. Diese Verbesserung zeigte sich bei allen Frauen, am stärksten jedoch bei denen, die kürzlich sexuell aktiv waren. Die Sexualität, die nicht nur die tatsächliche sexuelle Aktivität oder Befriedigung, sondern auch das sexuelle Interesse, die sexuelle Erregbarkeit, die sexuelle Spontaneität und die allgemeine sexuelle Zufriedenheit beinhaltet, scheint also am meisten von einer Verbesserung der Symptome zu profitieren.

Durch die kognitiven Interviews während der linguistischen Validierung kann die Inhaltsvalidität des deutschen DIVA als gegeben angesehen werden. Alle oben beschriebenen Validierungsergebnisse sind konsistent mit allen Validierungsstudien, die in der vorher beschriebenen systematischen Übersichtsarbeit identifiziert wurden (50). Die neuen Validierungsergebnisse stärken die Platzierung des DIVA in der COSMIN-Kategorie A, da keine ungenügenden Messeigenschaften gefunden werden konnten. Im Gegenzug zeigten sich eine hinreichende Strukturvalidität und interne Konsistenz.

Eine genaue Untersuchung der Strukturvalidität des DIVA mit der IRT bzw. Rasch-Analyse war gänzlich neu. Es wurde ersichtlich, dass es den Teilnehmerinnen dieser Stichprobe nicht immer gelungen war, sauber zwischen den fünf Antwortkategorien des DIVA zu unterscheiden. Eine Reduktion der Anzahl der Antwortkategorien von fünf auf drei konnte dieses Problem beheben. Trotzdem passte das Rasch-Modell nicht auf den DIVA als eine gesamte Skala, da die Annahmen der Eindimensionalität und der lokalen Unabhängigkeit verletzt wurden. Da alle Items des DIVA das Kriterium der Eindimensionalität weder in der Kurz- noch in der Langversion erfüllten, ist die Berechnung eines Scores für die beiden Gesamtskalen (kurz und lang) nicht zulässig. Das bedeutet, dass die Berechnung von DIVA-Gesamtskalen, wie es explorativ in der vorher beschriebenen Validierung und auch bei Nappi *et al.* (38) gemacht wurde, aufgrund der Multidimensionalität des DIVA nicht erlaubt ist. Jede Skala für sich erfüllte alle Annahmen des Rasch-Modells. Dadurch, dass nach der „Verfeinerung“ gemäß den COSMIN-Richtlinien alle vier Kriterien für eine hinreichende Strukturvalidität erfüllt waren (52), kann auch diese Messeigenschaft bei Verwendung von IRT bzw. des Rasch-Modells bei allen Skalen des DIVA als gegeben angenommen werden. Die zusätzlich adäquate interne Konsistenz für alle Skalen der „verfeinerten“ Version des DIVA bildet gemeinsam mit der hinreichenden Strukturvalidität, die sowohl durch KTT als auch durch IRT bzw. die Rasch-Analyse bestätigt wurde, eines von zwei Kriterien für eine Empfehlung gemäß den COSMIN-Richtlinien (52). Die Inhaltsvalidität der „verfeinerten“ Version sollte in einer qualitativen

Studie noch bestätigt werden, da durch die Eliminierung von zwei Items nicht ausgeschlossen werden kann, dass dadurch nun wichtige Aspekte fehlen (*comprehensiveness*). Erst wenn es Evidenz für eine hinreichende Inhaltsvalidität der „verfeinerten“ Version gibt, würde diese gemäß den COSMIN-Kriterien für den Einsatz in künftigen Studien empfohlen werden (52).

In der explorativen Analyse zum Einfluss von Hormonen zeigte sich, dass sich Frauen, die Hormone zur Behandlung von Wechseljahresbeschwerden einnahmen, weder in der Schwere urogenitaler Symptome noch in deren Auswirkung von Frauen, die keine Hormone einnahmen, unterschieden. Dieser Befund sollte unbedingt in künftigen Studien weiter untersucht werden.

Vor- und Nachteile der Rekrutierungsstrategie

Generell stellte es sich als sehr schwierig dar, eine adäquate Stichprobengröße zu erreichen. Die meisten Teilnehmerinnen konnten tatsächlich über Facebook-Gruppen rekrutiert werden, da man direkt auf den Link klicken konnte, um zur Befragung zu gelangen. Diese „Neugier“ bestätigten die fast 2000 Klicks. Bei der Facebook-Rekrutierung war die größte Hürde, überhaupt in die Gruppen aufgenommen zu werden, da es sich bei den meisten Gruppen um private Gruppen handelte, bei denen man bestimmte Eingangsfragen beantworten musste (Alter, Wechseljahre: ja/nein, Beschwerden etc.). Zudem war bei den meisten Gruppen in den Gruppenregeln aufgelistet, dass Eigen- bzw. Fremdwerbung verboten ist. Da zweifelhafte Methoden (wie beispielsweise ein Fake-Account) nicht in Frage kamen, wurde die Rekrutierung in Facebook durch die eigene Mutter unterstützt. Mit dieser informellen Vorgehensweise konnte tatsächlich eine große Menge an potentiellen Teilnehmerinnen erreicht werden, ohne konkrete Zahlen nennen zu können. Es kam nur selten vor, dass der Post von den Administratoren gelöscht wurde. Der Link wurde auch auf der offiziellen Facebook-Seite der Medizinischen Soziologie der Universität Regensburg gepostet und von deren Followern geteilt. Dieser formale und seriöse Rekrutierungsweg erbrachte aber kaum neue Teilnehmerinnen. Ohne absolute Zahlen nennen zu können, lief die Rekrutierung über die 5000 Flyer, die an unzählige Praxen und Kliniken im Raum Bayern geschickt wurden, eher mäßig, da viele Flyer vermutlich gar nicht ausgelegt wurden, in den Regalen der Wartezimmer keine Beachtung fanden oder der Aufwand einfach zu groß war, den Link extra einzutippen. Der Flyer wurde vielen FreundInnen, Bekannten und Verwandten mitgegeben, die entweder selbst schon in den Wechseljahren waren oder die als Arzt bzw. Ärztin oder ArzthelferIn in diversen Praxen und Kliniken arbeiteten, um ihn an Frauen in mittlerem oder höherem Alter zu verteilen. Als zweites wichtiges Rekrutierungsstandbein sind die Verteiler von Frau Prof. Stute, der leitenden Ärztin und stellvertretenden Chefärztin der Gynäkologischen Endokrinologie und

Reproduktionsmedizin der Universitätsklinik für Frauenheilkunde des Inselspitals Bern, zu nennen. Hierbei bestätigt sich die Vermutung, dass die persönliche Aufforderung durch einen Arzt oder eine Ärztin die Hemmschwelle senkt und doch viele bereit waren, diese Studie zu unterstützen. Im Nachhinein wäre es sehr interessant gewesen, zu erfahren, über welchen Kanal (Facebook, Flyer, Verteiler) die Teilnehmerinnen auf die Befragung aufmerksam geworden waren, um eine Aussage über künftige Rekrutierungsstrategien machen zu können. Nur ca. ein Drittel der Teilnehmerinnen hinterließ die E-Mail-Adresse am Ende der Befragung und stellte sich somit für eine erneute Teilnahme nach einer bzw. vier Wochen zur Verfügung. Von diesem Drittel nahmen allerdings dann wieder nur weniger als die Hälfte tatsächlich erneut an der Befragung teil, wodurch die Stichprobengröße zur Bestimmung der Test-Retest-Reliabilität und der Änderungssensitivität nicht groß genug war, um adäquate Analysen rechnen zu können. Die geringe Bereitschaft für eine erneute Teilnahme könnte auf einen mangelnden Anreiz (z. B. einen Gutschein) zurückzuführen sein. Des Weiteren enthält die E-Mail-Adresse häufig personenbezogene Informationen, die dem anonymen Charakter der Befragung gegenüberstehen. Da es sich ohnehin schon um ein sehr sensibles Thema handelte, könnte dieser Faktor die generell schon hohe Hemmschwelle noch verstärkt haben.

Limitationen der Studie

Die Stichprobe beinhaltete vor allem Frauen mit selbst berichteten vaginalen Symptomen, die laut deren Aussage im Zusammenhang mit der Menopause standen. Weder bei der Online-Stichprobe noch bei den Frauen, die die Papierversion ausfüllten, fand eine klinische Untersuchung der Frauen statt, wodurch weder deren Symptome bestätigt werden konnten noch die Ätiologie der Symptome bestimmt werden konnte. Somit konnten andere Ursachen der Symptome, wie z. B. vaginale Infektionen, Kontaktallergien oder sexuelle Dysfunktionen, nicht ausgeschlossen werden. Zudem kann die Frage zum Zyklus als Kriterium für das Erreichen der Menopause als kritisch und teilweise wenig aussagekräftig betrachtet werden, da viele Frauen parallel eine Hormonersatztherapie machten oder noch hormonell verhüteten, was auch den Zyklus beeinflusst. Außerdem ist aufzuführen, dass die Online-Datenerhebung einen Internetzugang sowie die Kompetenz, ein Smartphone bzw. einen Computer zu bedienen, voraussetzte und somit automatisch diejenigen ausschloss, die dies nicht besaßen.

Bei der Validierung wurden zwar zwei Datenquellen verwendet, jedoch wurden vor allem Frauen mit selbst berichteten Symptomen ohne klinische Diagnose eingeschlossen. Dies kann allerdings auch auf eine gute Repräsentativität der GSM-Population schließen lassen, da GSM-Symptome immer noch wenig berichtet werden und somit unterdiagnostiziert sind (25).

Als statistische Limitation ist aufzuführen, dass die explorativen Analysen mit dem gleichen Datensatz durchgeführt wurden wie die initialen Validierungsanalysen. In den explorativen Analysen wurden KFAs und eine EFA gerechnet. Wenn beide Analysen mit den gleichen Daten gerechnet werden, entsteht eine gewisse Abhängigkeit. In diesem Fall wäre eine Kreuzvalidierung, bei der man die Stichprobe teilt, vielleicht besser gewesen, jedoch hätte sich dadurch auch die Stichprobengröße für beide Hälften deutlich reduziert.

Ein letzter Punkt, der bisher noch gar nicht erwähnt wurde, ist die Tatsache, dass GSM-Symptome auch außerhalb der Menopause auftreten können. Somit kann auch ein Östrogenmangel, der nicht in Zusammenhang mit der Menopause steht, z. B. eine Scheidentrockenheit zur Folge haben. GSM-Symptome können zudem durch eine bilaterale Eierstockentfernung, Rauchen, Alkoholmissbrauch, sexuelle Abstinenz, seltenen Geschlechtsverkehr, Ovarialinsuffizienz, mangelnde Bewegung und keine vaginale Geburt begünstigt werden (24, 25). Für diese Population kann der DIVA nicht angewendet werden, da die Entwicklungs- und Validierungsstudien bisher nur bei symptomatischen, peri- und postmenopausalen Frauen durchgeführt wurden, obwohl die Auswirkung der Symptome natürlich ähnlich sein kann.

Ausblick

Künftige Validierungsstudien zum DIVA sollten sich auf die Bestimmung von sogenannten MID bzw. MIC *values* fokussieren. Bisher gibt es nur eine einzige Studie, die MID bzw. MIC *values* für jede Skala des DIVA nach zwölf Wochen bestimmt hat. Dieser Wert lag für die erste Skala bei -0,4, für die zweite Skala bei -0,8, für die dritte Skala bei -1,3 und für die vierte Skala bei -1,1 (102). Jedoch wäre es durchaus hilfreich, solche Werte schon für einen früheren Zeitpunkt, z. B. nach vier oder sechs Wochen, zu haben. Durch MID bzw. MIC *values* lassen sich Veränderungen über einen bestimmten Zeitraum besser interpretieren, da diese Werte Auskunft darüber geben, ob eine Veränderung in einem Score von einem Patienten oder einer Patientin überhaupt wahrgenommen wird und somit auch klinisch bedeutsam ist bzw. ob diese Veränderung nur auf den Messfehler zurückzuführen ist (103, 104). Würde gemäß den vorher genannten Schwellenwerten der Score für die erste Skala des DIVA nach zwölf Wochen um mindestens 0,4 besser ausfallen, würde die Patientin diese Veränderung als positiv wahrnehmen. In Amerika wurde außerdem vor kurzem in Kollaboration mit der NAMS die Entwicklung einer neuen, vielversprechenden PROM angestoßen. Eine vorläufige Version des *Vulvovaginal Atrophy Questionnaire* (VVAQ) wurde bereits mithilfe von Fokusgruppen mit Frauen mit bestätigtem GSM kreiert, jedoch steht eine umfassende Validierung noch aus (105).

4. Teil 2b: Exploration der urologischen Komponente des GSM mithilfe von Fokusgruppen (online und *face-to-face*)

4.1 Einleitung

Urologische Symptome als Teil des GSM, nämlich vermehrter bzw. verstärkter Harndrang, schmerzhaftes und/oder erschwertes Wasserlassen, Inkontinenz oder vermehrte Harnwegsinfektionen, sind immer noch wenig erforscht und verstanden hinsichtlich des postmenopausalen Östrogenmangels. Man weiß, dass eine Östrogengabe in der Menopause auch urologische Beschwerden verbessert, da sich überall in den Harnwegen Östrogenrezeptoren befinden (106). Die Prävalenz von vermehrtem bzw. verstärktem Harndrang wird auf 10 % bis 20 % bei postmenopausalen Frauen geschätzt (107, 108). Außerdem wird vermehrter bzw. verstärkter Harndrang mit Depression, Verlegenheit und Scham assoziiert, der soziale Aktivitäten, sexuelle Beziehungen, das Reisen, den Schlaf und die berufliche Funktionsfähigkeit einschränken kann (109). Dysurie wird von 7 % bis 13 % der postmenopausalen Frauen berichtet (67), wiederkehrende Harnwegsinfektionen werden von 8 % bis 10 % berichtet (110). Diese beiden Symptome gehen häufig miteinander einher und sind mit Einschränkungen in Bezug auf Aktivitäten, psychischem Stress, Depressionen und mehrfachen Arztbesuchen und Behandlungen assoziiert (111, 112). Von moderater Inkontinenz berichten ca. 15 % der Frauen im mittleren Alter, 10 % der Frauen berichten sogar von schwerer Inkontinenz (113). Obwohl negative Auswirkungen auf die Lebensqualität bekannt sind (114), begeben sich weniger als 40 % der betroffenen Frauen in ärztliche Behandlung (115).

Auch von der VVA bzw. der Menopause per se ist bereits bekannt, dass sich diese negativ auf das Leben von betroffenen Frauen auswirkt. Die VVA zeigte beispielsweise in einer qualitativen Studie negative Auswirkungen der Beschwerden auf alltägliche Aktivitäten, das emotionale Wohlbefinden, die sexuelle Funktionsfähigkeit, das Selbstkonzept und Körperbild und die zwischenmenschliche Beziehung und Kommunikation (65). Die Menopause generell steht z. B. in einem positiven Zusammenhang mit dem Auftreten depressiver Symptome, da mit der Menopause häufig negative Einstellungen gegenüber der Menopause an sich oder dem Älterwerden einhergehen (116).

Die Auswirkungen von Beschwerden lassen sich am besten mit qualitativer Forschung erfassen. Fokusgruppen stellen dabei eine ökonomische Datenerhebungsmethode dar, da mehrere Personen auf einmal erreicht werden können. Durch die Interaktion der TeilnehmerInnen entsteht eine einzigartige Datenquelle, da die TeilnehmerInnen ihre Erfahrungen miteinander

vergleichen können und durch die Ideen der anderen eigene Meinungen und Erfahrungen aktiviert werden können (117).

Die Auswirkungen der VVA bzw. der Menopause per se sind bereits gut erforscht und in der Literatur beschrieben. Auch zu den einzelnen Symptomen der urologischen Komponente gibt es Evidenz, dass diese das Leben der betroffenen Frauen beeinträchtigen. Jedoch wurden die Auswirkungen urologischer Symptome als Teil des GSM, d. h. in Kombination mit vaginalen Beschwerden, bisher kaum bis gar nicht untersucht. Da der DIVA vor der Namensänderung von VVA zu GSM entwickelt wurde, wurde in dessen Entwicklungsstudie die urologische Komponente des GSM auch nicht berücksichtigt (69). Ziel dieser Studie ist es nun, mithilfe von qualitativer Forschung herauszufinden, ob und inwieweit urologische Beschwerden als Teil des GSM peri- und postmenopausale Frauen beeinträchtigen.

4.2 Methode

4.2.1 Rekrutierung

Um ein möglichst breites Feld an betroffenen Frauen abzugreifen, wurde online und im Feld rekrutiert. Neben gewöhnlichen *face-to-face* Fokusgruppen wurde auch versucht, sogenannte Online-Fokusgruppen zu organisieren (118). Ein positives Ethikvotum wurde durch die Ethikkommission der Universität Regensburg erteilt (Aktenzeichen: 19-1314-101).

Die Rekrutierung für die Online-Fokusgruppen erfolgte über diverse Facebook-Gruppen für Frauen in den Wechseljahren. Ein kurzer Survey (über www.soscisurvey.de) wurde gepostet, um die Beschwerden der Frauen und deren Interesse an der Teilnahme an Online-Fokusgruppen abzufragen. Als Einschlusskriterium mussten mindestens ein vaginales (d. h. vaginale Trockenheit, Brennen, Juckreiz oder Schmerzen im Intimbereich) und mindestens ein urologisches Symptom (d. h. vermehrter Harndrang, schmerzhaftes und/oder erschwertes Wasserlassen, Inkontinenz oder vermehrte Harnwegsinfektionen) vorliegen. Frauen, die noch nicht in der Menopause waren, wurden ausgeschlossen. Allen Frauen, die die Einschlusskriterien erfüllten und Interesse an einer Online-Fokusgruppe hatten, wurde eine Teilnehmerinformation per E-Mail mit einem Link zum Online-Portal www.focusgroupit.com zugeschickt (vgl. Anhang H). *FocusGroupIt* ist ein geschütztes Portal, das speziell für die Durchführung von Online-Fokusgruppen erstellt wurde. Der Moderator bzw. die Moderatorin kann seine bzw. ihre Fragen posten und alle Teilnehmenden können ihre Antworten direkt darunter kommentieren. Dabei müssen nicht alle gleichzeitig online sein, sondern es kann

asynchron geantwortet werden. Über eine Nachricht per E-Mail wird man informiert, wenn ein neuer Kommentar gepostet wurde.

Der Online-Survey wurde von insgesamt 85 Frauen ausgefüllt, von denen 78 (91,8 %) angaben, bereits in der Menopause zu sein. Das Durchschnittsalter dieser 78 Frauen lag bei 51,5 Jahren (SD = 4,29; Median = 52; Range: 41 – 63). Vulvovaginale Beschwerden wurden von 40 (51,3 %) Frauen berichtet, urologische Beschwerden von 45 (57,7 %) Frauen. Die Häufigkeiten der berichteten Beschwerden sind in Abbildung 18 dargestellt.

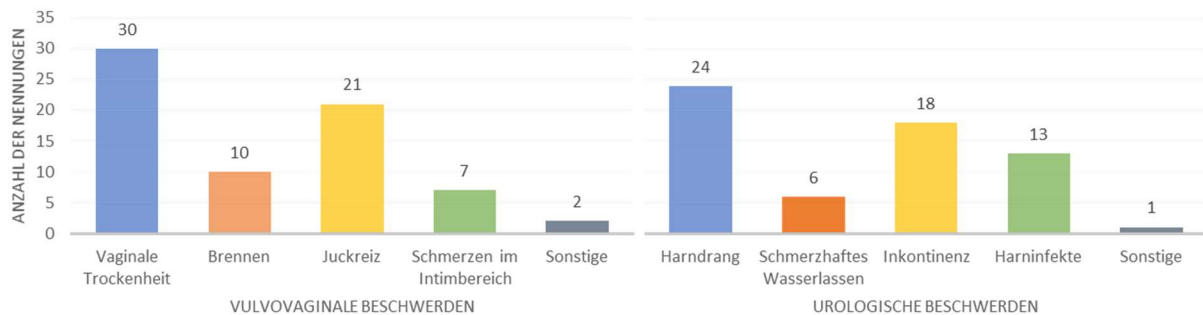


Abbildung 18 Häufigkeit berichteter vulvovaginaler (links, n = 40) und urologischer Beschwerden (rechts, n = 45)

Die Einschlusskriterien, nämlich mindestens ein vulvovaginales und mindestens ein urologisches Symptom, wurden von 25 (32,1 %) Frauen erfüllt. Diese 25 Frauen waren im Durchschnitt 51,6 Jahre alt (SD = 3,57; Median = 52; Range: 44 – 58). Von diesen äußerten elf Frauen Interesse an den Online-Fokusgruppen. Diesen wurde ein Einladungslink zugeschickt. Sechs Frauen im Alter von 46 bis 52 Jahren folgten der Einladung und diskutierten die Fragen im Portal. Der Ablauf der Rekrutierung und die Zusammensetzung der Stichprobe ist in Abbildung 19 dargestellt.

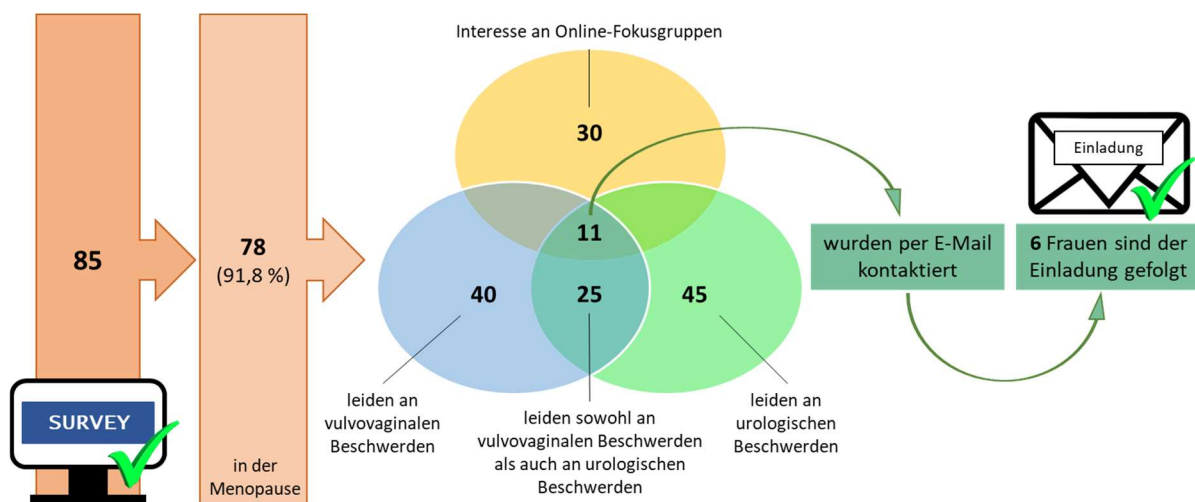


Abbildung 19 Ablauf der Rekrutierung und Zusammensetzung der Stichprobe

Warum die restlichen fünf Frauen, die ein Interesse angegeben hatten und auch die Einschlusskriterien erfüllten, der Einladung trotz mehrerer Erinnerungen nicht gefolgt waren, blieb unklar.

Die Rekrutierung der *face-to-face* Fokusgruppen erfolgte über die Universitätsklinik für Frauenheilkunde des Inselspitals Bern. Frauen mit urogenitalen Beschwerden wurden in der Sprechstunde von Frau Prof. Stute, über deren Verteiler und über das Intranet des Inselspitals Bern mithilfe eines Flyers (vgl. Anhang E, (b)) angesprochen und zur Teilnahme an den Fokusgruppen eingeladen. Es wurde außerdem Kontakt zu Herrn Prof. Ortmann, dem Direktor der Klinik für Frauenheilkunde und Geburtshilfe des Caritas-Krankenhauses St. Josef in Regensburg, aufgenommen. Ein Oberarzt hatte tatsächlich seine Unterstützung zugesichert, doch der Kontakt brach leider, unter anderem bedingt durch die Corona-Pandemie, ab.

Im Rahmen der Corona-Pandemie wurde der Fokus erneut verstärkt auf die Durchführung von Online-Fokusgruppen gelegt. Da insgesamt eine Stichprobengröße ≥ 25 angestrebt wurde, wurden ca. 60 Flyer (vgl. Anhang E, (c)) an die Firma proDERM (Institut für Angewandte Dermatologische Forschung) in Hamburg geschickt, die diese beim Zusatzscreening für eine klinische Studie der Firma Dr. August Wolff GmbH & Co. KG Arzneimittel verteilte. Zusätzlich wurde der Flyer von Frau Prof. Stute an zwei Kolleginnen der Klinik für Frauenheilkunde und Geburtshilfe des Universitätsklinikums Münster und des Universitätsklinikums Frankfurt (Goethe-Universität) zur Verteilung geschickt.

