**Key Words**
Pelvic floor, assessment, research, validity, reliability.



# Pelvic Floor Muscle Assessment: The PERFECT Scheme

**by** J Laycock
D Jerwood

**Summary**

Aims of study

1. To develop a digital technique to assess pelvic floor muscles (PFM).
2. To validate the technique and test for validity and reliability.
3. To translate the assessment into an exercise-based regimen.

**Method and Results** PERFECT is an acronym with P representing power (or pressure, a measure of strength using a manometric perineometer), E = endurance, R = repetitions, F = fast contractions, and finally ECT = every contraction timed. The scheme was developed to simplify and clarify PFM assessment. The pressure (referred to hereafter as power, although actually a surrogate measure of muscular strength) of a contraction was validated by examining perineal lift and perineometric pressure during a maximum voluntary contraction (MVC). Data analysis demonstrated significant positive correlations between power and both lift ($r = 0.864$; $p = 0.031$) and perineometric pressure ($r = 0.786$; $p = 0.001$); digitally assessed endurance correlated with the area under the pressure curve of a sophisticated perineometer ($r = 0.549$; $p = 0.001$). A study of inter-examiner reliability demonstrated highly significant positive correlations between two examiners for power ($r = 0.947$; $p < 0.001$), endurance ($r = 0.946$; $p < 0.001$), repetitions ($r = 0.730$; $p < 0.005$) and number of fast contractions ($r = 0.909$; $p < 0.001$). Scatter diagrams confirmed a lack of systematic bias between examiners.

Test-retest reliability produced highly significant correlations ($p < 0.001$) between power ($r = 0.929$) and endurance values ($r = 0.988$) recorded on two different occasions, with no convincing evidence of significant discrepancies between the pair of assessments.

**Conclusions** The PERFECT scheme has demonstrated reliability and validity as an assessment tool. Furthermore, it is proposed that this scheme provides guidelines for the planning of patient-specific exercise programmes which satisfy the principles of muscle training.

## Introduction

The pelvic floor muscles (PFM) consist of approximately 70% slow-twitch (type 1) and 30% fast-twitch (type 2) muscle fibres (Gilpin *et al*, 1989). Although the PFM are thought to work tonically and reflexly during routine daily activities, voluntary contractions are required for training.

PFM weakness is recognised as one of the problems encountered in patients with urinary and faecal incontinence, and re-education should address the perceived deficit, be it in the slow- or/and fast-twitch muscle fibres. There are several ways by which to assess the PFM, including digital palpation (Brink *et al*, 1989), using a pressure perineometer (Laycock and Jerwood, 1994), electromyography (EMG) (Haslam, 2002) ultrasound scanning (Vierhout and Jansen, 1989) and magnetic resonance imaging (MRI) (Khullar, 2002). Evaluation of muscle strength and endurance provides information on the severity of muscle weakness and forms the basis of patient-specific exercise programmes.

Research involving PFM assessment probably started with the work of Kegel (1948), who advocated clinical assessment of the pubococcygeus as part of a routine gynaecological examination, using one finger to palpate the pubococcygeus *per vaginam*. He maintained that two fingers placed the vaginal tissues under tension and distorted the anatomic relationships (Kegel, 1956). However, the premise that stretching the vaginal tissues during assessment may adversely affect outcomes was challenged by Chiarelli (1989) who maintained that stretching the muscle may produce an enhanced response. In support of this hypothesis, Jahnke *et al* (1989) showed that the initial phase of passive muscle stretching is associated with a rise in reflex tone.

The one-finger assessment technique

was again described by Hendrickson (1981), who defined a mild contraction as slight unsustained pressure on the examiner's finger; a moderate contraction as firm pressure held for one to three seconds; and a strong contraction as forceful pressure on the examiner's finger sustained for four seconds or more.

A different method was devised by Graber *et al* (1981) based on four components: control, sustained strength, atrophy and tone. Control is measured using a perineometer and indicates the patient's ability to contract and relax the PFM to command. Sustained strength is an indication of the time (up to 10 seconds) that a strong contraction could be held. Atrophy is used to describe the development and wastage of the muscle. This involved the technique of mapping, which necessitated palpating the PFM and charting areas of reduced muscle bulk. Tone was determined by the degree of resistance of the muscle against moderate pressure from the examining finger.

A further report (Worth *et al*, 1986) described another one-finger digital vaginal assessment with four components: pressure, duration, ribbing and position (of the examining finger).

- They defined pressure as the strength of contraction. If no pressure is felt, a score of 1 is given. If moderate pressure is felt, a score of 2 is given, and firm pressure is recorded as 3.
- Duration is based on the length of time a contraction can be sustained. If no contraction is felt or the contraction lasts no longer than 1 second, a score of 1 is assigned. A contraction held for 2 to 3 seconds is scored 2, and a score of 3 is assigned to a contraction held for 4 or more seconds.
- An assessment of ribbing refers to the tone and texture of the PFM during a contraction. If the muscle feels soft and flabby, it is assigned a score of 1; a score of 2 describes a muscle that feels different from the surrounding tissues but not ribbed, and a score of 3 is assigned if the muscle feels distinct, like rings of ribbing or ribbed muscle tissue.
- Assessment of position refers to the plane the examining finger is in, in relation to the vaginal introitus. If no force is exerted and the finger can easily slip out, a score of 1 is given. If the finger can be gripped somewhat but remains in the same position, a score of 2 is given, and if the finger is forcibly gripped, expelled, or pulled anteriorly, a score of 3 is given.

