

## 4. METHODEN

Im folgenden Kapitel wird das Studiendesign vorgestellt, die Stichprobe, an der die Effekte der dyadischen Planung überprüft werden sollen, die Materialien, die dazu herangezogen wurden, sowie die statistischen Auswertungsmethoden.

### 4.1 Design

#### 4.1.1 Messzeitpunkte

Um die Effekte der dyadischen Planung evaluieren zu können, wurden Patienten untersucht, die sich einer Prostatektomie unterzogen und in einer festen Partnerschaft lebten. Zur Kontrolle der postoperativ auftretenden Harninkontinenz wird Prostatektomiepatienten nach ihrer Operation empfohlen, ein neues ihnen bis dahin noch fremdes Verhalten aufzubauen: das Beckenbodentraining. Diese Studie war als Paarstudie angelegt, weil in Situationen, in denen Gesundheit wiederhergestellt oder Gesundheitsverhalten ausgeführt werden soll, nicht nur Personenmerkmale, sondern auch Aspekte des sozialen Umfeldes wie z.B. Partnerschaftsaspekte eine Rolle spielen. Sie hat ein experimentelles, längsschnittliches Studiendesign, bei dem sowohl die Patienten als auch ihre Partnerinnen über fünf Messzeitpunkte hinweg mittels Fragebogen zu sozialen Prozessen und verhaltensbezogenen Kognitionen sowie nur die Patienten zum Verhalten und zur Inkontinenz befragt wurden. In der vorliegenden Studie wurden aufgrund des hier gelegten Schwerpunkts auf die Wirkweise der dyadischen Planung in erster Linie die Patientendaten ausgewertet. Den Ablauf der Studie veranschaulicht Abbildung 4.1.



Abbildung 4.1: Studiendesign

Die ersten zwei Messzeitpunkte fielen in den Zeitraum des Klinikaufenthalts. Die Patienten und ihre Partnerinnen erhielten den ersten Fragebogen zur Baseline-Messung am Tag vor der Operation bei Aufnahme auf die Station (**t1**). Am ersten Tag nach der Operation wurde den Patienten eine Informationsbroschüre zur Erleichterung des Einstiegs in das

Beckenbodentraining ausgehändigt. Zusätzlich wurden sie am zweiten oder dritten Tag nach der Operation von den Physiotherapeuten der Klinik aufgesucht, um einige Beckenbodentrainingsübungen kennen zu lernen. Ein weiterer Fragebogen am zweiten Tag nach der Operation diente der Baseline-Messung für die sozial-kognitiven Variablen hinsichtlich des Beckenbodentrainings (**t2**). Diese sollten noch nicht am Tag vor der Operation erfragt werden, da erwartet wurde, dass sich die Patienten zu diesem Zeitpunkt auf die Operation vorbereiteten und sich erst in der anschließenden Erholungsphase intensiver mit den Konsequenzen des Eingriffs auseinandersetzen würden. Alle Fragebogen während des stationären Aufenthalts wurden von Studienmitarbeitern ausgeteilt und abgeholt.

Die **Planungsintervention** (s. Abschnitt 4.1.2) fand in der Regel am Tag der Dichtigkeitsprüfung (MCU) statt. Die MCU ist eine Röntgen-Kontrastmittel-Untersuchung. Sie wurde am letzten Tag des Klinikaufenthalts durchgeführt, um zu entscheiden, ob der Katheter entfernt werden konnte (s. Abschnitt 2.5.1). War die Operationswunde noch nicht verheilt, musste der Patient katheterisiert bleiben. Wurden Patienten mit Katheter entlassen, mussten sie sich nach einigen Tagen zur erneuten MCU vorstellen. Aus Gründen wie der Katheterisierung oder fester Reisezeiten der Studienteilnehmer wurde das Planungsheft 37 Studienteilnehmern (33%) vor Ort standardisiert erklärt und außerhalb der Klinik von den Studienteilnehmern bearbeitet. Studienteilnehmer, die das Planungsheft nicht in der Klinik bearbeiteten, waren gleichmäßig über alle Gruppen verteilt ( $\chi^2(3) = 5.57, n.s.$ ). Im Durchschnitt wurde das Planungsheft elf Tage ( $SD = 7.1$ ) nach der Operation ausgefüllt. Diejenigen, die das Planungsheft in der Klinik, d.h. am Entlassungstag oder – wenn erreichbar – am Tag der wiederholten MCU bearbeiteten, füllten das Planungsheft durchschnittlich zehn Tage ( $SD = 6.2$ ) nach der Operation aus, bei denjenigen, die es außerhalb der Klinik bearbeiteten, wurde es mit durchschnittlich 15 Tagen ( $SD = 7.5$ ) signifikant später ausgefüllt ( $F_{(1, 106)} = 15.45, p < .001$ ). Falls dieser Unterschied die Ergebnisse späterer Analysen veränderte, wurde dies berichtet.

Die Nachbefragung erstreckte sich über sechs Monate. Den Studienteilnehmern wurden zwei Wochen (**t3**), drei Monate (**t4**) und sechs Monate (**t5**) nach der Operation Fragebogen mit frankierten Rückumschlägen nach Hause zugeschickt. Da den Prostatektomiepatienten empfohlen wurde, nach dem stationären Aufenthalt eine Anschlussheilbehandlung in Anspruch zu nehmen, befand sich ein Teil der Patienten in den ersten Wochen nach der Entlassung aus der Klinik in einer Rehabilitationsklinik (**t3**:  $n = 14, 12.5\%$ ; **t4**:  $n = 5, 4.5\%$ ), so dass ihnen die Fragebogen dorthin geschickt wurden. Diese Patienten unterschieden sich zu keinem der beiden Messzeitpunkte in der Trainingsdauer (**t3**:  $t(110) = -1.41, n.s.$ ;

t4:  $t(110) = 0.99$ , *n.s.*), in der Trainingsintensität (t3:  $t(110) = -0.96$ , *n.s.*; t4:  $t(110) = -0.26$ , *n.s.*), in der dyadischen Handlungsplanung (t3:  $t(110) = 1.51$ , *n.s.*; t4:  $t(110) = -1.07$ , *n.s.*) oder in der dyadischen Bewältigungsplanung (t3:  $t(110) = 1.18$ , *n.s.*; t4:  $t(110) = -0.62$ , *n.s.*) des selbständig auszuführenden Beckenbodentrainings von denjenigen, die keine Anschlussheilbehandlung nutzten.

Alle Materialien wurden mit Kennziffern versehen, so dass die Daten pseudonymisiert waren. Die Teilnahme war freiwillig. Einverständniserklärungen der Patienten und ein positives Votum der Ethikkommission lagen vor.

#### 4.1.2 Planungsintervention

Die Planungsintervention erfolgte gemäß eines 2x2-Gruppen-Designs, um die dyadische Planung und die ihr folgenden volitionalen und sozialen Prozesse für die Steigerung des späteren Verhaltens zu überprüfen (s. Abbildung 4.2). Eine Hälfte der Patienten plante daher „dyadisch“, d.h. gemeinsam mit ihren Partnerinnen, wie das neue Gesundheitsverhalten in den Alltag eingebaut werden könnte, während die andere Hälfte der Patienten „individuell“, d.h. in Abwesenheit ihrer Partnerinnen planten. Diese Unterscheidung zwischen der Anleitung zu dyadischem und individuellem Planen bildete den ersten Interventionsfaktor *Planungssetting*. Um die Verhaltenseffekte tatsächlich auf das dyadische Planen eines spezifischen Verhaltens, nämlich des Beckenbodentrainings zurückführen zu können und nicht lediglich auf das Interventionssetting, sollten als Alternativtreatment zur Planung des „Beckenbodentrainings“ Pläne zu gesunder „Ernährung“ aufgestellt werden. Diese Unterscheidung hinsichtlich der *Planinhalte* bildete den zweiten Interventionsfaktor.

|                   |                     | <i>Planungssetting</i> |             |
|-------------------|---------------------|------------------------|-------------|
|                   |                     | dyadisch               | individuell |
| <i>Planinhalt</i> | Beckenbodentraining | $n = 28$               | $n = 29$    |
|                   | Ernährung           | $n = 29$               | $n = 26$    |

Abbildung 4.2: *Planinhalt x Planungssetting – Design*

Die Randbedingungen waren für die beiden das Beckenbodentraining planenden Gruppen dieselben wie für die beiden die Ernährung planenden Gruppen, während sich das zu planende Verhalten unterschied. Das 2x2-Gruppen-Design entsprach einer Kombination aus isolierter und vergleichender Evaluation. Einerseits konnten die Effekte des konkreten *Planinhalts* durch den Vergleich von Beckenbodentrainings- und Ernährungsplanung im Sinne einer isolierten Evaluation mit einem Kontrollprogramm untersucht werden. Andererseits konnten

---

auch die durch das *Planungssetting* angestoßenen sozialen Prozesse, die mit dem dyadischen Planen einhergehen, als Vergleich zwischen individuellem und dyadischem Planen im Sinne einer vergleichenden Evaluation mit einem Alternativprogramm kontrolliert werden (Hager, Patry, & Brezing, 2000).