4.2.2 Durchführung der Fokusgruppen

Der Leitfaden für die Durchführung der Fokusgruppen (vgl. Anhang J) wurde in Anlehnung an das konzeptuelle Modell (65), das der Entwicklung des DIVA zugrunde lag, erstellt. Zusätzlich zu den fünf Bereichen, in denen eine negative Auswirkung urologischer Beschwerden vermutet wurde, wurden zwei weitere Themen explorativ in den Leitfaden integriert: die Erwartungshaltung bezüglich der Entwicklung urologischer Beschwerden und die Aufklärung von ärztlicher Seite in Bezug auf urologische Beschwerden (vgl. Abbildung 20).



Abbildung 20 Konzeptuelles Modell (in Anlehnung an die Ergebnisse der Fokusgruppen, die zur Entwicklung des DIVA (65) durchgeführt wurden) mit den zwei zusätzlichen Themenblöcken „Erwartungshaltung“ und „Aufklärung von ärztlicher Seite“

Insgesamt wurde eine Online-Fokusgruppe (n = 6, ID01 – ID06) und eine *face-to-face* Fokusgruppe (n = 4, ID07 – ID10) an der Universitätsklinik für Frauenheilkunde des Inselspitals Bern durchgeführt. Den Teilnehmerinnen der *face-to-face* Fokusgruppe wurde vor der Gruppendiskussion eine Teilnehmerinformation, eine schriftliche Einverständniserklärung und ein kurzer Fragebogen zu demographischen Angaben, den aktuellen Beschwerden und zur Hormoneinnahme gegeben (vgl. Anhang I). Die gesamte Fokusgruppe wurde mit einem Aufnahmegerät aufgezeichnet. Das durchschnittliche Alter dieser Teilnehmerinnen lag bei 62,0 Jahren (SD = 16,19; Median = 58; Range: 47 – 85). Alle Teilnehmerinnen gaben leichte bis mittelgradige vaginale (Trockenheit, Brennen, Juckreiz, Schmerzen) und urologische Beschwerden (vermehrter Harndrang, Inkontinenz) an.

Auf den zusätzlichen Flyer (vgl. Anhang E, (c)) meldete sich nur eine weitere Frau im Alter von 53 Jahren (ID 11), die als urologische Beschwerde vermehrten Harndrang angab und online alle Fragen beantwortete und mit der Moderatorin diskutierte. Eine Übersicht über die eingeschlossene Stichprobe ist in Abbildung 21 dargestellt.

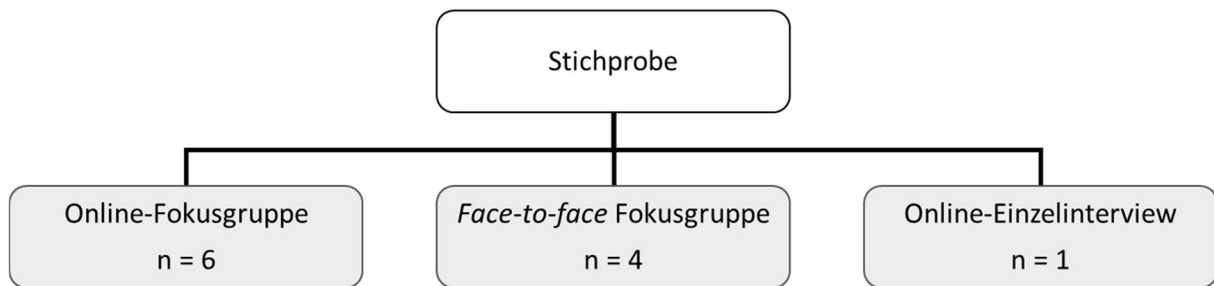


Abbildung 21 Übersicht über die eingeschlossene Stichprobe

4.2.3 Datenanalyse

Dadurch, dass bei der Online-Datenerhebung die Antworten getippt wurden, waren diese Transkripte bereits verfügbar. Die Aufnahme der *face-to-face* Fokusgruppe wurde wörtlich von einem professionellen Transkriptionsbüro (www.abtipper.de) transkribiert. Alle Transkripte wurden anschließend von zwei unabhängigen Reviewern mithilfe der qualitativen Datenanalyse-Software ATLAS.ti Version 8 kodiert. Qualitative Inhaltsanalyse (adaptiert nach Schreier (119)) wurde verwendet, um die Daten zu analysieren. Diskrepanzen zwischen den Reviewern in der Kodierung oder Kategorisierung wurden bis zum Konsensus diskutiert. Anhand des Leitfadens waren bereits deduktiv einige Kategorien vorgegeben, jedoch wurde im Datenmaterial auch induktiv nach neuen Kategorien gesucht. Eine theoretische Sättigung konnte aufgrund der geringen Stichprobengröße nur angenommen, jedoch nicht bestätigt werden (120).

4.3 Ergebnisse

Die fünf Kategorien der Auswirkungen urogenitaler Beschwerden entstanden deduktiv aus dem vorher entworfenen Leitfaden (vgl. die blauen Kästchen in Abbildung 22). Der Themenblock „Alltägliche Aktivitäten“ wurde Teil der etwas allgemeineren Kategorie „Alltag“. Aus den Themenblöcken „Erwartungshaltung“ und „Aufklärung von ärztlicher Seite“ entstand induktiv die Kategorie „Unerfüllte Bedürfnisse im Gesundheitswesen“. Völlig induktiv wurde der „Umgang mit den Beschwerden“ in den Daten gefunden, der ebenfalls als Kategorie mitaufgenommen wurde (vgl. die grünen Kästchen in Abbildung 22). Die jeweiligen Unterkategorien jeder Kategorie (vgl. die weißen Kästchen in Abbildung 22) beinhalten die während der Kodierung vergebenen Codes.

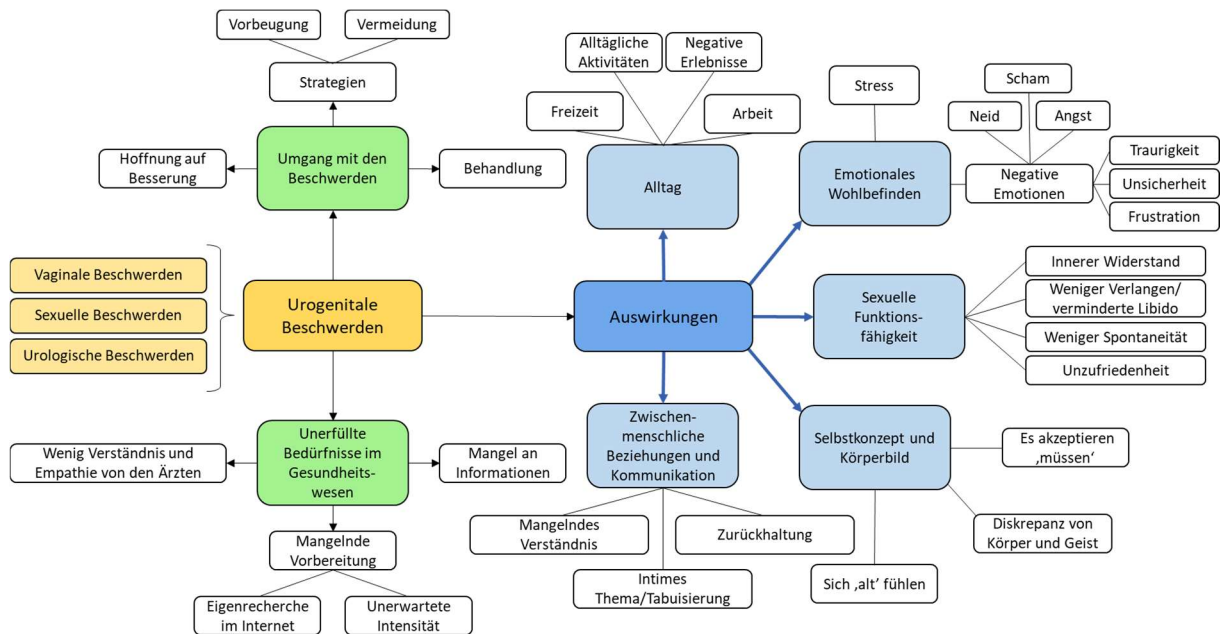


Abbildung 22 Themen-Cluster

Anmerkungen. Gelb: berichtete Beschwerden, grün: induktiv entstandene Kategorien, blau: deduktiv gefundene Kategorien, weiß: Unterkategorien.

4.3.1 Alltag

Der Alltag war von allen Teilnehmerinnen beeinträchtigt. Nicht nur alltägliche Aktivitäten, wie z. B. das Einkaufen, waren von urologischen Beschwerden betroffen, sondern auch ganz basale Verhaltensweise, wie z. B. das Lachen: „*Es gibt und gab schon Situationen, da musste ich mich zwingen, mit dem Lachen aufzuhören. Ich habe gemerkt, dass ich den Urin sonst nicht halten kann... und das wäre viel mehr als ein Tröpfchen gewesen. Nicht auszudenken! Ich lache gerne und finde das ganz schön blöd.*“ (ID06, 49 J.) Durch die Inkontinenz beispielsweise mussten besondere Vorbereitungen getroffen werden, wenn die Frauen das Haus für eine längere Zeit verließen. Es wurde als sehr mühsam empfunden, „*wenn man eben z. B. eine längere Autofahrt macht [...]. Da muss man viel anhalten und auf die Toilette gehen. Oder wenn ich einkaufen gehe in die Stadt, da muss ich immer schauen, wo gibt es eine Toilette, dass ich Wasserlassen kann.*“ (ID07, 58 J.) Auch andere Aktivitäten, wie z. B. ins Schwimmbad gehen, wurden vermieden, um Blasenentzündungen vorzubeugen. Urologische Beschwerden trugen auch dazu bei, dass betroffene Frauen ihren Hobbys, wie z. B. Wandern, Reisen oder ins Kino gehen, nicht mehr wie gewohnt nachgehen konnten. Eine Dame berichtete: „*Vor einer Woche war ich im Kino und da habe ich extra vorher nicht viel getrunken, weil die machen keine Pause im Kino. Und ich spürte schon gegen Ende [...] den Druck [...] und plötzlich ist das [der Urin] abgegangen. [...] Also ich hatte die ganze Hose nass.*“ (ID09, 58 J.) Da auch viele Frauen in die Wechseljahre kommen, bevor sie aufhören zu arbeiten, wirkten sich die Beschwerden auch

negativ auf deren Arbeit aus. Sie fühlten sich unwohl, wenn sie auf der Arbeit ständig zur Toilette gehen mussten und sie konnten bzw. wollten auch nicht einfach nach Hause gehen, wenn ihre Beschwerden wieder schlimmer waren. Vor allem im Umgang mit der Inkontinenz berichteten einige Frauen, auf Slipeinlagen oder Windeln zurückzugreifen. Um vor allem Blasenentzündungen zu vermeiden, versuchten einige Teilnehmerinnen ausreichend viel zu trinken. Das viele Trinken beugte zwar auf der einen Seite Infektionen vor, führte aber auch dazu, dass die Frauen noch häufiger zur Toilette mussten. Eine Teilnehmerin berichtete, immer eine Ersatzunterhose in ihre Tasche dabei zu haben.

4.3.2 Emotionales Wohlbefinden

Nachdem die Teilnehmerinnen gefragt wurden, welche Gefühle ihre Beschwerden in ihnen auslösen, wurden vor allem Ängste genannt, was noch auf sie zukommen würde. Ängste wären permanent da sowie damit verbundene Sorgen und Unsicherheit. Weitere Emotionen, die genannt wurden, waren Scham, Neid, Frustration und Traurigkeit. Vor allem Traurigkeit wurde in Bezug auf die sexuelle Funktionsfähigkeit erwähnt, da alles nicht mehr so funktionierte bzw. lief wie früher. Scham wurde oft mit negativen Situationen im Alltag in Verbindung gebracht. Vor allem durch die Inkontinenz wurden peinliche Erfahrungen gemacht: *„Letztens habe ich ins Bett gemacht! Ich habe geträumt, dass ich auf die Toilette gehe und habe losgelassen. Mann, das hat mich tagelang verfolgt!“ (ID05, 50 J.)* Durch die ständigen Gedanken an die Beschwerden fühlten sich einige Teilnehmerinnen mental gestresst. Sie ständen unter einer ständigen Nervosität und die Beschwerden erschienen auf Dauer frustrierend: *„Ich finde es sehr belastend, wenn ständig ein Druckgefühl besteht, ganz egal, ob ich wenig oder viel Wasserlassen kann. Der Druck macht übernervös. Auch sehr unangenehm ist, wenn plötzlich ganz starker Harndrang kommt und dann auch oft etwas Urin abgeht.“ (ID01, 52 J.)* Auch von Neid auf andere Frauen, die keine Beschwerden haben, wurde berichtet: *„Zugeben muss ich aber auch, dass ich auf die Damen, die keine oder kaum Beschwerden haben, schon ein wenig neidisch bin.“ (ID06, 49 J.)*

4.3.3 Sexuelle Funktionsfähigkeit

Negative Auswirkungen der Beschwerden auf die sexuelle Funktionsfähigkeit und das Sexualleben im Allgemeinen wurden vor allem von Frauen, die sich in Partnerschaften befanden, berichtet. Eine Single-Dame im Gegenzug empfand z. B. den Verlust der Libido als Entlastung, da sie sich dadurch nicht mehr selbst befriedigen musste. Bei vergebenen Frauen litt die Intimität in der Partnerschaft unter den Beschwerden: *„Ich denke, wenn ich alleine wäre, würde mich das weniger stressen [...]. Weil ich würde denken, ja das ist ja mein Bedürfnis und*

wenn es nicht vorhanden ist, ist es halt nicht vorhanden. Aber es ist ja nicht nur mein Bedürfnis, es ist auch ein Bedürfnis meines Ehemannes. Ein Bedürfnis, das entsteht, wenn man zusammen ist.“ (ID10, 47 J.) Aufgrund von urogenitalen Symptomen wie Inkontinenz, Trockenheit, Brennen, Libidoverlust und der Angst vor Infektionen wurde der Geschlechtsverkehr häufig nicht mehr als angenehm und befriedigend erlebt. Es wurden eher peinliche Erlebnisse und Sorgen damit assoziiert: „Seit ein paar Monaten ist meine Libido ziemlich auf dem Nullpunkt, was ich von mir nicht kenne. Das Gefühl beim Geschlechtsverkehr ist ein ganz anderes. Es fühlt sich teilweise wund an und es ist alles sehr empfindlich. [...] Meine Blase muss vorher entleert werden, sonst kommt mehr raus als ein Tröpfchen, was auch (wegen dem Geruch) sehr unangenehm ist. Ich muss zugeben, dass ich sogar einmal "danach" echt geweint habe, weil das Gefühl ein ganz anderes ist als früher. Eher unangenehm als schön. Die Orgasmus-Sicherheit ist nicht mehr da. Oft passiert da bei mir gar nichts. Das ist richtig schlimm für mich. Und die Angst vor einer abschließenden Blasenentzündung begleitet mich immer. Das ist psychisch echt stressig.“ (ID06, 49 J.) Das Sexualleben wäre durch die Beschwerden komplizierter und weniger spontan geworden. Vor allem das Zeigen von Verständnis des Partners bzw. der Partnerin wurde von den Teilnehmerinnen sehr wertgeschätzt und trug auch einen großen Teil dazu bei, dass die Intimität aufrechterhalten wurde: „Er ist auch sehr rücksichtsvoll. Aber mich stresst es mehr, weil ich finde, dass man ja gerne zusammen ist, aber dass es einfach schwieriger ist als früher. Und darüber bin ich schon traurig, weil es halt nicht mehr so spontan ist.“ (ID07, 58 J.) Nichtsdestotrotz führten die Beschwerden zu einer gewissen Zurückhaltung bis hin zu einer Überwindung zu sexueller Aktivität: „Ich muss mich auch manchmal überwinden. Das finde ich auch wichtig, dass ich nicht nur sage, ja ich habe halt keine Lust mehr [...]. Weil das stimmt nicht ganz, weil eigentlich im Herzen habe ich Lust, [...] meinem Mann zu lieben. Aber irgendwie kommt der Körper nicht mehr mit. Das ist das Schwierige.“ (ID10, 47 J.) Auch in Bezug auf die Sexualität wurde von Strategien berichtet, die z. B. Infektionen vorbeugen (nach dem Sex duschen/auf die Toilette gehen) oder ungewollten Urinabgang vermeiden (vor dem Sex auf die Toilette gehen).

4.3.4 Selbstkonzept und Körperbild

Durch die Beschwerden hätte sich bei den Teilnehmerinnen auch deren Selbstkonzept und Körperbild verändert. Eine Dame beschrieb die Symptome, die die Menopause begleiten „als Alterserscheinung, die schwer zu beeinflussen“ (ID11, 53 J.) ist. Man habe keine Wahl und müsse die Beschwerden akzeptieren: „[Ich] finde mich langsam damit ab, dass es so ist, wie es ist und ich da durch muss.“ (ID06, 49 J.) Die Beschwerden, die Einschränkungen, die mit den

Beschwerden einhergehen, aber auch die wahrgenommenen Veränderungen des Körpers trugen dazu bei, dass sich die Frauen ‚alt‘ fühlten. Vor allem die Inkontinenz sorgte häufig für einen Uringeruch, der bei den Frauen Unbehagen und Scham hervorrief und ihr Gefühl des Altseins bestätigte. Die meisten Teilnehmerinnen wurden von ihren Beschwerden überrascht und hatten demzufolge keine andere Wahl, als diese zu akzeptieren: *„Der Körper verändert sich und es ist auch normal, dass man es vermisst, dass man nicht mehr so ist wie vorher, als man sich wirklich jung und gut und wohl fühlte.“* (ID07, 58 J.) Vor allem Frauen in einer Partnerschaft berichteten von einer Divergenz zwischen Körper und Geist. Obwohl sie gerne noch in ihrer Partnerschaft intim werden wollten, ließ dies ihr Körper nicht mehr wie früher zu. Eine Teilnehmerin erzählte, dass sie kein inneres Gleichgewicht mehr habe: *„Ich finde, das ist das, was mich am traurigsten gemacht hat. Zu merken mein Herz möchte, mein Körper kommt nicht mehr mit. [...] Weil mein Herz ist voll Liebe eigentlich für meinen Ehemann und mein Körper nicht mehr, [...] diese Balance ist nicht mehr vorhanden. Und da irgendwo einen Weg finden [...]. Das ist schon eine Herausforderung für mich.“* (ID10, 47 J.)

4.3.5 Zwischenmenschliche Beziehungen und Kommunikation

Urogenitale Beschwerden stellten immer noch ein sehr intimes Thema dar, das häufig tabuisiert wird. Demzufolge erschien es wenig überraschend, dass betroffene Frauen häufig Hemmungen hatten, über dieses Thema zu sprechen, auch weil es emotional beladen ist: *„Es [ist] so intim [...] und privat. Weil [...] es auch etwas vom Herz betrifft, also das ist ja auch ein Schmerz irgendwie. Und ich finde es nicht einfach über einen Schmerz zu sprechen. Also es ist ja nicht nur [die] Vagina, dass es schmerzt, es schmerzt ja auch im Herzen.“* (ID10, 47 J.) In einer vertrauensvollen Partnerschaft, in der Familie oder mit engen FreundInnen fiel es einigen Teilnehmerinnen weniger schwer, über die Beschwerden zu sprechen: *„Ich spreche mit meinem Partner darüber. Die Beschwerden sind ja auch nicht zu übersehen. Er ist sehr verständnisvoll, was total guttut. Mit meinen engen Freundinnen bespreche ich das Thema auch. Sie leiden selbst unter Blasenentzündungen und können mich gut verstehen.“* (ID06, 49 J.) Verständnis von anderen wurde generell als sehr positiv und entlastend erlebt. Andere Damen hingegen berichteten von negativen Erfahrungen, wenn sie versuchten, über ihre Beschwerden zu sprechen: *„Ich lebe alleine und habe beschlossen, meine Beschwerden in Zukunft für mich zu behalten. Egal bei wem, Nachbarschaft, Verwandte, Bekannte, Ärzte ... überall nur Unverständnis! [...] Schlimm finde ich auch, dass man in der Realität nicht mal von Frauen besonderes Verständnis erhält, die sich auch im Wechsel befinden!“* (ID01, 52 J.) Für den Austausch wurde auch von Foren im Internet berichtet, an denen vor allem eine Frau deren

Anonymität schätzte. Sie hätte dadurch über ihre Beschwerden schreiben können und Hilfe bekommen: „*Ich wusste gar nichts über [die] Wechseljahre, bin durch die Internetsuche darauf gestoßen, weil ich wegen plötzlich schlimmen, mir unerklärbaren Beschwerden, dort Hilfe suchte.*“ (ID01, 52 J.) Alle Damen empfanden den Austausch im Rahmen dieser Studie als sehr angenehm. Eine Teilnehmerin der *face-to-face* Fokusgruppe meinte dazu: „*Also ich finde es spannend, sich jetzt in dieser Gruppe auszutauschen. Ich denke, das wäre manchmal ein Gewinn. [...] Ich finde es schade, dass man das zu wenig macht.*“ (ID10, 47 J.)

4.3.6 Unerfüllte Bedürfnisse im Gesundheitswesen

So gut wie alle Frauen berichteten, dass sie nicht erwartet hätten, zum einen so schlimme, zum anderen so früh Wechseljahresbeschwerden zu entwickeln: „*Also mich hat das Thema Wechseljahre völlig übermannt. Sicher wusste ich darüber Bescheid, aber nicht, dass es mich so treffen könnte. Dieses Ausmaß an Beschwerden der Wechseljahre, darauf war ich nicht vorbereitet.*“ (ID02, 48 J.) Eine andere Teilnehmerin sagte: „*Ich habe diese Beschwerden nicht so früh erwartet. Dass Harninkontinenz im Alter wahrscheinlich ist, wusste ich. Ich habe nicht darüber nachgedacht, dass ich diese Beschwerden bereits mit Mitte/Ende 40 bekommen könnte.*“ (ID06, 49 J.) Als Grund für die eigene Unwissenheit wurde zum einen angeführt, dass von ärztlicher Seite mit der Zielgruppe wenig über diese Beschwerden gesprochen wird. Zum anderen wird dieses Problem dadurch verschlimmert, dass von den behandelnden ÄrztInnen wenig Verständnis und Empathie entgegengebracht wird: „*Da wurde ich auch schlecht aufgeklärt, ich hatte keine Ahnung. Also ich hatte ganz, ganz starke Blutungen und hatte einen Arzt, der hat mir irgendwie Tabletten gegeben. Aber [er hat mir] nicht [erklärt], dass [es] einen Zusammenhang haben könnte, dass die Menopause kurz [...] bevorsteht.*“ (ID09, 58 J.) Des Weiteren wurde berichtet, dass die Frauen häufig Eigeninitiative ergreifen mussten, um ihre Beschwerden anzusprechen, was oftmals eine große Überwindung erforderte. Wenn sich die Frauen dazu entschieden hatten, Hilfe zu suchen, wurde die Erfahrung gemacht, dass ihre Beschwerden kaum besprochen wurden, die Frauen zum Urologen bzw. zur Urologin weitergeschickt wurden oder deren Beschwerden sogar von ihrem Gynäkologen bzw. ihrer Gynäkologin abgetan wurden: „*Eben als ich mit der Ärztin gesprochen habe, hat es mich viel Mut gekostet, weil ich finde, das ist etwas Intimes, etwas Persönliches. Und dann kommt so ein blöder Spruch, ja wissen Sie Frau [...]. Und so will sie [es] abtun und ja, vielleicht sind Sie ein bisschen gestresst und so, ja klar stresst es, wenn es [die Libido] plötzlich verschwunden ist von einem Tag auf den anderen.*“ (ID10, 47 J.) Auf die Frage, was die Teilnehmerinnen von ihren ÄrztInnen erwartet hätten, hatten viele eine klare Vorstellung: „*Jetzt, wo ich mich mitten*

in den Wechseljahren befinde, hätte ich mir ein Info-Gespräch von meinem Gynäkologen über die Wechseljahre gewünscht. Ich weiß, dass "nur" 1/3 der Frauen starke Beschwerden haben, 1/3 leichte und 1/3 gar keine. Trotzdem finde ich es wichtig, dass den Frauen, so Ende 30, eine kurze Info darüber gegeben wird, was eventuell auf sie zukommen könnte. Mich hat es mit fast allen Beschwerden voll erwischt. Einige davon hätte ich nicht den Wechseljahren zugeordnet, wenn ich nicht im Internet gelesen hätte.“ (ID06, 49 J.) Jedoch bestand nicht nur der Wunsch nach einer informativen Beratung, sondern auch nach einem „echten Interesse“ an ihrer Person: *„Ich finde eine kompetente Betreuung, die den ganzen Menschen anschaut und manchmal auch das psychische anspricht, wie geht es Ihnen so, finde ich extrem wichtig.“ (ID10, 47 J.)*