Test-retest reliability of this scoring system showed a statistically significant relationship of observation between ten patients examined and then re-examined 10 days later.

Another scoring system, this time linked with incontinence, was described by Brink *et al* (1989). This measure uses concepts of pressure (rated 1 to 4), time and displacement. These were evaluated by the index and middle fingers in the antero-posterior position (index finger resting on the middle finger) introduced 4 cm to 6 cm into the vagina. Test-retest for this scoring method was $r = 0.65$, $p < 0.01$, with inter-rater reliability $r = 0.91$, $p < 0.01$. A negative correlation between muscle strength and both urine loss and age was demonstrated.

The authors of all the aforementioned studies did not attempt to use any standard international muscle-grading scheme or to differentiate between slow- and fast-twitch muscle activity. Furthermore, they did not relate the assessment findings to an individual exercise programme. Moreover, co-contraction of the abdominal muscles was discouraged and this is now acknowledged to be inappropriate (Sapsford *et al*, 2001).

## Aims of Study

1. To develop a digital technique for quantitative assessment of the voluntary contractility of the PFM.
2. To validate the above technique and test for observer reliability.
3. To translate the measurements from the digital assessment into the planning of a patient-specific exercise programme.

## Methods and Materials

PERFECT is an acronym to remind all health professionals of the need to assess the main components of PFM contractility. This assessment scheme was developed to provide a simple, reliable method of PFM evaluation and involves four components, as shown in table 1.

**Laycock, J and Jerwood, D (2001)**. 'Pelvic floor muscle assessment: The PERFECT Scheme', *Physiotherapy*, **87**, 12, 631-642.

**Table 1: The PERFECT assessment scheme**

| | |
|---|---|
| P | Power (pressure) |
| E | Endurance |
| R | Repetitions |
| F | Fast |
| E | Every |
| C | Contraction |
| T | Timed |

Although the study was carried out on women by palpating the perivaginal muscles *per vaginam*, the PERFECT assessment can also be used for PFM assessment *per rectum* in men and women. To ensure reproducibility, the following factors were adopted throughout the study:

- The location and action of the pelvic floor muscles were described to the subjects in enough detail for adequate understanding of this muscle group.
- Whenever possible, subjects were positioned in supine with their head on two pillows. The hips were flexed and abducted, and the knees bent.
- The PFM were examined using the index finger placed approximately 4 cm to 6 cm inside the vagina and positioned at 4 o'clock and 8 o'clock to monitor muscle activity. Moderate pressure was applied over the muscle bulk to assist in the initiation of the appropriate muscle contraction.
- Verbal informed consent was obtained from all subjects.

## Power

Power is measured on a modified Oxford scale (table 2). The authors acknowledge that digital palpation during a maximal voluntary contraction (MVC) evaluates muscle strength, *not* power. However, with this caveat, the misnomer 'power' will be used throughout the text.

The following definitions are proposed:

*Grade 0* No discernible muscle contraction.

*Grade 1* A flicker or pulsation is felt under the examiner's finger.

*Grade 2* An increase in tension is detected, without any discernible lift.

*Grade 3* Muscle tension is further enhanced and characterised by lifting of the muscle belly and also elevation of the posterior vaginal wall. A grade 3 and stronger can be observed as an in-drawing of the perineum and anus.

*Grade 4* Increased tension and a good contraction are present which are capable of elevating the posterior vaginal wall against resistance (digital pressure applied to the posterior vaginal wall).

*Grade 5* Strong resistance can be applied to the elevation of the posterior vaginal wall; the examining finger is squeezed and drawn into the vagina (like a hungry baby sucking a finger).

**Table 2: Proposed modified Oxford grading scheme**

| *Grading* | *Muscle response* |
|---|---|
| 0 | Nil |
| 1 | Flicker |
| 2 | Weak |
| 3 | Moderate |
| 4 | Good |
| 5 | Strong |

Consequently, in a specific case, the power could simply be recorded as grade 3 for a moderate contraction. However, the registered grade is permitted to be augmented with a symbol + or –, when the need arises. Thus 3+ could be translated as there being more than a moderate contraction but less than a good contraction (grade 4). Similarly, a 3– is recorded when the contraction is less than a grade 3, but more than a grade 2. This augmentation is to allow for an element of doubt to be introduced and thereby to soften this (partly subjective) six-point ordinal scale. There will admittedly be a learning curve involved in assessing the strength of a PFM contraction.

### *Endurance*

Endurance is expressed as the length of time, up to 10 seconds, that an MVC can be sustained before the strength is reduced by 35% or more. In other words, the contraction is timed until the muscle starts to fatigue. A further possible indication of PFM fatigue may be the simultaneous contraction of hip adductors and glutei, and the stronger co-contraction of transversus abdominis. Breath-holding should be discouraged; if detected, the subject should be instructed to contract the pelvic floor on expiration.