Um potentielle Störfaktoren über die Gruppen hinweg zu minimieren, erfolgte die Zuordnung zu den vier Planungsgruppen randomisiert durch die Studienmitarbeiterinnen. Die Studienteilnehmer waren „verblindet“, sie wurden nicht darüber aufgeklärt, dass es vier verschiedene Interventionsgruppen gab und erfuhren den Inhalt ihrer Intervention erst zu Beginn des Planungsgesprächs. Die Studienmitarbeiterinnen nahmen die Zuordnung erst vor dem Planungsgespräch entsprechend einer fortlaufenden Liste vor. Basierend auf einer Teststärke von .90 zur Aufdeckung signifikanter Differenzen wurden mindestens 25 Teilnehmer in jede der vier Gruppen eingeschlossen. Die unterschiedlichen Gruppengrößen sind auf logistische Probleme parallel laufender Planungsgespräche und auf das Nichtantreffen sowie auf den nachträglichen Rückzug der Einverständniserklärung der Patienten oder ihrer Partnerinnen zurückzuführen.

Für die Planungsinterventionen wurde für jede der vier Gruppen ein Planungsheft entwickelt, mit dessen Hilfe die Studienteilnehmer standardisiert zur Aufstellung von Plänen angeleitet wurden. Die Gespräche waren für maximal eine halbe Stunde angelegt. Einleitend wurden die Paare gemeinsam darauf hingewiesen, dass bei einer Operation wie der Prostatektomie und aufgrund der zugrunde liegenden Krebserkrankung Verhaltensänderungen des Patienten empfohlen werden. So kann der Patient zur Kontrolle der Inkontinenz selbstständig Beckenbodentraining durchführen sowie durch eine gesunde Ernährung, d.h. „fett- und salzarm und viel Obst und Gemüse“ zusätzlich etwas für seine Gesundheit tun. Sie waren angehalten, lediglich die Instruktionen im Planungsheft zu wiederholen. Durch die Anwesenheit der Partnerin als auch durch dieselben Instruktionen während der Intervention sollten Unterschiede in der Interaktion oder in Beziehungsmerkmalen sowie in der konkreten Vorbereitung zwischen den Gruppen minimiert werden. Nach dem Austeilen und Vorstellen des Planungsheftes sollten die Studienteilnehmer selbständig ihre Pläne erarbeiten. Die trainierten Studienmitarbeiterinnen hielten sich für Verständnisfragen bereit, halfen aber nicht bei der inhaltlichen Formulierung der Pläne.

Im *dyadischen Setting* wurde den Paaren das Planungsheft überreicht. Sie wurden zudem darauf hingewiesen, das Planungsheft „gemeinsam“ zu bearbeiten, indem der Patient die Anleitungen und Beispiele laut vorlesen sollte, um anschließend mögliche Pläne miteinander zu diskutieren. Zusätzlich erhielt die Partnerin ein Protokollblatt, auf dem sie Ideen während

des gemeinsamen Planens festhalten sollte. Im *individuellen Setting* wurde die Partnerin gebeten, in einem separaten Raum einen weiteren kurzen Fragebogen zu bearbeiten. Erst dann erhielt der Patient das Planungsheft, das er allein ausfüllte. Auch er wurde angehalten, sich die Beispiele laut vorzulesen.

Die *Planungshefte* wurden mit einer Instruktion eingeleitet, wie das Heft ausgefüllt werden sollte. Auf der ersten Doppelseite wurden die Studienteilnehmer aufgeklärt, was bei der Aufstellung von *Handlungsplänen* zu beachten ist. Dazu erhielten sie auf der linken Seite ein Beispiel eines Handlungsplans je nach Gruppenzugehörigkeit zum Beckenbodentraining bzw. zur Ernährung. Auf der rechten Seite sollten die Studienteilnehmer dann bis zu drei eigene Handlungspläne generieren, indem sie:

- a. verschieden Pläne so konkret wie möglich überlegen sollten,
- b. mit ihrer Partnerin über die Ideen sprechen sollten (dyadisch) bzw. alternativ sich die Einzelheiten genau durch den Kopf gehen lassen sollten (individuell),
- c. die drei wichtigsten Pläne, die von den Patienten allein ausgeführt werden sollten, in die vorgegebene Tabelle mit Angaben zur Zeit (Wann?), zum Ort (Wo?) und zum Verhalten (Wie?) so konkret wie möglich eintragen sollten,
- d. die Selbstwirksamkeit zur Ausführung der Pläne einschätzen sollten und abschließend
- e. sich die Pläne gut einprägen sollten, indem sie sie sich bildlich vorstellten sowie sich fest vornehmen sollten, zukünftig in diesen Situationen auch wirklich genau so wie geplant zu handeln.

Im Durchschnitt stellten die Studienteilnehmer  $M = 2.68$  ( $SD = 0.69$ ) Handlungspläne auf, in der Klinik waren es  $M = 2.79$  ( $SD = 0.53$ ) Handlungspläne, außerhalb der Klinik dagegen  $M = 2.44$  ( $SD = 0.93$ ) Handlungspläne ( $t(42.97) = 2.03, p < .05$ ). Auf der nächsten Doppelseite erfolgte nach demselben Prozedere die Anleitung zur Aufstellung von bis zu drei *Bewältigungsplänen*. Auf der linken Seite erhielten die Studienteilnehmer erneut ein Beispiel. Auf der rechten Seite sollten sie wieder bis zu drei eigene Pläne generieren und in eine Tabelle eintragen. In eine Spalte wurden die kritischen Ereignisse, die es ihnen persönlich erschweren würden, das geplante Verhalten in die Tat umzusetzen eingetragen, in eine weitere Spalte ihre Strategie zum Umgang mit dem jeweiligen Hindernis. Die Anleitung dazu waren dieselben fünf Punkte a – e wie bei der Anleitung zur Handlungsplanung. Bis auf Punkt b (s.o.) waren die Instruktionen für das dyadische und individuelle Setting identisch. Im Durchschnitt stellten die Studienteilnehmer  $M = 1.96$  ( $SD = 0.94$ ) Bewältigungspläne auf, in der Klinik waren es  $M = 2.07$  ( $SD = 0.89$ ) Bewältigungspläne, außerhalb der Klinik  $M = 1.74$  ( $SD = 1.02$ ) Bewältigungspläne ( $t(107) = 1.72, n.s.$ ). Drei Patienten stellten weder Handlungs-

---

noch Bewältigungspläne in den Planungsheften auf. Abschließend sollten die Studienteilnehmer den Stichtag für den Beginn der Verhaltensausbübung notieren sowie das Datum des Tages, an dem das Planungsheft ausgefüllt wurde. In den meisten Fällen fielen die beiden Angaben auf denselben Tag. Für die Dokumentation der formulierten Ausführungs- und Bewältigungspläne wurden die Planungshefte von der Studienmitarbeiterin wieder eingesammelt. Die Planungshefte sind im Anhang N dargestellt. Die Auswertungen entsprachen „Intention-to-treat“-Analysen, da jeweils ein Teilnehmer der dyadischen Ernährungsgruppe, der individuellen Beckenbodentrainingsgruppe und der individuellen Ernährungsgruppe das Planungsheft unausgefüllt ließ.

