Prognostic Indicators of Poor Short-Term Outcome of Physiotherapy Intervention in Women With Stress Urinary Incontinence

Erik J.M. Hendriks,[1,2,3*,**] Alfons G.H. Kessels,[1,2†] Henrica C.W. de Vet,[4‡]
Arnold T.M. Bernards,[5§] and Rob A. de Bie[1,2,3¶]

[1]Department of Epidemiology, Maastricht University, Maastricht, The Netherlands
[2]School for Public Health and Primary Care, CAPHRI at Maastricht University, Maastricht, The Netherlands
[3]Centre for Evidence Based Physiotherapy, Maastricht University, Maastricht, The Netherlands
[4]EMGO Institute, VU University Medical Center, Amsterdam, The Netherlands
[5]Department of Research and Development, Dutch Institute for Allied Health Care, Amersfoort, The Netherlands

Aims: To identify prognostic indicators independently associated with poor outcome of physiotherapy intervention in women with primary or recurrent stress urinary incontinence (stress UI). **Methods:** A prospective cohort study was performed in physiotherapy practices in primary care to identify prognostic indicators 12 weeks after initiation of physiotherapy intervention. Patients were referred by general practitioners or urogynecologists. Risk factors for stress UI were examined as potential prognostic indicators of poor outcome. The primary outcomes were defined as poor outcome on the binary Leakage Severity scale (LS scale) and the binary global perceived effectiveness (GPE) score. **Results:** Two hundred sixty-seven women, with a mean age of 47.7 (SD = 8.3), with stress UI for at least 6 months were included. At 12 weeks, 43% and 59% of the women were considered recovered on the binary LS scale and the binary GPE score, respectively. Prognostic indicators associated with poor outcome included 11 indicators based on the binary LS scale and 8 based on the binary GPE score. The prognostic indicators shared by both models show that poor recovery was associated with women with severe stress UI, POP-Q stage > II, poor outcome of physiotherapy intervention for a previous UI episode, prolonged second stage of labor, BMI > 30, high psychological distress, and poor physical health. **Conclusions:** This study provides robust evidence of clinically meaningful prognostic indicators of poor short-term outcome. These findings need to be confirmed by replication studies. *Neurourol. Urodynam.* 29:336–343, 2010. © 2009 Wiley-Liss, Inc.

Key words: global perceived effectiveness; leakage severity scale; multivariable logistic regression; physiotherapy; prognostic indicators; prospective cohort study; stress urinary incontinence

INTRODUCTION

Stress urinary incontinence (stress UI) is a common health problem, which is not only a medical problem, but also a major burden on the health care system, implying increasing costs to society.[1,2] Prevalence figures in the Netherlands suggest that approximately 5−7% of adults have UI,[1] increasing among the elderly to 9% for men and 29% for women.[3,4]

Pelvic floor muscle training (PFMT) is generally considered as the first-line intervention for women with stress UI. It is relatively cheap and has hardly any adverse effects.[5−7] Systematic reviews of randomized clinical trials (RCTs) have shown that PFMT alone (or with adjunctive therapies) is an effective intervention for women with stress UI,[5−11] with "cure/improvement" rates ranging from 50% to 97%. However, definitions of "cure" or "improvement," as well as the populations, type and duration of interventions studied were highly variable, and the presence of prognostic factors may differ between and within these studies. There are still gaps in our knowledge on prognostic indicators associated with poor outcome.

Little research has been done to see whether risk factors for UI also have prognostic value and whether these are modifiable by targeted interventions. Although many risk factors have been suggested, the most consistently reported factors for stress UI in adult women are vaginal delivery, obesity,[12] limited physical activity, family history, ethnicity, genitourinary prolapse, depression, neurological disease (e.g., stroke and Parkinson), and age.[13−15] Stress UI might also be associated with obstetric risk factors such as a prolonged second stage of labor[16] and forceps-[17,18] or

vacuum-assisted delivery.[19,20] Episiotomy has not been found to increase the risk of stress UI, nor does it seem protective,[14] but is generally considered a risk factor for fecal incontinence.[13] Women who developed stress UI during pregnancy or after childbirth were also found to be at risk for long-term stress UI.[19] There is also support from the literature for other risk factors such as generalized disease processes (e.g., chronic obstructive pulmonary disease (COPD) and cardiovascular disease),[13,21] poor physical health[21−23] and psychological distress.[24−28] Recurrent stress UI could also arise after urogenital or incontinence surgery, with reported rates of 10−60%.[29] However, not all causes of recurrent stress UI are related to surgical techniques, type of material used,[30] or the experience of urogynecologists, as has been suggested.[31] Several mechanisms may coexist in one woman, hampering the identification of individual risk factors. The cause of stress UI is likely to be multi-factorial, thus attributable to a combination of risk factors simultaneously or in sequence, which may be different for different women.

To our knowledge, no study has examined prognostic indicators of stress UI of women receiving physiotherapy intervention. The objective of this prospective cohort study was to identify prognostic indicators of poor outcome 12 weeks after initiation of physiotherapy intervention in women with primary or recurrent stress UI episodes.

METHODS

Study Design and Patients

Patients enrolled in this prospective, multi-center, cohort study from August 2002 until December 2004. Data were collected among women with a primary or recurrent episode of stress UI referred for physiotherapy intervention by general practitioners (GPs) or urogynecologists. Stress UI was defined as involuntary leakage during effort, exertion or sneezing, and/or coughing.[32]

Seventy-one physiotherapists experienced in women's health, who had commented on the draft Dutch physiotherapy guideline on "Stress UI in adult women" (www.cebp.nl, guidelines) during its development, were invited to participate in this study.[33] Physiotherapists were eligible for participation if they worked in primary care, had (in their opinion) a good working relationship with the referring physicians and were willing to treat at least five consecutive patients during the enrolment period. All participating therapists were given a short training and were prepared for this study in two interactive educational meetings, during which the key recommendations of the guideline[33] were discussed. Instructions were given for using the standardized patient questionnaires, outcome measures, and physical examination to collect patient data.

