Can the PFDI or PFIQ be used to predict outcome in pelvic reconstructive surgery?☆

Le PFDI ou PFIQ permet-il de prédire les résultats en chirurgie reconstructive pelvienne?

V. Letouzey a,*, G. Mercier b, S. Adjoussou a, E. Bohoussou a, P. Mares a, R. de Tayrac a

a Gynecology and obstetrics department, Caremeau university hospital, place Pr R.-Debré, 30900 Nimes, France
b Medical informatics division, Arnaud de Villeneuve university hospital, 34000 Montpellier, France

Summary
Objective. — To determine a syndrome score threshold on PFDI or PFIQ predictive of a significant improvement in post-operative functional results.

Design. — A retrospective case review (Canadian Task Force Classification II-2).

Setting. — University and research hospital.

Population. — Women diagnosed with pelvic organ prolapse and repaired with synthetic vaginal mesh.

Methods. — Quality of life was arbitrarily considered to have improved significantly if the score decreases by more than 50% between pre-operatively and 36 months post-operatively. We investigated the pre-operative cut-off score predictive of no quality of life improvement at M36 from a prospective trial for surgical pelvic organ prolapse treatment.

Results. — The most accurate pre-operative cut-off score predicting a failure to improve quality of life at 36 months post-operatively was 62/300 (PFDI Score). This cut-off value had a positive predictive value of 83.6% and specificity of 62.1%. No significant threshold was obtained from the PFIQ score.

Conclusion. — The intensity of symptoms before surgery may interfere as a predictive factor for outcome.

© 2013 Published by Elsevier Masson SAS.

Abbreviations: AUC, Area under the curve; M0, pre-operative evaluation; M36, 36 months post-operatively; PFDI, Pelvic Floor Disorder Inventory; PFIQ, Pelvic Floor Impact Questionnaire; POP, Pelvic Organ Prolapse; POP-Q, International Pelvic Organ Prolapse staging system; QOL, Quality of life; ROC, Receiver operating characteristic.
☆ Level of evidence: 4.
* Corresponding author.
E-mail address: Vincent.go@wanadoo.fr (V. Letouzey).

1166-7087/$ — see front matter © 2013 Published by Elsevier Masson SAS.
http://dx.doi.org/10.1016/j.purol.2013.04.010
Introduction

Pelvic organ prolapse is seen in up to 43 to 76% of women presenting for routine gynecological care, and 3 to 6% involve descent beyond the hymen [1,2]. In the Women’s Health Initiative, 41% of women age 50 to 79 years showed some degree of pelvic organ prolapse (POP), including cystocele in 34%, rectocele in 19%, and uterine prolapse in 14% [3]. In a multicentre study of 1006 women age 18 to 83 years presenting for a routine gynecological examination, 24% had normal support and 38% had stage I, 35% stage II, and 2% stage III pelvic organ prolapse [4].

POP is a common disorder with prevalence, all stages combined, of 30%. After menopause, even though incidence is stable, a relationship is seen between the severity of the prolapse and patient age [4].

POP is not responsible for any severe morbidity or mortality, but it may have a marked functional impact on quality of life, self-esteem and sexuality [5–7]. As a result, POP must often be surgically repaired. Previous studies have demonstrated the feasibility of surgical treatment, by laparoscopy or vaginal route, and also its efficacy on both anatomic and functional symptoms [8,9]. Moreover, an improvement in quality of life (QOL) has been reported on the basis of validated questionnaires [10,11].

However, although the questionnaire score improved substantially in some women after surgery, others reported only a slight improvement. It would therefore be useful to be able to distinguish between these subgroups of women pre-operatively in order to select the best candidates for surgery. This functional surgery can also cause complications, and this is even more bothersome if patients are not necessarily improved.

As a prognostic factor, Gutman et al. found that vaginal descensus distal to the hymen accurately predicts bulging symptoms; however, they were unable to predict other pelvic floor symptoms [12]. This sign therefore appears to have limitations. Other markers potentially identifying good candidates for surgery therefore needed to be tested.