## **4.2 Stichprobe**

### *4.2.1 Rekrutierungsprozedere*

Die Teilnehmerrekrutierung fand im Zeitraum zwischen September 2005 und Juli 2006 in der Klinik für Urologie der Charité – Universitätsmedizin Berlin statt. Die Patienten und ihre Partnerinnen wurden am Aufnahmetag in den Räumen der Stationen der Klinik für Urologie von einer Studienmitarbeiterin kontaktiert. Ihnen wurden die Studienziele und der Studienablauf vorgestellt. Anhand eines ausführlichen Informationsblatts wurden sie über die Freiwilligkeit ihrer Teilnahme aufgeklärt (s. Anhang N). Erklärten sich die Patienten zur Teilnahme bereit, unterzeichneten sie eine Einverständniserklärung zur Verarbeitung und anonymisierten Veröffentlichung der personenbezogenen Daten im Rahmen des Forschungsvorhabens und zur Entbindung des behandelnden Klinikarztes von seiner Schweigepflicht für Angaben zum Gesundheitszustand des Patienten (s. Anhang N). Im Anschluss an die Einverständniserklärung wurden den Patienten und den Partnerinnen ihre ersten Fragebogen überreicht, so dass diese im Laufe des Tages ausgefüllt werden und zwei Tage später, d.h. am Tag nach der Operation wieder von der Studienmitarbeiterin abgeholt werden konnten.

Einschlusskriterien für die Teilnahme an der Studie waren, dass Patienten, die für eine laparoskopisch durchgeführte Prostatektomie in die Klinik aufgenommen wurden sowie der deutschen Sprache mächtig waren, um die Bearbeitung der Fragebogen zu gewährleisten. Als Weiteres sollten die Patienten in einer Partnerschaft leben und ihre Lebens- oder Ehepartnerinnen ebenfalls an der Studie teilnehmen.

### *4.2.2 Stichprobengröße*

In dieser Studie wurden Prostatektomiepatienten und ihre Partnerinnen hinsichtlich ihres Planungs- und Gesundheitsverhaltens untersucht. Die Gesundheitsverhaltensänderung betrifft

nicht nur den Patienten, sondern sollte im sozialen Kontext betrachtet werden. Sie ist von Interaktionsprozessen mit dem sozialen Umfeld beeinflusst. Deshalb wird der Patient als Mitglied der Dyade mit seiner Lebenspartnerin im sozialen Umfeld verstanden.

In dem genannten Zeitraum wurden  $N = 308$  Patienten für eine Prostatektomie in die Klinik aufgenommen. Ein Teil der Patienten ( $n = 45$ ) konnten aus logistischen Gründen (z.B. vorgezogene Voruntersuchungen, lange Arztkontakte) nicht erreicht und in die Studie aufgenommen werden. Bei 17 Patienten wurde keine laparoskopische Prostatektomie vorgenommen, sondern offen oder gar nicht operiert, so dass sie aus der Studie ausgeschlossen werden mussten. Etwa ein Sechstel der Patienten oder ihre Partnerinnen ( $n = 56$ ) lehnte von vornherein ab, an der Studie teilzunehmen oder zogen ihre Einwilligung ohne die Baseline-Fragebogen ausgefüllt zu haben wieder zurück. Einige Patienten ( $n = 41$ ) lebten allein, so dass sie das Partnerschaftskriterium zum Einschluss in die Studie nicht erfüllten. Die Hälfte der im genannten Zeitraum operierten Patienten ( $n = 149$ ) und ihre Partnerinnen konnten in die Studie aufgenommen werden. Bei 27 Paaren ergab es sich, dass die Partnerinnen nicht am Entlassungstag erscheinen konnten oder ihre Einverständniserklärungen zurückzogen, so dass das Einschlusskriterium in die Planungsintervention nicht erfüllt war. Die Ausgangsstichprobe, von der zum ersten Messzeitpunkt Daten vorlagen und die in die Planungsinterventionsgruppen eingeschlossen wurde, umfasste  $N = 112$  Dyaden von Prostatektomiepatienten und ihren Partnerinnen. Im Verlauf der Studie beendeten zwischen der Planungsintervention und t3 sechs Teilnehmer, zwischen t3 und t4 weitere sechs Teilnehmer und zwischen t4 und t5 noch einmal sechs Teilnehmer die Studienteilnahme. Für eine Übersicht siehe Tabelle 4.1. Zwölf Patienten gaben keine konkreten Gründe für den Studienabbruch an, drei Patienten konnten aufgrund ihres schlechten Gesundheitszustandes oder Komplikationen nicht mehr an der Studie teilnehmen und jeweils ein Patient gab den Tod eines Angehörigen, Überlastung oder den Widerruf der Teilnahmeeinwilligung der Partnerin an. Bis zum letzten Messzeitpunkt verblieben  $n = 94$  Teilnehmer in der Stichprobe. Dies entspricht einer Dropoutrate von 16 % der Ausgangsstichprobe.

*Tabelle 4.1: Anzahl der Studienteilnehmer im Verlauf der Studie*

| Stichprobe   | $N(n)$ | %     |
|--|--------|-------|
| Ausgangsstichprobe vor der OP (t1)                 | 112    | 100.0 |
| Daten bis einschließlich 2. Tag nach der OP (t2)   | 112    | 100.0 |
| Daten bis einschließlich 2 Wochen nach der OP (t3) | 106    | 94.6  |
| Daten bis einschließlich 3 Monate nach der OP (t4) | 100    | 89.3  |
| Daten bis einschließlich 6 Monate nach der OP (t5) | 94     | 83.9  |

---

Die Charakteristika der Teilnehmer, die nicht an allen Messzeitpunkten teilnahmen, werden in der folgenden Dropout-Analyse überprüft und sind im Anhang B den Charakteristika der in der Stichprobe verbleibenden Studienteilnehmer gegenüber gestellt.

#### 4.2.3 Dropoutanalyse

Im nächsten Schritt sollte untersucht werden, ob sich die Studienteilnehmer, die an allen Messzeitpunkten teilnahmen, von denen, die vorzeitig aus der Studie ausstiegen, hinsichtlich der soziodemographischen, sozial-kognitiven und sozialen Variablen zu den Baseline-Messungen zu t1 und t2 unterschieden. Der Dropout hängt nicht mit der Gruppenzugehörigkeit zusammen, sondern ist über alle vier Experimentalbedingungen gleich verteilt ( $\chi^2(3) = 3.28$ , n.s.). Mögliche Gruppenunterschiede wurden mittels mehrerer multivariater Varianzanalysen (MANOVA) sowie  $\chi^2$ -Tests untersucht. Um die Wahrscheinlichkeit eines  $\beta$ -Fehlers zu minimieren, wurde das Signifikanzniveau auf  $\alpha = .20$  festgelegt (Bortz, 1999). Die Ergebnisse sind tabellarisch im Anhang B dargestellt.

Studienteilnehmer, die über alle Messzeitpunkte an der Studie teilnahmen (Teilnehmer), und diejenigen, die vorzeitig aus der Studie ausschieden (Dropout), unterschieden sich lediglich in der Ängstlichkeit und Depressivität vor der Operation und im Nettohaushaltseinkommen. Mit dem Einkommen und der psychischen Belastung vor der Operation stieg auch die Wahrscheinlichkeit, die Studienteilnahme vorzeitig zu beenden. Dagegen unterschieden sich die beiden Gruppen nicht in den medizinischen Merkmalen, in den sozial-kognitiven und sozialen Variablen, in den Merkmalen der Partnerschaft sowie hinsichtlich ihres Alters oder weiterer soziodemographischer Merkmale. Aus diesen Ergebnissen ist zu schließen, dass das Risiko von Verzerrungen der Studienergebnisse als eher gering zu bewerten ist.

#### 4.2.4 Weitere Stichprobencharakteristika

Im Weiteren soll die Stichprobe mit Hilfe von soziodemographischen und medizinischen Angaben beschrieben werden. Die Ausgangsstichprobe bestand aus  $N = 112$  Studienteilnehmern im Alter von 46 bis 71 Jahren ( $M = 62.8$ ;  $SD = 6.0$ ), das Alter der Partnerinnen reichte von 27 bis 76 Jahren ( $M = 59.3$ ;  $SD = 8.8$ ) und lag damit im Durchschnitt 3.6 Jahre unter dem der Patienten.

Der Großteil der Teilnehmer ( $n = 100$ ; 89.3%) der Ausgangsstichprobe war verheiratet. Vier Teilnehmer gaben an, zwar ledig zu sein, aber in einer festen Partnerschaft zu leben



(3.6%). Fünf Teilnehmer, die gemeinsam mit einer Partnerin an der Studie teilnahmen, berichteten, sie seien geschieden (4.5%). Von drei Teilnehmern fehlten die Angaben zu ihrem Familienstand. Die meisten Studienteilnehmer lebten zusammen mit ihren Partnerinnen ( $n = 106$ ; 94.6%). Zwei (1.8%) Patienten lebten getrennt, einer davon nur zeitweise aus beruflichen Gründen. Von vier (3.6%) Patienten fehlten hierzu die Angaben. Des Weiteren berichtete der Großteil der Teilnehmer ( $n = 99$ ; 88.4%) Kinder zu haben (*Range*: 1-5;  $M = 2.0$  Kinder). Von sechs (5.4%) Teilnehmern fehlten diese Angaben.