The cohort of patients treated by these physiotherapists was followed from the initial visit to the end of physiotherapy intervention. No maximum or minimum number of sessions was prescribed, but numbers were tailored to the patient's individual needs. Physiotherapy intervention ended when no further improvement was expected. No effort was made by the researchers to alter or modify the course of physiotherapy intervention during the study, which was left to the therapists' discretion. Therapists were asked to record on a standard form the date and duration of each session, the goals and type of interventions applied in each session, and the reason to end the physiotherapy intervention.

Patients to be included in the cohort were restricted to women over 18 years of age with a diagnosis of stress UI that was urodynamically proven or based on history and physical examination, and a minimum duration of at least 6 months, to exclude women with a favorable natural course. Recurrent stress UI was defined as a symptom-free episode (less than once a month and only a small amount of urine loss) for at least 12 months preceding the current episode. Women had to be able to read and understand Dutch. Women were excluded when living in residential or nursing homes, having fecal incontinence, being pregnant or within 6 months after delivery, post-operative status within 6 months, exhibiting signs of active urinary tract infection, central neurological pathology (e.g., stroke and Parkinson's disease), diabetic neuropathy, congenital urological disorders, cognitive impairments or bladder cancer, using diuretics or having been classified as having mild stress UI. Patients having received additional non-physiotherapy related interventions within 6 months before entry or during the study were excluded. In our opinion, a restricted data set should be used for prognostic analyses when focusing on a particular intervention.

Patients reporting mixed UI, or where the therapist was unsure of the correct diagnosis, received a 4-day self-evaluation diary at baseline, including a weekend, to assess whether they had predominantly stress UI or predominantly urgency UI.[34,35] If they reported more urgency than stress UI episodes, they were classified as having predominantly urgency UI[34,35] and therefore excluded from this cohort. Urgency UI was defined as the involuntary loss of urine accompanied by or immediately preceded by urgency.[32,36]

Patients were informed about all aspects of this project by written and personal information at inclusion. Those who met the criteria and provided written consent were registered and their names were forwarded to the research team. The Medical Ethics Committee of the Deventer Hospitals, The Netherlands, approved the study.

Outcome Measures

We defined two primary outcomes: the "Leakage Severity" subscale of the PRAFAB-questionnaire (PRAFAB-Q) and the binary global perceived effectiveness (GPE) score. Patients were asked to complete the PRAFAB-Q at baseline, 12 weeks after initiation of the physiotherapy intervention, and at the end of physiotherapy intervention. The PRAFAB-Q consists of two subscales: the Leakage Severity (LS) scale (including the items *Pr*otection [use of pads], *A*mount of urine loss' and "*F*requency of UI") and the Perceived Impact scale, (including the item's *A*djustment [restriction of daily activities or participation due to the UI symptoms] and *B*ody or self-image [bother].[37] The PRAFAB-Q has excellent test−retest reliability, good construct validity, and excellent responsiveness to change in women with stress or predominantly stress UI.[35,38] A maximum of 4 points (min−max = 1−4) can be given per item. For this prognostic study we were only interested in the LS scale. Scores on the LS scale range from 3 to 12. A higher total score indicates more severe condition.[35,39] Recovery on the binary LS scale was defined as ≤4 points, which is almost the best attainable score for the items of frequency, amount, and protection.

At follow-up assessments, we determined participants' perceptions of the effectiveness of treatment and asked them to rate their overall change in health status from baseline, using the GPE score, by answering the question: "How does your current condition compare to what it was like before you started the physiotherapy?" Patients were classified into nine groups: (1) very much better, (2) much better, (3) moderately better, (4) slightly better, (5) unchanged, (6) slightly worse, (7) moderately worse, (8) much worse, and

(9) very much worse.[35] The GPE score was dichotomized into the binary GPE score with the categories "moderately to very much better" considered as recovered.

To obtain unbiased results for the main outcome at 12 weeks, patients were provided with a return envelope in which to place their completed PRAFAB-Q and GPE score, and were assured that their physiotherapist would not see their responses. Patients were not informed of their previous PRAFAB-Q scores when they completed the follow-up questionnaire. The binary LS scale provides relevant information on the best attainable score on LS items, whereas the GPE score indicates what patients consider to be a minimal important change.

Prognostic Indicators and Data Collection

The risk factors for a primary or recurrent stress UI episode that could potentially be used as prognostic indicators of poor recovery were examined in baseline assessments using a patient's self-administered questionnaire and standardized referral forms for physicians. The referral data and patients' individual responses to the questionnaire items were reviewed and checked for completeness by the physiotherapists at the initial visit, followed by an extended physical examination.

The following six categories of prognostic indicators were considered: (1) socio-demographic characteristics, (2) clinical history, (3) obstetric characteristics, (4) previous interventions, (5) general health characteristics, and (6) anti-incontinence medication (Table I). Self-assessed health was assessed by the question: "Compared to other women of your age, how do you perceive your physical health?" and was represented in the analyses as a dichotomous variable (excellent, very good, good, and fair vs. poor). Psychological distress was measured as suggested by Fultz and Herzog[25] by the questions: "Did you feel depressed most of the time during the past week? Did you feel lonely most of the time during the past week? Did you feel sad most of the time during the past week?" Answer categories were dichotomized into yes or no. High psychological distress was defined as all questions being answered affirmatively. To limit the number of potential prognostic indicators, different types of co-morbidity, forceps- or vacuum-assisted delivery, prolapse and incontinence surgeries, and co-interventions (i.e., anti-incontinence medication) were combined and clustered into single variables.