**Authors**

**J Laycock PhD FCSP** is a specialist continence physiotherapist in private practice who carrried out the research for this article.

**D Jerwood BSc PhD FSS** is head of mathematics, School of Computing and Mathematics, University of Bradford, and was responsible for the statistics.


This article was received on January 17, 2000, and accepted on August 22, 2001.


**Address for Correspondence**

Jo Laycock, The Culgaith Clinic, Pea Top Grange, Culgaith, Penrith CA10 1QW.

Example 1: 3/5 = grade 3 held for 5 seconds

*Repetitions*
The number of repetitions (up to 10) of the specific MVC (eg 3 as in example 1) is recorded, allowing four seconds rest between each contraction (which lasts for 5 seconds in example 1). The purpose of the PERFECT assessment is to determine the number of contractions necessary to overload the muscle, develop a practicable exercise programme and so produce a training effect.

Different 'rest' periods have been tested, and in the experience of the authors, four seconds will allow weak, easily fatigued muscles time to recover without permitting excessive rest periods for stronger muscles. It would be impractical to permit, say, 20 seconds rest between each contraction, as this would make the exercise session too long, and might never overload the PFM in some cases. Furthermore, in practical terms, once a subject can perform 10 repetitions of a 10-second maximum contraction, the rest time is reduced.

Limiting both repetitions and endurance to a maximum in this way is formally referred to as right-censoring of data, and its impact on results will be discussed later.

Example 2: 2/3/6 = grade 2 held for 3 seconds, and repeated 6 times (with 4 seconds rest between each contraction)

*Fast*
After a short rest (at least one minute), the number (up to 10) of one-second MVCs is assessed. Subjects are instructed to 'contract-relax' as quickly and strongly as possible, in their own time, until the muscles fatigue. Many patients can perform more than 10 fast contractions, but for practical reasons, the assessment stage should be limited to 10. Subsequent assessment during a treatment programme may identify a greater number of fast contractions, and that should be the number practised by those patients on a daily basis.

Example 3: 4/6/5//9 = good contraction, held for 6 seconds, repeated 5 times, followed by 9 fast contractions

*Every Contraction Timed*
This completes the acronym and reminds the examiner to time and record the above sequence of events.

**Validity Study 1**
Since an increase in pressure on the examiner's finger may be misinterpreted as pressure exerted by an increase in intra-abdominal pressure (rather than by perivaginal muscle contraction), it was necessary to validate the source of activity. It is well documented (Bø *et al*, 1989) that a moderate to strong PFM contraction also incorporates an element of lift which would not be manifest during an increase in abdominal pressure alone. Validation of the PFM contraction was therefore ratified by measuring perineal lift.

Eight women, mean age 43.9 years (range 21 to 61 years) and mean parity 2.5 (range 0 to 5), randomly recruited from a gynaecology outpatient clinic, agreed to take part in this study. A lightweight vaginal probe with a thin plastic rod extension was introduced into the distal 5 cm of the vagina. The rod was placed alongside a ruler and cephalad movement of the rod during an MVC was observed and measured.

**Validity Study 2**
A consecutive sample of 233 women attending a gynaecology outpatient clinic were recruited into this study, of whom 147 were diagnosed as having stress incontinence. Power and endurance were assessed both digitally (as described above) and, after a two-minute rest, using a sophisticated pressure perineometer previously tested for reproducibility and reliability (Laycock and Jerwood, 1994). This perineometer recorded maximum pressure and endurance was assessed by calculating the area under the pressure curve of a maximum 10-second contraction. The same patient position (see above) was used for each method.

**Inter-examiner Reliability Study**
Ten patients were independently assessed by two examiners at the same visit. Demographically, all the women were incontinent, mean age 47.6 years (range 26 to 58), and mean parity 2.7 (range 1 to 4). Each subject was tested for strength (scale 0 to 5), endurance (up to 10 seconds), number of repetitions (up to 10) and number of fast contractions (up to 10), with at least two minutes rest between each test and three minutes rest between each examination. The order of examination between physiotherapists was randomised in order to eliminate

systematic bias. Inter-examiner reliability over the two assessments was determined using correlation and graphical techniques. Each examiner recorded the values of the assessment independently, without knowledge of the other's result.

### Test-retest Reliability

Twenty incontinent women entered this study. Assessment was carried out before cystometry and repeated at the first physiotherapy appointment two to five weeks later by the same examiner, evaluating power and endurance. To conduct this test responsibly, the results of the first assessment were not available until after completion of the second examination. Test-retest reliability was determined using correlation and graphical techniques for the two results.

### Exercise Programme Planning

The third purpose of the study was to formulate a method for translating the PERFECT assessment into a patient-specific exercise programme. Not only has such a regimen to satisfy general muscle training principles of overload and specificity, but it must also be practicable within the context of the individual patient since, without a high degree of compliance, the therapy is unlikely to be effective.