4.3.7 Umgang mit den Beschwerden

Jede Frau berichtete von ihren eigenen Strategien, um mit den Beschwerden umzugehen. Einige der Strategien wurden bereits in den Sektionen „Alltag“ und „sexuelle Funktionsfähigkeit“ aufgegriffen. Im Großen und Ganzen ließen sich die Strategien in Präventions- und Vermeidungsstrategien unterteilen. Zu den Präventionsstrategien konnte z. B. das viele Trinken oder das Nutzen von Slipeinlagen und Windeln gezählt werden, zu den Vermeidungsstrategien z. B. das Vermeiden langer Autofahrten oder spontaner sexueller Aktivität. Beide Strategien wurden jedoch als Einschränkung der Lebensqualität von den betroffenen Frauen wahrgenommen. Neben diesen Strategien wurde von zwei weiteren Aspekten berichtet, die auch zum Umgang mit den Beschwerden gehören. Darunter fiel zum einen die medikamentöse Behandlung, vor allem die Hormontherapie. Besonders das Finden der richtigen Kombination und Dosis wurde zu Beginn der Behandlung als stressig erlebt. Als zweiter Aspekt tauchte die innere Einstellung gegenüber der aktuellen Situation in den Diskussionen auf. Neben negativen Emotionen, die bereits weiter vorne beschrieben wurden, wurde dennoch auch versucht, positiv zu bleiben und dankbar zu sein: *„Ich gebe die Hoffnung nicht auf, dass es mir eines Tages wieder annähernd so geht, wie vor der Menopause.“ (ID06, 49 J.)* Eine andere Dame ergänzte: *„Ich finde auch, die eigene Einstellung ist wichtig. Also für mich ist [es] zum Beispiel jetzt die Dankbarkeit [für Medikation].“ (ID08, 85 J.)* All diese Aspekte, wie Strategien, Medikamente und eine positive Einstellung, gaben den Frauen zu einem gewissen Grad wieder ein Gefühl von Kontrolle über ihren Körper und über ihr Leben. Eine Frau kam zu dem Schluss: *„Solange ich eine Toilette habe irgendwo und mein Wasser [...], geht es mir gut.“ (ID09, 58 J.)*

4.4 Diskussion

In dieser qualitativen Studie zeigte sich, dass urologische Beschwerden als Teil des GSM Auswirkungen auf den Alltag, das emotionale Wohlbefinden, die sexuelle Funktionsfähigkeit, das Selbstkonzept und Körperbild und auf die zwischenmenschliche Beziehung und Kommunikation haben. So gut wie alle Teilnehmerinnen hätten nicht erwartet, derart intensive Beschwerden zu bekommen: „*Nein, ich habe es nicht erwartet, dass es so in einem solchen Maß kommt und so früh.*“ (ID10, 47 J.) Im Umgang mit den Beschwerden hatte jede Teilnehmerin ihren eigenen Weg gefunden, der von praktischen Strategien über eine medikamentöse Behandlung bis hin zu einfacher Hoffnung auf Besserung reichte. Vor allem im Hinblick auf das Gesundheitssystem wurde von vielen Hindernissen berichtet, wie einer mangelnden Aufklärung über die Menopause, Unverständnis und wenig Empathie von ärztlicher Seite und einem hohen Maß an Eigeninitiative, um überhaupt in den ärztlichen Austausch über die Beschwerden zu kommen. In einer kürzlich publizierten, australischen Interviewstudie zeigte sich, dass Frauen vor allem wenig über die langfristigen Konsequenzen der Menopause wissen (121). Huang *et al.* (65) postulierten in deren qualitativen Studie, dass viele Frauen zögern, sich Hilfe für ihre Beschwerden zu suchen und Ärzte bzw. Ärztinnen deshalb aktiv nach Symptomen fragen sollten, was zeigt, dass sich im letzten Jahrzehnt nicht viel im Umgang mit urogenitalen Beschwerden verändert hat. In dieser Studie wurde außerdem berichtet, dass sich vaginale Beschwerden negativ auf das tägliche Leben auswirkten, z. B. erlebten betroffene Frauen Einschränkungen bei der Wahl ihrer Kleidung. Zudem beeinflusste die Symptomatik deren Stimmung auf eine negative Art und Weise, sie fühlten sich in ihrer Sexualität eingeschränkt und beschrieben einen Verlust ihrer „Weiblichkeit“ (65).

Urologische Beschwerden als Teil des GSM haben also ähnliche negative Auswirkungen auf das Leben von betroffenen Frauen, wie es bereits von der VVA (36-38, 65) und von jedem urologischen Symptom einzeln bekannt ist (109, 111, 112, 114). In der Literatur gibt es außerdem Hinweise darauf, dass urologische Beschwerden in Kombination mit vaginalen Beschwerden einen kumulativen Effekt haben und die Lebensqualität betroffener Frauen weiter beeinträchtigen (39). Umso wichtiger ist es auch, dass die Auswirkungen des GSM in seiner Gänze weiter untersucht werden, um darauf aufbauend entweder den DIVA zu erweitern oder eine ganz neue PROM zu entwickeln, die die Auswirkungen des GSM vollständig abbildet.

In dieser qualitativen Studie wurde nicht nur über urologischen Beschwerden und deren Auswirkungen gesprochen, sondern auch über vaginale Beschwerden und die Menopause bzw. die Wechseljahre *per se*. Hinsichtlich der Häufigkeiten vaginaler Symptome konnte anhand der

quantitativen Screening-Befragung bestätigt werden, dass vaginale Trockenheit das meistberichtete Symptom postmenopausaler Frauen ist (75). Von allen Teilnehmerinnen wurde die Menopause als eine leidvolle Veränderung wahrgenommen, die mit verschiedenen Beschwerden einhergeht. Schon die französische Schriftstellerin und Feministin Simone de Beauvoir beschrieb den Eintritt in die Menopause als „schlagartigen Verlust der Weiblichkeit“. Ihre pessimistischen Einschätzungen gingen sogar noch weiter: *„Solche wie Pubertät, erste geschlechtliche Erfahrungen, Klimakterium geben sich durch viel entscheidendere Krisen zu erkennen als beim Mann. Während dieser kontinuierlich altert, wird der Frau die Weiblichkeit schlagartig genommen. Noch verhältnismäßig jung verliert sie den erotischen Anreiz und die Fruchtbarkeit, aus denen sie in den Augen der Gesellschaft und in ihren eigenen Augen die Rechtfertigung ihrer Existenz und ihre Glücksmöglichkeiten ableitete: Ihrer ganzen Zukunft beraubt, hat sie etwa die Hälfte ihres Lebens als Erwachsene vor sich ...“* In ihren Augen erfolgt jede Umbruchsphase im Leben einer Frau „mit gefährlicher Brutalität“ (122, 123). Nicht ganz so drastisch wie die Aussagen von Simone de Beauvoir, jedoch durchaus auch als heftig wurde der Eintritt in die Menopause von einer Fokusgruppen-Teilnehmerin beschrieben: *„Für mich sind die Wechseljahre ein Albtraum, ich bin regelrecht davon geschockt, wie es so eine leidvolle (natürliche) Veränderung überhaupt geben kann!!! [...] Ich bin immer von Ängsten umgeben, irgendwie wird der Körper zum Feind der Seele. Man kann nur schwer beschreiben, wie man sich fühlt, wenn durch das Hormon-Chaos im ganzen Körper die Steuerung fehlgeleitet ist!! Es ist, wie wenn bei einem Auto sich Gas-Pedal und Bremse verselbstständigen und es immerzu hin und her wechselt zwischen Stopp und Höchstgeschwindigkeit...“* (ID01, 52 J.)

Die Rekrutierung der Zielpopulation, nämlich postmenopausale Frauen mit VVA und zusätzlichen urologischen Beschwerden, erwies sich als sehr schwierig, wie bereits im Methodenteil beschrieben. Die Rekrutierungsschwierigkeiten stützen den Verdacht, dass es sich bei vaginalen und urologischen Beschwerden immer noch um ein Tabuthema handelt. Selbiges wurde auch von den Teilnehmerinnen berichtet. Aus diesem Grund wurden große Hoffnungen in Online-Fokusgruppen gesetzt, da diese sich eben vor allem für sehr sensible Themen eignen, wie die eigene Sexualität. Im Online-Setting blieben die Teilnehmerinnen anonym und konnten genau durchdenken, was sie schreiben wollten. Zudem erlauben Online-Fokusgruppen eine zeitliche und örtliche Flexibilität, d. h. die Datenerhebung findet lokal unabhängig statt und eine Terminkoordination ist nicht notwendig. Nachteil dieser Methode war, dass nur der „jüngere“ Teil der Zielpopulation erreicht wurde. Die 25 Frauen, die die Einschlusskriterien erfüllten, waren im Schnitt 51,6 Jahre alt, die jüngste 44, die älteste nur 58

Jahre alt. In der Menopause befinden sich aber vor allem Frauen ab 50 Jahren. Diese Beobachtung deutet darauf hin, dass ältere Frauen online eher schlecht bis gar nicht erreicht wurden, da diese entweder keinen Zugang zum Internet bzw. zu einem PC oder Smartphone hatten und somit auch nicht auf Facebook registriert waren. Da bei diesem Beschwerdebild allerdings davon auszugehen ist, dass sich die Beschwerden mit dem Alter verschlimmern, war es sehr wohl wichtig, auch die Erfahrungen von älteren Damen zu hören. Diese Schwäche sollten *face-to-face* Fokusgruppen mit „echten“ Patientinnen ausgleichen, was durchaus gelang. Die Teilnehmerinnen waren zwischen 47 und 85 Jahren alt. Der Altersdurchschnitt lag bei 62,0 Jahren, der Median bei 58 Jahren.

Limitationen der Studie

Trotz zahlreicher Rekrutierungsversuche, wie vorher beschrieben, konnte nicht die gewünschte Stichprobengröße von ≥ 25 erreicht werden. Dadurch, dass aufgrund der Rekrutierungsschwierigkeiten die Datenerhebung gestoppt werden musste, hatte sich eine theoretische Sättigung nur angedeutet. Als weitere Limitation ist die Tatsache aufzuführen, dass nur Frauen mit selbst berichteten Symptomen eingeschlossen wurden und keine klinische Untersuchung zur Bestätigung der Symptome vorher stattfand. Somit konnte auch nicht ausgeschlossen werden, dass die Symptome in Verbindung mit anderen Erkrankungen standen.

Ausblick

Vor allem die Ergebnisse zu den unerfüllten Bedürfnissen im Gesundheitswesen können als Anstoß für Verbesserungen in der Gesundheitsversorgung verwendet werden. Viele Frauen gaben an, dass es ihnen unangenehm war, ihre Beschwerden von sich aus anzusprechen, und selbst wenn sie sich dazu überwunden hatten, wurde ihnen mit Unverständnis oder Abweisung begegnet. Zudem wurde auch deutlich, dass unter den Frauen immer noch große Unwissenheit über das Thema „Menopause“ und die damit einhergehenden Beschwerden herrscht. Auch hier sollte eine stärkere Aufklärung, zum einen über die Ätiologie der Symptome, zum anderen aber auch über Behandlungsmöglichkeiten, ins Auge gefasst werden.

5. Allgemeine Diskussion

In dieser Arbeit wurden die Ziele verfolgt, 1) alle GSM- bzw. VVA-spezifischen PROMs zu identifizieren und deren Qualität zu bewerten, 2) ein präferiertes Kandidateninstrument in einer deutschen Population zu validieren und 3) die Auswirkungen der urologischen Komponente in einer qualitativen Studie genauer zu explorieren.

In der systematischen Übersichtsarbeit wurden zwei PROMs gefunden, VSQ und DIVA, die viele adäquate Messeigenschaften und keinerlei ungenügende Messeigenschaften aufwiesen. Diese beiden PROMs konnten gemäß den COSMIN-Richtlinien für den Einsatz in künftigen klinischen GSM-Studien empfohlen werden. Der DIVA wurde dem VSQ gegenüber sogar favorisiert, da bei dessen Item-Generierung die Zielpopulation, postmenopausale Frauen mit der VVA bzw. dem GSM, miteinbezogen wurden.

Demzufolge wurde die deutsche Version des DIVA anschließend validiert. In dieser Validierungsstudie konnten mehrere exzellente Messeigenschaften des DIVA bestätigt werden. Zudem wurde erstmals eine Rasch-Analyse mit dem DIVA zur Überprüfung der Strukturvalidität durchgeführt. Daraus resultierte, dass der DIVA als eine multidimensionale Skala aufgefasst werden muss. Jede Subskala erfüllte das Kriterium der Eindimensionalität und beschreibt somit nur ein einziges Konstrukt. Die deutsche Version des DIVA wurde außerdem „verfeinert“, indem die Items und Antwortkategorien genau analysiert wurden, die Anzahl an Antwortoptionen von fünf auf drei reduziert wurde und zwei weniger gut funktionierende Items ausgeschlossen wurden. Dadurch erzielte der DIVA eine noch bessere Passung auf das Modell. Validierungslücken zur Änderungssensitivität und Interpretierbarkeit des DIVA, die in der systematischen Übersichtsarbeit identifiziert wurden, konnte in der Zwischenzeit bereits von einer anderen Forschergruppe geschlossen werden (102).

In den Fokusgruppen wurde sehr deutlich, dass sich auch die urologische Komponente des GSM, ähnlich wie die VVA (65), negativ auf viele Bereiche des Lebens von betroffenen Frauen auswirkt. So gut wie alle Teilnehmerinnen hätten nicht erwartet, derart schlimme Wechseljahresbeschwerden zu entwickeln. Neben Strategien zum Umgang mit den Beschwerden wurden auch einige Schwachstellen im Gesundheitswesen identifiziert. Urogenitale Beschwerden stellen immer noch ein tabuisiertes und schambehaftetes Thema dar. Sie werden nach wie vor häufig von betroffenen Frauen einfach hingenommen und wenig bei ihrem Arzt bzw. ihrer Ärztin thematisiert. Wenn die Beschwerden beim Arzt oder der Ärztin besprochen werden, werden sie eher abgetan. Dadurch steigt die Hemmschwelle an, diese

erneut anzusprechen, was oftmals in Resignation endet und neben den körperlichen Beschwerden auch psychische Konsequenzen nach sich ziehen kann (84, 116).

Dadurch, dass immer häufiger PROs, wie z. B. die Lebensqualität, in klinischen Studien erfasst werden, ist auch ein Bedarf an qualitativ hochwertigen PROMs gegeben, die diese PROs messen können. Auf *clinicaltrials.gov* beispielsweise sind viele klinische Interventionsstudien registriert, die den DIVA, aber auch andere PROMs, wie z. B. den VSQ oder den FSFI, zur Erfassung ihrer Endpunkte einsetzen. PROs werden dabei sowohl als primäre als auch als sekundäre Endpunkte erfasst. Die Qualität einer PROM kann wiederum nur mithilfe von Validierungsstudien überprüft werden und je mehr gute Studien zu Messeigenschaften von PROMs durchgeführt werden, desto höher ist die Evidenz, dass die Qualität einer PROM als gesichert angesehen werden kann. Durch die Replikation der Messeigenschaften konnte zum einen die Qualität des DIVA bestätigt werden, zum anderen konnten mehrere exzellente Messeigenschaften des DIVA auf die deutsche Version ausgeweitet werden. Bereits in vier verschiedenen Sprachen konnte der DIVA nun schon stabile Ergebnisse vorweisen (38, 69, 75, 76). Das liegt unter anderem daran, dass der Entwicklung des DIVA eine fundierte, qualitative Forschung zugrunde liegt, in die die Zielpopulation, postmenopausale Frauen mit vaginalen Beschwerden, miteinbezogen wurde (65).

Implikationen für Forschung und Praxis

Es konnte eine klare Empfehlung für den Einsatz des DIVA in künftigen klinischen GSM-Studien ausgesprochen werden. Auch die deutsche Version zeigte viele hervorragende Messeigenschaften und erfüllt die COSMIN-Kriterien für den Einsatz in klinischer Forschung. Die Auswirkungen der urologischen Komponente werden jedoch im DIVA nicht berücksichtigt. In der qualitativen Studie zeigte sich allerdings sehr deutlich, dass auch urologische Beschwerden als Teil des GSM negative Auswirkungen auf unterschiedliche Bereiche des Lebens von betroffenen Frauen haben. In einem nächsten Schritt wäre es sinnvoll, diese Erkenntnisse zu nutzen und aus den Aussagen der Teilnehmerinnen Items zu formulieren, aus denen entweder eine separate PROM entstehen könnte oder eben ein zusätzliches urologisches „Modul“, das direkt in den DIVA als „Add-on“ integriert werden könnte. Eine Überprüfung der Inhaltsvalidität, sowie eine neue, umfassende Validierungsstudie wäre dann wieder vonnöten. Für die klinische Forschung ist der Einbezug der subjektiven Komponente, die mithilfe von PROMs erfasst werden kann, maßgeblich, da PROs kaum mit objektiven Parametern, wie z. B. mit dem vaginalen pH-Wert, korrelieren. Die vier Skalen des DIVA zeigten z. B. in einer kürzlich publizierten Studie quasi keine Korrelation ($r = -0,08 - 0,02$) mit

dem vaginalen pH-Wert (100). Das heißt, ohne den Einschluss subjektiver Aussagen würden wichtige Auswirkungen und Einschränkungen der Beschwerden verborgen bleiben.

In Bezug auf die Praxis wurde deutlich, dass viele Frauen von der Menopause übermannt werden und sich hilflos dieser natürlichen Veränderung ihres Körpers ausgeliefert fühlen. Ein offenes Ohr eines Arztes bzw. einer Ärztin und das Gefühl, ernst genommen zu werden, können bereits kurzfristig entlastend wirken. Wie bereits erwähnt, stellen urogenitale Beschwerden im Rahmen der Menopause immer noch ein mit Scham behaftetes Thema dar und die Enttabuisierung dieser Thematik sollte durch eine transparente Aufklärung, auch im Hinblick auf Behandlungsmöglichkeiten, angestrebt werden. So steht es auch im aktuellsten Statement der NAMS geschrieben, dass durch die Aufklärung, die Diagnose und den angemessenen Umgang mit dem GSM viele beunruhigende Symptome behoben werden können und somit die Lebensqualität und sexuelle Gesundheit von betroffenen Frauen verbessert werden kann (40). In der qualitativen Studie wurde auch der Wunsch geäußert, dass GynäkologInnen und auch Ärzte bzw. Ärztinnen allgemein proaktiv auf ihre Patientinnen zugehen, nachfragen und aufklären sollten. Nur wenn man sensibel und respektvoll mit dieser Thematik umgeht, lässt sich das Vertrauen der Patientinnen gewinnen. Allein die Aussage, dass GSM-Symptome in den Wechseljahren sehr üblich sind, kann helfen, ein Gespräch einzuleiten. Mithilfe von offenen Fragen können die Beschwerden der Patientinnen erfasst werden. Auch die Vorgeschichte der Patientinnen sollte beleuchtet werden, um potentielle Risikofaktoren für das GSM eruieren zu können (24). Oftmals bleibt im Praxisalltag wenig Zeit, um auf jede Patientin adäquat einzugehen und deren Vertrauen gewinnen zu können. Um einen ersten Überblick zu bekommen, können PROMs, wie der DIVA, eine gute und schnelle Möglichkeit bieten, Beeinträchtigungen der Lebensqualität von Frauen mit urogenitalen Beschwerden im Rahmen der Menopause aufzudecken und einen Einstieg in ein vertrauensvolles Gespräch zu finden. In einer kürzlich publizierten Studie von Mastroianni *et al.* (124) wurde überprüft, ob durch den DIVA auch die Identifikation und Diagnosestellung des GSM bei Frauen ≥ 45 Jahren verbessert werden kann. Es wurde tatsächlich ein statistisch signifikanter Unterschied in der GSM-Diagnose bei Patientinnen, die den DIVA ausfüllten, im Vergleich zu den Patientinnen, die den Fragebogen nicht ausfüllten, beobachtet. Der DIVA stellt also auch ein Screening-Instrument für Frauen im mittleren Alter dar, das die Identifikation und somit auch die Behandlung des GSM in der Routineversorgung steigern kann. Dadurch, dass das GSM immer noch viel zu wenig diagnostiziert wird (24, 40), könnte man so dieser hohen Dunkelziffer entgegenwirken und das Leid der betroffenen Frauen schneller lindern.

Kritische Auseinandersetzung mit dem Begriff „GSM“

Die Einführung des Begriffs „GSM“ im Jahre 2014 sollte zudem auch kritisch betrachtet werden. Bei der Diskussion auf der Konsensus-Konferenz war kein(e) VertreterIn der urologischen Gesellschaft anwesend, obwohl der urologischen Komponente eine große Bedeutung zugeschrieben wurde. Zudem sind in der dort generierten Symptomliste sexuelle Dysfunktionen mit vier von acht Symptomen stark überrepräsentiert (15), wobei Symptome sexueller Dysfunktion in allen Bevölkerungsschichten auftreten können und nicht spezifisch für die Menopause sein müssen. In dieser Definition des GSM wird außerdem nicht angegeben, wie viele Symptome oder Anzeichen vorhanden sein müssen, um eine entsprechende Diagnose geben zu können. Es besteht die Gefahr, dass viele Frauen den Begriff GSM im Internet lesen und somit eine „leichte“ Erklärung für ihre Beschwerden finden werden. Dies könnte dazu führen, dass weniger Frauen einen Arzt bzw. eine Ärztin aufsuchen und somit auch andere Erkrankungen mit ähnlichen Symptomen nicht ausgeschlossen werden können. Dadurch, dass urologische Symptome in das Syndrom eingeschlossen wurden, wird die Behandlung erschwert, da man anstatt einer spezifischen Diagnose mit einer passenden Behandlung nun eine allgemeinere Diagnose mit einer möglicherweise weniger spezifischeren Behandlung bekommen kann (49). Nichtsdestotrotz soll mit diesem neuen, akkurateren, umfassenderen, öffentlich akzeptierbaren und weniger unangenehmen Begriff die Kommunikation zwischen ÄrztInnen und ihren Patientinnen verbessert werden. In der Forschungsliteratur konnte sich dieser Begriff bereits gut etablieren, in der klinischen Praxis scheint das GSM noch nicht Einzug gefunden zu haben. Eine einheitliche deutsche Übersetzung des GSM gibt es noch nicht, man liest jedoch immer häufiger vom „urogenitalen Menopause(n)-Syndrom“ oder vom „urogenitalen Syndrom der Menopause“ (125, 126).

6. Zusammenfassung

Das *Genitourinary Syndrome of Menopause* (GSM) zeigt deutliche Auswirkungen auf das Leben von betroffenen Frauen. Durch die hohe Prävalenz der Beschwerden unter postmenopausalen Frauen und die Tatsache, dass urogenitale Beschwerden entgegen vieler Annahmen von alleine nicht wieder verschwinden, ist weiterhin eine große klinische Forschungsaktivität in diesem Bereich zu erwarten.

In der systematischen Übersichtsarbeit wurden insgesamt vier *patient-reported outcome measures* (PROMs) identifiziert, von denen zwei für die Verwendung in künftigen klinischen GSM-Studien gemäß den COSMIN-Richtlinien empfohlen werden konnten. Nur beim *Day-to-Day Impact of Vaginal Aging* (DIVA) *Questionnaire* wurde die Zielpopulation, postmenopausale Frauen mit vulvovaginaler Atrophie bzw. dem GSM, bei dessen Entwicklung miteinbezogen.

Aus diesem Grund wurde eine weitere Validierungsstudie mit der deutschen Version des DIVA durchgeführt, bei der mehrere exzellente Messeigenschaften des DIVA bestätigt werden konnten. Durch die erstmalige Verwendung der probabilistischen Testtheorie konnte nun noch ein komplett neuer Blick auf die einzelnen Items sowie auf die Antwortoptionen geworfen werden, um die Multidimensionalität des DIVA aufzuzeigen und den ohnehin schon sehr hochwertigen DIVA in seiner Qualität weiter zu verbessern. Zudem resultierte aus diesen Ergebnissen eine „verfeinerte“ Version des deutschen DIVA, die ebenfalls mehrere exzellente Messeigenschaften zeigen konnte. Der DIVA kann guten Gewissens weiterhin in klinischen Studien verwendet werden und auch die deutsche Version des DIVA liefert reliable und valide Ergebnisse.

Da der DIVA nur Items zu der genitalen und sexuellen Komponente des GSM enthält, wurde die urologische Komponente in einer qualitativen Untersuchung genauer exploriert. Es zeigte sich, dass auch urologische Beschwerden als Teil des GSM negative Auswirkungen auf das Leben von betroffenen Frauen haben und urogenitale Beschwerden im Rahmen der Menopause generell ein immer noch stark tabuisiertes Thema darstellen. Diese qualitative Studie legte offen, dass es auch einer PROM bedarf, die die Auswirkungen der urologischen Komponente des GSM erfasst, da der DIVA dieser Anforderung nicht gerecht wird. Darauf sollte sich künftige Forschung fokussieren, gemäß dem Motto „Qualität vor Quantität“.