Statistical Analysis

All data were anonymously processed by the research physician (A.B.) and checked by our research assistant (J.G.) and the department secretary. Descriptive statistics were used to summarize demographic, clinical, and treatment characteristics of the study sample and the two binary outcome measures, the LS scale and GPE score.

Multivariable logistic regression analyses with specified prognostic indicators were applied to the two binary outcomes. We explored linearity between the continuous prognostic indicators (age and parity) and the main outcomes. Multicollinearity was examined using the variance inflation factor (VIF), the reciprocal tolerance statistic (1/VIF), and a correlation matrix of all prognostic indicator variables. VIF values indicate whether an indicator has a strong linear association with the other indicators. The average VIF value of all indicators should not be substantially greater than 1 and the reciprocal tolerance statistics should not be lower than 0.20.[40]

TABLE I. Baseline Characteristics of Prognostic Cohort (N = 267): Number (%) Unless Otherwise Stated

Potential prognostic indicators	Number (%)
(I) Socio-demographic	
Age (year) [mean (SD)]	47.7 (8.3)
≤45	89 (33.3)
46−63	178 (66.7)
Education[a]	
Low	67 (25.1)
Intermediate	126 (47.2)
High	74 (27.7)
(II) Clinical history	
Type of UI	
Stress	228 (85.4)
Mixed (dominant stress)	39 (14.6)
Stress UI severity [mean (SD)]	12.2 (2.4)
Severe (≥14 points)[b]	81 (30.3)
Menopausal status (peri- and menopausal)	101 (37.8)
Pelvic organ prolapse stage[c]	
0	184 (68.9)
1−2	65 (24.3)
3	18 (6.7)
(III) Obstetric	
Parity [mean (SD)]	1.64 (1.3)
Null parity	81 (30.3)
1	20 (7.5)
2	106 (39.7)
≥3	60 (22.5)
Prolonged 2nd stage of labor (>90 min)	39 (14.6)
Forceps- or vacuum-assisted delivery (yes)	14 (5.2)
Perineal laceration[d]	
≤1st degree	172 (64.4)
2nd degree	56 (21.0)
3rd degree	39 (14.6)
(IV) Previous intervention	
Previous physiotherapy[e]	
Good outcome	50 (18.7)
Poor outcome	25 (9.4)
Previous urogynecological surgery[e]	
Stress UI surgery (yes)	44 (16.5)
Prolapse surgery (yes)	46 (17.2)
Stress UI and prolapse surgery (yes)	21 (7.9)
(V) General health	
Weight [mean (SD)]	26.4 (5.1)
Body mass index (≥30)	26 (9.7)
Psychological distress (yes)	50 (18.7)
Physical health status (poor)	55 (20.6)
Co-morbidity[f] (yes)	53 (19.9)
(VI) Co-intervention	
Anti-incontinence medication (yes)	31 (11.6)

[a]Education: low = no education, primary school or lower vocational school; intermediate = lower general secondary or middle vocational school; high = higher vocational school or university.
[b]A total PRAFAB-Q score ≥14 points was defined as severe stress UI.[37]
[c]Based on the method of the Pelvic Organ Prolapse Quantification (POP-Q) system.
[d]Including episiotomy in at least one delivery.
[e]Recurrent stress UI.
[f]Co-morbidity included: COPD (n = 31) and cardiovascular problems (n = 22).

For the multivariable logistic regression analyses we followed Steyerberg's rules for relatively small data sets.[41−43] The models were constructed for each of the two binary outcomes, using all specified prognostic indicators, and forced into a full multivariable logistic regression model to control for confounding or inter-correlations between indicators, and calculate mutually adjusted odds ratios

(OR) for both binary outcomes of recovery. Indicators with a P-value ≥0.50 were then omitted one by one from the model using the indicator with the highest P-value first (non-automated backward selection).[42,44] This procedure was repeated, as recommended by Steyerberg et al.,[42] until all indicators with a P-value <0.50 were retained in the final model. A liberal P-value <0.50 increases the chances of identifying true predictors, limiting the bias in selected coefficients.[41,42] We examined the stability of the models and forced back the excluded prognostic indicators one at a time to examine significant change in prognostic indicators and regression coefficients (followed by consideration of precision). Adjustments were made for anti-incontinence medication. We did not examine interaction terms but relied on the main effects of the prognostic indicators. The results are presented as OR with corresponding 95% confidence interval (95% CI). OR smaller than 1 indicate an increased risk of poor outcome relative to the reference category. Prognostic indicators with a significant contribution in the models were based on a P-value ≤0.1. Univariate OR were also calculated and presented to reflect the strength of each relationship, together with the corresponding 95% CI.

The analyses were performed using SPSS software for Windows (version 15.1, 2007). All patient data were included in the analysis. Occasional missing values on baseline and outcome variables will be completed using the expectation maximization method available in SPSS software. Reported P values were two-tailed.

RESULTS

Study Population

Thirty-six of the 71 eligible physiotherapists (51%), working in 34 different private practices, agreed to participate. Major reasons for not participating were the expected workload (n = 16), problems experienced in the communication with physicians, and expected difficulty in data collection for additional information (n = 14) and lack of time (n = 5). Participating physiotherapists were all women with ample experience in women's health (mean = 11.2 years [SD = 3.3]; range = 8–17 years).