No mechanism has yet been developed that correlates anatomical abnormalities with the impact of genital prolapse on quality of life. And, in order to identify good candidates for genital prolapse surgery, it is clear that we need criteria that are able to predict post-surgical changes in quality of life.

The aim of this preliminary study was to use validated questionnaires, before and after surgery for genital prolapse, to assess symptoms (PFDI) and quality of life (PFIQ) in an attempt to highlight a threshold above which there is a likelihood of improving post-surgical quality of life [5,13,14].

Population and methods

A prospective study was conducted in 230 women diagnosed with pelvic organ prolapse in the gynecology departments of 13 public and private hospitals in France [15]. They were aged 18 or more and underwent POP repair. All gave their informed consent to take part in the study.

The pre-operative evaluation (MO) included medical history and a urogynaecological examination. All departments used a standardized case report form. Physical evaluations were performed using the International Pelvic Organ Prolapse staging system (POP-Q), assessing anterior and posterior vaginal wall prolapse, and uterine or vault prolapse on maximum Valsalva effort in the seated semi-lithotomy position. Quality of life was assessed using translated French versions of the Pelvic Floor Disorder Inventory (PFDI) and Pelvic Floor Impact Questionnaire (PFIQ) [5]. PFDI and PFIQ are self-administered questionnaires and were completed by patients without the presence of medical staff. They cover urinary, colo-recto-anal and pelvic/vaginal symptoms related to pelvic floor disorders. A psychometric evaluation showed a good correlation between PFDI/PIFQ scores and physical symptoms.

Each patient recruited presented with at least a symptomatic vaginal wall prolapse staged 2 to 4 by the International Pelvic Organ Prolapse staging system (Ba or Bp ≥ –1) and a decreased quality of life. No threshold in the Quality of Life Questionnaire was used as an indication for surgery.

In order to avoid bias due to the surgical procedure, cystocele repair was performed using the same vaginal
synthetic mesh procedure. All patients were operated on by the vaginal route using a prosthetic mesh (Ugytex™, Sofradim, Trévoux, France). This is a low-weight, polypropylene monofilament mesh offering appropriate tissue in growth [16]. Posteriorly, the mesh was implanted with two arms sutured to the sacropinous ligaments, as previously described by the authors [17]. Concomitant procedures such as vaginal hysterectomy were performed in accordance with each surgeon’s preferred technique. When the patient suffered from associated pre-operative stress urinary incontinence, a mid-urethral sling was put in place after anterior vaginal skin closure, by a separate incision.

Patients were instructed to rest for 2 weeks after the surgery, they were allowed to return to work after 4 weeks, and to take part in sport or have sexual intercourse after 6 weeks. Follow-up visits were scheduled for 6 weeks, 6 months, and 1, 2 and 3 years (M36). A urogynecological examination was performed at each visit using the POP-Q system. Post-operative functional results for symptoms and quality of life were evaluated using the PFDI and PFQI.

The main endpoint for each follow-up step was improvement in quality of life at M36 defined as a decrease of more than 50% in the score at M36 compared with the baseline score. On this basis, patients were considered as “improved” (decrease ≥ 50%) or “non-improved” (decrease < 50% or increase). No institutional review board or ethics committee has been solicited, because this specific retrospective statistical analysis has concerned validated questionnaire score on a non-randomized patient’s cohort.

The sample size was calculated in order to achieve a precision of 0.1 for a sensitivity of 80%. Assuming a 95% confidence level and a 67% prevalence of non-improved patients, a total sample size of 93 was needed.

The epidemiological and clinical characteristics of the “improved” and “non-improved” patients were compared using the Chi² test or Fisher’s exact test for categorical variables, and Student’s t test or the Wilcoxon rank-sum test for continuous variables.