Anhang A Suchstrategie für MEDLINE

Population: Genitourinary Syndrome of Menopause

((atrophic vaginitis/) or
((exp vagina/ or vaginal diseases/ or vaginitis/ or exp vulva/ or vulvar diseases/ or pruritus vulvae/ or
vulvitis/ or vulvar vestibulitis/ or vulvovaginitis/) and (atrophy/ or menopause/ or perimenopause/
or postmenopause/)) or
((vaginitis or vaginosis or vulvitis or vulvovaginitis or vestibulitis).ti,ab,kf.) or
((vulv* or vagin* or vulvovagin* or "vulvo vagin*" or genitourin* or "genito urin*" or urogenit* or
"uro genit*" or genital) adj3 (atroph* or dystroph* or kraurosis or syndrome* or symptoms or dry*
or odor or soreness or itching or prurit#s or irritat* or burning or stinging or soreness or pain* or
hurting or discomfort or fissuring or ageing).ti,ab,kf.) or
("genitourinary syndrome of menopause".ti,ab,kf.) or
(dyspareunia/ or vulvodinia/) or
((vulvar vestibulitis syndrome or vulvodinia* or vestibulodynia* or dyspareunia* or clitorodynia* or
algopareuni* or vulvismus).ti,ab,kf.) or
(((pain* or unpleasant or discomfort) adj3 (coitus or coital or intercourse or penetrat* or cohabit* or
sex or sexual or vulv* or vagin* or genital)).ti,ab,kf.))
OR
(((exp Vaginal Discharge/ or (vaginal discharge* or Leukorrhea*).ti,ab,kf.) OR
(Uterine Prolapse/ or ((uter* or vagin*) adj3 prolapse).ti,ab,kf.) OR
(((Vagina* or post-coital or postcoital or coital or coitus or intercourse) adj3 (Hemorrhage* or
Bleeding)).ti,ab,kf.) OR
("Sexual Dysfunction, Physiological"/ or "Sexual Dysfunctions, Psychological"/ or (((sex or sexual or
psychosexual) adj2 (dysfunction* or dysfunction* or disorder* or symptoms or function* or health
or wellness or well-being or satisfact* or dissatisfact* or pleasure or enjoyment or joy or problem*
or complaint* or difficult* or symptom* or distress or stress or morbidity or activit* or life or lifes
or lives or aversion)).ti,ab,kf.)) OR
((sexuality or intimacy or female genital self-image or frigid*).ti,ab,kf.) OR
(Libido/ or (Libido or arousal or ((sex or sexual) adj (desire* or drive* or interest* or disinterest*)) or
(desire adj (dysfunction* or dysfunction* or disorder* or problem*))).ti,ab,kf.) OR
((lubrication or arousal or (sexual adj (response or stimulation or excitement))).ti,ab,kf.) OR
((orgasm* or anorgas* or orgasmic).ti,ab,kf.) OR
("Urinary Bladder, Overactive"/ or "Urinary Bladder, Neurogenic"/ or (Overactive adj (Bladder or
detrusor)).ti,ab,kf.) OR
(Cystitis/ or (Cystiti* or Painful Bladder Syndrome).ti,ab,kf.) OR

(Urinary Tract Infections/ or (Urinary Tract Infection*).ti,ab,kf.) OR
(Urinary Incontinence/ or (Urin* adj3 (Incontinen* or urgency)).ti,ab,kf.) OR
(Dysuria/ or (Dysuria).ti,ab,kf.) OR
(Nocturia/ or (Nocturia or Nycturia).ti,ab,kf.) OR
(Urination Disorders/ or ((urination or micturition) adj disorder*).ti,ab,kf.))

AND

(Menopause/ or Perimenopause/ or Postmenopause/ or (menopaus* or perimenopaus* or peri-
menopaus* or postmenopaus* or post-menopaus* or climacterium).ti,ab,kf.))

Construct: Patient-reported outcome measures

For optimal sensitivity the search strategy of this search element is based on the PubMed filter "Quality of life (QoL)" of R. de Vries (<https://blocks.bmi-online.nl/catalog/294>), the PubMed filter "Patient reported outcome measures (PROMs)" of E.P. Jansma and R. de Vries (<https://blocks.bmi-online.nl/catalog/248>) and additional search terms from the "PROM group construct & instrument type filter" (Mackintosh *et al.* 2010) available from the COSMIN website (<http://www.cosmin.nl/images/upload/files/PROM%20Gp%20filtersOCTOBER%202010FINAL.pdf>).

All filters were last checked on March 19, 2018.

("Quality of Life"/) OR
((HR-PRO or HRPRO or qol or QL or hrql or hrqol).ti,ab,kf.) OR
((quality adj2 (life or living)).ti,ab,kf.) OR
(exp "Activities of Daily Living"/) OR
((activit* adj2 (daily adj (life or living))).ti,ab,kf.) OR
("adl".ti,ab,kf.) OR
("chronic limitation o# activity".ti,ab,kf.) OR
("self care*".ti,ab,kf.) OR
("Health Status"/) OR
((health adj2 (status or state or level* or index* or indices or profile*)).ti,ab,kf.) OR
("Patient Reported Outcome Measures"/) OR
((prom or proms or pro or pros).ti,ab,kf.) OR
(((patient OR self OR carer OR proxy) adj3 (appraisal* or appraised or report OR reported OR reporting OR rated OR rating OR ratings OR based OR assessed OR assessment OR assessments)).ti,ab,kf.) OR
(((disability OR function OR functional OR functions OR subjective OR utility OR utilities OR wellbeing OR well being) adj3 (outcome OR outcomes OR index OR indices OR instrument OR instruments OR measure OR measures OR questionnaire OR questionnaires OR profile OR profiles OR scale OR scales OR score OR scores OR status OR survey OR surveys)).ti,ab,kf.) OR

((satisfaction or emotions or life-impact or sexual-impact or functioning or well-being).ti,ab,kf.)

Measurement properties

Sensitive filter of: Terwee CB, Jansma EP, Riphagen II, de Vet HCW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009;18:1115-1123.

We use the translation to Ovid MEDLINE by Alberta University (Canada) (available from <http://www.cosmin.nl/images/upload/files/Translated%20search%20filter%20for%20measurement%20properties%20for%20OVID.pdf>) with minor corrections and syntactical tweaking:

- * We deleted the age filter that was not existent in Terwee *et al.*'s original filter.
- * We combined the search statements to a single-line search strategy.

((instrumentation OR methods).sh.) OR
((Validation Studies OR Comparative Study).pt.) OR
(exp Psychometrics/) OR
(psychometr*.ti,ab.) OR
((clinimetr* OR clinometr*).tw.) OR
(exp "Outcome Assessment (Health Care)") OR
(outcome assessment.ti,ab.) OR
(outcome measure*.tw.) OR
(exp Observer Variation/) OR
(observer variation.ti,ab.) OR
(exp Health Status Indicators/) OR
(exp "Reproducibility of Results") OR
(reproducib*.ti,ab.) OR
(exp Discriminant Analysis/) OR
((reliab* OR unreliab* OR valid* OR coefficient OR homogeneity OR homogeneous OR "internal consistency").ti,ab.) OR
((cronbach* and (alpha OR alphas)).ti,ab.) OR
((item and (correlation* OR selection* OR reduction*)).ti,ab.) OR
((agreement OR precision OR imprecision OR "precise values" OR test-retest).ti,ab.) OR
((test and retest).ti,ab.) OR
((reliab* and (test OR retest)).ti,ab.) OR
((stability OR interrater OR inter-rater OR intrarater OR intra-rater OR intertester OR inter-tester OR intratester OR intra-tester OR interobserver OR inter-observer OR intraobserver OR intra-observer OR intertechnician OR inter-technician OR intratechnician OR intra-technician OR interexaminer

OR inter-examiner OR intraexaminer OR intra-examiner OR interassay OR inter-assay OR
 intraassay OR intra-assay OR interindividual OR inter-individual OR intraindividual OR intra-
 individual OR interparticipant OR inter-participant OR intraparticipant OR intra-participant OR
 kappa OR kappa's OR kappas OR repeatab*).ti,ab.) OR
 (((replicab* OR repeated) and (measure OR measures OR findings OR result OR results OR test OR
 tests)).ti,ab.) OR
 ((generaliza* OR generalisa* OR concordance).ti,ab.) OR
 ((intraclass and correlation*).ti,ab.) OR
 ((discriminative OR "known group" OR factor analysis OR factor analyses OR dimension* OR
 subscale*).ti,ab.) OR
 ((multitrait and scaling and (analysis OR analyses)).ti,ab.) OR
 ((item discriminant OR interscale correlation* OR error OR errors OR "individual variability").ti,ab.)
 OR
 ((variability and (analysis OR values)).ti,ab.) OR
 ((uncertainty and (measurement OR measuring)).ti,ab.) OR
 (("standard error o# measurement" OR sensitiv* OR responsive*).ti,ab.) OR
 (((minimal OR minimally OR clinical OR clinically) and (important OR significant OR detectable)
 and (change OR difference)).ti,ab.) OR
 ((small* and (real OR detectable) and (change OR difference)).ti,ab.) OR
 ((meaningful change OR "ceiling effect" OR "floor effect" OR "Item response model" OR IRT OR
 Rasch OR "Differential item functioning" OR DIF OR "computer adaptive testing" OR "item bank"
 OR "cross-cultural equivalence").ti,ab.)

Feasibility

This is based on the search terms for the concept feasibility (search statement #1) in Heintl, Daniel; Prinsen, Cecilia A. C.; Drucker, Aaron M.; Ofenloch, Robert; Humphreys, Rosemary; Sach, Tracey *et al.* (2016): Measurement properties of quality of life measurement instruments for infants, children and adolescents with eczema. Protocol for a systematic review. In: Systematic reviews 5, S. 25. DOI: 10.1186/s13643-016-0202-z.

(accepta* OR "ease o# use" OR practica* OR feasib*).ti,ab.

Exclusion filter

Exclusion filter of: Terwee CB, Jansma EP, Riphagen II, de Vet HCW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res 2009;18:1115-1123.

This filter was modified in order to be run as a separate search statement.

("addresses" OR "biography" OR "case reports" OR "comment" OR "directory" OR "editorial" OR "festschrift" OR "interview" OR "lectures" OR "legal cases" OR "legislation" OR "letter" OR "news" OR "newspaper article" OR "patient education handout" OR "popular works" OR "congresses" OR "consensus development conference" OR "consensus development conference, nih" OR "practice guideline").pt. OR (exp "animals"/ NOT "humans"/)

Additional search for individual measurement instruments

We will search for the names of relevant measurement instruments. This is the list of instruments already known.

("vulvovaginal symptoms questionnaire").mp. OR
(VSQ and (vulv* or vagin*)).mp. OR
("Day-to-Day Impact of Vaginal Aging Questionnaire").mp. OR
("diva questionnaire").mp. OR
("Vaginal symptom bothersomeness scale").mp. OR
("Female sexual problems measure").mp. OR
("Female Sexual Function Index").mp. OR
(FSFI).mp. OR
("Female Genital Self Image Scale").mp. OR
(FGSIS).mp. OR
("Sexual Adjustment and Body Image Scale").mp. OR
(SABIS).mp. OR
("German Questionnaire on Feelings of Inadequacy in Social and Sexual Situations").mp. OR
("FUSS social scale").mp. OR
("FUSS sexual scale").mp. OR
("Geriatric Sexuality Inventory").mp.

Filter for publication language

(English or French or German or Italian).lg.

Anhang B Suchstrategie für *EMBASE*

Population: Genitourinary Syndrome of Menopause

((atrophic vaginitis/ or vagina atrophy/) or
((exp vagina/ or vagina disease/ or vagina pain/ or vaginal burning sensation/ or vaginal dryness/ or
vaginal pruritus/ or vulva disease/ or vulva pruritus/ or vulvitis/ or vulvar vestibulitis/ or
vulvovaginitis/) and (atrophy/ or exp "menopause and climacterium"/)) or
((vaginitis or vaginosis or vulvitis or vulvovaginitis or vestibulitis).ti,ab,kw.) or
((vulv* or vagin* or vulvovagin* or "vulvo vagin*" or genitourin* or "genito urin*" or urogenit* or
"uro genit*" or genital) adj3 (atroph* or dystroph* or kraurosis or syndrome* or symptoms or dry* or
odo?r or soreness or itching or prurit#s or irritat* or burning or stinging or soreness or pain* or hurting
or discomfort or fissuring or ageing).ti,ab,kw.) or
("genitourinary syndrome o# menopause".ti,ab,kw.) or
(dyspareunia/ or vulvodynia/ or vulvovaginal discomfort/ or vaginal discomfort/ or vulva kraurosis/)
or
((vulvar vestibulitis syndrome or vulvodynia* or vestibulodynia* or dyspareunia* or clitorodynia* or
algopareuni* or vulvismus).ti,ab,kw.) or
(((pain* or unpleasant or discomfort) adj3 (coitus or coital or intercourse or penetrat* or cohabit* or
sex or sexual or vulv* or vagin* or genital)).ti,ab,kw.))
OR
(((leukorrhoea/ or vagina discharge/ or (vaginal discharge* or leukorrhoea*).ti,ab,kw.) OR
(vaginal vault prolapse/ or uterus prolapse/ or ((uter* or vagin*) adj3 prolapse).ti,ab,kw.) OR
(vagina bleeding/ or ((vagina* or post-coital or postcoital or coital or coitus or intercourse) adj3
(h?emorrhage* or bleeding)).ti,ab,kw.) OR
("sexual dysfunction"/ or "female sexual dysfunction"/ or (((sex or sexual or psychosexual) adj2
(d#sfunction* or d#s function* or disorder* or symptoms or function* or health or wellness or well-
being or satisfact* or dissatisfact* or pleasure or enjoyment or joy or problem* or complaint* or
difficult* or symptom* or distress or stress or morbidity or activit* or life or lifes or lives or
aversion)).ti,ab,kw.)) OR
(frigidity/ or (sexuality or intimacy or female genital self-image or frigid*).ti,ab,kw.) OR
(libido/ or "libido disorder"/ or "hypoactive sexual desire disorder"/ or sexual arousal disorder/ or
(libido or arousal or ((sex or sexual) adj (desire* or drive* or interest* or disinterest*)) or (desire adj
(d#sfunction* or d#s function* or disorder* or problem*))).ti,ab,kw.) OR
(lubrication/ or vaginal lubrication/ or (lubrication or arousal or (sexual adj (response or stimulation or
excitement))).ti,ab,kw.) OR
(anorgasmia/ or orgasm disorder/ or (orgasm* or anorgas* or orgastic).ti,ab,kw.) OR

("overactive bladder"/ or "neurogenic bladder"/ or ((overactive or neurogenic) adj (bladder or detrusor)).ti,ab,kw.) OR
(exp cystitis/ or (cystiti* or painful bladder syndrome).ti,ab,kw.) OR
(urinary tract infection/ or (urinary tract infection*).ti,ab,kw.) OR
(urine incontinence/ or urge incontinence/ or (urin* adj3 (incontinen* or urgency)).ti,ab,kw.) OR
(dysuria/ or (dysuria).ti,ab,kw.) OR
(nocturia/ or (nocturia or nycturia).ti,ab,kw.) OR
(micturition disorder/ or ((urination or micturition) adj disorder*).ti,ab,kw.))

AND

(exp "menopause and climacterium"/ or (menopaus* or perimenopaus* or peri-menopaus* or postmenopaus* or post-menopaus* or climacterium).ti,ab,kw.))

Construct: Patient-reported outcome measures

("quality of life"/) OR
((HR-PRO or HRPRO or qol or QL or hrql or hrqol).ti,ab,kw.) OR
((quality adj2 (life or living)).ti,ab,kw.) OR
("daily life activity"/) OR
((activit* adj2 (daily adj (life or living))).ti,ab,kw.) OR
("adl".ti,ab,kw.) OR
("chronic limitation o# activity".ti,ab,kw.) OR
("self care*".ti,ab,kw.) OR ("health status"/) OR
((health adj2 (status or state or level* or index* or indices or profile*)).ti,ab,kw.) OR
("patient-reported outcome"/) OR
((prom or proms or pro or pros).ti,ab,kw.) OR
(((patient OR self OR carer OR proxy) adj3 (appraisal* or appraised or report OR reported OR reporting OR rated OR rating OR ratings OR based OR assessed OR assessment OR assessments)).ti,ab,kw.) OR
(((disability OR function OR functional OR functions OR subjective OR utility OR utilities OR wellbeing OR well being) adj3 (outcome OR outcomes OR index OR indices OR instrument OR instruments OR measure OR measures OR questionnaire OR questionnaires OR profile OR profiles OR scale OR scales OR score OR scores OR status OR survey OR surveys)).ti,ab,kw.) OR
((satisfaction or emotions or life-impact or sexual-impact or functioning or well-being).ti,ab,kw.)

Measurement properties

Sensitive filter of: Terwee CB, Jansma EP, Riphagen II, de Vet HCW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res 2009;18:1115-1123.

We use the version translated to EMBASE (Ovid) from Heintl, Daniel; Prinsen, Cecilia A. C.; Drucker, Aaron M.; Ofenloch, Robert; Humphreys, Rosemary; Sach, Tracey *et al.* (2016): Measurement properties of quality of life measurement instruments for infants, children and adolescents with eczema. Protocol for a systematic review. In: Systematic reviews 5, S. 25. DOI: 10.1186/s13643-016-0202-z.

exp instrumentation/ or exp validation study/ or exp reproducibility/ or reproducib\$.mp. or exp psychometry/ or psychometr\$.mp. or clinimetr*.mp. or clinimetr\$.mp. or exp observer variation/ or observer variation.mp. or exp discriminant analysis/ or exp reliability/ or reliab\$.mp. or exp Validity/ or valid\$.mp. or coefficient.mp. or internal consistency.mp. or (cronbach\$ and (alpha or alphas)).mp. or item correlation.mp. or item correlations.mp. or item selection.mp. or item selections.mp. or item reduction.mp. or item reductions.mp. or agreement.mp. or precision.mp. or imprecision.mp. or precise values.mp. or (test-retest or (test and retest) or (reliab\$ and (test or retest)) or stability or interrater or inter-rater or intrarater or intra-rater or intertester or inter-tester or intratester or intra-tester or interobserver or inter-observer or intraobserver or intra-observer or intertechnician or intertechnician or intratechnician or intra-technician or interexaminer or inter-examiner or intraexaminer or intra-examiner or interassay or inter-assay or intraassay or intra-assay or interindividual or inter-individual or intraindividual or intra-individual or interparticipant or inter-participant or intraparticipant or intra-participant or kappa or kappa\$ or coefficient of variation or repeatab\$ or ((replicap\$ or repeated) and (measure or measures or findings or result or results or test or tests))).mp. or (generaliza\$ or generalisa\$ or concordance or (intraclass and correlation\$) or discriminative or known group or factor analysis or factor analyses or factor structure or factor structures or dimensionality or subscale\$ or multitrait scaling analysis or multitrait scaling analyses or item discriminant or interscale correlation or interscale correlations or ((error or errors) and (measure\$ or correlat\$ or evaluat\$ or accuracy or accurate or precision or mean)) or individual variability or interval variability or rate variability or variability analysis or (uncertainty and (measurement or measuring)) or standard error of measurement or sensitiv\$ or responsive\$ or (limit and detection) or minimal detectable concentration or interpretab\$ or (small\$ and (real or detectable) and (change or difference)) or meaningful change or minimal important change or minimal important difference or minimally important change or minimally important difference or minimal detectable change or minimal detectable difference or minimally detectable change or minimally detectable difference or minimal real change or minimal real difference or minimally real change or minimally real difference).mp. or (ceiling effect or floor effect or Item response model or IRT or Rasch or Differential item functioning or DIF or computer adaptive testing or item bank or cross-cultural equivalence).mp.

Feasibility

This is based on the search terms for the concept feasibility (search statement #1) in Heintl, Daniel; Prinsen, Cecilia A. C.; Drucker, Aaron M.; Ofenloch, Robert; Humphreys, Rosemary; Sach, Tracey *et*

al. (2016): Measurement properties of quality of life measurement instruments for infants, children and adolescents with eczema. Protocol for a systematic review. In: *Systematic reviews* 5, S. 25. DOI: 10.1186/s13643-016-0202-z.

(accepta* OR "ease o# use" OR practica* OR feasib*).ti,ab.

Exclusion filter

We use the version for EMBASE (Ovid) from Heintl, Daniel; Prinsen, Cecilia A. C.; Drucker, Aaron M.; Ofenloch, Robert; Humphreys, Rosemary; Sach, Tracey *et al.* (2016): Measurement properties of quality of life measurement instruments for infants, children and adolescents with eczema. Protocol for a systematic review. In: *Systematic reviews* 5, S. 25. DOI: 10.1186/s13643-016-0202-z.

(conference abstract or conference paper or conference review or editorial or erratum or letter or note).pt.

Additional search for individual measurement instruments

We will search for the names of relevant measurement instruments. This is the list of instruments already known.

("vulvovaginal symptoms questionnaire").mp. OR
(VSQ and (vulv* or vagin*)).mp. OR
("Day-to-Day Impact o# Vaginal Aging Questionnaire").mp. OR
("diva questionnaire").mp. OR
("Vaginal symptom bothersomeness scale").mp. OR
("Female sexual problems measure").mp. OR
("Female Sexual Function Index").mp. OR
(FSFI).mp. OR
("Female Genital Self Image Scale").mp. OR
(FGSIS).mp. OR
("Sexual Adjustment and Body Image Scale").mp. OR
(SABIS).mp. OR
("German Questionnaire on Feelings o# Inadequacy in Social and Sexual Situations").mp. OR
("FUSS social scale").mp. OR
("FUSS sexual scale").mp. OR
("Geriatric Sexuality Inventory").mp.

Filter for publication language

(English or French or German or Italian).lg.

Anhang C Suchstrategie für *Web of Science*

ANPASSUNG AN WOS:

DONE:

- * Feldbezeichner
- * quote phrases
- * Maskierung, Trunkierung: \$ -> *, ? -> \$, # -> ?
- * Weiteres: Was?

Population: Genitourinary Syndrome of Menopause

((TS=(vaginitis or vaginosis or vulvitis or vulvovaginitis or vestibulitis)) or
(TS=((vulv* or vagin* or vulvovagin* or "vulvo vagin*" or genitourin* or "genito urin*" or urogenit*
or "uro genit*" or genital) NEAR/2 (atroph* or dystroph* or kraurosis or syndrome* or symptoms or
dry* or odor\$ or soreness or itching or prurit?s or irritat* or burning or stinging or soreness or pain* or
hurting or discomfort or fissuring or ageing))) or
(TS=("genitourinary syndrome of menopause")) or
(TS=("vulvar vestibulitis syndrome" or vulvodynia* or vestibulodynia* or dyspareunia* or
clitorodynia* or algopareuni* or vulvismus)) or
(TS=((pain* or unpleasant or discomfort) NEAR/2 (coitus or coital or intercourse or penetrat* or
cohabit* or sex or sexual or vulv* or vagin* or genital))))
OR
(((TS=("vaginal discharge*" or leukorrhea*)) OR
(TS=((uter* or vagin*) NEAR/2 prolapse)) OR
(TS=((vagina* or "post-coital" or postcoital or coital or coitus or intercourse) NEAR/2 (h\$emorrhage*
or bleeding))) OR
(TS=((sex or sexual or psychosexual) NEAR/1 (d?sfunction* or "d?s function*" or disorder* or
symptoms or function* or health or wellness or well-being or satisfact* or dissatisfact* or pleasure or
enjoyment or joy or problem* or complaint* or difficult* or symptom* or distress or stress or
morbidity or activit* or life or lifes or lives or aversion)))) OR
(TS=(sexuality or intimacy or "female genital self-image" or frigid*)) OR
(TS=(libido or arousal or ((sex or sexual) NEAR/0 (desire* or drive* or interest* or disinterest*)) or
(desire NEAR/0 (d?sfunction* or "d?s function*" or disorder* or problem*)))) OR
(TS=(lubrication or arousal or (sexual NEAR/0 (response or stimulation or excitement)))) OR
(TS=(orgasm* or anorgas* or orgasmic)) OR
(TS=(overactive NEAR/0 (bladder or detrusor))) OR
(TS=(cystiti* or "painful bladder syndrome")) OR

(TS=("urinary tract infection*")) OR
 (TS=(urin* NEAR/2 (incontinen* or urgency))) OR
 (TS=(dysuria)) OR
 (TS=(nocturia or nycturia)) OR
 (TS=((urination or micturition) NEAR/0 disorder*))
 AND
 (TS=(menopaus* or perimenopaus* or peri-menopaus* or postmenopaus* or post-menopaus* or climacterium))

Construct: Patient-reported outcome measures

For optimal sensitivity the search strategy of this search element is based on the PubMed filter "Quality of life (QoL)" of R. de Vries (<https://blocks.bmi-online.nl/catalog/294>), the PubMed filter "Patient reported outcome measures (PROMs)" of E.P. Jansma and R. de Vries (<https://blocks.bmi-online.nl/catalog/248>) and additional search terms from the "PROM group construct & instrument type filter" (Mackintosh *et al.* 2010) available from the COSMIN website (<http://www.cosmin.nl/images/upload/files/PROM%20Gp%20filtersOCTOBER%202010FINAL.pdf>). All filters were last checked on March 19, 2018.