Two hundred seventy-nine women enrolled in the study, 12 of whom were excluded from the analyses because they were classified by the PRAFAB-Q as having mild UI (≤7 points). All 267 women included in the cohort completed the baseline questionnaire items, PRAFAB-Q scores, and the outcome measurements 12 weeks after the initiation of physiotherapy intervention. The mean number of women included per physiotherapist was 7.8 (SD = 2.4; min−max = 4−12). Fifty-four percent of the women had been referred by GPs, the others by urogynecologists. The study sample included primarily women of Caucasian origin (97%) and 29% were employed. The median time span between the first symptoms of stress UI and enrolment was 2 years (1st−3rd IQ range = 1−10), with about 60% enrolling within 3 years. The demographic characteristics and prognostic indicators are presented in Table I. The majority of included women (85%) were classified as having stress UI and 15% as mixed but predominantly stress UI. About two-thirds of the included women were over 45 years of age. The mean baseline UI severity in terms of total PRAFAB-Q score was 12.2 (SD = 2.4; min−max 8−18 points) while 30% (n = 81) were classified as having severe baseline stress UI (≥14 points on the PRAFAB-Q).

Physiotherapy Intervention

The physiotherapists recorded a mean number of 9.5 sessions (SD = 3.2) per patient. Two hundred thirty-five (88%) patients ended the physiotherapy intervention within 12 weeks (mean = 9.2 weeks). The remaining 32 patients ended the intervention within 16 weeks (mean = 13.6 weeks). The sessions lasted approximately 30–40 min. There were no significant differences in mean number of sessions between improved and non-improved women. The relative contributions of the two most common types of intervention showed that PFMT was used in 94% of all sessions, followed by a combination of patient information, including written materials, and some type of educational intervention (92%). The relative contribution of exercises to improve posture and PFM control during activities of daily life was 54%, that of exercises to improve physical health 11%. Adjunctive therapies, like biofeedback or electrical stimulation, were rarely used (<8%). All patients were provided with a home exercise program. None of the patients reported adverse effects.

Outcomes

About 43% (n = 116) of the patients were recovered as defined on the binary LS scale, and 59% (n = 158) as defined on the binary GPE score. The kappa coefficient for agreement between these binary outcome measures was 0.61 (or Spearman's r was 0.64).

Preparation for the Prognostic Analysis

The association between age (in years) as a continuous indicator and the log odds of poor outcome showed no association until the age of 45, and a continuous linear effect above this age. Consequently, age ≤45 years was defined as the reference group. The same pattern was also found for the number of deliveries: no association with the log odds of a poor outcome until two deliveries, but a linear increase in women with three or more deliveries. We dichotomized parity into <3 versus ≥3 vaginal deliveries for the logistic regression analyses. We also explored the inter-correlations between the prognostic indicators and found that menopausal status was moderately correlated with age (Spearman correlation [or Kendall's Tau-b] of 0.46). The mean VIF value of 1.12 (min−max = 1.05−1.34), and corresponding Tolerance statistics all >0.74 indicate no serious threat of multi-collinearity between prognostic indicators. Most of the selected prognostic indicators were univariately associated with a poor outcome in both the outcome measures, and shown in Tables II and III.

Prognostic Indicators of Poor Outcome

The final multivariable logistic regression analyses resulted in 12 prognostic indicators associated with recovery based on the binary LS scale (P-value ≤0.1), 11 of which were associated with a poor outcome (Table II). They included lower education compared to those with a high education, mixed UI, severe stress UI, peri- and post-menopausal status, POP-Q stage III, poor outcome of physiotherapy intervention for a previous UI episode, prolonged second stage of labor (>90 min), BMI greater than 30, high psychological distress, poor physical health, and co-morbidity. Two indicators with a P-value ≥0.50 were not included in the final binary LS scale model, viz. "previous incontinence and prolapse surgery" and parity.

TABLE II. Univariable and Multivariable Logistic Regression Analyses Associated With Poor Outcome on the Binary LS-Scale[†] 12 Weeks After Initiation of Physiotherapy Intervention in 267 Women With Stress UI

Prognostic indicators	Univariable analyses OR (95% CI)	Multivariable analyses[a] OR (95% CI)
Age (>45 years)[b]	0.98 (0.97−0.99)**	0.99 (0.97−1.01)
Education (low-intermediate vs. high)	0.35 (0.20−0.60)**	**0.37 90.16−0.86)**
Type of UI (mixed UI)	0.12 (0.10−0.34)**	**0.04 90.01−0.15)**
Severe stress UI (≥14 points)	0.19 (0.10−0.34)**	**0.09 (0.03−0.21)**
Menopausal status (peri- and menopausal)	0.77 (0.47−1.28)	**0.41 (0.17−0.98)*
Pelvic organ prolapse stage		
0	1	1
1−2	1.52 (0.86−2.68)	0.95 (0.42−2.16)
3	0.08 (0.01−0.58)**	**0.10 (0.01−1.05)*
Previous physiotherapy[c]		
No	1	1
Yes, good outcome	0.77 (0.41−1.45)	1.80 (0.67−4.65)
Yes, poor outcome	0.09 (0.02−0.40)**	**0.05 (0.01−0.32)**
Previous incontinence and prolapse surgery[c] (yes)	0.49 (0.19−1.32)	—
Parity (≥3)	0.58 (0.32−1.06)*	—
Prolonged 2nd stage (>90 min)	0.61 (0.29−1.24)	**0.17 (0.05−0.56)**
Forceps/vacuum (yes)	0.34 (0.09−1.24)*	0.27 (0.04−1.98)
Perineal laceration		
≤1st degree	1	1
2nd degree	1.46 (0.79−2.67)	2.31 (0.89−6.05)*
3rd degree	0.44 (0.20−0.95)*	0.48 (0.15−1.58)
Body mass index (≥30)	0.55 (0.33−1.30)	**0.28 (0.08−0.94)**
Psychological distress (yes)	0.26 (0.13−0.55)**	**0.29 (0.11−0.89)**
Physical health status (poor)	0.16 (0.07−0.36)**	**0.32 (0.11−0.87)**
Co-morbidity (yes)	0.49 (0.26−0.94)**	**0.40 (0.15−1.01)*