**Patient 1**

| *P* | *E* | *R* | *F* | |
|---|---|---|---|---|
| 3 | 5 | 4 | 7 | This assessment describes a patient with a moderate (grade 3) contraction, held for 5 seconds and repeated 4 times; followed by 7 fast contractions |

This patient would be instructed to practise 4 (R) of her strongest contractions lasting 5 seconds (E) (with 4 seconds rest between) at each exercise session, and 7 fast (F) strong contractions at other times. She should aim to increase the number of repetitions of the long MVC over subsequent weeks up to 10 repetitions. After this, over the succeeding weeks, she should aim to increase the hold time to 10 seconds. In addition, the number of fast contractions should be progressively increased to the maximum number possible (many patients can do up to 50 fast contractions). The aim is to be able to perform 10 repetitions of a 10-second MVC at regular intervals during the day, and 10+ fast contractions at other times during the day. Equal numbers of sessions of slow and fast contractions are encouraged, up to six per day.

**Patient 2**

| *P* | *E* | *R* | *F* | |
|---|---|---|---|---|
| 2 | 2 | 3 | – | This assessment describes a patient with a weak (grade 2) contraction, held for 2 seconds and repeated 3 times |

This patient needs to practise 3 (R) of her strongest contractions lasting 2 seconds (E) (with 4 seconds rest between each contraction) as many times as possible during the day. In view of the weak state of the PFM, assessing the number of fast contractions would be confusing, and this should be done at a later date when endurance has improved to 4 seconds. Treatment should progress in the first instance by aiming to hold the contraction for 3 (then 4) seconds. Once this is achieved, strength should be addressed and the patient encouraged to contract harder. Once a grade 3 has been achieved, endurance is once again addressed, and the subject encouraged to increase the length of contraction.

In all cases, regular PFM re-assessment is required to monitor progress and establish a 'new' exercise programme.

## Results

### Validity Study 1

This study measured perineal lift (in centimetres) during an MVC (P – scale 0 to 5) and the results are displayed in table 3. Six of the eight subjects examined during an MVC exhibited cephalad movement of the vaginal probe greater

**Table 3: Results of study to measure perineal lift**

| *Subject* | *Age (years)* | *Parity* | *P* | *Continence status* | *Lift (cm)* |
|---|---|---|---|---|---|
| 1 | 61 | 2 | 2+ | Incontinent | 0.3 |
| 2 | 51 | 3 | 2+ | Incontinent | 0.4 |
| 3 | 21 | 0 | 3+ | Continent | 0.5 |
| 4 | 36 | 2 | 5 | Continent | 0.7 |
| 5 | 53 | 4 | 2 | Incontinent | 0.0 |
| 6 | 45 | 2 | 3+ | Incontinent | 1.0 |
| 7 | 51 | 5 | 2+ | Incontinent | 0.1 |
| 8 | 31 | 2 | 4 | Incontinent | 0.7 |

than 0.1 cm. The subjects who failed to demonstrate such a movement (patients 5 and 7, both incontinent) had had delivered *per vaginam* four and five babies respectively, compared with a mean of 1.8 babies for the remaining subjects. Over all subjects, a degree of negative correlation between perineal lift and parity was detected ($r = -0.664$; $p = 0.073$) (Spearman's rho for ordinal data).

Any correlation coefficient in the range 0.61 to 0.80 can be considered 'good', and with $r = -0.664$ the coefficient of variation is 44% ($100\ \rho^2$). In other words, 44% of the variation in perineal lift can be explained by parity alone. It should be noted in passing that this evidence has been detected in a sample of only eight patients and therefore fails to reach levels at which statistical significance can be claimed, although the result may still be clinically significant. More pertinent is that examination of the data for all eight subjects showed a positive value for Spearman's coefficient of correlation between P (modified Oxford grading, scale 0 to 5) and perineal lift ($r = 0.864$; $p = 0.031$). This coefficient is in excess of 0.80 and therefore can be considered 'very good', however once again this result must be viewed in the context of sample size ($n = 8$). The coefficient of variation now rises to almost 75% and confirms the sympathetic relationship between perineal lift and the digital assessment scale discussed earlier. This adds some credence to the hypothesis that digital assessment can be used to measure power (P) of a pelvic floor muscle contraction.

**Table 4: Results of Validity Study 2. Demographic variables of age, parity and BMI**

| *Variable* | *Continent (n = 86)* | | *Incontinent (n = 147)* | | *Difference* |
|---|---|---|---|---|---|
| | *No* | *(SD)* | *No* | *(SD)* | |
| Age | 38.5 | (16.8) | 45.4 | (11.9) | p < 0.00005 |
| Parity | 1.5 | (1.4) | 2.5 | (1.5) | p < 0.00005 |
| BMI | 23.3 | (3.1) | 25.9 | (4.0) | p < 0.00005 |
| P | 3.6 | (1.1) | 3.1 | (1.2) | p < 0.00005 |
| Max. press. | 52 | (24) | 32 | (21) | p < 0.00005 |
| E | 7.0 | (2.9) | 6.3 | (3.3) | p = 0.103 |
| Area | 346 | (198) | 216 | (156) | p < 0.00005 |

SD = standard deviation
P = power (scale 0 to 5) digital assessment
Maximum pressure (mm Hg) perineometric assessment
E = endurance (seconds) digital assessment
Area (mm Hg.s) measured with perineometer

### Validity Study 2

The results of the digital and perineometric testing are shown in table 4 together with a selected number of demographic details. It is now relevant to note the large number of subjects in each group (86 continent and 147 incontinent), so that statistical comparisons between groups are not so sensitive to choice of test procedure (parametric or non-parametric).