Etwa die Hälfte der Patienten ( $n = 63$ ; 56.2%) und 41 Partnerinnen (36.6%) gingen 12 bis 13 Jahre zur Schule (d.h. Fachhochschule und Abitur), während 41 Patienten (36.6%) und 61 Partnerinnen (54.5%) als höchsten Schulabschluss die Mittlere Reife (Realschule) oder einen Haupt- bzw. Volksschulabschluss berichteten. Je ein Patient und eine Partnerin (0.9%) hatten keinen Schulabschluss. Sieben (6.3%) Patienten und neun (8.0%) Partnerinnen machten hierzu keine Angaben. Zum Zeitpunkt der Operation waren 39 Patienten (34.8%) und 45 Partnerinnen (40.2%) berufstätig, sieben Patienten (6.3%) und sechs Partnerinnen (5.4%) waren arbeitslos und 61 Patienten (54.5%) und 51 Partnerinnen (45.5%) waren zur Zeit der Studiendurchführung berentet. Von sechs (5.4%) Patienten und zehn (8.9%) Partnerinnen fehlten diese Angaben. Das Haushaltsnettoeinkommen von 42 Teilnehmern (37.5%) lag bei mehr als 2000 Euro im Monat, von 47 Teilnehmern (42.0%) lag es darunter. Etwa ein Fünftel der Teilnehmer ( $n = 23$ ; 20.5%) wollte dazu keine Angaben machen.

Hinsichtlich der medizinischen Angaben lässt sich die Stichprobe wie folgt beschreiben: Etwas weniger als die Hälfte der Patienten ( $n = 45$ , 40.2%) wurde nerverhaltend operiert. Von zwei Patienten (1.8%) fehlten dazu die Angaben. Die durchschnittliche Katheterliegedauer betrug 13.0 Tage ( $SD = 11.4$ ). Zwei Wochen nach der Operation, d.h. zum ersten postalischen Nachbefragungstermin ( $t_3$ ), mussten 24.1% der Patienten ( $n = 27$ ) noch einen Katheter tragen. Die psychische Belastung vor der Operation kann durch das Wissen um die Höhe des PSA-Wertes beeinflusst sein, weil den meisten Patienten vermittelt wird, dass das Vorhandensein von prostataspezifischen Antigenen im Blut steigt, wenn ein Prostatakarzinom vorliegt. Der PSA-Wert ist aber nicht prostatakarzinomspezifisch, so dass ein erhöhter Wert nicht gleichbedeutend mit dem Vorliegen eines Prostatakarzinoms ist, sondern auch in der gesunden Vorsteherdrüse oder bei einer benignen prostatistischen Hyperplasie vorkommen kann. Gleichzeitig schließt ein PSA-Wert innerhalb des Referenzbereichs das Vorhandensein eines Prostatakarzinoms nicht aus. Bei auffälligen PSA-Werten wird daher zur weiteren Abklärung eine Stanzbiopsie, d.h. Gewebeentnahmen vorgenommen, um mehr Sicherheit über das Vorliegen malignen Gewebes zu erhalten. Als Normalwert im Blut wird ein PSA-

Wert von  $\leq 4$  ng/ml angegeben, wobei die Referenzbereiche in Abhängigkeit vom Alter variieren: 0-2,5 ng/ml (40-49 Jahre), 0-3,5 ng/ml (50-59 Jahre), 0-4,5 ng/ml (60-69 Jahre), 0-6,5 ng/ml (70-79 Jahre) (Oesterling, Jacobsen, & Chute, 1993). In dieser Stichprobe lag der PSA-Wert durchschnittlich bei 7.5 ( $SD = 5.4$ ,  $Range$  1.4 – 50.0) und damit im erhöhten Bereich. Von zwei Patienten (1.8%) konnten die Angaben hierzu nicht ermittelt werden. Ein Maß für den Schweregrad der Krebserkrankung ist das Tumorstaging T. Wie bereits beschrieben (s. Abschnitt 2.5.1), kann es Werte zwischen T1 und T4 annehmen. Darüber hinaus wird die Metastasierung mit N und M kodiert. Letztgenannte Angaben lagen für einen Großteil der Patienten nicht vor, da eine Prostatektomie in diesem Stadium der Erkrankung nicht durchgeführt wird. Angaben zum Tumorstaging fehlten von zwei Patienten, drei Patienten (2.7%) befanden sich im ersten, der Großteil der Patienten ( $n = 85$ , 75.9%) im zweiten und 22 Patienten (19.6%) im dritten Stadium. Kein Patient wurde im Stadium T4 diagnostiziert, in dem das Tumorgewebe schon benachbartes Gewebe infiltriert hätte, da in diesem Stadium üblicherweise keine Prostatektomie durchgeführt wird. Anhand der ASA-Klassifikation wurde der körperliche Allgemeinzustand des Patienten vom behandelnden Arzt eingeschätzt, indem Werte von 1 = „sehr wenig beeinträchtigt“ bis 3 = „sehr stark beeinträchtigt“ vergeben wurden. Durchschnittlich wurde ein Wert von 2.0 vergeben, wobei 13 Patienten (11.6%) als kaum (= 1), 84 Patienten (75.0%) als etwas (= 2) und elf Patienten (9.8%) als stark (= 3) beeinträchtigt eingeschätzt wurden. Von vier Patienten (3.6%) fehlten dazu die Angaben.

### 4.3 Materialien und Instrumente

Im Folgenden werden die Instrumente zur Erfassung der zentralen Konstrukte aufgeführt. Für die genaue Beschreibung der Planungsintervention wurden die verwendeten Planungshefte bereits in Abschnitt 4.1.2 beschrieben.

#### 4.3.1 Fragebogeninstrumente

Die Fragebogen beinhalteten Skalen, die die sozial-kognitiven Komponenten des HAPA-Modells, soziale Prozesse sowie Verhaltensweisen abfragten (s. Anhang A und N). Die Deskriptiva wie Mittelwerte, Standardabweichungen, Schiefe und Exzess sind im Anhang C für alle Messzeitpunkte aufgelistet. Primär sollte die Anwendung der dyadischen Handlungs- und Bewältigungsplanung durch die Planungsintervention gesteigert werden. Als weitere zentrale Maße galten das berichtete Beckenbodentraining und die Inkontinenz. Zur weiteren

Evaluation des dyadischen Planens wurden darüber hinaus die Handlungskontrolle und die phasenspezifischen Selbstwirksamkeitserwartungen als volitionale Konstrukte sowie die verhaltensspezifische soziale Kontrolle und Unterstützung als soziale Konstrukte betrachtet.

#### 4.3.1.1 Sozial-kognitive Variablen

Als Variablen des HAPA-Modells wurden neben der Intention als motivationale Konstrukte die Risikowahrnehmung, die positiven und negativen Handlungs-Ergebnis-Erwartungen und die motivationale Selbstwirksamkeitserwartung erfasst. Neben der Planung wurden als weitere volitionale Konstrukte die Aufrechterhaltungs- und die Wiederherstellungselbstwirksamkeitserwartung erhoben. Bislang gibt es noch keine Studie, in der die Aufnahme des Beckenbodentrainings vor dem Hintergrund des HAPA-Modells untersucht wurde. Deshalb wurden die HAPA-Skalen zur körperlichen Aktivität (Sniehotta, Scholz, & Schwarzer, 2005; Sniehotta, Schwarzer et al., 2005) auf das Beckenbodentraining adaptiert. Im Folgenden werden die zentralen Variablen, die zur Überprüfung der Hypothesen herangezogen wurden, beschrieben. Die Operationalisierungen weiterer HAPA-Variablen sind in Anhang A angeführt.