OR, odds ratio; 95% CI, 95% confidence interval.

*0.05 ≤ P-value ≤ 0.10.

**P-value <0.05.

[†]Recovery was defined on the binary LS scale as ≤4 points (min−max LS-scale = 3−12 points).

[a]Full model of prognostic indicators with a P-value <0.50. Indicators associated with poor outcome and a P-value ≤0.1 are in bold. Prognostic indicators with a P-value ≥0.5, which were deleted one by one from the final model, were "Previous incontinence and prolapse surgery": OR (95% CI) = 0.26 (0.04−1.95) and "Parity (≥3)": OR (95% CI) = 0.77 (0.34−2.21). Final model was adjusted for anti-incontinence medication: OR (95% CI) = 2.27 (0.66−7.82).

[b]Piecewise transformation in which ≤45 years of age was defined as the reference group and a continuous linear effect above this age.

[c]Recurrent stress UI.

The results of the final multivariable logistic regression analyses of factors associated with a poor outcome on the GPE score comprised eight prognostic indicators, seven of which coincided with those found on the binary LS scale (Table III). Indicators associated with a poor outcome were severe stress UI, POP-Q stage III, poor outcome of physiotherapy intervention for a previous UI episode, three or more vaginal deliveries, prolonged second stage of labor (>90 min), BMI > 30, high psychological distress, and poor physical health. Two indicators with a P-value ≥0.50 were not included in the multivariable logistic regression model based on the binary GPE score, viz. co-morbidity and age.

DISCUSSION

The recovery rate 12 weeks after initiation of the physiotherapy intervention was 43% on the binary LS scale and 59% on the binary GPE score. A comprehensive set of frequently reported risk factors was examined for its association with poor outcome of physiotherapy intervention. The results of this prospective cohort study demonstrate strong negative associations between reported risk factors and recovery, though the associations vary between outcome measures. We identified 11 prognostic indicators associated with poor outcome based on the binary LS scale, 7 of which coincided

with those found based on the binary GPE score. The difference between the two outcome measures is plausible and may be related to the less rigorous definition of recovery, that is, moderately, much and very much better, used for the binary GPE score. The binary LS scale was defined as almost the best attainable score (≤4 points; min−max = 3−12) for the items of frequency, amount, and protection. Hardly any of the patients classified as "recovered" were using pads, illustrating their confidence and ability to control their continence.[35] We consider both outcome measures to be relevant when providing caregivers and patients with information on the potential prognostic indicators associated with poor recovery, to allow informed decisions.

Prognostic Indicators

The prognostic indicators identified in our study on the basis of both outcome measures show that women with severe stress UI at baseline, POP-Q stage > II, poor outcome of physiotherapy for a previous episode, prolonged second stage of labor (>90 min), BMI > 30, high psychological distress, and poor physical health are more likely to show poor recovery. The last three of these indicators might be modifiable by additional interventions (e.g., life-style changes and motivational interviewing) or by targeted interventions by other

TABLE III. Univariable and Multivariable Logistic Regression Analyses Associated With Poor Outcome on the Binary GPE Score[†] 12 Weeks After Initiation of Physiotherapy Intervention in 267 Women With Stress UI

Prognostic indicators	Univariable analyses OR (95% CI)	Multivariable analyses[a] OR (95% CI)
Age (>45 years)[b]	0.99 (0.98–1.00)*	—
Education (low-intermediate vs. high)	0.43 (0.24–0.78)**	0.62 (0.29–1.34)
Type of UI (mixed UI)	0.68 (0.34–1.36)	0.61 (0.26–1.44)
Severe stress UI (≥14 points)	0.38 (0.23–0.64)**	**0.48 (0.23–0.98)**
Menopausal status (peri- and menopausal)	1.02 (0.61–1.68)	0.68 (0.35–1.34)
Pelvic organ prolapse stage		
0		1
1–2	1.99 (1.06–3.72)*	1.70 (0.77–3.75)
3	0.88 (0.02–0.39)**	**0.10 (0.02–0.53)**
Previous physiotherapy[c]		
No		1
Yes, good outcome	0.45 (0.24–0.85)**	0.63 (0.28–1.41)
Yes, poor outcome	0.12 (0.04–0.34)**	**0.10 (0.03–0.34)**
Previous incontinence and prolapse surgery[c] (yes)	0.44 (0.19–1.20)	0.61 (0.20–1.88)
Parity (≥3)	0.43 (0.24–0.78)**	**0.37 (0.17–0.78)**
Prolonged 2nd stage (>90 min)	0.48 (0.24–0.45)**	**0.39 (0.16–0.95)**
Forceps/vacuum (yes)	0.36 (0.12–1.12)*	0.43 (0.11–1.75)
Perineal laceration		
≤1st degree		1
2nd degree	0.81 (0.44–1.19)	0.86 (0.38–1.98)
3rd degree	0.35 (0.17–0.72)**	0.53 (0.21–1.34)
Body mass index (≥30)	0.39 (9.17–0.90)**	**0.21 (0.07–0.60)**
Psychological distress (yes)	0.31 (0.16–0.59)**	**0.41 (0.17–0.99)**
Physical health status (poor)	0.14 (0.07–0.28)**	**0.20 (0.09–0.45)**
Co-morbidity (yes)	0.66 (0.36–1.20)	—

OR, odds ratio; 95% CI, confidence interval.