Logistic regression was used to determine the baseline PFDI cut-off that predicted an improvement at M36. A receiver operating characteristic (ROC) curve was constructed by plotting sensitivity against 1—the specificity of each possible cut-off. The area under the ROC curve (AUC) was calculated along with its 95% confidence interval. The AUC is equal to the probability that the classifier will rank a randomly chosen positive patient higher than a randomly chosen negative one. The cut-off was determined to optimize both sensitivity and specificity. Statistical measures of the performance of the cut-off point (sensitivity, specificity, positive predictive value and negative predictive value) were calculated along with their 95% confidence intervals. Model performance was quantified by the AUC (the closer to 1, the better). The goodness-of-fit of the model was evaluated using the Hosmer and Lemeshow (lack of fit if P < 0.05). No validation was used due to the small sample size. Spearman correlation coefficients were tested for correlation between Quality of Life Scores (M0, M36, delta) and symptoms. Statistical significance was set at 5% and all statistical tests performed using SAS software 9.1 (SAS Institute, Cary, NC).

Results

Baseline characteristics

Two hundred and thirty consecutive patients with symptomatic anterior or posterior vaginal wall prolapse were recruited in a prospective multicenter study involving 13 French gynecological or urological departments (four university, three public and six private hospitals) and 18 surgeons experienced in vaginal prolapse repair. One hundred and twenty-four were following up at 36 months. Questionnaires of 114 women were analyzed. PFDI and PFQI scores were established at M0 and M36 for 109 and 67 patients, respectively. Median baseline pre-operative PFDI and PFQI scores corresponded to 75 (range: 0–229.6) and 41 (range: 2.7–297.2), respectively (Fig. S1). The baseline characteristics of the “PFDI” population corresponded to mean parity of 2.8 (SD = 1.6), age 64 years (SD = 11.6) and a body mass index of 25 (SD = 3.7). Other baseline epidemiological and clinical characteristics are presented in Tables S1, S2 and Table 1.

Comparison between “improved” and “non-improved” patients

At M36, 67 patients completed the PFQI, giving a median score of 17.6 (range: 0–223.8). This PFQI indicated that 53 (79.1%; 95CI 68—87) had experienced an improvement in quality of life. No difference was found between improved and non-improved patients for median baseline PFQI score which was 38.8 (range: 3.1–297.2) and

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Mean scores for Quality of Life measurements.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Symptoms</strong></td>
<td><strong>PFDI (n=109)</strong></td>
</tr>
<tr>
<td><strong>Scales</strong></td>
<td><strong>Pre-op</strong></td>
</tr>
<tr>
<td>Mean UDI</td>
<td>86.7</td>
</tr>
<tr>
<td>Mean CRADI</td>
<td>73.8</td>
</tr>
<tr>
<td>Mean POPDI</td>
<td>107.6</td>
</tr>
</tbody>
</table>

UDI: Urinary Distress Inventory; UIQ: Urinary Impact Questionnaire; CRADI: Colo-Recto-Anal Distress Inventory; CRAIQ: Colo-Recto-Anal Impact Questionnaire; POPIQ: Pelvic Organ Prolapse Impact Questionnaire; M0: pre-operative evaluation; M36: 36 months post-operatively; PFDI: Pelvic Floor Disorder Inventory; PFQI: Pelvic Floor Impact Questionnaire.
56.8 (range: 2.7–124.2), respectively (P = 0.52, Wilcoxon rank-sum test). Thus, baseline PFIQ was not considered as a potential predictor of an improvement at M36, and the analysis was halted (Fig. S2.a).

At M36, 109 patients completed the PFDI, giving a median score of 9.6 (range: 0–193.8). This PFDI indicated that 80 (73.4%, 95CI 64–81) had experienced an improvement in quality of life. The median baseline PFIQ score was significantly higher in improved than in non-improved patients: 81.1 (range: 4–229.6) and 55.4 (range: 0–179.7), respectively (P = 0.02, Wilcoxon rank-sum test) (Fig. S2.b). Based on the maximization of the sum of the estimated sensitivity and specificity, the best baseline PFDI cut-off predictive of an improvement at M36 was 62. This resulted in sensitivity of 0.7 (95CI: 0.59–0.80) and specificity of 0.62 (95CI: 0.42–0.79) (Table 2). Overall performance was deemed acceptable: AUC = 0.64 (95CI: 0.53–0.76; P = 0.02) (Fig. 1) and the fit of the model was good: P = 0.49 (Hosmer and Lemeshow test).