*Intention zum Beckenbodentraining.* Die Intention zum Beckenbodentraining wurde mit drei Items entsprechend des Vorgehens von Sniehotta, Scholz und Schwarzer (2005) erfasst. Die Studienteilnehmer sollten auf einer vierstufigen Ratingskala (1 = „stimmt nicht“, 4 = „stimmt genau“) bezogen auf die vergangene Woche angeben, ob sie sich vorgenommen haben, „mehrmals wöchentlich Beckenbodentraining auszuführen“, „regelmäßig Beckenbodentraining auszuführen“, sich „genau an das zu halten, was in der Klinik für das Beckenbodentraining geraten wurde“ und „an einer Beckenbodentrainingsgruppe teilzunehmen“. Weil nicht alle Patienten vorhatten, eine Anschlussheilbehandlung in einer Rehabilitationsklinik in Anspruch zu nehmen, ging das letzte Item nicht mit den anderen Items in den Skalenwert ein. Die drei Items besaßen interne Konsistenzen zwischen  $\alpha = .76$  und  $\alpha = .92$  zu den Messungen am zweiten Tag bis sechs Monate nach der Operation. Alle Werte der internen Konsistenz sind im Anhang C aufgelistet. Die Retest-Reliabilität war dagegen eher moderat mit Stabilitätswerten zwischen  $r = .23$  und  $r = .43$  (s. Anhang C).

*Planung des Beckenbodentrainings.* Die Anwendung der Planung als Vorbereitung des Beckenbodentrainings in der jeweils vergangenen Woche wurde mit Hilfe von vier Skalen erfasst: der *individuellen Handlungsplanung und Bewältigungsplanung* sowie der danach entwickelten *dyadischen Handlungs- und Bewältigungsplanung*. Die individuellen Planungsskalen unterschieden sich in ihrem Stamm („Ich habe bereits konkret geplant, ...“) von denen

zur Erfassung der dyadischen Planung („Ich habe bereits gemeinsam mit meiner Partnerin konkret geplant, ...“). Mit den Vier-Item-Skalen zur individuellen und dyadischen Handlungsplanung wurde erfasst, ob innerhalb der letzten sieben Tage selbständig geplant wurde, „wann“, „wo“, „wie“ und „wie oft“ das Beckenbodentraining ausgeführt werden sollte. Dies entspricht den Items von Sniehotta, Scholz und Schwarzer (2005) zur Planung der körperlichen Aktivität, hier adaptiert auf das Beckenbodentraining. Die beiden Fünf-Item-Bewältigungsplanungsskalen erfassten, inwiefern innerhalb der letzten sieben Tage der Umgang mit Hindernissen selbständig geplant wurde. Die vorgegebenen Hindernisse und Strategien wurden entsprechend der Vorgaben von Sniehotta, Schwarzer et al. (2005) eher allgemein gehalten: „was ich tue, wenn einmal etwas dazwischen kommt“, „wie ich damit umgehe, wenn ich mal ausgesetzt habe“, „was ich in schwierigen Situationen tun kann, um meinen Vorsätzen treu zu bleiben“, „bei welchen günstigen Gelegenheiten ich meinen Beckenboden trainieren kann“ und „wann ich besonders aufpassen muss, um nicht auszusetzen“. Alle Antworten auf die Items konnten auf einer vierstufigen Ratingskala (1 = „stimmt nicht“, 4 = „stimmt genau“) abgestuft werden. Die *dyadische Handlungsplanungsskala* besaß interne Konsistenzen zwischen  $\alpha = .93$  und  $\alpha = .98$  zu den Messungen am zweiten Tag bis sechs Monate nach der Operation. Die Retest-Reliabilität lag zwischen  $r = .28$  und  $r = .48$  (s. Anhang C). Die *dyadische Bewältigungsplanungsskala* besaß interne Konsistenzen zwischen  $\alpha = .97$  und  $\alpha = .99$  zu den Messungen am zweiten Tag bis sechs Monate nach der Operation. Die Retest-Reliabilität lag zwischen  $r = .26$  und  $r = .52$  (s. Anhang C). Die *individuelle Handlungsplanungsskala* besaß interne Konsistenzen zwischen  $\alpha = .92$  und  $\alpha = .98$  zu den Messungen am zweiten Tag bis sechs Monate nach der Operation. Die Retest-Reliabilität war zum Teil gering mit Stabilitätswerten zwischen  $r = .05$  und  $r = .45$  (s. Anhang C). Die *individuelle Bewältigungsplanungsskala* besaß interne Konsistenzen zwischen  $\alpha = .90$  und  $\alpha = .96$  zu den Messungen am zweiten Tag bis sechs Monate nach der Operation. Die Werte für alle Messzeitpunkte sind im Anhang C aufgelistet. Die Retest-Reliabilität lag zwischen  $r = .23$  und  $r = .57$  (s. Anhang C).

*Aufrechterhaltungsselbstwirksamkeitserwartungen.* Die Aufrechterhaltungsselbstwirksamkeit zum Beckenbodentraining wurde mit fünf Items bezogen auf die vergangene Woche erfasst, die dem Stamm „Ich bin mir sicher, dass ich dauerhaft regelmäßig Beckenbodentraining ausführen kann, auch wenn ...“ folgten und unterschiedliche Situationen abdeckten: „lange brauche, bis das Verhalten zur Gewohnheit geworden ist“, „eigentlich mehr Lust zu etwas anderem habe“, „es jedes Mal Überwindung kostet“, „es einem mal nicht so gut geht“ und „sich die Inkontinenz dadurch nicht sofort verbessert“. Zur Beantwortung der Items stand

den Studienteilnehmern wieder eine vierstufige Rating-Skala (1 = „stimmt nicht“, 4 = „stimmt genau“) zur Verfügung. Die Items entsprachen denen von Sniehotta, Scholz und Schwarzer (2005) zur Erfassung der Aufrechterhaltungsselbstwirksamkeit zur körperlichen Aktivität und wurden inhaltlich aber auf das Beckenbodentraining adaptiert. Die internen Konsistenzen lagen zwischen  $\alpha = .86$  und  $\alpha = .91$  zu den Messungen am zweiten Tag bis sechs Monate nach der Operation. Die Werte für alle Messzeitpunkte sind im Anhang C aufgelistet. Die Retest-Reliabilität variierte zwischen  $r = .11$  und  $r = .43$  (s. Anhang C).

*Wiederherstellungsselbstwirksamkeitserwartungen.* Die letzte phasenspezifische Selbstwirksamkeitserwartung ist die Wiederherstellungsselbstwirksamkeit als Überzeugung, nach Rückfällen oder Aussetzern das Verhalten erneut initiieren und ausführen zu können (Schwarzer & Renner, 2000). Sie wurde bezogen auf die letzte Woche mit drei Items erfasst, wie sicher sich die Studienteilnehmer sind, „das Beckenbodentraining wiederaufnehmen“ zu können, auch wenn vom Studienteilnehmer die „konkreten Pläne mehrmals verschoben“ wurden, er sich „einmal nicht aufraffen konnte“ oder er „mehrmals ausgesetzt“ hat. Diese Items wurden ebenfalls nach Sniehotta, Scholz und Schwarzer (2005) auf das Beckenbodentraining adaptiert und konnten auf der vierstufigen Ratingskala von 1 = „stimmt nicht“ bis 4 = „stimmt genau“ beantwortet werden. Die Skala besaß interne Konsistenzen zwischen  $\alpha = .90$  und  $\alpha = .98$  zu den Messungen am zweiten Tag bis sechs Monate nach der Operation. Die Werte für alle Messzeitpunkte sind im Anhang C aufgelistet. Die Retest-Reliabilität variierte zwischen  $r = .11$  und  $r = .52$  (s. Anhang C).

*Handlungskontrolle.* Über die Konstrukte des HAPA-Modells hinaus wurde als weitere selbstregulative Komponente die Handlungskontrolle erfasst (Carver & Scheier, 1981). In Anlehnung an Sniehotta und Kollegen (Sniehotta, Nagy, Scholz, & Schwarzer, 2006; Sniehotta, Scholz, & Schwarzer, 2005) wurden als die drei Facetten der Handlungskontrolle mit jeweils zwei Items das Bewusstsein über die eigenen Standards, die Selbstbeobachtung und das Regulationsbemühen erfasst. Auch hier konnten die Items wieder auf einer vierstufigen Ratingskala (1 = „stimmt nicht“, 4 = „stimmt genau“) bezogen auf die vergangenen sieben Tage beantwortet werden. Als Maß für die Handlungskontrolle wurde der Gesamtskalenmittelwert als allgemeine selbstregulative Strategie mit internen Konsistenzen zwischen  $\alpha = .87$  und  $\alpha = .92$  zu den Messungen zwei Wochen bis sechs Monate nach der Operation herangezogen. Die Werte für alle Messzeitpunkte sind im Anhang C aufgelistet. Die Retest-Reliabilität variierte zwischen  $r = .03$  und  $r = .40$  (s. Anhang C).