**P-value <0.05.

*0.05 ≤ P-value ≤ 0.10.

[†]Recovery was defined on the binary GPE score as "moderately to very much better".

[a]Full model of prognostic indicators with a P-value <0.50. Indicators associated with poor outcome and a P-value ≤0.1 are in bold. Prognostic indicators with a P-value ≥0.5, which were deleted one by one from the final model, were "Co-morbidity": OR (95% CI) = 1.12 (0.52–2.75) and "Age": OR (95% CI) = 1.00 (0.98–1.01), Final model was adjusted for anti-incontinence medication: OR (95% CI) = 1.61 (0.53–4.89).

[b]Piecewise transformation in which ≤ 45 years of age was defined as the reference group and a continuous linear effect above this age.

[c]Recurrent stress UI.

experts (e.g., dieticians and psychologists), but are probably hardly modifiable during a 12-week physiotherapy intervention. The negative association between a prolonged second stage of labor and poor outcome may be explained by biological causes resulting in problems with anatomical structures and physiological function (e.g., pelvic floor fascial support), pudendal nerves (supplying the pelvic organs), and the pelvic floor. More generally, if the severity of stress UI, prolapse, or symptoms (bother) is not reduced by a 12-week physiotherapy intervention, patients should be referred to an urogynecologist for further diagnostics or management. Both our prognostic models suggest that age, perineal laceration up to third degree, and forceps- or vacuum-assisted delivery are not associated with poor outcome. These prognostic indicators might have a favorable natural course or have the potential to be improved to a satisfactory level by physiotherapy intervention.

The simplest method to predict poor outcome might be based on those factors on which both models agree, or on the prognostic indicators associated with poor outcome in terms of GPE score. Although the GPE score may not be perfect as a gold standard, it has high face validity and is strongly associated with both the total PRAFAB-Q[35] and the LS scale, as was demonstrated by the high kappa coefficient for agreement in this and a previous study.[37]

Strengths and Weaknesses

The strength of this multi-center study is the completeness of the data at baseline and at 12 weeks and the blinded outcome measurements. The prognostic indicators and outcome measures we used were simple and easy to collect and did not require extensive training for the therapists. Despite its completeness, our study may have had some limitations, including patient selection bias, potential recall bias, and misclassification of the prognostic indicators and outcome measurements. Recruitment of eligible participants was left to the physiotherapists, and it is possible that not all newly referred and eligible patients were consecutively included. Recall bias for outcome measurements is not very likely because patients completed the simple and easy outcome questionnaires at 12 weeks, returning them within 7 days by return envelope, so that the physiotherapists were unaware of the results. Completion of all forms in the study protocol required a good relationship and easy communication between physiotherapists and referring physicians. The therapists checked the baseline questionnaire for completeness and consistency with referral data. Although it was easy for them to contact the referring physicians for additional information or discussion, this does not guarantee that aspects like the degree of perineal laceration are correctly

coded in the physician's records, as was demonstrated by Brubaker et al.[45] The same uncertainty applies to the location of episiotomy and whether it was elective or restricted to selected cases. The impact of various obstetric interventions on the development of stress UI, including forceps- or vacuum-assisted delivery and episiotomy, which might lead to serious perineal trauma, is still controversial but seems not to be associated with poor outcome.

Prognostic research aims to find causal associations. A prognostic indicator is believed to be causally related to an outcome if the association remain when confounding is excluded. Nevertheless, our findings need to be confirmed in other studies and other populations. The key difference between RCTs and observational studies is that cohort studies are subject to confounding, which might threaten their validity, so careful attention to confounders is a critical aspect. Our physiotherapy intervention was not strictly protocolized, as is often the case in RCTs, but based on a clinical guideline for the diagnosis and treatment of women with stress UI.[33] This may have resulted in more treatment heterogeneity or confounding by indication, which is difficult to adjust in observational research. However, the prognostic indicators were studied in a sample of community-dwelling women, using a comprehensive set of potential risk factors and data collected by a large number of therapists and centers. We therefore believe that the results are more generalizable to other community-based cohorts than those based on highly selected patients as in RCTs. Patients in RCTs do not really reflect the many variable conditions, co-morbidities, and other health status characteristics of patients in everyday clinical practice. The estimated improvement rates in our study were in line with those based on systematic reviews of RCTs, and are not overoptimistic. However, this does not guarantee unbiased estimates of the treatment effects. Furthermore, other prognostic indicators not measured in this study (e.g., smoking behavior or a family history) may have affected the natural or clinical course of recovery. It should be noted that the prognostic models might be less applicable for the younger (<30 year) and older patient groups (>60 year), as both groups were underrepresented in this cohort (about 4% and 5%, respectively). In addition, we cannot compare the results of our cohort with those found in other study populations because, as far as we know, no similar study has been done before. Hence, application of our models in other community-based populations and different clinical settings is necessary to confirm the generalizability of these prognostic indicators of poor outcome.