In other words, 83.6% (72.5%–91.5%) of patients with a baseline PFDI score of more than 62 were improved at M36 while 42.9% (27.7%–59.0%) of patients with a baseline PFDI score of less than 62 were not improved; and 70% (58.7%–79.7%) of patients improved at M36 had a baseline PFDI score of more than 62 while 62.1% (42.3%–79.3%) of patients not improved at M36 had a baseline PFDI score of less than 62. The prediction was right for 67.9% (58.3%–76.5%) of the 109 patients.

### Discussion

This study showed that pre-operative assessment of the PFDI score could predict the improvement experienced after vaginal surgery for genital prolapse repair using polypropylene mesh.

Like Redwine and Wright [18] for endometriosis surgery, we found that women with higher pre-operative symptom scores showed a greater improvement in post-operative symptom scores. But vaginal prolapse literature did not include any pre-operative symptom assessment that is able to predict the post-operative improvement in symptoms.

One of the major challenges for physicians treating women with pelvic organ prolapse was to select the best candidates for synthetic mesh repair. Current guidelines propose using surgical treatment for advanced prolapse and a non-surgical approach for early stages. These criteria were questionable since no studies have shown any strict relationship between prolapse stage and improvement in symptoms and quality of life. In fact, this was a procedure that exposed women to a small but real risk of potentially severe complications such as mesh erosion and de novo dyspareunia, with no guarantee of an improvement in symptoms and QOL.

The definition of clinical improvement, i.e. a 50% reduction, should be established both by a statistically significant difference between pre-operative and post-operative questionnaires and by the clinical course. A 50% improvement in symptoms was likely to constitute a significant improvement and is routinely used in the literature [19]. We arbitrarily defined significant improvement in quality of life. It led to different requirements of symptom relief for improvement based on baseline scores such that those with largest symptom burden (highest scores) at baseline are required to have the greater improvement in terms of absolute change in score than those with lower symptom burden (lower scores) to be called “improved”. However, with a low pre-operative score it would be difficult to improve but it was always true when you have no complaint. Furthermore, we did not use any subjective questionnaire to fix our improvement definition. Barber et al. estimate the minimum important difference (MID) for subscales of the PFDI [20]. They use anchor- and distribution-based approaches. Anchors included a subjective and non-referenced global index of change.

In this preliminary study, even though all patients presented advanced prolapse, very substantial differences in symptoms and their impact on quality of life were observed. For instance, the scores obtained in the PFDI 20 questionnaire ranged from 0 to 224 on a scale of 0–300. Similarly, the PFIQ 7 questionnaire showed that pre-operative values ranged from 5 to 297, showing that some patients reported no impairment of quality of life associated with genital prolapse. However, it is interesting to note that the questionnaires had the advantage of giving a wide distribution of pre-operative values.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Sensitivity and specificity of the best baseline PFDI cut-off (i.e. 62) predicting an improvement at M36.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity</td>
</tr>
<tr>
<td></td>
<td>%</td>
</tr>
<tr>
<td></td>
<td>(58.7–79.7)</td>
</tr>
</tbody>
</table>

PFDI: Pelvic Floor Disorder Inventory; PPV: Positive predictive value; NPV: Negative Predictive value.

Figure 1. The best baseline PFDI cut-off predicting an improvement at M36 was a score of 62. Overall model performance was deemed acceptable: AUC = 0.64 (95CI: 0.53 to 0.76; P = 0.02) as was its goodness of fit: P = 0.49 (Hosmer and Lemeshow test).
The choice of this threshold gave the following diagnostic values: 83.6% of patients with a PFDI score of more than 62 at M0 were “improved” at M36, and 62.1% of patients who were not improved at M36 had score of less than 62 at M0. Patients with a score of more than 62 (PFDI-M0) were significantly more likely to be improved at M36 (83.6% versus 57.1%, P=0.002). If the mistake is made of not operating, POP will deteriorate and the procedure will be required later. It now remains to define the reevaluation date (1 year). Regarding patients with a PFDI score of less than 62 at M0, medical (hormonal or non-hormonal treatments) should be preferred, pending an increase in the score to the surgery-triggering threshold (Table 2). QOL questionnaire and threshold may not be a factor indicating surgical procedure by itself.