(TS=(HR-PRO or HRPRO or qol or QL or hrql or hrqol)) OR
 (TS=(quality NEAR/1 (life or living))) OR
 (TS=(activit* NEAR/1 (daily NEAR/0 (life or living)))) OR
 (TS=("adl")) OR
 (TS=("chronic limitation of activity")) OR
 (TS=("self care*")) OR
 (TS=(health NEAR/1 (status or state or level* or index* or indices or profile*))) OR
 (TS=(prom or proms or pro or pros)) OR
 (TS=((patient OR self OR carer OR proxy) NEAR/2 (appraisal* or appraised or report OR reported OR reporting OR rated OR rating OR ratings OR based OR assessed OR assessment OR assessments))) OR
 (TS=((disability OR function OR functional OR functions OR subjective OR utility OR utilities OR wellbeing OR "well being") NEAR/2 (outcome OR outcomes OR index OR indices OR instrument OR instruments OR measure OR measures OR questionnaire OR questionnaires OR profile OR profiles OR scale OR scales OR score OR scores OR status OR survey OR surveys))) OR
 (TS=(satisfaction or emotions or life-impact or sexual-impact or functioning or well-being))

Measurement properties

Sensitive filter of: Terwee CB, Jansma EP, Riphagen II, de Vet HCW. Development of a methodological

PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res 2009;18:1115-1123.

Adaptation to Web of Science

(TS=(instruments OR methodology)) OR
(TS=(psychometr*)) OR
(TS=(clinimetr* OR clinometr*)) OR
(TS=("outcome assessment")) OR
(TS=("outcome measure*")) OR
(TS=("observer variation")) OR
(TS=(reproducib*)) OR
(TS=(reliab* OR unreliab* OR valid* OR coefficient OR homogeneity OR homogeneous OR "internal consistency")) OR
(TS=(cronbach* and (alpha OR alphas))) OR
(TS=(item and (correlation* OR selection* OR reduction*))) OR
(TS=(agreement OR precision OR imprecision OR "precise values" OR test-retest)) OR
(TS=(test and retest)) OR
(TS=(reliab* and (test OR retest))) OR
(TS=(stability OR interrater OR inter-rater OR intrarater OR intra-rater OR intertester OR inter-tester OR intratester OR intra-tester OR interobserver OR inter-observer OR intraobserver OR intra-observer OR intertechnician OR inter-technician OR intratechnician OR intra-technician OR interexaminer OR inter-examiner OR intraexaminer OR intra-examiner OR interassay OR inter-assay OR intraassay OR intra-assay OR interindividual OR inter-individual OR intraindividual OR intra-individual OR interparticipant OR inter-participant OR intraparticipant OR intra-participant OR kappa OR "kappa's" OR kappas OR repeatab*)) OR
(TS=((replicab* OR repeated) and (measure OR measures OR findings OR result OR results OR test OR tests))) OR
(TS=(generaliza* OR generalisa* OR concordance)) OR
(TS=(intraclass and correlation*)) OR
(TS=(discriminative OR "known group" OR "factor analysis" OR "factor analyses" OR dimension* OR subscale*)) OR
(TS=(multitrait and scaling and (analysis OR analyses))) OR
(TS=("item discriminant" OR "interscale correlation*" OR error OR errors OR "individual variability")) OR
(TS=(variability and (analysis OR values))) OR
(TS=(uncertainty and (measurement OR measuring))) OR

(TS=("standard error of measurement" OR sensitiv* OR responsive*)) OR
(TS=((minimal OR minimally OR clinical OR clinically) and (important OR significant OR detectable) and (change OR difference))) OR
(TS=(small* and (real OR detectable) and (change OR difference))) OR
(TS=("meaningful change" OR "ceiling effect" OR "floor effect" OR "Item response model" OR IRT OR Rasch OR "Differential item functioning" OR DIF OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence"))

Feasibility

This is based on the search terms for the concept feasibility (search statement #1) in Heinl, Daniel; Prinsen, Cecilia A. C.; Drucker, Aaron M.; Ofenloch, Robert; Humphreys, Rosemary; Sach, Tracey *et al.* (2016): Measurement properties of quality of life measurement instruments for infants, children and adolescents with eczema. Protocol for a systematic review. In: Systematic reviews 5, S. 25. DOI: 10.1186/s13643-016-0202-z.

TS=(accepta* OR "ease of use" OR practica* OR feasib*)

Additional search for individual measurement instruments

We will search for the names of relevant measurement instruments. This is the list of instruments already known.

(TS=("vulvovaginal symptoms questionnaire")) OR
(TS=(VSQ and (vulv* or vagin*))) OR
(TS=("Day-to-Day Impact of Vaginal Aging Questionnaire")) OR
(TS=("diva questionnaire")) OR
(TS=("Vaginal symptom bothersomeness scale")) OR
(TS=("Female sexual problems measure")) OR
(TS=("Female Sexual Function Index")) OR
(TS=(FSFI)) OR
(TS=("Female Genital Self Image Scale")) OR
(TS=(FGSIS)) OR
(TS=("Sexual Adjustment and Body Image Scale")) OR
(TS=(SABIS)) OR
(TS=("German Questionnaire on Feelings of Inadequacy in Social and Sexual Situations")) OR
(TS=("FUSS social scale")) OR
(TS=("FUSS sexual scale")) OR
(TS=("Geriatric Sexuality Inventory"))

Exclusion filter

Filter for unwanted document types, then NOT them out:

DOCUMENT TYPES: (Art Exhibit Review OR Biographical-Item OR Chronology OR Dance Performance Review OR Editorial Material OR Fiction, Creative Prose OR Film Review OR Hardware Review OR Letter OR Meeting Abstract OR Meeting Summary OR Music Performance Review OR Music Score OR Music Score Review OR News Item OR Poetry OR Proceedings Paper OR TV Review, Radio Review OR TV Review, Radio Review Video OR Theater Review)

Filter for publication language

Filter: LANGUAGE: (English OR French OR German OR Italian)

Anhang D Deutsche Version des DIVA

MESSINSTRUMENT FÜR DIE ALLTÄGLICHEN AUSWIRKUNGEN DES VAGINALEN ALTERNS

Fragebogen zu den alltäglichen Auswirkungen der vaginalen Atrophie (des vaginalen Alterns)

Wir sind daran interessiert, die Auswirkungen vaginaler Symptome (Symptome im Bereich der Scheide), wie Trockenheit, Schmerzen, Reizung und Juckreiz, auf Ihren Alltag zu verstehen. Kreuzen Sie bitte unten für jede Frage diejenige Antwort an, die am besten beschreibt, wie Ihre Aktivitäten, Beziehungen und Gefühle während der vergangenen vier Wochen durch diese Symptome beeinträchtigt wurden.

Teil A. Wie sehr haben während der vergangenen vier Wochen vaginale Symptome, wie Trockenheit, Schmerzen, Reizung oder Juckreiz, es für Sie unangenehm gemacht oder Ihre Fähigkeit beeinträchtigt, Folgendes zu tun:

1. So schnell wie sonst zu gehen?	<input type="checkbox"/> ₀ Überhaupt nicht	<input type="checkbox"/> ₁ Ein wenig	<input type="checkbox"/> ₂ Mäßig	<input type="checkbox"/> ₃ Ziemlich	<input type="checkbox"/> ₄ Extrem
2. Die Kleidung oder Unterwäsche Ihrer Wahl zu tragen?	<input type="checkbox"/> ₀ Überhaupt nicht	<input type="checkbox"/> ₁ Ein wenig	<input type="checkbox"/> ₂ Mäßig	<input type="checkbox"/> ₃ Ziemlich	<input type="checkbox"/> ₄ Extrem
3. Zur Toilette zu gehen oder sich nach dem Toilettengang abzuwischen?	<input type="checkbox"/> ₀ Überhaupt nicht	<input type="checkbox"/> ₁ Ein wenig	<input type="checkbox"/> ₂ Mäßig	<input type="checkbox"/> ₃ Ziemlich	<input type="checkbox"/> ₄ Extrem
4. Länger als eine Stunde zu sitzen?	<input type="checkbox"/> ₀ Überhaupt nicht	<input type="checkbox"/> ₁ Ein wenig	<input type="checkbox"/> ₂ Mäßig	<input type="checkbox"/> ₃ Ziemlich	<input type="checkbox"/> ₄ Extrem
5. Erholenden Schlaf zu bekommen?	<input type="checkbox"/> ₀ Überhaupt nicht	<input type="checkbox"/> ₁ Ein wenig	<input type="checkbox"/> ₂ Mäßig	<input type="checkbox"/> ₃ Ziemlich	<input type="checkbox"/> ₄ Extrem

Teil B. Wie oft haben Sie sich während der vergangenen vier Wochen wegen vaginaler Symptome, wie Trockenheit, Schmerzen, Reizung oder Juckreiz, wie folgt gefühlt:

6. Depressiv oder niedergeschlagen?	<input type="checkbox"/> ₀ Nie	<input type="checkbox"/> ₁ Selten	<input type="checkbox"/> ₂ Manchmal	<input type="checkbox"/> ₃ Ziemlich oft	<input type="checkbox"/> ₄ Sehr oft
7. Verlegen?	<input type="checkbox"/> ₀ Nie	<input type="checkbox"/> ₁ Selten	<input type="checkbox"/> ₂ Manchmal	<input type="checkbox"/> ₃ Ziemlich oft	<input type="checkbox"/> ₄ Sehr oft
8. Frustriert oder verbittert?	<input type="checkbox"/> ₀ Nie	<input type="checkbox"/> ₁ Selten	<input type="checkbox"/> ₂ Manchmal	<input type="checkbox"/> ₃ Ziemlich oft	<input type="checkbox"/> ₄ Sehr oft
9. Sich selbst bemitleidend?	<input type="checkbox"/> ₀ Nie	<input type="checkbox"/> ₁ Selten	<input type="checkbox"/> ₂ Manchmal	<input type="checkbox"/> ₃ Ziemlich oft	<input type="checkbox"/> ₄ Sehr oft

Teil C. Bei den folgenden Fragen geht es um die Auswirkungen Ihrer Symptome auf vaginalen Geschlechtsverkehr sowie andere Arten sexueller Aktivität wie Selbstbefriedigung. Haben während der vergangenen vier Wochen vaginale Symptome, wie Trockenheit, Schmerzen, Reizung oder Juckreiz, Folgendes beeinträchtigt:

10. Ihren Wunsch nach oder Ihr Interesse an Geschlechtsverkehr oder anderen Arten sexueller Aktivität (einschließlich Selbstbefriedigung)?	<input type="checkbox"/> ₀ Überhaupt nicht	<input type="checkbox"/> ₁ Ein wenig	<input type="checkbox"/> ₂ Mäßig	<input type="checkbox"/> ₃ Ziemlich	<input type="checkbox"/> ₄ Extrem
11. Wie häufig Sie Geschlechtsverkehr oder andere Arten sexueller Aktivität hatten (einschließlich Selbstbefriedigung)?	<input type="checkbox"/> ₀ Überhaupt nicht	<input type="checkbox"/> ₁ Ein wenig	<input type="checkbox"/> ₂ Mäßig	<input type="checkbox"/> ₃ Ziemlich	<input type="checkbox"/> ₄ Extrem
12. Ihre Fähigkeit, während sexueller Aktivität erregt zu werden (einschließlich Selbstbefriedigung)?	<input type="checkbox"/> ₀ Überhaupt nicht	<input type="checkbox"/> ₁ Ein wenig	<input type="checkbox"/> ₂ Mäßig	<input type="checkbox"/> ₃ Ziemlich	<input type="checkbox"/> ₄ Extrem
13. Ihre Fähigkeit, hinsichtlich sexueller Aktivität spontan zu sein (einschließlich Selbstbefriedigung)?	<input type="checkbox"/> ₀ Überhaupt nicht	<input type="checkbox"/> ₁ Ein wenig	<input type="checkbox"/> ₂ Mäßig	<input type="checkbox"/> ₃ Ziemlich	<input type="checkbox"/> ₄ Extrem
14. Ihre Fähigkeit, sich zu entspannen und sexuelle Aktivität zu genießen (einschließlich Selbstbefriedigung)?	<input type="checkbox"/> ₀ Überhaupt nicht	<input type="checkbox"/> ₁ Ein wenig	<input type="checkbox"/> ₂ Mäßig	<input type="checkbox"/> ₃ Ziemlich	<input type="checkbox"/> ₄ Extrem
15. Das Ausmaß der sexuellen Befriedigung (zum Höhepunkt kommen), die Sie während sexueller Aktivität empfunden haben (einschließlich Selbstbefriedigung)?	<input type="checkbox"/> ₀ Überhaupt nicht	<input type="checkbox"/> ₁ Ein wenig	<input type="checkbox"/> ₂ Mäßig	<input type="checkbox"/> ₃ Ziemlich	<input type="checkbox"/> ₄ Extrem
16. Ihren Wunsch nach oder Ihr Interesse an einer sexuellen Beziehung?	<input type="checkbox"/> ₀ Überhaupt nicht	<input type="checkbox"/> ₁ Ein wenig	<input type="checkbox"/> ₂ Mäßig	<input type="checkbox"/> ₃ Ziemlich	<input type="checkbox"/> ₄ Extrem
17. Ihre Zuversicht, dass Sie einen Partner sexuell befriedigen könnten?	<input type="checkbox"/> ₀ Überhaupt nicht	<input type="checkbox"/> ₁ Ein wenig	<input type="checkbox"/> ₂ Mäßig	<input type="checkbox"/> ₃ Ziemlich	<input type="checkbox"/> ₄ Extrem
18. Ihre Zufriedenheit mit Ihrem Sexualleben insgesamt?	<input type="checkbox"/> ₀ Überhaupt nicht	<input type="checkbox"/> ₁ Ein wenig	<input type="checkbox"/> ₂ Mäßig	<input type="checkbox"/> ₃ Ziemlich	<input type="checkbox"/> ₄ Extrem

Teil D. Die folgenden Aussagen beschreiben Arten und Weisen, wie sich Ihre vaginalen Symptome möglicherweise auf Ihre Gefühle hinsichtlich Ihrer eigenen Person und Ihres Körpers ausgewirkt haben. Geben Sie bitte an, wie zutreffend jede der folgenden Aussagen während der vergangenen vier Wochen für Sie gewesen ist.

- | | | | | | |
|---|---|---|--|--|--|
| 19. Meine vaginalen Symptome geben mir das Gefühl, alt zu werden. | <input type="checkbox"/> ₀ Trifft überhaupt nicht zu | <input type="checkbox"/> ₁ Trifft etwas zu | <input type="checkbox"/> ₂ Trifft zum Teil zu | <input type="checkbox"/> ₃ Trifft weitgehend zu | <input type="checkbox"/> ₄ Trifft voll zu |
| 20. Ich fühle mich wegen meiner vaginalen Symptome nicht begehrenswert. | <input type="checkbox"/> ₀ Trifft überhaupt nicht zu | <input type="checkbox"/> ₁ Trifft etwas zu | <input type="checkbox"/> ₂ Trifft zum Teil zu | <input type="checkbox"/> ₃ Trifft weitgehend zu | <input type="checkbox"/> ₄ Trifft voll zu |
| 21. Wenn ich an meine vaginalen Symptome denke, habe ich das Gefühl, ich habe etwas verloren. | <input type="checkbox"/> ₀ Trifft überhaupt nicht zu | <input type="checkbox"/> ₁ Trifft etwas zu | <input type="checkbox"/> ₂ Trifft zum Teil zu | <input type="checkbox"/> ₃ Trifft weitgehend zu | <input type="checkbox"/> ₄ Trifft voll zu |
| 22. Meine vaginalen Symptome geben mir das Gefühl, dass mein Körper abbaut. | <input type="checkbox"/> ₀ Trifft überhaupt nicht zu | <input type="checkbox"/> ₁ Trifft etwas zu | <input type="checkbox"/> ₂ Trifft zum Teil zu | <input type="checkbox"/> ₃ Trifft weitgehend zu | <input type="checkbox"/> ₄ Trifft voll zu |
| 23. Ich fühle mich wegen meiner vaginalen Symptome weniger sexuell attraktiv. | <input type="checkbox"/> ₀ Trifft überhaupt nicht zu | <input type="checkbox"/> ₁ Trifft etwas zu | <input type="checkbox"/> ₂ Trifft zum Teil zu | <input type="checkbox"/> ₃ Trifft weitgehend zu | <input type="checkbox"/> ₄ Trifft voll zu |

Anhang E Flyer für die Rekrutierung

(a) Flyer zur Rekrutierung für die quantitative Befragung (Validierung des DIVA)

♀ **Liebe Frauen,**

Sie sind in den Wechseljahren und leiden unter vaginalen Beschwerden wie Trockenheit, Brennen, Juckreiz oder Schmerzen im Intimbereich?

Dann nehmen Sie sich bitte  **10 - 15 Minuten** Zeit und füllen nachfolgenden **Fragebogen** online aus.

Im Rahmen meiner Doktorarbeit interessiere ich mich für **Ihre Beschwerden**, vor allem aber wie sich diese auf **Ihre Lebensqualität** auswirken.  **Ihre Perspektive** ist dabei wichtig!

Über folgenden Link gelangen Sie zum Fragebogen:

 www.soscisurvey.de/Auswirkungen_Menopause/ 

Vielen Dank für Ihre Mithilfe!
Bei Rückfragen dürfen Sie mich gerne kontaktieren.

Michaela Gabes
michaela.gabes@klinik.uni-regensburg.de

 **Danke!** 

 Universität Regensburg

(b) Flyer zur Rekrutierung der face-to-face Fokusgruppe(n) am Inselspital Bern

Liebe Frauen,

Sie sind in den **Wechseljahren** und leiden neben **vaginalen Beschwerden** (wie Trockenheit, Brennen oder Juckreiz) auch an **urologischen Beschwerden** (wie schmerzhaftem oder vermehrtem Wasserlassen, Inkontinenz oder vermehrten Harnwegsinfekten)?

Im Rahmen meiner Doktorarbeit interessiere ich mich für Ihre Beschwerden, vor allem aber wie sich diese auf Ihre Lebensqualität auswirken.


Ich würde Sie gerne dazu einladen, über Ihre Beschwerden mit anderen betroffenen Frauen im kleinen Kreis zu diskutieren. Dafür sollten Sie 80 - 90 Minuten einplanen.


Wann?	Am 31.01.2020 um 08:30 Uhr bzw. 10:30 Uhr
Wo?	Im Kursraum B171a , Inselspital Bern

Bei Interesse melden Sie sich bitte bei **Frau Prof. Stute** (petra.stute@insel.ch oder 031-632-1310) für die Gruppendiskussion an. Als Dank erhält jede Teilnehmerin an der Gruppendiskussion einen **15€-Amazon-Gutschein**.

Ich freue mich auf spannende Diskussionen mit Ihnen!

Michaela Gabes
michaela.gabes@klinik.uni-regensburg.de

 Universität Regensburg



(c) Flyer zur Rekrutierung der Online-Fokusgruppen

Sie sind in den **Wechseljahren** und leiden neben **vaginalen Beschwerden** (wie Trockenheit, Brennen oder Juckreiz) auch an **urologischen Beschwerden** (wie schmerzhaftem oder vermehrtem Wasserlassen, Inkontinenz oder häufigen Harnwegsinfekten)?

Im Rahmen meiner Doktorarbeit interessiere ich mich für Ihre Beschwerden, vor allem aber wie sich **Ihre Beschwerden auf Ihr Leben und Ihre Lebensqualität** auswirken.

Ich würde Sie gerne dazu einladen, sich online über Ihre Beschwerden mit bis zu 10 weiteren betroffenen Frauen auszutauschen. Dazu hätte ich ein paar Fragen auf einem geschützten Portal für Sie vorbereitet.

Wenn Sie Lust haben, **von Ihren Erfahrungen zu berichten** und dazu **beitragen** möchten, **dass Ihre Probleme auch als solche wahrgenommen werden**, dann melden Sie sich bitte für weitere Informationen bei mir.

Michaela Gabes: 0941/944-5227 michaela.gabes@klinik.uni-regensburg.de

Als kleines Dankeschön erhält jede Teilnehmerin an der Online-Diskussion einen **15€-Amazon-Gutschein**.

Ich freue mich auf spannende Diskussionen mit Ihnen!

Michaela Gabes



Anhang F Mplus Syntax zur Überprüfung der Strukturvalidität des DIVA

Für alle Frauen:

```
TITLE: CFA;
DATA: FILE IS Mplus_full_sample_300720_resp_transf.dat;
VARIABLE: NAMES ARE
    DI02 DI03 DI04 DI05 DI06
    DI07 DI08 DI09 DI10
    DI11 DI12 DI13 DI14 DI15 DI16 DI17 DI18 DI19
    DI20 DI21 DI22 DI23 DI24;
    USE VARIABLES ARE
    DI02 DI03 DI04 DI05 DI06
    DI07 DI08 DI09 DI10
    DI11 DI12 DI17 DI18 DI19
    DI20 DI21 DI22 DI23 DI24;
MODEL: dom_1 BY DI02 DI03 DI04 DI05 DI06;
    dom_2 BY DI07 DI08 DI09 DI10;
    dom_3_short BY DI11 DI12 DI17 DI18 DI19;
    dom_4 BY DI20 DI21 DI22 DI23 DI24;
OUTPUT: STDYX;
```

Für kürzlich sexuell aktive Frauen:

```
TITLE: CFA;
DATA: FILE IS Mplus_sex_sample_300720_resp_transf.dat;
VARIABLE: NAMES ARE
    DI02 DI03 DI04 DI05 DI06
    DI07 DI08 DI09 DI10
    DI11 DI12 DI13 DI14 DI15 DI16 DI17 DI18 DI19
    DI20 DI21 DI22 DI23 DI24;
    USE VARIABLES ARE
    DI02 DI03 DI04 DI05 DI06
    DI07 DI08 DI09 DI10
    DI11 DI12 DI13 DI14 DI15 DI16 DI17 DI18 DI19
    DI20 DI21 DI22 DI23 DI24;
MODEL: dom_1 BY DI02 DI03 DI04 DI05 DI06;
    dom_2 BY DI07 DI08 DI09 DI10;
    dom_3_long BY DI11 DI12 DI13 DI14 DI15
    DI16 DI17 DI18 DI19;
    dom_4 BY DI20 DI21 DI22 DI23 DI24;
OUTPUT: STDYX;
```

Anhang G SPSS Syntax für die Aufbereitung der Daten und die Validierung des DIVA

```
RECODE SD02_01 (40 thru 49=2) (50 thru 59=3) (60 thru 69=4) (70 thru Highest=5) (Lowest thru 39=1)
  INTO age_category.
```

```
VARIABLE LABELS age_category 'age_category'.
```

```
EXECUTE.
```

```
*****
```

*Aufbereitung der DIVA Scores

```
RECODE DI13 DI14 DI15 DI16 (6=0) (1 thru 5=1) INTO sex_activity sex_activity_2 sex_activity_3
  sex_activity_4.
```

```
VARIABLE LABELS sex_activity 'sex_activity' /sex_activity_2 'sex_activity_2' /sex_activity_3
  'sex_activity_3' /sex_activity_4 'sex_activity_4'.
```

```
EXECUTE.
```

```
COMPUTE sexual_activity=sex_activity = 1 OR sex_activity_2 = 1 OR sex_activity_3 = 1 OR
  sex_activity_4 = 1.
```

```
EXECUTE.
```

```
*****
```

```
RECODE DI02 DI03 DI04 DI05 DI06 DI07 DI08 DI09 DI10 DI11 DI12 DI13 DI14 DI15 DI16 DI17 DI18 DI19
  DI20 DI21 DI22 DI23 DI24 (1=0) (2=1) (3=2) (4=3) (5=4) (6=0).
```

```
EXECUTE.
```

```
RECODE PH01_01 PH01_02 PH01_03 PH01_04 (-9=0) (1=0) (2=1) (3=2) (4=3).
```

```
EXECUTE.
```

```
COMPUTE Domain1=(DI02 + DI03 + DI04 + DI05 + DI06) / 5.
```

```
EXECUTE.
```

```
COMPUTE Domain2=(DI07 + DI08 + DI09 + DI10) / 4.
```

```
EXECUTE.
```

```
USE ALL.
```

```
COMPUTE filter_$=(sexual_activity = 1).
```

```
VARIABLE LABELS filter_$ 'sexual_activity = 1 (FILTER)'.
```

```
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
```

```
FORMATS filter_$ (f1.0).
```

```
FILTER BY filter_$.
```

```
EXECUTE.
```

```
COMPUTE Domain3_long=(DI11 + DI12 + DI13 + DI14 + DI15 + DI16 + DI17 + DI18 + DI19) / 9.
```

```
EXECUTE.
```

```
COMPUTE Total_long=(DI02 + DI03 + DI04 + DI05 + DI06 + DI07 + DI08 + DI09 + DI10 + DI11 + DI12 +
  DI13 + DI14 + DI15 + DI16 + DI17 + DI18 + DI19 + DI20 + DI21 + DI22 + DI23 + DI24) / 23.
```

```
EXECUTE.
```

```
FILTER OFF.
```

```
USE ALL.
```

```
EXECUTE.
```

```
COMPUTE Domain3_short=(DI11 + DI12 + DI17 + DI18 + DI19) / 5.
```

```
EXECUTE.
```

```
COMPUTE Total_short=(DI02 + DI03 + DI04 + DI05 + DI06 + DI07 + DI08 + DI09 + DI10 + DI11 + DI12 +
  DI17 + DI18 + DI19 + DI20 + DI21 + DI22 + DI23 + DI24) / 19.
```

```
EXECUTE.
```

```
COMPUTE Domain4=(DI20 + DI21 + DI22 + DI23 + DI24) / 5.
```

```
EXECUTE.
```