#### 4.3.1.2 Soziale Variablen

*Negative soziale Kontrolle.* Zur Erfassung der sozialen Kontrolle durch die Lebenspartnerin innerhalb der letzten sieben Tage konnten die Patienten auf einer Drei-Item-Skala in Anlehnung an Lewis und Rook (1999) angeben, wie sehr ihre Partnerinnen negative soziale Kontrolltaktiken anwendeten. Diese Items umfassten den Versuch, ein schlechtes Gewissen zu machen, das Drängen zum Beckenbodentraining und Vorwürfe, nicht genug zu trainieren. Auf einer vierstufigen Ratingskala (1 = „stimmt nicht“, 4 = „stimmt genau“) konnten die Studienteilnehmer ihre Einschätzung abstufen. Die Skala wies interne Konsistenzen zwischen  $\alpha = .60$  und  $\alpha = .71$  zu den Messungen am zweiten Tag bis sechs Monate nach der Operation auf. Die Werte für alle Messzeitpunkte sind im Anhang C aufgelistet. Die Retest-Reliabilität variierte zwischen  $r = .25$  und  $r = .67$  (s. Anhang C).

*Soziale Unterstützung.* Wie viel soziale Unterstützung der Patient von seiner Partnerin bei der Durchführung des regelmäßigen Beckenbodentrainings innerhalb der letzten Woche erhielt, wurde mit sechs Items in Anlehnung an Schulz und Schwarzer (2003) gemessen. Diese umfassten Angaben zur Unterstützung wie Trost, Komplimente, Hilfestellungen beim Training, gemeinsames Trainieren, Erinnerungen und Bestärkungen bei der Aufnahme des Beckenbodentrainings. Erneut konnten die Patienten auf einer vierstufigen Ratingskala (1 = „stimmt nicht“, 4 = „stimmt genau“) ihre Einschätzung abgeben. Diese Skala besaß interne Konsistenzen zwischen  $\alpha = .81$  und  $\alpha = .86$  zu den Messungen zwei Wochen bis sechs Monate nach der Operation. Die Werte für alle Messzeitpunkte sind im Anhang C aufgelistet. Die Retest-Reliabilität variierte zwischen  $r = .45$  und  $r = .46$  (s. Anhang C).

#### 4.3.1.3 Verhaltensvariablen

*Beckenbodentraining.* Das tatsächlich ausgeführte Beckenbodentraining wurde mit den Angaben der Studienteilnehmer, an wie vielen Tagen der letzten Woche sie das Training durchgeführt hatten, wie häufig (durchschnittlich) pro Tag trainiert wurde und wie lange (durchschnittlich) eine Trainingseinheit in Minuten gedauert hat, erfasst. Dieses Vorgehen zur Erfassung von körperlicher Aktivität wurde bereits in verschiedenen Studien genutzt, um individuelle Angaben zur Quantität des Trainings zu erhalten (Booth, 2000; Scholz, Sniehotta, & Schwarzer, 2005; Sniehotta, Scholz, & Schwarzer, 2005). Zur Baseline-Messung wurden die Studienteilnehmer gefragt, ob sie vor der Operation schon einmal das Training durchgeführt hatten und, falls sie schon regelmäßig trainiert hatten, wie häufig sie das in den letzten sieben Tagen taten und wie lange die Einheiten (in Minuten) andauerten. Zu den Nachbefragungsterminen wurden die Patienten nach ihrem Trainingsverhalten jeweils

innerhalb der vergangenen sieben Tage gefragt. Dabei konnten sie unterscheiden zwischen Training, das sie selbständig ausgeführt hatten, und solchem, zu dem sie in einem Kurs während eines Reha-Aufenthalts oder bei einem ambulanten Physiotherapeuten angeleitet worden waren.

Als Maß des Beckenbodentrainings kann zum einen die *Trainingsdauer* in Minuten pro Woche als Produkt aus der Anzahl der Tage, an denen pro Woche trainiert wurde, der durchschnittlichen Anzahl von Trainingseinheiten pro Tag und der durchschnittlichen Dauer einer Einheit, oder zum anderen die *Trainingsintensität* als durchschnittliche Dauer pro Einheit betrachtet werden. Diese Maße spiegeln sowohl die Quantität des Trainings basierend auf der Häufigkeit und Dauer als auch die Intensität einer durchschnittlichen Trainingseinheit wieder. Die Retest-Reliabilität für die Trainingsdauer variierte zwischen  $r = -.01$  und  $r = .41$ , die für die Trainingsintensität zwischen  $r = -.05$  und  $r = .60$  (s. Anhang C).

*Ernährung.* In einer isolierten Evaluation müssen als Kriteriumsmaße nur die Ziele des Experimentaltreatments, also das Beckenbodentraining überprüft werden, so dass die in Anlehnung an das Beckenbodentraining erfasste Trinkmenge sowie der Obst- und Gemüsekonsum pro Woche nicht näher in die Analyse einbezogen wurden (Hager et al., 2000).

*Weitere Verhaltensmaße.* Zur weiteren Überprüfung des Verhaltens wurde im ersten Fragebogen erfasst, ob Beckenbodentraining den Patienten überhaupt bekannt ist. In den späteren Fragebogen konnten die Patienten angeben, ob sie „aus gesundheitlichen Gründen auf das Training verzichten mussten“. Diese Angabe ist relevant, wenn es um Rückfälle bei der Aufrechterhaltung des Verhaltens geht. Auf beide Items konnte mit einer Zustimmung („ja“) oder Ablehnung („nein“) geantwortet werden.

#### 4.3.1.4 Inkontinenz

Um neben den selbstberichteten Verhaltensmaßen objektivere Maße für die Notwendigkeit des Beckenbodentrainings zu besitzen, wurden die Patienten gebeten, die Stärke ihrer Inkontinenz anzugeben. Dazu wurde die deutsche Version des ICIQ-Fragebogen in seiner kurzen Version (International Consultation on Incontinence Questionnaire – Short Form, ICIQ-SF) der International Consultation on Incontinence (Avery et al., 2004; Hampel, Gillitzer, Wiesner, & Thüroff, 2007) eingesetzt (s. Anhang N). Dieser Fragebogen erfasst mit zwei Items die subjektive Häufigkeit, d.h. wie oft es in den vergangenen sieben Tagen durchschnittlich zu unwillkürlichem Harnverlust kam (0 = „nie“ bis 5 = „ständig“) und die subjektive Stärke, d.h. wie hoch der Harnverlust eingeschätzt wurde (0 = „kein Urinverlust“, 2 = „geringe Menge“, 4 = „mittelgroße Menge“ und 6 = „große Menge“) bezogen auf die

vergangene Woche. Auf einem weiteren Item zur Lebensqualität mit einer elfstufigen Antwortskala (0 = „gar nicht“ bis 10 = „stark“) konnten die Studienteilnehmer zudem angeben, „wie stark ihr Leben durch den Urinverlust beeinträchtigt“ war. Die drei Angaben wurden aufsummiert, so dass der Summenwert mit Werten zwischen 0 und 21 als globales Maß die Inkontinenz als psychische und physische Beeinträchtigung widerspiegelt. Dieser ICIQ-Index besaß interne Konsistenzen zwischen Cronbach's  $\alpha = .40$  und  $\alpha = .79$ . Die Retest-Reliabilität variierte zwischen  $r = .01$  und  $r = .81$  (s. Anhang C).

Der ICIQ-SF ist trotz seiner Kürze aussagekräftig und gut validiert. So wurden von Karantanis, Fynes, Moore und Stanton (2004) die subjektiven Angaben zur Häufigkeit und Stärke der Inkontinenz sowie zur Beeinträchtigung durch die Inkontinenz mit einem 24-h-Vorlagen-Test verglichen, indem die Menge bzw. das Gewicht des in Vorlagen aufgefangenen Harns von inkontinenten Personen mit ihren Angaben im ICIQ-SF korreliert wurde. Kontrolliert für das Alter der Versuchspersonen ergaben sich eine signifikante Korrelation des 24-h-Vorlagen-Tests mit dem ICIQ-Index sowie den drei Items zwischen  $r = .28$  und  $r = .58$ . Ebenfalls signifikante Zusammenhänge fanden sich mit den Angaben in einem Inkontinenztagebuch zur durchschnittlichen Häufigkeit der Inkontinenzepisoden (Kendalls tau-b = .33 bis tau-b = .44). Unter Berücksichtigung dieser Angaben zur Validität wurde in dieser Studie nur der ICIQ-SF eingesetzt, so dass den Patienten die Durchführung eines Vorlagen-Tests erspart werden konnte.

Über diese Angaben hinaus wurde erfasst, wie viele Vorlagen von den Patienten in der vergangenen Woche durchschnittlich pro Tag benötigt wurden.