We used a logistic regression to determine the baseline PFDI cut-off predictive of an improvement. This approach is particularly relevant when the objective is to maximize the accuracy of the prediction, i.e. to maximize both sensitivity and specificity [19,21]. The overall performance and goodness-of-fit of the regression model were acceptable. In addition, the data base was multicenter and therefore included the diversity of patient recruitment by different operators (gynecologists, urologists, private, public and university institutions). We did not perform any cross validation of the rule. However, a prospective validation study is ongoing, which is the recommended method to validate prediction rules [22].

Although questionnaires covered a broad field and validate evidence related to the vaginal prolapse, it should be noted that a patient with very specific POP-related symptoms might see little improvement in her questionnaire score after surgery. This was especially the case with the short-form questionnaires that are always preferred. The fact that no statistically significant improvement was found between M0 and M36 for the PFIQ questionnaire probably underlines the limits of this questionnaire, or simply the size of the sample.

Finally, the results of this study were established at the 3-year follow-up point, which is valid for assessing the impact of surgery. However, this lead to a lost of follow-up rate to almost 50% (109/230). This could be likely to bias the results but no difference, except history of prolapse, was detected between the two groups of patients (Table S3). Repeat surgery and anatomical results were not different for the patients answering the questionnaire or not.

Certain limitations to this study should be underlined. Firstly, our results only concerned women with advanced stage POP where there is a greater risk of complications after surgical repair. Moreover, our study only concerned women undergoing vaginal repair, and therefore may not be relevant to (often younger) women undergoing laparoscopy.

The impact of POP cannot be fully assessed without an evaluation of the patient’s sexuality. This item was especially important given that prosthetic surgery is associated with almost 10% of de novo dyspareunia. However, this questionnaire could be validated only for sexually active patients.

In order to evaluate the surgical route used for POP surgical repair using mesh, the pre-operative symptoms threshold should also be evaluated after laparoscopy treatment [23].

Our cohort studied may be considered as too small, and as we sought to obtain the advantages of the questionnaire for patient selection, we are currently conducting a study in a larger population. This was why a prospective validation study with higher population is ongoing.

Based on the results of this study, we believe that QOL and symptoms questionnaires should be extensively used as methods to evaluate POP before any treatment, and ensure long-term follow-up in efforts to identify changes in certain quality of life parameters. Goals factor as resolution of urinary symptoms, ability to perform daily activities, and sexual function goals were at least as important as resolution of prolapse symptoms and may be the reason for seeking care [24]. This was why, instead of using a simpler set of questionnaires, we have planned to use an electronic form of the questionnaires, for instance on a tablet PC [25], to help practitioners using our model.

Research into predictive factors was necessary if QOL in a functional disease such as pelvic organ prolapse, with surgical repair, has to be improved. Impact on QOL before surgery would seem to hold some promise. A cut-off pre-operative score, if validated on PFDI (> 62), should improve POP treatment, and a prospective multicentre evaluation was ongoing to validate this cut-off. This validated questionnaire, which was available in several languages, provided a simple tool that can be used to select and inform women who might benefit from vaginal repair for POP.

Disclosure of interest

No conflict of interest for any authors except Pr R. de T who is consultant for Boston Sci.

Acknowledgments


Funding statement: this study was not support by any grant or industry.

Appendix A. Supplementary material

Supplementary materials (Tables S1—S3, Fig. S1 and S2 and French version of the text) associated with this article can be found at http://www.sciencedirect.com, at doi:10.1016/j.purol.2013.04.010.

References