Appealing to the central limit theorem for induced asymptotic normality, it becomes reasonable to apply t-tests for unrelated samples for all inter-group comparisons. The incontinent subjects were found to be older on average (by 6.9 years), had borne one more child on average, and were associated with a higher BMI (by 2.6 units on average). All these differences were found to be statistically significant with $p < 0.00005$. It should be said that large sample sizes generally produce much reduced standard errors for the differences in means, and statistical power rises accordingly. In these cases, statistical significance is often easier to demonstrate.

From a clinical point of view, on average, the continent subjects exhibited 16% more power (digitally assessed), and 63% higher maximum pressure, and generated 60% more area under the perineometric pressure curve. Each of these differences proved statistically significant (with $p < 0.00005$). Only endurance (also assessed by digital palpitation), which registered an average increase of 11% (of continent subjects over incontinents) failed to reach levels of statistical significance ($p = 0.103$).

The relationship between the digital assessment of pressure (P) and the maximum perineometric pressure ($r = 0.786$) implies that 62% of the variation in perineometric pressure could be explained by the variation in P. The relationship between area under the pressure curve and endurance (E) was less convincing ($r = 0.549$), whereby only 30% of the variation in area could be explained by the variation in E.

Agreement between digital and perineometric measurements of pressure adds further evidence of the validity of this form of assessment for pelvic floor muscles.

## Inter-examiner Reliability Study

Reliability between different examiners involved two physiotherapists (JL and SHD) each performing complete PERFECT assessments on ten subjects on the same visit, and the results are shown in table 5. Analysis shows highly significant positive correlation coefficients between the two examiners for P ($r = 0.947$; $p < 0.001$), E ($r = 0.946$; $p < 0.001$), R ($r = 0.730$; $p < 0.005$) and F ($r = 0.909$; $p < 0.001$).

When testing for significance of correlation coefficients, the P values relate to rejection of a null hypothesis claiming that the true correlation is zero which is not a useful concept within validity studies. Subjects (such as patients 3 and 4 in table 5) who are performing at the boundary of data censorship (that is, holding an MVC for at least 10 seconds, and being able to produce at least 10 fast contractions) will enhance any correlation coefficients. Removing patients 3 and 4 from these correlation studies will reduce every one of the coefficients given in table 5, the greatest impact being on R, for which the coefficient falls dramatically from $r = 0.730$ to $r = 0.530$.

Especially when assessing agreement between two examiners, it should be appreciated that a high coefficient of correlation is certainly necessary, but not (in itself) sufficient. Correlation is a concept which describes only linearity between variables – not agreement. It is therefore perfectly feasible for one assessor to display a systematic bias (compared with another assessor) with such consistency, that the results are still highly correlated. For 'agreement', further supportive evidence is required (Bland and Altman, 1999). This can be provided using a number of techniques, but perhaps the simplest is the scatter diagram (see figures 1 and 2). The digital assessment of pressure (P) of figure 1 shows six agreements, three over-assessments (by SHD compared to JL) by only a half unit of grading, and only one (full unit) under-assessment. In nett terms, the mean difference over ten subjects is 0.05 which, when considered together with a correlation coefficient of $r = 0.947$, suggests a lack of systematic bias and a strong (linear) relationship between assessors.

Similar plots result from the two

**Table 5: Analysis of results of digital assessment of ten patients by two examiners (SHD and JL) on the same visit**

| *Patient* | *P* | | *E* | | *R* | | *F* | |
|---|---|---|---|---|---|---|---|---|
| | *SHD* | *JL* | *SHD* | *JL* | *SHD* | *JL* | *SHD* | *JL* |
| 1 | 3.0 | 2.5 | 7 | 6 | 6 | 5 | 7 | 6 |
| 2 | 4.0 | 4.0 | 9 | 10 | 3 | 6 | 8 | 7 |
| 3 | 5.0 | 5.0 | 10 | 10 | 8 | 6 | 10 | 10 |
| 4 | 4.5 | 4.0 | 10 | 10 | 8 | 7 | 10 | 10 |
| 5 | 2.0 | 2.0 | 4 | 3 | 3 | 2 | 6 | 5 |
| 6 | 2.0 | 3.0 | 3 | 2 | 4 | 5 | 5 | 6 |
| 7 | 3.0 | 3.0 | 3 | 3 | 6 | 5 | 5 | 6 |
| 8 | 1.0 | 1.0 | 2 | 2 | 1 | 2 | 3 | 2 |
| 9 | 1.5 | 1.0 | 4 | 5 | 3 | 3 | 4 | 5 |
| 10 | 4.0 | 4.0 | 3 | 3 | 4 | 5 | 7 | 7 |
| Spearman's correlation | r = 0.947 | | r = 0.946 | | r = 0.730 | | r = 0.909 | |
| Coefficient significance level | p < 0.001 | | p < 0.001 | | p < 0.005 | | p < 0.001 | |