#### **4.4 Umgang mit fehlenden Werten**

Fehlende Werte sind darauf zurückzuführen, dass (a) einzelne Items nicht beantwortet wurden (Item-non-response) sowie dass (b) Studienteilnehmer vorzeitig die Studienteilnahme beendeten (s.o.) oder (c) einzelne Messzeitpunkte ausließen (unit-non-response). Acht Personen ließen entweder den Fragebogen am zweiten Tag, zwei Wochen oder drei Monate nach der Operation aus, füllten die folgenden Fragebogen aber wieder aus. Ein Patient ließ die beiden Fragebogen zwei Wochen und drei Monate nach der Operation unbeantwortet. Die fehlenden Werte, die sowohl durch Unit-non-response als auch durch Item-non-response entstanden, wurden durch Schätzwerte ersetzt. Für 17.5% der zu vergebenden Werte fehlten über alle Personen und alle Messzeitpunkte hinweg Angaben. In dieser Studie wurden die fehlenden Werte mit Hilfe des EM-Algorithmus des Programms NORM – Version 2.02 (Schafer, 1998) ersetzt. Dieser Algorithmus schätzt die fehlenden Werte so, dass „die gesamte



---

Information im Datensatz widerspruchsfrei und maximal plausibel ist“ (Wirtz, 2004). EM ist ein zweischrittiges, iteratives Maximum-Likelihood-Verfahren. Im ersten Schritt (*estimation*) werden die fehlenden Werte aufgrund der aus den vorhandenen Werten berechneten Parametern geschätzt. Im zweiten Schritt (*maximization*) werden aus den im ersten Schritt ermittelten Werten der Mittelwertvektor und die Kovarianzmatrix berechnet, die zur erneuten Kalkulation der Parameter für den ersten Schritt dienen. Auf diese Weise wird das Vorgehen so lange wiederholt, bis die Kovarianzmatrizen zweier aufeinander folgender Durchgänge konvergieren. Das EM-Verfahren hat das weniger strenge „missing at random“-Kriterium als Voraussetzung (Little & Rubin, 1987). Gegenüber dem z. B. listenweisen oder paarweisen Fallausschluss hat das EM-Verfahren den Vorteil, dass sich die Stichprobe und damit die Teststärke nicht sowie die Varianz der Ausprägungen durch die Schätzung nur wenig reduzieren.

Die fehlenden Werte wurden auf dem Itemlevel je Skala geschätzt, so dass die bestehenden Skalenitems sowie das Alter der Studienteilnehmer als Prädiktoren in die Berechnung eingingen. Für die längsschnittliche Schätzung der fehlenden Werte zum Messzeitpunkt  $t$  ging zudem der jeweilige Skalenwert zum Messzeitpunkt  $t-1$  in die Schätzung mit ein. Auf diese Weise konnten die Analysen mit einem kompletten Datensatz zu allen Messzeitpunkten berechnet werden.

Eine häufig geäußerte Kritik am EM-Algorithmus ist das Risiko der Homogenisierung. Vor allem wenn keine adäquaten Prädiktoren eingesetzt wurden, könnten die geschätzten Werte zu nah am Mittelwert liegen, so dass die Varianz reduziert wird (Little & Rubin, 1987). Das hier beschriebene Schätzvorgehen auf Itemebene versucht jedoch dieses Risiko zu minimieren.

#### **4.5 Auswertungsmethoden**

Die statistischen Analysen wurden mit Hilfe der Statistik-Programme „SPSS 15.0 für Windows“ und „AMOS 7.0“ (Arbuckle & Wothke, 1999) durchgeführt. Für die Dropout-Analyse wurden multivariate einfaktorielle Varianzanalysen und  $\chi^2$ -Test durchgeführt. Die Validität der dyadischen Planung wurde mittels Hauptkomponentenanalysen, Pearson-Korrelationen, univariaten Varianzanalysen und  $\chi^2$ -Tests überprüft. Der Randomisierungskontrolle erfolgte mittels multivariater zweifaktorieller Varianzanalysen, um die Nullhypothesen zu überprüfen, dass die Interventionsgruppen sich nicht in den zur Baseline gemessenen Merkmalen unterscheiden (Weinfurt, 2000). Auch hier wurde zum Schutz vor einem  $\beta$ -Fehler das Signifikanzniveau auf  $p > .20$  festgelegt (Bortz, 1999). Die zwei wichtigsten Analysearten

zur Überprüfung der Hypothesen waren zwei- bzw. dreifaktorielle Varianzanalysen ohne und mit Messwiederholungen sowie multiple Regressionsanalysen als Standard- oder manifeste Pfadanalysen.

*Varianzanalysen mit Messwiederholungen.* Varianzanalysen mit Messwiederholungen lassen Veränderungen eines Merkmals über die Zeit in Abhängigkeit von verschiedenen Gruppenzugehörigkeiten (Faktorausprägungen) zu, indem die Veränderung der Mittelwerte über alle Personen hinweg (Zeitfaktor) als auch zwischen den Personen betrachtet werden (Weinfurt, 2000). Neben der Beachtung gleicher Zellbesetzungen ist es wichtig, Varianzhomogenität zwischen den einzelnen Faktorausprägungen mittels Box's M-Tests zu prüfen. Aufgrund der gleichbesetzten Gruppen wurden kaum Abweichungen in den Kovarianzmatrizen erwartet. Die *multivariaten* F-Statistiken zur Überprüfung von Veränderungen über die Zeit erfordern keine Sphärizität, d.h. keine Homogenität der Interkorrelationen aufeinander folgender Messzeitpunkte. Um die Baseline-Unterschiede der motivationalen Selbstwirksamkeit, der negativen Kontrolle oder der nerverhaltenden Operationstechnik zu berücksichtigen, wurden zusätzlich Kovarianzanalysen durchgeführt. Die Überprüfungen der Annahmen der Normalverteilung, der Linearität, der Varianz- und Regressionshomogenität sowie der Reliabilität der Kovariaten waren zufrieden stellend. Die Ergebnisse der Kovarianzanalyse wurden dann berichtet, wenn sie von den nicht adjustierten Ergebnissen abwichen.

*Pfadanalysen.* Mit Multi-Gruppen-Pfadanalysen können komplexe Analysen zur Wirkung der Planungsintervention berechnet werden. Zur Schätzung und Testung der Modellparameter wurde der in der Software AMOS 7.0 (Arbuckle & Wothke, 1999) implementierte Maximum-Likelihood-Ansatz verwendet (Hoogland & Boomsma, 1998). In der vorliegenden Studie wurde der Interventionsfaktor *Planinhalt* (Beckenbodentraining = 1 vs. gesunde Ernährung = 0) neben weiteren Prädiktoren zur Vorhersage des Beckenbodentrainings eingesetzt, während der zweite Interventionsfaktor *Planungssetting* als Moderator diente. Als kategoriale Variable bestimmte dieser Faktor die beiden Gruppen (dyadisch = 1 vs. individuell = 0), für die die Analysen separat berechnet und miteinander verglichen wurden. So konnte überprüft werden, ob sich die verglichenen Gruppen hinsichtlich ihrer Vorhersagemuster unterschieden. Zur Überprüfung der Bedeutsamkeit einzelner Pfade wurde das *Critical Ratios* (C.R.)-Kriterium herangezogen, das sich aus dem Verhältnis des geschätzten unstandardisierten Parameterwertes und dessen Standardfehler ergibt. Entsprechend der t-Werte-Verteilung wurde die Nullhypothese „C.R. = 0“ überprüft. Werte von  $C.R. \geq 1,96$  sprachen mit einer Irrtumswahrscheinlichkeit von 5 % für bedeutsame Beiträge zur Varianzaufklärung.