Another point needing attention is that we included many indicators in the models compared to the frequency of the outcomes. Overfitting may decrease the stability of the model and is mainly determined by the number of events in the smallest outcome category of patients.[41–43,46,47] Overfitting bias was minimized by using Steyerberg rules[42,43] by excluding prognostic indicators with a P-value ≥0.50. The stability and precision of the regression coefficients of the final models were examined by forcing back the excluded prognostic indicators one at a time to examine significant change in the regression coefficients, which we did not observe. The rule of thumb that logistic models should be used with a minimum of 10 events per variable is based on a few simulation studies,[46] in which only the number of events was varied and not the sample size and distribution of effects of the binary predictors. Vittinghof and McCulloch[46] conducted a large simulation study of other influences on confidence intervals, type I error, relative bias, and model performance measures. They concluded that the rule of thumb of 10 can be relaxed. The

prevalence of the smallest outcome category was 41%, which is rather high. In this respect, the standardized response mean[48] (ratio of individual change to the SD of that change) of both outcome measures were large (>2.0, data not shown) and we did not find a serious threat of multi-collinearity, which may have prevented serious bias and unduly optimistic results.[47,49]

Relevance of the Findings

If our prognostic indicators prove to be valid, the results of our study may have several implications for clinical practice and research. First, the results can be used to inform patients about their prognosis and to schedule additional targeted interventions for related problems that also need to be addressed. These indicators could also be used in randomized controlled trials to achieve a balance between control and intervention groups. Differences in the prevalence of prognostic factors might also explain the variable cure or improvement rates reported in systematic reviews.[5–8] Our findings underscore the importance of developing and incorporating other, targeted interventions for potentially modifiable risk factors, with greater emphasis on patient education to improve self-control and self-efficacy, behavioral strategies, life-style changes, physical health or combined interventions, together with other health professionals, to address psychological distress (depression), body weight, and adequate management of co-morbidities like COPD and cardiovascular diseases.

CONCLUSION

Identifying patients at risk of poor outcome is crucial for clinical practice and research. This study provided robust evidence for clinically meaningful prognostic indicators of poor short-term outcome of physiotherapy for stress UI. However, these findings need to be confirmed in other community-based populations and clinical settings. Further research should also evaluate specific treatment strategies based on these prognostic indicators of poor outcome.

REFERENCES

1. CVZ. Urine-incontinentie bij volwassenen (Urinary incontinence in adults). Diemen, The Netherlands: College voor Zorgverzekeringen (CVZ); 2003.
2. CVZ. Attentie voor Incontinentie (Attention to Incontinence). Diemen, The Netherlands: College voor Zorgverzieningen (CVZ); 2006.
3. Teunissen TA, van den Bosch WJ, van den Hoogen HJ, et al. Prevalence of urinary, fecal and double incontinence in the elderly living at home. Int Urogynecol J Pelvic Floor Dysfunct 2004;15:10–3; discussion 13.
4. Teunissen TAM. Urinary incontinence in the elderly in general practice [Dissertation]. Radboud University, 2006.
5. Hay-Smith EJ, Bo Berghmans LC, Hendriks HJ, et al. Pelvic floor muscle training for urinary incontinence in women. Cochrane Database Syst Rev 2008;(3):CD001407.
6. Neumann PB, Grimmer KA, Deenadayalan Y. Pelvic floor muscle training and adjunctive therapies for the treatment of stress urinary incontinence in women: A systematic review. BMC Womens Health 2006;6:11.

7. Shamliyan TA, Kane RL, Wyman J, et al. Systematic review: Randomized, controlled trials of nonsurgical treatments for urinary incontinence in women. Ann Intern Med 2008;148:459−73.

8. Hay-Smith EJ, Dumoulin C. Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women. Cochrane Database Syst Rev 2006 Jan 25;(1):CD005654.

9. Latthe PM, Foon R, Khan K. Nonsurgical treatment of stress urinary incontinence (SUI): grading of evidence in systematic review. BJOG 2008; 115:435−44.

10. Rogers RG. Urinary Stress Incontinence in Women. N Engl J Med 2008; 358:1029−36.

11. Dumoulin C, Hay-Smith J. Pelvic floor muscle training versus no treatment for urinary incontinence in women. A Cochrane systematic review. Eur J Phys Rehabil Med 2008;44:47−63.

12. Hunskaar S. A systematic review of overweight and obesity as risk factors and targets for clinical intervention for urinary incontinence in women. Neurourol Urodyn 2008;27:749−57.

13. Landefeld CS, Bowers BJ, Feld AD, et al. National Institutes of Health state-of-the-science conference statement: Prevention of fecal and urinary incontinence in adults. Ann Intern Med 2008;148:449−58.

14. Norton P, Brubaker L. Urinary incontinence in women. Lancet 2006;367:57−67.

15. Dietz HP, Wilson PD. Childbirth and pelvic floor trauma. Best Pract Res Clin Obstet Gynaecol 2005;19:913−24.

16. Farrell SA, Allen VM, Baskett TF. Parturition and urinary incontinence in primiparas. Obstet Gynecol 2001;97:350−6.

17. Arya LA, Jackson ND, Myers DL, et al. Risk of new-onset urinary incontinence after forceps and vacuum delivery in primiparous women. Am J Obstet Gynecol 2001;185:1318−23; discussion 1323−4.

18. Burgio KL, Zyczynski H, Locher JL, et al. Urinary incontinence in the 12-month postpartum period. Obstet Gynecol 2003;102:1291−8.

19. Viktrup L, Lose G. The risk of stress incontinence 5 years after first delivery. Am J Obstet Gynecol 2001;185:82−7.

20. MacLennan AH, Taylor AW, Wilson DH, et al. The prevalence of pelvic floor disorders and their relationship to gender, age, parity and mode of delivery. BJOG 2000;107:1460−70.