***Aufbereitung des PHQ-4**

```
COMPUTE PHQ_depression=PH01_01 + PH01_02.  
EXECUTE.
```

```
COMPUTE PHQ_anxiety=PH01_03 + PH01_04.  
EXECUTE.
```

```
COMPUTE PHQ_4_total=PHQ_depression + PHQ_anxiety.  
EXECUTE.
```

***Aufbereitung des SF-12**

```
RECODE SF01 (1=1) (2=2) (3=3) (4=4) (5=5) (-9=-9) INTO ghp1.  
VARIABLE LABELS ghp1 'ghp1'.  
EXECUTE.
```

```
RECODE SF02 (1=1) (2=2) (3=3) (-9=-9) INTO pfi02.  
VARIABLE LABELS pfi02 'pfi02'.  
EXECUTE.
```

```
RECODE SF02 (1=1) (2=2) (3=3) (-9=-9) INTO pfi04.  
VARIABLE LABELS pfi04 'pfi04'.  
EXECUTE.
```

```
RECODE SF04_01_1 SF04_02_1 SF04_03_1 SF04_04_1 SF04_05_1 SF04_06_1 SF04_07_1 SF04_08_1  
SF04_09_1  
SF04_10_1 (1=0) (2=1) INTO SF04_1 SF05_1 SF06_1 SF07_1 SF08_1 SF09_1 SF010_1 SF011_1 SF012_1  
SF013_1.  
VARIABLE LABELS SF04_1 'SF04_1'/SF05_1 'SF05_1'/SF06_1 'SF06_1'/SF07_1 'SF07_1'/SF08_1  
'SF08_1'/SF09_1 'SF09_1'/SF010_1 'SF010_1'/SF011_1 'SF011_1'/SF012_1 'SF012_1'/SF013_1  
'SF013_1'.  
EXECUTE.
```

```
RECODE SF04_01_2 SF04_02_2 SF04_03_2 SF04_04_2 SF04_05_2 SF04_06_2 SF04_07_2 SF04_08_2  
SF04_09_2  
SF04_10_2 (1=0) (2=2) INTO SF04_2 SF05_2 SF06_2 SF07_2 SF08_2 SF09_2 SF010_2 SF011_2 SF012_2  
SF013_2.  
VARIABLE LABELS SF04_2 'SF04_2'/SF05_2 'SF05_2'/SF06_2 'SF06_2'/SF07_2 'SF07_2'/SF08_2  
'SF08_2'/SF09_2 'SF09_2'/SF010_2 'SF010_2'/SF011_2 'SF011_2'/SF012_2 'SF012_2'/SF013_2  
'SF013_2'.  
EXECUTE.
```

```
RECODE SF04_01_3 SF04_02_3 SF04_03_3 SF04_04_3 SF04_05_3 SF04_06_3 SF04_07_3 SF04_08_3  
SF04_09_3  
SF04_10_3 (1=0) (2=3) INTO SF04_3 SF05_3 SF06_3 SF07_3 SF08_3 SF09_3 SF010_3 SF011_3 SF012_3  
SF013_3.  
VARIABLE LABELS SF04_3 'SF04_3'/SF05_3 'SF05_3'/SF06_3 'SF06_3'/SF07_3 'SF07_3'/SF08_3  
'SF08_3'/SF09_3 'SF09_3'/SF010_3 'SF010_3'/SF011_3 'SF011_3'/SF012_3 'SF012_3'/SF013_3  
'SF013_3'.  
EXECUTE.
```

```
RECODE SF04_01_4 SF04_02_4 SF04_03_4 SF04_04_4 SF04_05_4 SF04_06_4 SF04_07_4 SF04_08_4  
SF04_09_4  
SF04_10_4 (1=0) (2=4) INTO SF04_4 SF05_4 SF06_4 SF07_4 SF08_4 SF09_4 SF010_4 SF011_4 SF012_4  
SF013_4.  
VARIABLE LABELS SF04_4 'SF04_4'/SF05_4 'SF05_4'/SF06_4 'SF06_4'/SF07_4 'SF07_4'/SF08_4  
'SF08_4'/SF09_4 'SF09_4'/SF010_4 'SF010_4'/SF011_4 'SF011_4'/SF012_4 'SF012_4'/SF013_4  
'SF013_4'.  
EXECUTE.
```

```
RECODE SF04_01_5 SF04_02_5 SF04_03_5 SF04_04_5 SF04_05_5 SF04_06_5 SF04_07_5 SF04_08_5
SF04_09_5
SF04_10_5 (1=0) (2=5) INTO SF04_5 SF05_5 SF06_5 SF07_5 SF08_5 SF09_5 SF010_5 SF011_5 SF012_5
SF013_5.
```

```
VARIABLE LABELS SF04_5 'SF04_5' /SF05_5 'SF05_5' /SF06_5 'SF06_5' /SF07_5 'SF07_5' /SF08_5
'SF08_5' /SF09_5 'SF09_5' /SF010_5 'SF010_5' /SF011_5 'SF011_5' /SF012_5 'SF012_5' /SF013_5
'SF013_5'.
EXECUTE.
```

```
COMPUTE mhi4=SF05_1 + SF05_2 + SF05_3 + SF05_4 + SF05_5.
EXECUTE.
```

```
COMPUTE mhi3=SF06_1 + SF06_2 + SF06_3 + SF06_4 + SF06_5.
EXECUTE.
```

```
COMPUTE vital2=SF07_1 + SF07_2 + SF07_3 + SF07_4 + SF07_5.
EXECUTE.
```

```
COMPUTE pain2=SF08_1 + SF08_2 + SF08_3 + SF08_4 + SF08_5.
EXECUTE.
```

```
COMPUTE rolph2=SF09_1 + SF09_2 + SF09_3 + SF09_4 + SF09_5.
EXECUTE.
```

```
COMPUTE rolph3=SF010_1 + SF010_2 + SF010_3 + SF010_4 + SF010_5.
EXECUTE.
```

```
COMPUTE rolem2=SF011_1 + SF011_2 + SF011_3 + SF011_4 + SF011_5.
EXECUTE.
```

```
COMPUTE rolem3=SF012_1 + SF012_2 + SF012_3 + SF012_4 + SF012_5.
EXECUTE.
```

```
COMPUTE soc2=SF013_1 + SF013_2 + SF013_3 + SF013_4 + SF013_5.
EXECUTE.
```

```
FREQUENCIES VARIABLES=ghp1 pfi02 pfi04 mhi4 mhi3 vital2 pain2 rolph2 rolph3 rolem2 rolem3 soc2
/ORDER=ANALYSIS.
```

```
*****
```

```
*** STEP 1: Data cleaning and reverse coding (where necessary) ***.
```

```
*** Version 2: ROLPH2, ROLPH3, ROLEM2, ROLEM3: 5 cats (version 1: 2 cats) ***.
```

```
*** Version 2: VITAL2, MHI3, MHI4: 5 cats (version 1: 6 cats) ***.
```

```
*****
```

```
recode ROLPH2, ROLPH3, ROLEM2, ROLEM3 (lo thru 0.9, 5.1 thru hi,sysmis=9).
```

```
add val lab ROLPH2, ROLPH3, ROLEM2, ROLEM3 9 "k. Angabe".
```

```
mis val ROLPH2, ROLPH3, ROLEM2, ROLEM3 (9).
```

```
fre ROLPH2, ROLPH3, ROLEM2, ROLEM3.
```

```
recode PFI02, PFI04 (lo thru 0.9, 3.1 thru
```

```
hi,sysmis=9).
```

```
add val lab PFI02, PFI04 9 "k. Angabe".
```

```
mis val PFI02, PFI04 (9).
```

```
fre PFI02, PFI04.
```

```
recode GHP1, PAIN2, SOC2 (lo thru 0.9, 5.1 thru
```

```
hi,sysmis=9).
```

```
add val lab GHP1, PAIN2, SOC2 9 "k. Angabe".
```

```
mis val GHP1, PAIN2, SOC2 (9).
```

```
fre GHP1, PAIN2, SOC2.
```

```
recode VITAL2, MHI3, MHI4 (lo thru 0.9, 5.1 thru hi,sysmis=9).
```

```
add val lab VITAL2, MHI3, MHI4 9 "k. Angabe".
```

```

mis val VITAL2, MHI3, MHI4 (9).
fre VITAL2, MHI3, MHI4.
*****
*** Invert vars / reverse coding ***.
*** remark: PAIN2 is already asked in inversed matter in SOEP04 questionnaire ***.
*** (vs SF12v2 standard) ***.
*** reverse coding: GHP1, VITAL2, MHI3 ***.
*** recalibration GHP1 (General health) is not done due to missing calibration data ***.
*****
COMPUTE RGHP1 = 6 - GHP1.
COMPUTE RVITAL2 = 6 - VITAL2.
COMPUTE RMHI3 = 6 - MHI3.
des var = rghp1 rvital2 rmhi3.
*****
*** STEP 2: Building 8 Subscales of SF12v2 ***.
*** 8 Subscales out of 12 Items (4*1 Item, 4*2 Items) ***.
*** a) 8 Sumscores mit Transformation 0-100 ***.
*** b) Transformation of 8 Sumscores using SOEP04 results ***.
*** remark: (no complex missing data estimation, since algorithms are not available in ***.
*** SF12v2-manual or from sf-website) ***.
*** All cases with missing data are excluded from analysis (from the beginning) ***.
*****
compute pf100 = ((mean.2(pfi02,pfi04)-1)/2)*100.
compute rp100 = ((mean.2(rolph2,rolph3)-1)/4)*100.
compute re100 = ((mean.2(rolem2,rolem3)-1)/4)*100.
compute mh100 = ((mean.2(mhi4,rmhi3)-1)/4)*100.
compute bp100 = ((mean.1(pain2)-1)/4)*100.
compute gh100 = ((mean.1(rghp1)-1)/4)*100.
compute vt100 = ((mean.1(rvital2)-1)/4)*100.
compute sf100 = ((mean.1(soc2)-1)/4)*100.
VARIABLE LABELS pf100 "Physical functioning (0-100 score value, 2 vars)".
VARIABLE LABELS rp100 "Role physical (0-100 score value, 2 vars)".
VARIABLE LABELS bp100 "Bodily pain (0-100 score value, 1 var)".
VARIABLE LABELS gh100 "General health (0-100 score value, 1 var)".
VARIABLE LABELS vt100 "Vitality (0-100 score value, 1 var)".
VARIABLE LABELS sf100 "Social functioning (0-100 score value, 1 var)".
VARIABLE LABELS re100 "Role emotional (0-100 score value, 2 vars)".
VARIABLE LABELS mh100 "Mental health (0-100 score value, 2 vars)".
des var = pf100 rp100 re100 mh100 bp100 gh100 vt100 sf100.
*****
*** transformation of values of 8 scales using means and sd's of SOEP04 norm population ***.
compute pf_z = (pf100 - 73.14) / 32.15.
compute rp_z = (rp100 - 74.29) / 26.46.
compute bp_z = (bp100 - 73.27) / 27.14.
compute gh_z = (gh100 - 60.57) / 24.08.
compute vt_z = (vt100 - 52.94) / 22.84.
compute sf_z = (sf100 - 83.22) / 23.57.
compute re_z = (re100 - 81.94) / 22.35.
compute mh_z = (mh100 - 61.96) / 20.47.
des var = pf_z rp_z bp_z gh_z vt_z sf_z re_z mh_z.
*****
*** transformation of 8 z-transformed scales using norm based scoring (mean=50, sd=10) ***.
compute pf_nbs = (pf_z*10)+50.
compute rp_nbs = (rp_z*10)+50.
compute bp_nbs = (bp_z*10)+50.
compute gh_nbs = (gh_z*10)+50.
compute vt_nbs = (vt_z*10)+50.
compute sf_nbs = (sf_z*10)+50.
compute re_nbs = (re_z*10)+50.
compute mh_nbs = (mh_z*10)+50.

```



```

VARIABLE LABELS pf_nbs "Physical functioning (NBS)".
VARIABLE LABELS rp_nbs "Role physical (NBS)".
VARIABLE LABELS bp_nbs "Bodily pain (NBS)".
VARIABLE LABELS gh_nbs "General health (NBS)".
VARIABLE LABELS vt_nbs "Vitality (NBS)".
VARIABLE LABELS sf_nbs "Social functioning (NBS)".
VARIABLE LABELS re_nbs "Role emotional (NBS)".
VARIABLE LABELS mh_nbs "Mental health (NBS)".
des var = pf_nbs rp_nbs bp_nbs gh_nbs vt_nbs sf_nbs mh_nbs re_nbs .
*****
*** STEP 3:Calculation of aggregate scores PCS and MCS ***.
*** 1.calculation of summary scores using factor score coefficients from factor analysis ***.
*** (PCA, Varimax) with SOEP04 population ***.
*** Note: calculation of scores for SF12v2 is different from SF12v1 ***.
*** transformation of summary scores on norm based scoring. (mean=50, sd=10) ***.
*****
compute agg_phys = (0.414 * pf_z) + (0.279 * rp_z) + (0.331 * bp_z) + (0.330 * gh_z) + (-0.041 * vt_z) + (-0.068
* sf_z) + (-0.110 * re_z) + (-0.244 * mh_z).
compute agg_ment = (-0.209 * pf_z) + (-0.021 * rp_z) + (-0.105 * bp_z) + (-0.103 * gh_z) + (0.258 * vt_z) +
(0.333 * sf_z) + (0.378 * re_z) + (0.489 * mh_z).
des var = agg_phys agg_ment.
compute pcs = (agg_phys*10)+50.
compute mcs = (agg_ment*10)+50.
VARIABLE LABELS pcs "PCS: Summary scale Physical (NBS)".
VARIABLE LABELS mcs "MCS: Summary scale Mental (NBS)".
des var = pcs mcs.
*****

```

*Validierungsanalysen

*internal consistency

**domain 1

```

RELIABILITY
/VARIABLES=DI02 DI03 DI04 DI05 DI06
/SCALE('ALL VARIABLES') ALL
/MODEL=ALPHA
/STATISTICS=DESCRIPTIVE SCALE
/SUMMARY=TOTAL.

```

**domain 2

```

RELIABILITY
/VARIABLES=DI07 DI08 DI09 DI10
/SCALE('ALL VARIABLES') ALL
/MODEL=ALPHA
/STATISTICS=DESCRIPTIVE SCALE
/SUMMARY=TOTAL.

```

```

USE ALL.
COMPUTE filter_$=(sexual_activity = 1).
VARIABLE LABELS filter_$ 'sexual_activity = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.

```

**domain 3 long-form

```

RELIABILITY
/VARIABLES=DI11 DI12 DI13 DI14 DI15 DI16 DI17 DI18 DI19
/SCALE('ALL VARIABLES') ALL
/MODEL=ALPHA

```

```
/STATISTICS=DESCRIPTIVE SCALE
/SUMMARY=TOTAL.
```

****total long-form**

RELIABILITY

```
/VARIABLES=DI02 DI03 DI04 DI05 DI06 DI07 DI08 DI09 DI10 DI11 DI12 DI13 DI14 DI15 DI16 DI17 DI18
DI19 DI20 DI21 DI22
DI23 DI24
/SCALE('ALL VARIABLES') ALL
/MODEL=ALPHA
/STATISTICS=DESCRIPTIVE SCALE
/SUMMARY=TOTAL.
```

```
FILTER OFF.
USE ALL.
EXECUTE.
```

****domain 3 short-form**

RELIABILITY

```
/VARIABLES=DI11 DI12 DI17 DI18 DI19
/SCALE('ALL VARIABLES') ALL
/MODEL=ALPHA
/STATISTICS=DESCRIPTIVE SCALE
/SUMMARY=TOTAL.
```

****domain 4**

RELIABILITY

```
/VARIABLES=DI20 DI21 DI22 DI23 DI24
/SCALE('ALL VARIABLES') ALL
/MODEL=ALPHA
/STATISTICS=DESCRIPTIVE SCALE
/SUMMARY=TOTAL.
```

****total short-form**

RELIABILITY

```
/VARIABLES=DI02 DI03 DI04 DI05 DI06 DI07 DI08 DI09 DI10 DI11 DI12 DI17 DI18 DI19 DI20 DI21 DI22
DI23 DI24
/SCALE('ALL VARIABLES') ALL
/MODEL=ALPHA
/STATISTICS=DESCRIPTIVE SCALE
/SUMMARY=TOTAL.
```

***convergent validity**

CORRELATIONS

```
/VARIABLES=MR01_08 MR01_09 MR01_10 AN02 AN03 AN04 AN05 AN06 Domain1 Domain2
Domain3_short Domain4 Total_short PHQ_depression PHQ_anxiety PHQ_4_total pf_nbs rp_nbs
bp_nbs gh_nbs vt_nbs sf_nbs re_nbs mh_nbs pcs mcs
/PRINT=TWOTAIL NOSIG
/STATISTICS DESCRIPTIVES
/MISSING=PAIRWISE.
```

NONPAR CORR

```
/VARIABLES=MR01_08 MR01_09 MR01_10 AN02 AN03 AN04 AN05 AN06 Domain1 Domain2
Domain3_short Domain4 Total_short PHQ_depression PHQ_anxiety PHQ_4_total pf_nbs rp_nbs
bp_nbs gh_nbs vt_nbs sf_nbs re_nbs mh_nbs pcs mcs
/PRINT=SPEARMAN TWOTAIL NOSIG
/MISSING=PAIRWISE.
```

```
USE ALL.
COMPUTE filter_$=(sexual_activity = 1).
```

```
VARIABLE LABELS filter_$ 'sexual_activity = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

CORRELATIONS

```
/VARIABLES=MR01_08 MR01_09 MR01_10 AN02 AN03 AN04 AN05 AN06 Domain3_long Total_long
  PHQ_depression PHQ_anxiety PHQ_4_total pf_nbs rp_nbs bp_nbs gh_nbs vt_nbs sf_nbs re_nbs mh_nbs pcs
  mcs
/PRINT=TWOTAIL NOSIG
/STATISTICS DESCRIPTIVES
/MISSING=PAIRWISE.
```

NONPAR CORR

```
/VARIABLES=MR01_08 MR01_09 MR01_10 AN02 AN03 AN04 AN05 AN06 Domain3_long Total_long
  PHQ_depression PHQ_anxiety PHQ_4_total pf_nbs rp_nbs bp_nbs gh_nbs vt_nbs sf_nbs re_nbs mh_nbs pcs
  mcs
/PRINT=SPEARMAN TWOTAIL NOSIG
/MISSING=PAIRWISE.
```

```
FILTER OFF.
USE ALL.
EXECUTE.
```

*discriminant validity

**by age category

NPART TESTS

```
/K-W=Domain1 Domain2 Domain3_short Domain4 Total_short BY age_category(1
  5)
/MISSING ANALYSIS.
```

USE ALL.

```
COMPUTE filter_$(sexual_activity = 1).
VARIABLE LABELS filter_$ 'sexual_activity = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

NPART TESTS

```
/K-W=Domain3_long Total_long BY age_category(1
  5)
/MISSING ANALYSIS.
```

```
FILTER OFF.
USE ALL.
EXECUTE.
```

**by disease severity (MRS item 10)

NPART TESTS

```
/K-W=Domain1 Domain2 Domain3_short Domain4 Total_short BY MR01_10(1 5)
/MISSING ANALYSIS.
```