Darüber hinaus ermöglichten die pfadanalytischen Multi-Gruppen-Analysen die Berechnung moderierter Mediationen bzw. moderierter indirekter Effekte. Von einer *moderierten Mediation* war dann auszugehen, wenn die Stärke eines indirekten Effekts von der Moderatorvariablen abhing (Preacher, Rucker, & Hayes, 2007). Für diese Analysen wurde wie folgt vorgegangen:

Im ersten Schritt wurde aufgrund theoretischer Überlegungen ein Pfadmodell für beide Gruppen geschätzt (Maruyama, 1998). Zur Überprüfung möglicher Moderationen bzw. Settingunterschiede dienten hierarchische Modellvergleiche im Sinne von *nested models*. Dazu wurden alle Pfade für beide Gruppen gleichgesetzt. Indem dieses restringierte Modell in ein weniger restringiertes Modell verschachtelt (*nested models*) wurde, stellte es einen Spezialfall des allgemeineren Modells dar, für das alle Parameter in beiden Gruppen separat geschätzt wurden. Da ein weniger restringiertes Modell weniger Freiheitsgrade besitzt als ein restringierteres Modell, weist es immer einen schlechteren Modellfit auf. Daraus folgt, dass mit Hilfe des  $\chi^2$ -Differenz-Test ( $\Delta\chi^2$ ) – als Differenz der  $\chi^2$ -Werte<sup>1</sup> beider Modelle gemessen an der Differenz ihrer Freiheitsgrade – geprüft werden konnte, ob sich der Modellfit durch die Restriktion signifikant verschlechterte. Ein signifikantes Ergebnis dieses Tests gab an, dass das restringiertere Modell die Daten schlechter als das Kontrollmodell mit weniger Restriktionen erklärte. Einzelne Pfadkoeffizienten mussten somit als verschieden für beide Gruppen angenommen werden. Es lagen Moderationen einzelner Zusammenhänge vor (Homburg & Giering, 2001).

War dies der Fall, wurden im zweiten Schritt Hypothesen-geleitete Modelle erstellt, in denen jeweils ein Pfad für beide Gruppen gleichgesetzt wurde. Diese teilweise restringierten Modelle wurden erneut gegen das unrestringierte Modell getestet. Mit Hilfe dieser Vergleiche ließen sich diejenigen Pfade identifizieren, die ursächlich für die Gruppenunterschiede im Gesamtmodell waren (Maruyama, 1998). Für den abschließenden Analyseschritt wurden die Pfade, für die sich keine Gruppenunterschiede ergaben, gleichgesetzt.

Im dritten Schritt wurde das bestehende Modell darauf hin überprüft, ob im Modell Mediationen bzw. moderierte Mediationen vorkamen. Den neueren Ansätzen (Shrout & Bolger, 2002; Taylor, MacKinnon, & Tein, 2008) folgend ist ein signifikanter totaler Effekt des Prädiktors auf das Kriterium in distalen Modellen keine Voraussetzung mehr für die Mediationsprüfung – im Vergleich zu Baron und Kenny's (1986) schrittweisem Vorgehen. Gründe für diese weniger konservative Annahme sind, dass der totale Effekt nicht signifikant

---

<sup>1</sup> Der  $\chi^2$ -Wert als globales Maß der Modellgüte berechnet sich über Minimum der Diskrepanzfunktion, d.h. der maximalen Ähnlichkeit zwischen Modellvorhersage und Empirie gewichtet durch die Stichprobengröße.

---

werden muss, weil die Mediation im Fall von distalen, längsschnittlichen Modellen mehr statistische Power besitzt (Shrout & Bolger, 2002) oder aufgrund von Suppressionseffekten im Fall, dass die Mediation und der direkte Effekt entgegen gesetzte Vorzeichen aufweisen (Taylor et al., 2008). Ob der indirekte Effekt als signifikant angenommen werden konnte, wurde mittels des Tests der gemeinsamen Signifikanz (*Joint Significance Test*; MacKinnan, Lockwood, Hoffman, West, & Sheets, 2002) überprüft. Dieser bezieht sich darauf, ob sich die jeweiligen am indirekten Effekt beteiligten Pfade signifikant von Null unterscheiden. Eine Schätzung der Größe des indirekten Effekts oder des direkten Effekts ermöglicht diese Methode dagegen nicht. Ursprünglich war dieser Test für indirekte Effekte mit zwei Pfaden vorgesehen, doch kann er im Gegensatz zu anderen Signifikanztests auch leicht auf den Fall mit drei Pfaden angewandt werden. In Modellen mit indirekten Effekten bestehend aus drei Pfaden zeigt der Test der gemeinsamen Signifikanz eine Mediation an, wenn jeder der drei Pfade signifikant von Null abweicht (Taylor et al., 2008). Der Test gemeinsamer Signifikanz wird von Taylor und Kollegen (2008), die sechs Methoden der Mediationstestung miteinander verglichen, gemeinsam mit zwei Bootstrapverfahren (percentile bootstrap, bias-corrected bootstrap) zur Testung von Drei-Pfad-Mediationseffekten empfohlen. Indem AMOS als statistisches Auswertungsprogramm in der vorliegenden Arbeit gewählt wurde, konnte das Bootstrapverfahren nicht zum Zweck des Signifikanztests herangezogen werden. Mit AMOS kann lediglich der indirekte Effekt des Gesamtmodells getestet werden, nicht aber der Drei-Pfad-Effekt allein. Würden alternativ alle nicht am direkten oder indirekten Drei-Pfad-Effekt der Planungsintervention auf das Beckenbodentraining auf Null gesetzt, wäre dieses Modell mit dem eigentlichen Modell hinsichtlich der Pfadkoeffizienten und damit auch hinsichtlich des Drei-Pfad-Effekts nicht mehr zu vergleichen. Die Empfehlung der konservativen Methode, des Tests gemeinsamer Signifikanz, beruht auf Ergebnissen zur Schnelligkeit und Einfachheit der Anwendung bei gleichzeitig geringer  $\alpha$ -Fehler-Wahrscheinlichkeit und hoher Teststärke (Taylor et al., 2008).

*Umgang mit Ausreißern.* Die Variablenverteilungen wurden auf uni- und multivariate Ausreißer untersucht. Zunächst wurden univariate Ausreißer als solche Werte definiert, die weiter als drei Standardabweichungen vom Mittelwert entfernt lagen (Tabachnick & Fidell, 2001). In den Verteilungen der sozialen und sozial-kognitiven Variablen traf dies für vier Variablen der Baselinemessungen zu, auf denen eine bis drei Personen besonders niedrige Werte aufwiesen. Sie wurden aufgrund dessen, dass sie innerhalb des normalen Wertebereichs lagen und an die übrige Stichprobe angeschlossen waren, nicht verändert. Die Verteilungen des Beckenbodentrainings hinsichtlich der Trainingsintensität einer Trainingseinheit und

---

wöchentlichen Gesamttrainingsdauer zeigten dagegen zu allen Messzeitpunkten deutlich extremere Ausreißer, bei denen zwischen einer und sieben Personen Werte aufwiesen, die mehr als drei Standardabweichungen über dem Stichprobenmittel lagen. Ihr Einfluss wurde reduziert, indem ihre Werte auf die nächst höhere Einheit (bbtshw: + 1x, bbtsti: + 5 min, bbtsmin: + 5 min) oberhalb des Wertes, der noch an die Stichprobe angeschlossen war, angepasst wurden.

Auf multivariate Ausreißer wurden die Daten mit Hilfe der Mahalanobis-Distanz in vorangeschalteten Regressionsanalysen oder getrennt für die Moderator- bzw. Settinggruppen mit AMOS untersucht. Die Mahalanobis-Distanz identifiziert ungewöhnliche Wertekombinationen der unabhängigen Variablen sowie Fälle, die einen großen Einfluss auf das Regressionsmodell ausüben. Multivariate Ausreißer sind durch einen signifikanten  $\chi^2$ -Wert ( $p < .001$ ) bei Freiheitsgraden in Höhe der Prädiktorenzahl des Regressionsmodells gekennzeichnet (Tabachnick & Fidell, 2001). Es wurden keine multivariaten Ausreißer ermittelt.

*Effektgrößen.* Um die Effekte der Intervention besser einschätzen zu können, wurden für die Analysen entsprechende Effektstärken angegeben. Die Effektstärken von Gruppenunterschieden entsprechend der beiden Planungsinterventionsfaktoren wurden mittels des Differenzmaßes  $d$ , als standardisierte Differenz der Mittelwerte zweier Gruppen quantifiziert, wobei  $d = .20$  einen kleinen,  $d = .50$  einen mittleren und  $d = .80$  einen großen Effekt angibt. Pearson-Korrelationen von  $r = .10$  wurden als kleiner, von  $r = .30$  als mittlerer und von  $r = .50$  als großer Effekt betrachtet. Den Beitrag der einzelnen Haupt- und Interaktionseffekte zur Varianzaufklärung in der Varianzanalyse wurden mittels  $\eta^2$  angegeben, wobei  $\eta^2 = .10$  als kleiner,  $\eta^2 = .25$  als mittlerer und  $\eta^2 = .40$  als großer Effekt gilt (Bortz & Döring, 1995; Cohen, 1988).