21. Brown JS, Grady D, Ouslander JG, et al. Prevalence of urinary incontinence and associated risk factors in postmenopausal women. Heart & Estrogen/Progestin Replacement Study (HERS) Research Group. Obstet Gynecol 1999;94:66−70.

22. Nakanishi N, Tatara K, Naramura H, et al. Urinary and fecal incontinence in a community-residing older population in Japan. J Am Geriatr Soc 1997;45: 215−9.

23. Waetjen LE, Liao S, Johnson WO, et al. Factors associated with prevalent and incident urinary incontinence in a cohort of midlife women: A longitudinal analysis of data: Study of women's health across the nation. Am J Epidemiol 2007;165:309−18.

24. Melville JL, Delaney K, Newton K, et al. Incontinence severity and major depression in incontinent women. Obstet Gynecol 2005;106:585−92.

25. Fultz NH, Herzog AR. Self-reported social and emotional impact of urinary incontinence. J Am Geriatr Soc 2001;49:892−9.

26. McGrother CW, Donaldson MM, Hayward T, et al. Urinary storage symptoms and comorbidities: A prospective population cohort study in middle-aged and older women. Age Ageing 2006;35:16−24.

27. Melville JL, Katon W, Delaney K, et al. Urinary incontinence in US women: A population-based study. Arch Intern Med 2005;165:537−42.

28. Nygaard I, Turvey C, Burns TL, et al. Urinary incontinence and depression in middle-aged United States women. Obstet Gynecol 2003;101:149−56.

29. Amaye-Obu FA, Drutz HP. Surgical management of recurrent stress urinary incontinence: A 12-year experience. Am J Obstet Gynecol 1999;181:1296−307; discussion 1307−9.

30. Daneshgari F, Moore C, Frinjari H, et al. Patient related risk factors for recurrent stress urinary incontinence surgery in women treated at a tertiary care center. J Urol 2006;176:1493−9.

31. Schraffordt Koops SE, Bisseling TM, van Brummen HJ, et al. What determines a successful tension-free vaginal tape? A prospective multicenter cohort study: Results from The Netherlands TVT database. Am J Obstet Gynecol 2006;194:65−74.

32. Abrams P, Cardozo L, Fall M, et al. The standardisation of terminology of lower urinary tract function: Report from the Standardisation Sub-committee of the International Continence Society. Neurourol Urodyn 2002;21:167−78.

33. Berghmans BC, Bernards AT, Hendriks HJM, et al. Physiotherapeutic management for genuine stress incontinence. Phys Ther Rev 1998;3:133−147.

34. Brown JS, McNaughton KS, Wyman JF, et al. Measurement characteristics of a voiding diary for use by men and women with overactive bladder. Urology 2003;61:802−9.

35. Hendriks EJ, Bernards AT, Berghmans BC, et al. The psychometric properties of the PRAFAB-questionnaire: A brief assessment questionnaire to evaluate severity of urinary incontinence in women. Neurourol Urodyn 2007;26:998−1007.

36. Abrams P, Artibani W, Cardozo L, et al. Reviewing the ICS 2002 Terminology Report: The Ongoing Debate. Neurourol Urodyn 2006;25:293.

37. Hendriks EJ, Bernards AT, Staal JB, et al. Factorial validity and internal consistency of the PRAFAB questionnaire in women with stress urinary incontinence. BMC Urol 2008;8:1.

38. Hendriks EJ, Bernards AT, de Bie RA, et al. The minimal important change of the PRAFAB questionnaire in women with stress urinary incontinence: Results from a prospective cohort study. Neurourol Urodyn 2008;27: 379−87.

39. Vierhout ME. Meting van ongewenst urineverlies bij de vrouw [Measurement of undesirable urine loss in women]. Ned Tijdschr Geneeskd 1990; 134:1837−40.

40. Field A. Discovering statistics using SPSS. 2nd edition. London: Sage Publications Ltd; 2005.

41. Steyerberg EW, Eijkemans MJ, Habbema JD. Stepwise selection in small data sets: A simulation study of bias in logistic regression analysis. J Clin Epidemiol 1999;52:935−42.

42. Steyerberg EW, Eijkemans MJ, Harrell FE, Jr., et al. Prognostic modelling with logistic regression analysis: A comparison of selection and estimation methods in small data sets. Stat Med 2000;19:1059−79.

43. Steyerberg EW, Eijkemans MJ, Van Houwelingen JC, et al. Prognostic models based on literature and individual patient data in logistic regression analysis. Stat Med 2000;19:141−60.

44. Mushkudiani NA, Hukkelhoven CW, Hernandez AV, et al. A systematic review finds methodological improvements necessary for prognostic models in determining traumatic brain injury outcomes. J Clin Epidemiol 2008;61: 331−43.

45. Brubaker L, Bradley CS, Handa VL, et al. Anal sphincter laceration at vaginal delivery: Is this event coded accurately? Obstet Gynecol 2007; 109:1141−5.

46. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol 2007;165:710−8.

47. Babyak MA. What you see may not be what you get: A brief, nontechnical introduction to overfitting in regression-type models. Psychosom Med 2004;66:411−21.

48. Norman GR, Stratford P, Regehr G. Methodological problems in the retrospective computation of responsiveness to change: The lesson of Cronbach. J Clin Epidemiol 1997;50:869−79.

49. Green S. How many subjects does it take to do a regression analysis? Multivar Behav Res 1991;26:499−510.