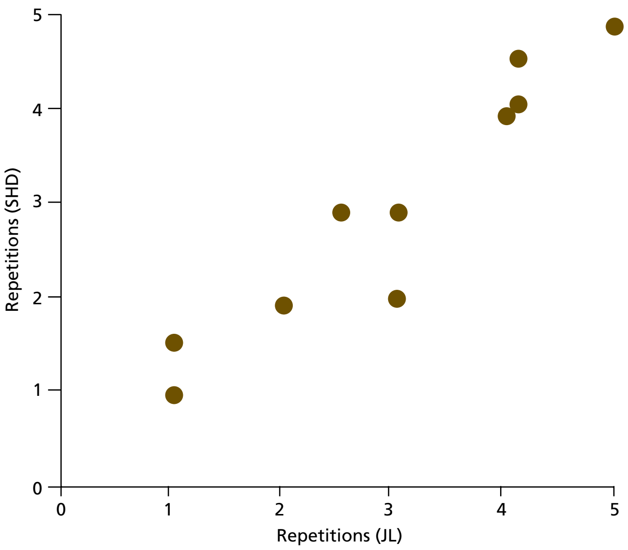


**Fig 1: Pressure of PFM contractions assessed by JL and SHD with line of equality**

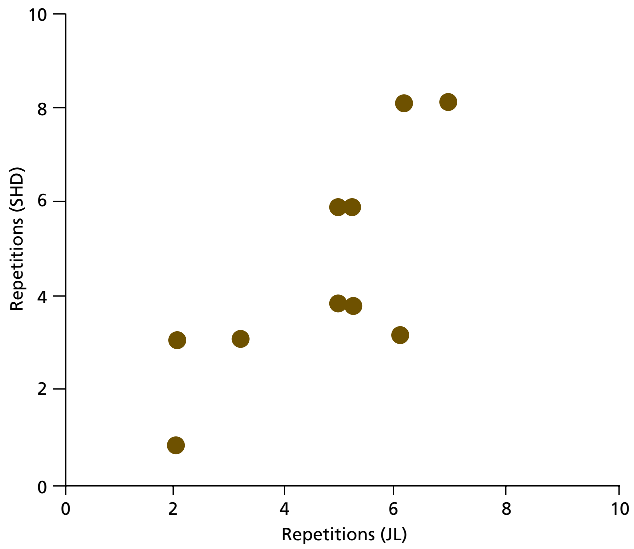


**Fig 2: Numbers of repetitions recorded by JL and SHD with line of equality**

assessments of endurance (E) and the two counts of fast contractions (F) (not reproduced here), however the weakest inter-assessor agreement was found with repetitions (R) and this is illustrated in figure 2. Now, there is only one (perfect) agreement, with four over-assessments (totalling 5 units) and three under-assessments (also totalling 5 units); the greatest discrepancies occurr with two subjects who recorded 6 repetitions for JL. The mean nett difference is now zero but (with a correlation coefficient of only 0.730), the impression is given of a lack of systematic bias between examiners with less convincing consistency. Notice, however, when recording repetitions the censoring limit of 10 which had been imposed by the physiotherapist was never actually attained for any of these ten subjects, although other imposed maxima were regularly attained by subjects 3 and 4. This could imply that the R component was found to be a particularly fatiguing exercise by those subjects, and its implementation introduced additional sources of intra-patient variation to be confounded with inter-assessor variation.

The inter-examiner assessments of number of repetitions (R) apart, taken over the remaining three digital scores, this study demonstrated 46.7% exact agreement of the digital scores. However, disagreement in all but two cases (both in the repetition test) did not exceed one unit of measurement. These encouraging results provide global evidence in support of the inter-examiner reliability of these components of the digital assessment of the PFM.

**Table 6: Test-retest results of digital assessment of power (P) and endurance (E)**

| *Patient* | *Power* | | *Endurance* | |
|---|---|---|---|---|
| | *Test 1* | *Test 2* | *Test 1* | *Test 2* |
| *1* | *3.0* | *4.0* | *2* | *2* |
| 2 | 0.0 | 0.0 | 0 | 0 |
| 3 | 3.0 | 3.0 | 3 | 3 |
| 4 | 4.0 | 3.5 | 10 | 10 |
| 5 | 3.0 | 3.5 | 5 | 6 |
| 6 | 3.5 | 4.0 | 10 | 10 |
| 7 | 2.5 | 2.0 | 5 | 6 |
| 8 | 2.0 | 2.0 | 2 | 3 |
| 9 | 0.0 | 0.0 | 0 | 0 |
| 10 | 3.0 | 3.5 | 7 | 7 |
| 11 | 2.5 | 2.5 | 4 | 4 |
| 12 | 2.5 | 3.0 | 2 | 3 |
| 13 | 4.0 | 4.5 | 8 | 9 |
| 14 | 4.0 | 4.0 | 10 | 10 |
| 15 | 4.5 | 4.0 | 4 | 6 |
| 16 | 2.5 | 2.5 | 8 | 8 |
| 17 | 2.0 | 2.0 | 10 | 10 |
| 18 | 3.0 | 3.5 | 10 | 10 |
| 19 | 5.0 | 5.0 | 5 | 5 |
| 20 | 3.5 | 4.0 | 10 | 10 |
| Spearman's correlation | $r = 0.929$ | | $r = 0.988$ | |
| Coefficient significance level | $p < 0.001$ | | $p < 0.001$ | |