USE ALL.

```
COMPUTE filter_$(sexual_activity = 1).
VARIABLE LABELS filter_$ 'sexual_activity = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

EXECUTE.

NPAR TESTS

/K-W=Domain3_long Total_long BY MR01_10(1 5)
/MISSING ANALYSIS.

FILTER OFF.

USE ALL.

EXECUTE.

****by disease severity (AN02)**

NPAR TESTS

/K-W=Domain1 Domain2 Domain3_short Domain4 Total_short BY AN02(1 5)
/STATISTICS DESCRIPTIVES
/MISSING ANALYSIS.

USE ALL.

COMPUTE filter_\$=(sexual_activity = 1).

VARIABLE LABELS filter_\$ 'sexual_activity = 1 (FILTER)'.
VALUE LABELS filter_\$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_\$ (f1.0).
FILTER BY filter_\$.

FORMATS filter_\$ (f1.0).

FILTER BY filter_\$.

EXECUTE.

NPAR TESTS

/K-W=Domain3_long Total_long BY AN02(1 5)
/STATISTICS DESCRIPTIVES
/MISSING ANALYSIS.

FILTER OFF.

USE ALL.

EXECUTE.

****by country (SD07)**

NPAR TESTS

/K-W=Domain1 Domain2 Domain3_short Domain4 Total_short BY SD07(1 3)
/MISSING ANALYSIS.

USE ALL.

COMPUTE filter_\$=(sexual_activity = 1).

VARIABLE LABELS filter_\$ 'sexual_activity = 1 (FILTER)'.
VALUE LABELS filter_\$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_\$ (f1.0).
FILTER BY filter_\$.

FORMATS filter_\$ (f1.0).

FILTER BY filter_\$.

EXECUTE.

NPAR TESTS

/K-W=Domain3_long Total_long BY SD07(1 3)
/MISSING ANALYSIS.

FILTER OFF.

USE ALL.

EXECUTE.

Anhang H Teilnehmerinformation für die Online-Fokusgruppe

Liebe Frauen,

vielen Dank, dass Sie sich für die Teilnahme an einer Gruppendiskussion über Beschwerden im Rahmen der Wechseljahre interessieren. Sicherlich fragen Sie sich, wie so eine Diskussion online ablaufen soll.

Hier ein paar Informationen zur Diskussion an sich:

- Dauer: Die Fragen sind über mehrere Tage im Portal für Sie zugänglich, d. h. Sie sind mit der Beantwortung und Diskussion der Fragen zeitlich flexibel.
- Teilnehmer: 5-8 Frauen in den Wechseljahren + Moderator (Michaela Gabes)
- Alle Fragen und Antworten werden getippt.
- Ich werde Fragen zu Ihren persönlichen Erfahrungen ins Portal stellen, die jede Teilnehmerin direkt unter der Frage beantworten kann. Ich würde mich freuen, wenn Sie die Fragen möglichst zeitnah, vollständig und ausführlich beantworten könnten. Bei Unklarheiten würde ich einfach nachfragen!
- Sie müssen sich nicht bei Ihrem echten Namen nennen! Aus datenschutzrechtlichen Gründen wäre es sogar wünschenswert, wenn Sie unter einem Pseudonym/Decknamen an der Gruppendiskussion teilnehmen würden. Alles, was Sie schreiben, wird vertraulich behandelt, d. h. weder Ihr Arzt bzw. Ihre Ärztin noch irgendjemand anderes wird über diese Diskussion informiert. Es kann sein, dass ich einzelne Zitate von Ihnen benutzen werde, allerdings nicht in Zusammenhang mit Ihrem Namen. Bitte machen Sie sich klar, dass jede Teilnehmerin der Diskussion sehen kann, was Sie schreiben. Versuchen Sie bitte, vertrauliche Informationen, wie z. B. Ihre Adresse etc., nicht zu benutzen.
- Bitte respektieren Sie sich gegenseitig in der Gruppe. Ich hoffe, dass durch meine Fragen eine gemeinsame Diskussion entsteht. Es gibt keine richtigen oder falschen Antworten – ich interessiere mich einfach für Ihre ganz persönlichen Erfahrungen und Meinungen!

Mit diesem Link gelangen Sie zum Portal:

<https://focusgroupit.com/groups/7f9ea43b>

Klicken Sie auf „Dieser Fokusgruppe beitreten“ und melden sich an. Nachdem Sie sich angemeldet haben, können Sie meine Fragen sehen und beantworten! Ich freue mich auf Ihre Beiträge! 😊

Bei technischen Problemen schreiben Sie mir bitte eine E-Mail (michaela.gabes@klinik.uni-regensburg.de).

Wenn alle Fragen ausreichend diskutiert sind und es keine weiteren Fragen mehr gibt, werde ich Ihnen als Dank einen 15€ Amazon-Gutschein per E-Mail zukommen lassen!



Anhang I Teilnehmerinformation, Einverständniserklärung und kurzer Fragebogen für die Durchführung der *face-to-face* Fokusgruppen

Information für Teilnehmerinnen der Gruppendiskussion „Urogenitale Beschwerden im Rahmen der Menopause“

Liebe Teilnehmerin,

vielen Dank, dass Sie sich für die Teilnahme an einer Gruppendiskussion, einer sogenannten Fokusgruppe, über Beschwerden im Rahmen der Wechseljahre interessieren. Mithilfe dieser Fokusgruppe soll ersichtlich werden, wie urogenitale Beschwerden Ihren Alltag beeinflussen. Mich interessieren vor allem Ihre persönlichen Erfahrungen, Einschätzungen und Ansichten. Ich würde mich freuen, wenn Sie meine Fragen möglichst vollständig und ausführlich beantworten.

Durch die Teilnahme an der Gruppendiskussion entstehen für Sie keine wesentlichen Belastungen oder Risiken. Alle von Ihnen gemachten Angaben dienen ausschließlich wissenschaftlichen Zwecken und werden nur im Rahmen meiner Doktorarbeit verwendet.

Das Gespräch wird ca. **70 – 80 Minuten** dauern.

Alles, was Sie sagen, wird **vertraulich** behandelt, d. h. weder Ihr Arzt bzw. Ihre Ärztin noch irgendjemand anderes wird über diese Diskussion informiert. Es kann sein, dass ich einzelne Zitate von Ihnen benutzen werde, allerdings nicht in Zusammenhang mit Ihrem Namen. Bitte machen Sie sich klar, dass jeder in dieser Diskussion hören kann, was Sie sagen. Versuchen Sie bitte, vertrauliche Informationen, wie z. B. Ihre Adresse etc., nicht zu benutzen.

Die Fokusgruppendifkussion wird audiographisch mit Hilfe eines Aufnahmegerätes aufgezeichnet und anschließend in anonymer Form verschriftlicht.

Die gesammelten Informationen (Daten) aus dem Gespräch werden zur Auswertung in einer Datenbank zusammengefasst. Angaben, mit denen ein Rückschluss auf Ihre Person möglich ist, werden getrennt von den erhobenen Daten gespeichert. Eine Weitergabe zu anderen Zwecken als meiner Doktorarbeit erfolgt nicht.

Ihre Teilnahme ist **freiwillig** und kann jederzeit (auch während des Gesprächs) abgebrochen werden. Die Nichtteilnahme an der Studie hat keinerlei negative Folgen für Sie. Ebenso können Sie Ihre Einwilligung zur Teilnahme an der Studie jederzeit ohne Angabe von Gründen widerrufen, ohne dass Ihnen daraus irgendwelche Nachteile entstehen.

Bitte **respektieren Sie sich** gegenseitig in der Gruppe. Ich hoffe, dass durch meine Fragen eine gemeinsame Diskussion entsteht. Es gibt keine richtigen oder falschen Antworten – ich interessiere mich einfach für Ihre ganz persönlichen Erfahrungen und Meinungen!

Bevor die Diskussion beginnen kann, brauche ich Ihr **Einverständnis**. Ich würde zudem gerne einige persönliche Informationen zu Ihnen erfassen (Alter, Beschwerden, etc.). Bitte füllen Sie die nachfolgenden Fragen wahrheitsgemäß aus. Nachdem diese Gruppendiskussion beendet ist, werde ich Ihnen als Dank einen 15€-Amazon-Gutschein per E-Mail zukommen lassen! Bei Rückfragen können Sie sich gerne bei mir melden (michaela.gabes@klinik.uni-regensburg.de).

Bitte willigen Sie ein, dass Sie mit der Teilnahme und der Datenspeicherung einverstanden sind.

Vielen Dank!

Hiermit willige ich,

Vorname, Nachname

in die Teilnahme an einer Fokusgruppe für die o.g. Studie ein.

Ich bestätige, dass meine Teilnahme an der Fokusgruppe freiwillig erfolgt.

Ich bestätige, dass ich die Informationen zur Europäischen Datenschutz-Grundverordnung erhalten und verstanden habe.

Ich erkläre mich damit einverstanden, dass im Rahmen dieser Fokusgruppe mich betreffende personenbezogene Daten/Angaben erhoben, verschlüsselt (pseudonymisiert) auf elektronischen Datenträgern aufgezeichnet und verarbeitet werden dürfen. Ich bin damit einverstanden, dass die Studienergebnisse in nicht rück-entschlüsselbarer (anonymer) Form veröffentlicht werden.

Mir ist bekannt, dass ich jederzeit die Löschung meiner Daten verlangen kann.

Ort, Datum, Unterschrift der Teilnehmerin

1. Ihr **Alter** in Jahren: _____
2. Ihre **Muttersprache**: Deutsch Sonstige: _____
3. Bitte beschreiben Sie Ihren **Menstruationszyklus** (= Zeitraum von einer bis zur nächsten Blutung):
 - Mein Zyklus ist regelmäßig, d. h. die Periode kommt ca. 1-mal pro Monat
 - Mein Zyklus ist unregelmäßig, d. h. kürzer als 21 Tage oder länger als 35 Tage, aber keine Pausen von mehr als 2 Monaten
 - Mein Zyklus ist unregelmäßig, d. h. es treten Pausen von mindestens 2 Monaten auf
 - Ich habe seit mindestens 1 Jahr keine Periode mehr gehabt, obwohl meine Gebärmutter da ist
 - Ich habe keine Gebärmutter mehr

4. Nehmen Sie **Hormone** ein?
 - Nein
 - Ja, zur Schwangerschaftsverhütung (z. B. „Pille“, Hormonstäbchen, Dreimonatsspritze, Hormonspirale)
 - Ja, zur Behandlung von Wechseljahresbeschwerden (z. B. Tablette, Pflaster, Gel, Spray)

5. Leiden Sie an **vulvovaginalen Beschwerden** im Rahmen der Wechseljahre?

Zu vulvovaginalen Beschwerden zählen z. B. vaginale Trockenheit, Brennen, Juckreiz oder Schmerzen im Intimbereich.

- ja
- nein

falls ja:

- An welchen vulvovaginalen Beschwerden leiden Sie? (mehrere Antworten möglich)

- vaginale Trockenheit
- Brennen
- Juckreiz
- Schmerzen im Intimbereich
- Sonstige: _____

- Wie würden Sie den Schweregrad Ihrer vulvovaginalen Beschwerden einschätzen?

- leicht
- mittel
- stark
- sehr stark

6. Leiden Sie an **urologischen Beschwerden** im Rahmen der Wechseljahre?

Zu urologischen Beschwerden zählen z. B. vermehrter Harndrang, schmerzhaftes und/oder erschwertes Wasserlassen, Inkontinenz oder vermehrte Harnwegsinfektionen.

- ja
- nein

falls ja:

- An welchen urologischen Beschwerden leiden Sie? (mehrere Antworten möglich)
 - vermehrter Harndrang (ständiges Bedürfnis, Wasser lassen zu müssen)
 - schmerzhaftes und/oder erschwertes Wasserlassen
 - Inkontinenz
 - vermehrte Harnwegsinfektionen
 - Sonstige: _____
- Wie würden Sie den Schweregrad Ihrer urologischen Beschwerden einschätzen?
 - leicht
 - mittel
 - stark
 - sehr stark

Wenn ich Ihnen als Dankeschön einen 15€-Amazon-Gutschein zusenden darf, hinterlassen Sie bitte Ihre E-Mail-Adresse:

Vielen Dank!

Anhang J Leitfaden für die Durchführung der Fokusgruppen

Allgemeine Einführung	
<p>Hallo, mein Name ist Michaela Gabes. Ich werde die Diskussion leiten.</p> <p>Ich würde Sie darum bitten, sich gegenseitig in der Gruppe zu respektieren. Ich hoffe, dass durch meine Fragen eine gemeinsame Diskussion entsteht. Es gibt keine richtigen oder falschen Antworten – ich interessiere mich einfach für Ihre ganz persönlichen Erfahrungen und Meinungen! Alles, was Sie sagen, wird vertraulich behandelt.</p> <p>Sollten Sie eine Frage nicht verstehen, fragen Sie gerne nach!</p>	
Einführung ins Thema	
<p>In meiner Doktorarbeit interessiere ich mich dafür, welche Beschwerden bei Frauen in den Wechseljahren auftreten und wie diese sich auf ihre Lebensqualität auswirken. Mit Ihnen möchte ich vor allem urologische Beschwerden diskutieren. Zu den urologischen Beschwerden gehören vermehrter Harndrang (ständiges Bedürfnis, Wasser lassen zu müssen), schmerzhaftes und/oder erschwertes Wasserlassen, Inkontinenz und vermehrte Infektionen der Harnwege. In dieser Diskussion sind Sie als Betroffene gefragt!</p>	
Erfahrungen der Patientinnen	
	<p>Persönliche Erfahrungen mit urologischen Beschwerden</p> <p>Unter welchen urologischen Beschwerden (vermehrter Harndrang, schmerzhaftes und/oder erschwertes Wasserlassen, Inkontinenz und vermehrte Infektionen der Harnwege) leiden Sie seit Beginn der Wechseljahre und welche dieser Beschwerden schränken Sie am meisten ein?</p>
Thema 1	<p>Alltägliche Aktivitäten:</p> <p>Wie sehr beeinträchtigen Ihre urologischen Beschwerden Ihre alltäglichen Aktivitäten? Welche alltäglichen Aktivitäten werden besonders beeinträchtigt?</p> <p><i>(Trigger: längeres Sitzen, spazieren gehen, Fahrrad fahren, Kleiderauswahl; Rituale, um die Beschwerden zu lindern/zu ertragen)</i></p>
Thema 2	<p>Emotionales Wohlbefinden:</p> <p>Wie sehr beeinträchtigen Ihre urologischen Beschwerden Ihr emotionales Wohlbefinden? Wie fühlen Sie sich aufgrund Ihrer Beschwerden?</p> <p><i>(Trigger: Depression, Frustration, Scham, Angst, Sorgen)</i></p>
Thema 3	<p>Sexuelle Funktionsfähigkeit:</p> <p>Wie sehr beeinträchtigen Ihre urologischen Beschwerden Ihre sexuelle Funktionsfähigkeit? Haben Sie noch Interesse an sexuellen Aktivitäten? Erleben Sie</p>

	<p>noch Befriedigung bei jeglicher sexuellen Aktivität? Sind Sie mit Ihrem Sexualleben zufrieden? Was wünschen Sie sich?</p> <p><i>(Trigger: Schuldgefühle, Vermeidungsstrategien, Überspielen der Beschwerden)</i></p>
Thema 4	<p>Selbstkonzept und Körperbild:</p> <p>Wie beeinflussen Ihre urologischen Beschwerden Ihr Körperbild? Wie nehmen Sie sich selbst wahr?</p> <p><i>(Trigger: alt werden, abnehmende „Weiblichkeit“, sozialer Druck)</i></p>
Thema 5	<p>Zwischenmenschliche Beziehungen und Kommunikation:</p> <p>Wie sehr beeinträchtigen Ihre urologischen Beschwerden Ihre zwischenmenschlichen Beziehungen (Freundschaften, Partnerschaft, Familie)? Sprechen Sie über Ihre Beschwerden? Wenn ja, mit wem? Wenn nein, warum nicht?</p> <p><i>(Trigger: Überwindung; Angst davor, ausgelacht/nicht ernst genommen zu werden; Scham, Isolation)</i></p>
Thema 6	<p>Erwartungshaltung bezüglich urologischer Beschwerden:</p> <p>Haben Sie erwartet, solche Beschwerden zu entwickeln und warum? Inwieweit wussten Sie zu Beginn der Wechseljahre über diese Beschwerden Bescheid? Was glauben Sie, sind die Ursachen für Ihre Beschwerden?</p>
Thema 7	<p>Aufklärung von ärztlicher Seite:</p> <p>Ist es Ihnen schwergefallen, Ihre Beschwerden bei einem Arzt bzw. Ihrer Ärztin anzusprechen? Hatten Sie das Gefühl, dass sich Ihr Arzt bzw. Ihre Ärztin mit den Beschwerden auskennt? Fühlten Sie sich ernst genommen?</p>
Ende	
<p>Vielen Dank, dass Sie an dieser Diskussion teilgenommen haben. Gibt es noch etwas, was Sie der Diskussion hinzufügen möchten? Haben Sie noch das Gefühl, etwas mitteilen zu möchten, bevor wir uns verabschieden? Ihre Meinungen sind mir sehr wichtig. Als Dank werde ich Ihnen allen einen 15€-Amazon-Gutschein zuschicken, den Sie per E-Mail in den nächsten Wochen erhalten werden. Auf Wiedersehen!</p>	

Literaturverzeichnis

1. Berwick DM. Medical associations: guilds or leaders? *BMJ*. 1997;314(7094):1564-5.
2. Mercieca-Bebber R, King MT, Calvert MJ, Stockler MR, Friedlander M. The importance of patient-reported outcomes in clinical trials and strategies for future optimization. *Patient Relat Outcome Meas*. 2018;9:353-67.
3. US Food and Drug Administration. Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Available at: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims>. 2009.
4. Kamudoni P, Nutjaree J, Salek S. *Living with Chronic Disease: Measuring Important Patient-Reported Outcomes*. Singapore: Springer; 2018.
5. Au HJ, Ringash J, Brundage M, Palmer M, Richardson H, Meyer RM. Added value of health-related quality of life measurement in cancer clinical trials: the experience of the NCIC CTG. *Expert Rev Pharmacoecon Outcomes Res*. 2010;10(2):119-28.
6. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. *Journal of Evaluation in Clinical Practice*. 2006;12(5):559-68.
7. Nelson EC, Eftimovska E, Lind C, Hager A, Wasson JH, Lindblad S. Patient reported outcome measures in practice. *BMJ*. 2015;350:g7818.
8. Hostettler S, Kraft E, Bosshard C. Patient-reported outcome measures: Die Patientensicht zählt. *Schweiz Ärztztg*. 2018;99(40):1348-52.
9. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet*. 2014;383(9912):166-75.
10. Mookink LB, Prinsen CA, Patrick DL, Alonso J, Bouter LM, De Vet HC, et al. COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs) - user manual. Available at: <https://www.cosmin.nl/tools/guideline-conducting-systematic-review-outcome-measures/>. 2018.
11. De Vet HC, Terwee CB, Mookink LB, Knol DL. *Measurement in Medicine*. New York: Cambridge University Press; 2011.
12. Terwee CB, Prinsen CA, Chiarotto A, de Vet HC, Bouter LM, Alonso J, et al. COSMIN methodology for assessing the content validity of Patient-Reported Outcome Measures (PROMs). User manual. Available at: <https://www.cosmin.nl/tools/guideline-conducting-systematic-review-outcome-measures/>. 2017.
13. Mookink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737-45.
14. Coyne KS, Wyrwich KW. ISPOR Task Force For Clinical Outcomes Assessment: Clinical Outcome Assessments: Conceptual Foundation-Report of The ISPOR Clinical Outcomes Assessment - Emerging Good Practices For Outcomes Research Task Force. *Value Health*. 2015;18(6):739-40.
15. Portman DJ, Gass ML. Genitourinary syndrome of menopause: new terminology for vulvovaginal atrophy from the International Society for the Study of Women's Sexual Health and the North American Menopause Society. *Menopause*. 2014;21(10):1063-8.
16. Mishra GD. Menopause, A Stage in the Life of Women. In: Cano A, editor. *Menopause: A Comprehensive Approach*: Springer International Publishing; 2017. p. 3-8.
17. Zablocki-Thomas P. The post-reproductive lifespan: evolutionary perspectives. *BioSciences Master Reviews*. 2013.
18. Casper RF. Clinical manifestations and diagnosis of menopause. In: Post TW, editor. *UpToDate*. Waltham, MA. 2020.
19. Takahashi M, Singh RS, Stone J. A Theory for the Origin of Human Menopause. *Front Genet*. 2016;7:222.
20. GBD Mortality Collaborators. Global, regional, and national age-sex-specific mortality and life expectancy, 1950-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018;392(10159):1684-735.
21. Panay N. Genitourinary syndrome of the menopause--dawn of a new era? *Climacteric*. 2015;18 Suppl 1:13-7.
22. Faubion SS, Sood R, Kapoor E. Genitourinary Syndrome of Menopause: Management Strategies for the Clinician. *Mayo Clin Proc*. 2017;92(12):1842-9.
23. Parish SJ, Nappi RE, Krychman ML, Kellogg-Spaldt S, Simon JA, Goldstein JA, et al. Impact of vulvovaginal health on postmenopausal women: a review of surveys on symptoms of vulvovaginal atrophy. *Int J Womens Health*. 2013;5:437-47.
24. Farrell AM E. Genitourinary syndrome of menopause. *Aust Fam Physician*. 2017;46(7):481-4.

25. Gandhi J, Chen A, Dagur G, Suh Y, Smith N, Cali B, et al. Genitourinary syndrome of menopause: an overview of clinical manifestations, pathophysiology, etiology, evaluation, and management. *Am J Obstet Gynecol.* 2016;215(6):704-11.
26. Nappi RE, Palacios S. Impact of vulvovaginal atrophy on sexual health and quality of life at postmenopause. *Climacteric.* 2014;17(1):3-9.
27. Nappi RE, Lachowsky M. Menopause and sexuality: prevalence of symptoms and impact on quality of life. *Maturitas.* 2009;63(2):138-41.
28. Palma F, Volpe A, Villa P, Cagnacci A. Vaginal atrophy of women in postmenopause. Results from a multicentric observational study: The AGATA study. *Maturitas.* 2016;83:40-4.
29. Bachmann GA. Vulvo-vaginal complaints. In: Lobo R, editor. *Treatment of the Postmenopausal Woman.* New York. Raven Press; 1994.
30. Nappi RE, Martini E, Cucinella L, Martella S, Tiranini L, Inzoli A, et al. Addressing Vulvovaginal Atrophy (VVA)/Genitourinary Syndrome of Menopause (GSM) for Healthy Aging in Women. *Frontiers in endocrinology.* 2019;10:561.
31. von Wolff M, Stute P. *Gynäkologische Endokrinologie und Reproduktionsmedizin: Das Praxisbuch.* Stuttgart: Schattauer; 2013. 249-70.
32. Davila GW, Singh A, Karapanagiotou I, Woodhouse S, Huber K, Zimberg S, et al. Are women with urogenital atrophy symptomatic? *Am J Obstet Gynecol.* 2003;188(2):382-8.
33. Thomas HN, Hess R, Thurston RC. Correlates of Sexual Activity and Satisfaction in Midlife and Older Women. *Ann Fam Med.* 2015;13(4):336-42.
34. Thornton K, Chervenak J, Neal-Perry G. Menopause and Sexuality. *Endocrinol Metab Clin North Am.* 2015;44(3):649-61.
35. Lonnee-Hoffmann RA, Dennerstein L, Leher P, Szoeko C. Sexual function in the late postmenopause: a decade of follow-up in a population-based cohort of Australian women. *J Sex Med.* 2014;11(8):2029-38.
36. Simon JA, Kokot-Kierepa M, Goldstein J, Nappi RE. Vaginal health in the United States: results from the Vaginal Health: Insights, Views & Attitudes survey. *Menopause.* 2013;20(10):1043-8.
37. Simon JA, Nappi RE, Kingsberg SA, Maamari R, Brown V. Clarifying Vaginal Atrophy's Impact on Sex and Relationships (CLOSER) survey: emotional and physical impact of vaginal discomfort on North American postmenopausal women and their partners. *Menopause.* 2014;21(2):137-42.
38. Nappi RE, Palacios S, Bruyniks N, Particco M, Panay N, investigators ES. The burden of vulvovaginal atrophy on women's daily living: implications on quality of life from a face-to-face real-life survey. *Menopause.* 2019;26(5):485-91.
39. Hunter MM, Nakagawa S, Van Den Eeden SK, Kuppermann M, Huang AJ. Predictors of impact of vaginal symptoms in postmenopausal women. *Menopause.* 2016;23(1):40-6.
40. The 2020 genitourinary syndrome of menopause position statement of The North American Menopause Society. *Menopause.* 2020;27(9):976-92.
41. Aguiar LB, Politano CA, Costa-Paiva L, Juliato CRT. Efficacy of Fractional CO₂ Laser, Promestriene, and Vaginal Lubricant in the Treatment of Urinary Symptoms in Postmenopausal Women: A Randomized Clinical Trial. *Lasers Surg Med.* 2020;52(8):713-20.
42. Bumphenkiatikul T, Panyakhamlerd K, Chatsuwat T, Ariyasriwatana C, Suwan A, Taweeapolcharoen C, et al. Effects of vaginal administration of conjugated estrogens tablet on sexual function in postmenopausal women with sexual dysfunction: a double-blind, randomized, placebo-controlled trial. *BMC Womens Health.* 2020;20(1):173.
43. Cruz VL, Steiner ML, Pompei LM, Strufaldi R, Fonseca FLA, Santiago LHS, et al. Randomized, double-blind, placebo-controlled clinical trial for evaluating the efficacy of fractional CO₂ laser compared with topical estriol in the treatment of vaginal atrophy in postmenopausal women. *Menopause.* 2018;25(1):21-8.
44. Paraiso MFR, Ferrando CA, Sokol ER, Rardin CR, Matthews CA, Karram MM, et al. A randomized clinical trial comparing vaginal laser therapy to vaginal estrogen therapy in women with genitourinary syndrome of menopause: The VeLVET Trial. *Menopause.* 2020;27(1):50-6.
45. Politano CA, Costa-Paiva L, Aguiar LB, Machado HC, Baccaro LF. Fractional CO₂ laser versus promestriene and lubricant in genitourinary syndrome of menopause: a randomized clinical trial. *Menopause.* 2019;26(8):833-40.
46. Ribeiro AE, Monteiro NES, Moraes AVG, Costa-Paiva LH, Pedro AO. Can the use of probiotics in association with isoflavone improve the symptoms of genitourinary syndrome of menopause? Results from a randomized controlled trial. *Menopause.* 2018;26(6):643-52.
47. Sokol ER, Karram MM. An assessment of the safety and efficacy of a fractional CO₂ laser system for the treatment of vulvovaginal atrophy. *Menopause.* 2016;23(10):1102-7.
48. Tovar-Huamani J, Mercado-Olivares F, Grandez-Urbina JA, Pichardo-Rodriguez R, Tovar-Huamani M, García-Perdomo H. Efficacy of fractional CO₂ laser in the treatment of genitourinary syndrome of menopause in Latin-American Population: First Peruvian experience. *Lasers Surg Med.* 2019;51(6):509-15.

49. Vieira-Baptista P, Marchitelli C, Haefner HK, Donders G, Perez-Lopez F. Deconstructing the genitourinary syndrome of menopause. *Int Urogynecol J*. 2017;28(5):675-9.
50. Gabes M, Knuttel H, Stute P, Apfelbacher CJ. Measurement properties of patient-reported outcome measures (PROMs) for women with genitourinary syndrome of menopause: a systematic review. *Menopause*. 2019;26(11):1342-53.
51. Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res*. 2017.
52. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1147-57.
53. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42.
54. Prinsen CA, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" - a practical guideline. *Trials*. 2016;17(1):449.
55. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924-6.
56. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*. 2015;350:g7647.
57. Vissers T, Vries Rd. Quality of life (QoL) search blocks 2018 [updated 28.02.2018]. Available from: <https://blocks.bmi-online.nl/catalog/294>.
58. Jansma EP, Vries Rd. Patient reported outcome measures (PROMs) search blocks 2017 [updated 21.04.2017]. Available from: <https://blocks.bmi-online.nl/catalog/248>.