## Test-retest Reliability Study

The results of re-testing 20 women two to five weeks after the initial assessment are shown in table 6. These results show that for muscle strength (P), 45% exact agreement was demonstrated. Furthermore, of the 11 measures that differed, ten of these differed by only 0.5 grade and the remainder by one grade. In addition, it is shown that 8/11 demonstrated an increase in strength whereas 3/11 demonstrated a decrease.

Scatter graphs (similar to figures 1 and 2, but not reproduced here) illustrate a slight tendency for some patients to improve over the period between the tests. However, the mean nett shift of only +0.125 per patient is well within the minimum unit of graduation (of 0.5).

Regarding endurance, 70% exact agreement was demonstrated, with a variation of only one second in all except one, which was two seconds. In all cases, endurance remained the same or increased, despite the fact that these women were not instructed to practise pelvic floor exercises. The complete absence of patients with shorter endurance periods is surprising, but the mean nett shift is only 0.4 seconds per patient and (again) well within the minimum unit of evaluation (of one second).

Evaluating Spearman's rank correlation coefficients for these test-retest results, the coefficient for power (P) is found to be 0.929 ($p < 0.001$) and for endurance (E) stands even higher at 0.988. These extremely high levels of correlation together with mean discrepancies which cannot be reduced by systematic adjustments provide convincing evidence that, for an experienced physiotherapist, the digital evaluation of power and endurance of PFM contractions following a short delay is a reliable technique.

## Planning Exercise Programmes

The procedure for planning patient-specific exercise programmes described earlier follows recognised muscle training methods and has proved to be successful in clinical practice. However, in the past, many women have admitted to neglecting to practise their prescribed pelvic floor exercises and so realistic regimens need to be agreed with each patient. Previously, the majority of patients have complained that pelvic floor exercises were not interesting, but with the introduction of long, and fast and short contractions, and the setting of targets, more women are co-operating in this self-help rehabilitation.

Although data have not been made available here, a number of interesting facts have emerged as a result of this methodology. For example, a woman's ability to contract the PFM appears to be subject to a learning curve, and many women can begin to produce a stronger and longer contraction before any possible physiological change in the muscle fibres could have taken place. In addition, some women report an improvement in continence symptoms without any measurable improvement in their PFM contractility. This was discussed in a study by Miller *et al* (1996) who described 'the knack', a precisely-timed anticipatory pelvic floor contraction before and during a cough. Using 'the knack', significant reduction in involuntary urine loss during a cough was demonstrated.

## Conclusions and Discussion

In physiotherapy practice, palpation of a muscle can detect changes in resting tone, identify areas of atrophy and differentiate between the state of contraction and relaxation, and this is certainly the case with the pelvic floor muscles. During a vaginal examination using the distal pad of the index finger to palpate the perivaginal muscles, a definite bulging and lifting of the muscles are felt during a contraction, and in both the contracted and relaxed state, areas of atrophy can be detected.

Furthermore, a moderate to strong contraction of the levator ani muscles has both a squeeze and a lift component, and the proposed digital assessment scoring system provides a means of calibrating this physical effect.

The PERFECT assessment scheme, although a subjective method, has been shown to be reliable and reproducible. The first component P (power/pressure, but actually evaluating strength) graded the pressure exerted on the finger from 0 to 5, with intermediate scoring permitted (for example 3+), giving a greater flexibility than the scoring of 1 to 4 suggested by Brink *et al* (1989). The Oxford grading system is well understood in physical medicine and the modifications described herein have proved to be an effective way of assessing PFM strength. Regarding different muscle components (that is, fast- and slow-twitch fibres), it has been shown that fast fibres are recruited only during activities involving speed and/or power (Edwards, 1978) and if the response to a command to contract the PFM maximally is sluggish, then it is probable that only slow-twitch muscle fibres are being used. Jones and Rutherford (1987) reported that in some untrained muscles, the fast-firing muscle fibres are never recruited, and so these may need targeting in some other way during PFM rehabilitation. It is postulated that an MVC will recruit both fast- and slow-twitch muscle fibres (if the contraction is strong and brisk) and so P would represent recruitment of both fibre types.