59. Mackintosh A, i Comabella CC, Hadi M, Gibbons E, Fitzpatrick R, Roberts N. PROM group construct & instrument type filters 2010 [19.03.2018]. Available from: <http://www.cosmin.nl/images/upload/files/PROM%20Gp%20filtersOCTOBER%202010FINAL.pdf>.
60. Greenhalgh J, Long AF, Flynn R. The use of patient reported outcome measures in routine clinical practice: lack of impact or lack of theory? *Social Science & Medicine*. 2005;60(4):833-43.
61. Terwee CB, Jansma EP, Riphagen II, de Vet HCW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality of Life Research*. 2009;18(8):1115-23.
62. Heini D, Prinsen CA, Drucker AM, Ofenloch R, Humphreys R, Sach T, et al. Measurement properties of quality of life measurement instruments for infants, children and adolescents with eczema: protocol for a systematic review. *Syst Rev*. 2016;5:25.
63. Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. *J Med Libr Assoc*. 2016;104(3):240-3.
64. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *Open Med*. 2009;3(3):e123-30.
65. Huang AJ, Luft J, Grady D, Kuppermann M. The day-to-day impact of urogenital aging: perspectives from racially/ethnically diverse women. *J Gen Intern Med*. 2010;25(1):45-51.
66. Erekson EA, Yip SO, Wedderburn TS, Martin DK, Li FY, Choi JN, et al. The Vulvovaginal Symptoms Questionnaire: a questionnaire for measuring vulvovaginal symptoms in postmenopausal women. *Menopause*. 2013;20(9):973-9.
67. Erekson EA, Li FY, Martin DK, Fried TR. Vulvovaginal symptoms prevalence in postmenopausal women and relationship to other menopausal symptoms and pelvic floor disorders. *Menopause*. 2016;23(4):368-75.
68. Fernandez-Alonso AM, Alcaide-Torres J, Fernandez-Alonso IM, Chedraui P, Perez-Lopez FR. Application of the 21-item Vulvovaginal Symptoms Questionnaire in postmenopausal Spanish women. *Menopause*. 2017;24(11):1295-301.
69. Huang AJ, Gregorich SE, Kuppermann M, Nakagawa S, Van Den Eeden SK, Brown JS, et al. Day-to-Day Impact of Vaginal Aging questionnaire: a multidimensional measure of the impact of vaginal symptoms on functioning and well-being in postmenopausal women. *Menopause*. 2015;22(2):144-54.
70. McKenna SP, Whalley D, Renck-Hooper U, Carlin S, Doward LC. The development of a quality of life instrument for use with post-menopausal women with urogenital atrophy in the UK and Sweden. *Qual Life Res*. 1999;8(5):393-8.
71. Chen RQ, Davis SR, Wong CM, Lam TH. Validity and cultural equivalence of the standard Greene Climacteric Scale in Hong Kong. *Menopause*. 2010;17(3):630-5.
72. Chren MM, Lasek RJ, Quinn LM, Mostow EN, Zyzanski SJ. Skindex, a quality-of-life measure for patients with skin disease: reliability, validity, and responsiveness. *J Invest Dermatol*. 1996;107(5):707-13.
73. Schmitt J, Apfelbacher C, Spuls PI, Thomas KS, Simpson EL, Furue M, et al. The Harmonizing Outcome Measures for Eczema (HOME) roadmap: a methodological framework to develop core sets of outcome measurements in dermatology. *J Invest Dermatol*. 2015;135(1):24-30.

74. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: version 1.0. *Trials*. 2017;18(3):280.
75. Palacios S, Nappi RE, Bruyniks N, Particco M, Panay N, Investigators ES. The European Vulvovaginal Epidemiological Survey (EVES): prevalence, symptoms and impact of vulvovaginal atrophy of menopause. *Climacteric*. 2018;21(3):286-91.
76. Moral E, Delgado JL, Carmona F, Caballero B, Guillan C, Gonzalez PM, et al. The impact of genitourinary syndrome of menopause on well-being, functioning, and quality of life in postmenopausal women. *Menopause*. 2018;25(12):1418-23.
77. Kroenke K, Spitzer RL, Williams JB, Lowe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics*. 2009;50(6):613-21.
78. Nübling M, Andersen HH, Mühlbacher A. Entwicklung eines Verfahrens zur Berechnung der körperlichen und psychischen Summenskalen auf Basis der SOEP-Version des SF 12 (Algorithmus). DIW Berlin, German Institute for Economic Research, 2006.
79. Potthoff P, Heinemann LA, Schneider HP, Rosemeier HP, Hauser GA. [The Menopause Rating Scale (MRS II): methodological standardization in the German population]. *Zentralbl Gynakol*. 2000;122(5):280-6.
80. Heinemann K, Assmann A, Mohner S, Schneider HP, Heinemann LA. [Reliability of the Menopause Rating Scale (MRS): Investigation in the German population]. *Zentralbl Gynakol*. 2002;124(3):161-3.
81. Heinemann LA, DoMinh T, Strelow F, Gerbsch S, Schnitker J, Schneider HP. The Menopause Rating Scale (MRS) as outcome measure for hormone treatment? A validation study. *Health Qual Life Outcomes*. 2004;2:67.
82. Schneider HP, Heinemann LA, Rosemeier HP, Potthoff P, Behre HM. The Menopause Rating Scale (MRS): comparison with Kupperman index and quality-of-life scale SF-36. *Climacteric*. 2000;3(1):50-8.
83. Ofenloch RF, Weisshaar E, Apfelbacher C. New evidence for construct validity and interpretability of the German Quality of Life in Hand Eczema Questionnaire (QOLHEQ). *Contact Dermatitis*. 2020;83(3):189-95.
84. Freeman EW, Sammel MD, Lin H, Nelson DB. Associations of hormones and menopausal status with depressed mood in women with no history of depression. *Archives of general psychiatry*. 2006;63(4):375-82.
85. TNS Infratest Sozialforschung. 2016. SOEP 2016 – Erhebungsinstrumente 2016 (Welle 33) des Sozioökonomischen Panels: Personenfragebogen, Stichproben A-L3. SOEP Survey Papers 345: Series A. Berlin: DIW/SOEP.
86. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34(3):220-33.
87. Moosbrugger H, Schermelleh-Engel K. Exploratorische (EFA) und Konfirmatorische Faktorenanalyse (CFA). In: Moosbrugger H, Kelava A, editors. *Testtheorie und Fragebogenkonstruktion*. Berlin Heidelberg, Springer; 2012. p. 326-43.
88. Brown TA. *Confirmatory Factor Analysis for Applied Research*. New York: Guilford Press; 2015.
89. Schmitt TA. Current Methodological Considerations in Exploratory and Confirmatory Factor Analysis. *Journal of Psychoeducational Assessment*. 2011;29(4):304-21.
90. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61(2):102-9.
91. Stover AM, McLeod LD, Langer MM, Chen WH, Reeve BB. State of the psychometric methods: patient-reported outcome measure development and refinement using item response theory. *J Patient Rep Outcomes*. 2019;3(1):50.
92. Petrillo J, Cano SJ, McLeod LD, Coon CD. Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: a comparison of worked examples. *Value Health*. 2015;18(1):25-34.
93. Kamudoni P, Mueller B, Salek MS. The development and validation of a disease-specific quality of life measure in hyperhidrosis: the Hyperhidrosis Quality of Life Index (HidroQOL(c)). *Qual Life Res*. 2015;24(4):1017-27.
94. Pett M, Lackey NR, Sullivan J. *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Thousand Oaks: Sage; 2006.
95. Gabes M, Chamlin SL, Lai JS, Cella D, Mancini AJ, Apfelbacher CJ. Development of a validated short-form of the Childhood Atopic Dermatitis Impact Scale, the CADIS-SF15. *J Eur Acad Dermatol Venereol*. 2020;34(8):1773-8.
96. Linacre JM. Investigating rating scale category utility. *J Outcome Meas*. 1999;3(2):103-22.
97. Andrich D, Humphry SM, Marais I. Quantifying Local, Response Dependence Between Two Polytomous Items Using the Rasch Model. *Applied Psychological Measurement*. 2012;36(4):309-24.
98. Masters GN. A rasch model for partial credit scoring. *Psychometrika*. 1982;47(2):149-74.
99. Shea TL, Tennant A, Pallant JF. Rasch model analysis of the Depression, Anxiety and Stress Scales (DASS). *BMC Psychiatry*. 2009;9(1):21.

100. Hunter MM, Guthrie KA, Larson JC, Reed SD, Mitchell CM, Diem SJ, et al. Convergent-Divergent Validity and Correlates of the Day-to-Day Impact of Vaginal Aging Domain Scales in the MsFLASH Vaginal Health Trial. *J Sex Med.* 2020;17(1):117-25.
101. Isidori AM, Pozza C, Esposito K, Giugliano D, Morano S, Vignozzi L, et al. Development and validation of a 6-item version of the female sexual function index (FSFI) as a diagnostic tool for female sexual dysfunction. *J Sex Med.* 2010;7(3):1139-46.
102. Gibson CJ, Huang AJ, Larson JC, Mitchell C, Diem S, LaCroix A, et al. Patient-centered change in the day-to-day impact of postmenopausal vaginal symptoms: results from a multicenter randomized trial. *Am J Obstet Gynecol.* 2020;223(1):99.e1-.e9.
103. Brunner HI, Higgins GC, Klein-Gitelman MS, Lapidus SK, Olson JC, Onel K, et al. Minimal clinically important differences of disease activity indices in childhood-onset systemic lupus erythematosus. *Arthritis care & research.* 2010;62(7):950-9.
104. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. *Controlled Clinical Trials.* 1989;10(4):407-15.
105. Shifren JL, Zincavage R, Cho EL, Magnavita A, Portman DJ, Krychman ML, et al. Women's experience of vulvovaginal symptoms associated with menopause. *Menopause.* 2019;26(4):341-9.
106. Mitchell CM, Waetjen LE. Genitourinary Changes with Aging. *Obstet Gynecol Clin North Am.* 2018;45(4):737-50.
107. Coyne KS, Sexton CC, Vats V, Thompson C, Kopp ZS, Milsom I. National community prevalence of overactive bladder in the United States stratified by sex and age. *Urology.* 2011;77(5):1081-7.
108. Irwin DE, Milsom I, Hunskaar S, Reilly K, Kopp Z, Herschorn S, et al. Population-based survey of urinary incontinence, overactive bladder, and other lower urinary tract symptoms in five countries: results of the EPIC study. *Eur Urol.* 2006;50(6):1306-14; discussion 14-5.
109. Kinsey D, Pretorius S, Glover L, Alexander T. The psychological impact of overactive bladder: A systematic review. *J Health Psychol.* 2016;21(1):69-81.
110. Perrotta C, Aznar M, Mejia R, Albert X, Ng CW. Oestrogens for preventing recurrent urinary tract infection in postmenopausal women. *Cochrane Database Syst Rev.* 2008(2):Cd005131.
111. Renard J, Ballarini S, Mascarenhas T, Zahran M, Quimper E, Choucair J, et al. Recurrent Lower Urinary Tract Infections Have a Detrimental Effect on Patient Quality of Life: a Prospective, Observational Study. *Infect Dis Ther.* 2014;4(1):125-35.
112. Wagenlehner F, Wullt B, Ballarini S, Zingg D, Naber KG. Social and economic burden of recurrent urinary tract infections and quality of life: a patient web-based study (GESPRIT). *Expert Rev Pharmacoecon Outcomes Res.* 2018;18(1):107-17.
113. Sampselle CM, Harlow SD, Skurnick J, Brubaker L, Bondarenko I. Urinary incontinence predictors and life impact in ethnically diverse perimenopausal women. *Obstet Gynecol.* 2002;100(6):1230-8.
114. Botlero R, Bell RJ, Urquhart DM, Davis SR. Urinary incontinence is associated with lower psychological general well-being in community-dwelling women. *Menopause.* 2010;17(2):332-7.
115. Waetjen LE, Xing G, Johnson WO, Melnikow J, Gold EB. Factors associated with reasons incontinent midlife women report for not seeking urinary incontinence treatment over 9 years across the menopausal transition. *Menopause.* 2018;25(1):29-37.
116. Vivian-Taylor J, Hickey M. Menopause and depression: is there a link? *Maturitas.* 2014;79(2):142-6.
117. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, et al. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1--eliciting concepts for a new PRO instrument. *Value Health.* 2011;14(8):967-77.
118. Howells LM, Chalmers JR, Cowdell F, Ratib S, Santer M, Thomas KS. 'When it goes back to my normal I suppose': a qualitative study using online focus groups to explore perceptions of 'control' among people with eczema and parents of children with eczema in the UK. *BMJ Open.* 2017;7(11):e017731.
119. Schreier M. *Qualitative Content Analysis in Practice.* Los Angeles; London; New Delhi; Singapore; Washington DC: SAGE; 2012.
120. Fusch P, Ness L. Are We There Yet? Data Saturation in Qualitative Research. *Qualitative Report.* 2015;20:1408-16.
121. Herbert D, Bell RJ, Young K, Brown H, Coles JY, Davis SR. Australian women's understanding of menopause and its consequences: a qualitative study. *Climacteric.* 2020:1-7.
122. Wunderlich D. *EigenSinnige Frauen: Zehn Porträts.* München, Zürich: Piper; 2015.
123. Bigalke K. Die Suche nach dem richtigen Weg durch die Wechseljahre [Internet]. *Deutschlandfunk Kultur;* 2019 [cited 13/08/2020]. Available from: https://www.deutschlandfunkkultur.de/menopause-die-suche-nach-dem-richtigen-weg-durch-die.976.de.html?dram:article_id=448368.

124. Mastroianni J, Thompson JA, Shifren JL, Zuckerman AL, Pereira K. Improving the identification of genitourinary syndrome of menopause through the utilization of the Day-to-Day Impact of Vaginal Aging questionnaire. *Menopause*. 2020;27(11):1295-301.
125. Wenderlein J. Endokrinologie. Urogenitales Menopause-Syndrom mehr beachten. *Geburtshilfe und Frauenheilkunde*. 2015;75(05):424–7.
126. Wolff F. Das urogenitale Syndrom der Menopause. *gynäkologie + geburtshilfe*. 2018;23(5):48-9.

Danksagung

Ich möchte mich herzlich bei Herrn Prof. Apfelbacher für die Betreuung meiner Arbeit und Unterstützung in jeglicher Hinsicht bedanken. Auch, dass Sie als zweiter Reviewer das COSMIN *Risk of Bias Rating* übernommen haben, ist nicht selbstverständlich.

Des Weiteren möchte ich Herrn Prof. Koller und Frau Prof. Stute für die Übernahme des Mentorats, die Teilnahme an den Kolloquien und den fachlichen Input danken. Großer Dank geht vor allem an Frau Prof. Stute für ihre Mithilfe bei der Rekrutierung und die Möglichkeit, eine Fokusgruppe vor Ort bei ihr am Inselspital in Bern durchführen zu können.

Ebenso möchte ich auch Herrn Dr. Knüttel besonderen Dank für die Unterstützung bei der systematischen Literatursuche im Rahmen der systematischen Übersichtsarbeit aussprechen.

Ein herzliches Dankeschön geht auch an Peter Werkmann, der alle Titel, Abstracts und Volltexte als zweiter Reviewer für die systematische Übersichtsarbeit gescreent hat.

Auch dem Team der Medizinischen Soziologie möchte ich meinen Dank für den ständigen Austausch und die motivierenden Worte während der gesamten Promotionsphase aussprechen.

Großer Dank geht auch an meine Hilfskraft Gesina, die alle Papierfragebögen eingegeben hat und meine Doktorarbeit formal korrigiert hat.

Auch meiner baldigen Schwägerin Angelika möchte ich für die formale Korrektur der Doktorarbeit danken.

Ich möchte mich auch bei meinem Verlobten bedanken, der mich bei der Rekrutierung unterstützt hat und immer an mich geglaubt hat. Danke, dass du immer hinter mir stehst.

Ein großes Dankeschön möchte ich meiner kleinen Schwester Daniela aussprechen, die mich bei der Rekrutierung unterstützt hat und die Doktorarbeit formal korrigiert hat. Sie hat mich auch dann motiviert, wenn ich selbst gezweifelt habe.

Besonderer Dank gilt auch meinem lieben Opa, der die Briefe für die Rekrutierung sorgfältig bestückt und beschriftet hat. Ohne ihn wäre ich nicht so weit gekommen, wie ich heute bin. Er hat mich mein ganzes Leben über immer motiviert, sich für meine Fortschritte interessiert und meine Erfolge honoriert. Ihm möchte ich diese Doktorarbeit widmen.

Nicht zuletzt geht besonderer Dank an meine Eltern, ohne die mir das Psychologiestudium und demzufolge auch die Promotion nicht möglich gewesen wären. Vor allem meiner Mutter möchte ich danken, dass sie mich bei der Rekrutierung unterstützt hat.

Lebenslauf

Persönliche Daten:

Vor- und Nachname: Michaela Gabes
Anschrift: Wallnerstr. 44; 93102 Pfatter
E-Mail-Adresse: michaela.gabes@klinik.uni-regensburg.de
Geburtsdatum/-ort: 4. Februar 1994; Regensburg
Staatsangehörigkeit: deutsch

Ausbildung:

Sept. 2000 bis Juli 2004 Grundschule Mintraching
Sept. 2004 bis Juni 2012 Gymnasium Neutraubling
Juni 2012 Abitur (1,1)
Okt. 2012 bis Juli 2015 Bachelor of Science in Psychologie, Universität Regensburg (1,6)
Okt. 2015 bis Januar 2018 Master of Science in Psychologie, Universität Regensburg (1,4)
Sept. 2015 bis Januar 2016 Auslandssemester, Université Blaise Pascal, Clermont-Ferrand, Frankreich
Sept. 2017 bis Januar 2018 Auslandssemester, Università degli Studi di Firenze, Florenz, Italien

Praktika/Nebentätigkeiten:

Sept. 2012 bis Okt. 2016 Aushilfe bei „New Yorker“ im Donaeinkaufszentrum, Regensburg
Februar/März 2014 6-wöchiges Praktikum in der **Jugend- und Familientherapeutischen Beratungsstelle** der Stadt Regensburg
– Pflichtpraktikum im Rahmen des Psychologie-Studiums (B.Sc.)
– Protokollführung in Erstgesprächen; Teilnahme an Außen-terminen (Eltern-/Lehrerabende); Durchführung und Auswertung von Intelligenztests
Februar/März 2016 6-wöchiges Praktikum am **Bezirksklinikum Mainkofen**
– Pflichtpraktikum im Rahmen des Psychologie-Studiums (M.Sc.)
– Teilnahme an Gruppentherapien, Chefarzt-Visiten und Einzelgesprächen; Kennenlernen typischer psychiatrischer Erkrankungen; eigenständiges Testen von Patienten (z. B. Konzentrations-/Demenztestungen)
Sept. 2016 bis Feb. 2017 Wissenschaftliche Hilfskraft am **Lehrstuhl für Biologische, Klinische und Rehabilitationspsychologie**, Universität Regensburg
– Arbeit im Tierlabor; Recherchetätigkeiten; Mithilfe bei Projekten
November 2016 2-wöchiges Praktikum in der **Justizvollzugsanstalt Straubing**
– freiwilliges Praktikum im Rahmen des Psychologie-Studiums
– Teilnahme an Therapieprogrammen und Insassengesprächen; Kennenlernen verschiedener Beschäftigungsbereiche (Regelvollzug, psychiatrische Abteilung, Krankenhaus); Begleitung einer Ausführung

Feb. 2017 bis Sept. 2017	Wissenschaftliche Hilfskraft im Studienzentrum Regensburg , Uniklinikum Regensburg <ul style="list-style-type: none"> – Datenerhebung im Rahmen der NaKo-Gesundheitsstudie: Durchführung von Interviews, Gedächtnistests, Hörtests und Handgreifkraftmessungen – Fragebogeneingabe und Ergebnisbriefherstellung
April 2017 bis Juli 2017	3-monatiges Praktikum in der Abteilung für Forensische Psychiatrie und Psychotherapie am Bezirksklinikum Regensburg <ul style="list-style-type: none"> – freiwilliges Praktikum im Rahmen des Psychologie-Studiums – Mitarbeit am Präventionsprojekt „<i>Kein Täter werden</i>“ der Sexualwissenschaftlichen Ambulanz der Universität Regensburg – Durchführung von klinischen Interviews und Intelligenztests an den Patienten; Protokollierung bei Explorationen – Mitarbeit bei Begutachtungen; Vorbereitung einer Tagung (Summer Conference 2017)

Beruflicher Werdegang

Feb. 2018 – Sept. 2019	Wissenschaftliche Mitarbeiterin in der Medizinischen Soziologie , Institut für Epidemiologie und Präventivmedizin, Uniklinikum Regensburg <ul style="list-style-type: none"> – Validierung von Fragebögen zu verschiedenen Hauterkrankungen (Neurodermitis, Hyperhidrose etc.) – Vorstellung der Forschungsergebnisse auf nationalen und internationalen Kongressen – Lehre im Rahmen des Seminars „Medizinische Soziologie“
Seit Mai 2018	Promotion zum Dr. sc. hum., Universität Regensburg
Seit Sept. 2018	Ausbildung zur Psychologischen Psychotherapeutin am Ausbildungsinstitut für Verhaltenstherapie in Regensburg (AVR)
Seit April 2019	Mitarbeit in der AVR Institutsambulanz in der Lehrpraxis Dr. Mohr , Zeitlarn <ul style="list-style-type: none"> – Psychotherapeutische Betreuung ambulanter Patienten im Rahmen der Psychotherapie-Ausbildung
Seit Okt. 2019	Wissenschaftliche Mitarbeiterin am Institut für Sozialmedizin und Gesundheitssystemforschung der Otto-von-Guericke Universität Magdeburg

Zusatzqualifikationen/Weiterbildungen

April 2013	Teilnahme an einem von der Universität Regensburg organisierten Hypnose-Workshop
Feb. 2017 bis Juli 2017	Fortbildung „Versuchstiere und Tierschutz (FELASA B)“
Juli 2017	Teilnahme an einer universitären Exkursion ins Bezirksklinikum Straubing
April 2018	Fortbildung „Literaturrecherche in PubMed/Medline“, Uniklinikum Regensburg
Seit Okt. 2018	Gutachtertätigkeit beim <i>British Journal of Dermatology</i>

- Okt. 2019 Workshop: How Rasch and Classical Test Theory Can Complement Each Other When Assessing the Reliability and Validity of Clinical Outcome Assessment, ISOQOL 2019, San Diego
- Nov. 2019 Workshop: Rhetorik und Präsentation, Zentrum für Hochschuldidaktik, Universität Regensburg
- Okt. 2020 Workshop: Gute Wissenschaftliche Praxis, Universität Regensburg

Kongresse

- Okt. 2018 Poster-Präsentation auf der ISOQOL (International Society of Quality of Life Research) 2018, Dublin
- März 2019 Poster-Präsentation und Vortrag bei einer AG-Sitzung bei der ADF-Tagung (Arbeitsgemeinschaft Dermatologische Forschung), München
- April 2019 Vortrag beim HOME (Harmonising Outcome Measures for Eczema) Meeting VII, Tokio
- Okt. 2019 Poster-Präsentation auf der ISOQOL (International Society of Quality of Life Research) 2019, San Diego
- Okt. 2020 Poster-Präsentation bei der virtuellen EADV (European Academy of Dermatology and Venereology) 2020

Fachliche Kompetenz

- Fremdsprachenkenntnisse: Englisch (Schule: 6 Jahre; Sprachkurs; fließend)
 Französisch (Schule: 5 Jahre; Sprachkurs; Auslandssemester; fließend)
 Latein (7 Jahre in der Schule)
 Italienisch (Schule: 3 Jahre; Sprachkurse; Auslandssemester; fließend)
- Auswertungsprogramme: Fundierte Kenntnisse in SPSS, SAS, Mplus, Winsteps und ATLAS.ti

Publikationen

- Gabes M**, Stute P, Apfelbacher CJ. Refinement of the German Day-to-Day Impact of Vaginal Aging questionnaire in peri- and postmenopausal women using Item Response Theory and Classical Test Theory. *Menopause* (New York, NY). Forthcoming 2021.
- Thomas K, Apfelbacher CJ, Chalmers JR, Simpson E, Spuls PI, Gerbens LAA, Williams HC, Schmitt J, **Gabes M**, Howells L, Stuart BL, Grinich E, Pawlitschek T, Burton T, Howie L, Gadkari A, Eckert L, Ebata T, Boers M, Saeki H, Nakahara T, Katoh N. Recommended core outcome instruments for health-related quality of life, long-term control and itch intensity in atopic eczema trials: results of the HOME VII consensus meeting. *Br J Dermatol*. 2020 Nov 11 [online ahead of print].
- Gabes M**. Are existing quality-of-life instruments in cutaneous T-cell lymphoma of sufficient quality? *Br J Dermatol*. 2020 Sep 1 [online ahead of print].
- Gabes M**, Apfelbacher C, quality of life working group of the Harmonising Outcome Measures for Eczema (HOME) initiative. IDQoL, CDLQI and the 45-item CADIS received a sufficient content validity rating during the HOME VII meeting in Japan: a group discussion study. *J Eur Acad Dermatol Venereol*. 2020 Aug 3 [online ahead of print].
- Gabes M**, Chamlin SL, Lai JS, Cella D, Mancini AJ, Apfelbacher CJ. Development of a validated short-form of the Childhood Atopic Dermatitis Impact Scale, the CADIS-SF15. *J Eur Acad Dermatol Venereol*. 2020 Aug;34(8):1773-8.

- Gabes M**, Jourdan C, Schramm K, Masur C, Abels C, Kamudoni P, Salek S, Apfelbacher C. Hyperhidrosis Quality of Life Index (HidroQoL©): further validation and clinical application in patients with axillary hyperhidrosis using data from a phase III randomized controlled trial. *Br J Dermatol*. 2020 Jun 8 [online ahead of print].
- Gabes M**, Chamlin SL, Lai JS, Cella D, Mancini AJ, Apfelbacher CJ. Evaluation of responsiveness and estimation of smallest detectable change and minimal important change scores for the Childhood Atopic Dermatitis Impact Scale. *Br J Dermatol*. 2020 Feb;182(2):348-54.
- Gabes M**, Tischer C, Apfelbacher C. Measurement properties of quality-of-life outcome measures for children and adults with eczema: An updated systematic review. *Pediatr Allergy Immunol*. 2020 Jan;31(1):66-77.
- Gabes M**. Appearance in acne: an often-overlooked concept. *Br J Dermatol*. 2019 Dec;181(6):1121-2.
- Gabes M**, Knüttel H, Stute P, Apfelbacher CJ. Measurement properties of patient-reported outcome measures (PROMs) for women with genitourinary syndrome of menopause: a systematic review. *Menopause (New York, NY)*. 2019 Nov;26(11):1342-53.

Selbstständigkeitserklärung

„Ich, Michaela Gabes, geboren am 04.02.1994 in Regensburg, erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe.

Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Insbesondere habe ich nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.“

Ort, Datum

Michaela Gabes