It is further hypothesised that the endurance (E) and repetitions (R) of a PFM contraction are measures of slow-twitch muscle fibre activity. Gosling *et al* (1981) described the PFM slow-twitch fibres as being responsible for maintaining continuous muscle activity over prolonged periods, with the fast-twitch fibres recruited reflexly during sudden increases in intra-abdominal pressure, for example when coughing. A reduction in slow-twitch activity would therefore manifest itself in a short duration contraction and few repetitions, and PFM in this category may thus be responsible for decreased support of the proximal urethra and reduced urethral occlusion. A reduction in fast-twitch fibre activity may result in a slower, weaker reflex response to increases in intra-abdominal pressure. Logically, one should assess both the fast- and slow-twitch muscle fibre activity before planning an exercise regimen.

The third component in the PERFECT assessment method, the number of repetitions a subject could perform before the muscles fatigued, was incorporated into the scheme to provide information on the number of contractions an individual should perform at each exercise session to predict 'overload'. Previous studies have suggested a wide range of daily contractions; Benvenuti *et al* (1987) proposed 120 daily PFM contractions whereas Bø *et al* (1989) used 8 to 12 daily contractions, and these recommendations appeared to be made at random. The present study proposes patient-specific exercise programmes, as practised in other physiotherapy regimens for muscle dysfunction. The number of repetitions gives further information on fatigue; Edwards (1978) states that fatigue is failure to maintain the expected force with continued or repeated contractions.

It is postulated that the last component, F – number of fast contractions, provides a measure of fast-fibre activity. Millard (1987) was possibly the first to recognise the importance of practising fast and slow PFM contractions.

Compared with other studies (Hendrickson, 1981; Graber *et al*, 1981; Worth *et al*, 1986; Brink *et al*, 1989), the PERFECT method gives the examiner more flexibility and is less ambiguous.

### Validity Study 1

PFM digital assessment of strength is represented by P (power/pressure), which is thought to manifest itself with both squeeze pressure and lift. Consequently, it was decided to examine the 'lift' component of an MVC in this study. The sample number is small ($n = 8$) and so the results should be interpreted with caution, but analysis of the data showed a positive and significant correlation between perineal lift and digital assessment scores ($r = 0.864$). This theory would endorse the postural function of the PFM in supporting the proximal urethra, with reduced support, causing bladder neck descent on coughing, precipitating incontinence. It is postulated that continuous activity (resting tone) of the slow-twitch fibres should maintain the advantageous position of the bladder neck and proximal urethra, with the fast-twitch fibres reflexly recruited to provide a quick, strong re-enforcement of urethral lift and squeeze, during, for example, coughing. Bladder neck lift was also reported by Vierhout and Jansen (1989) using transrectal ultrasound on 17 women (15 with stress incontinence and two with other bladder problems). They reported a mean lift of 5.5 mm with a PFM contraction which compares well with the 5.0 mm lift observed in incontinent women in this study (6.0 mm for continent subjects).

### Validity Study 2

This study aimed to validate the subjective digital assessment of P (pressure/power) and E (endurance) and to compare the sensitivity between digital and perineometric techniques in a large sample of women ($n = 233$), most of whom were incontinent.

Accuracy of the digital method was demonstrated by the highly significant correlation of P with maximum pressure ($r = 0.786$; $p < 0.001$), and E with the area under the pressure curve ($r = 0.549$; $p < 0.001$) as recorded by a perineometer. This evidence lends support to the validity of the digital method. These findings are reinforced by the study by Brink *et al* (1989) of 388 older women (mean age 67.5), who showed significant positive correlation between digital pressure scores and electromyography (EMG) scores ($r = 0.60$; $p < 0.01$).

Further support for the digital technique is given in a study ($N = 263$) comparing digital scores (modified Oxford scale) and pressure (using the pelvic floor exerciser/perineometer – PFX), where good agreement between the two techniques was demonstrated (Isherwood and Rane, 2000).

To assess endurance digitally the examiner must be able to detect a reduction in pressure during a 10-second MVC, at which point the time (in seconds) is recorded. Digital measures are admittedly less sensitive than perineometric measures, due largely to the subjective nature of the test and the difficulty in gauging small changes in pressure. In addition, the area under the pressure curve is a record of total pressure recorded during the ten-second MVC (even though it may fluctuate), whereas E is an estimation of the time an MVC can be held at the maximum level. This disparity of measurement is reflected in the reduced (but still highly significant) correlation ($r = 0.549$; $p < 0.001$). The highly significant differences in both P and E values between continent and incontinent women suggest that these measures are important in the maintenance of continence.

**Reliability Studies**

Inter-examiner reliability was tested on ten subjects examined by JL and SHD on the same visit and showed significantly high coefficients of correlation between the two examiners for all four components of the assessment with the greatest discrepancy in assessing the number of repetitions. This discrepancy may be explained by the different levels of rapport between patient and examiner leading to a variable effort on the part of the patient. Furthermore, there is a learning curve involved with any new assessment technique and variability may be due to differing length of experience between the two examiners.

**Exercise Programme Planning**

The plethora of PFM exercise regimens in the literature indicates a lack of standardisation and, when compared with general rehabilitation methodology, the reports show lack of scientific application. There is a wide range of muscular strength and endurance across any female population, and so it is postulated that a uniform, standard regimen is not appropriate. Instead, assessment of the fast- and slow-twitch components, leading to an individual exercise programme, has been proposed, to target specifically the weakness of each individual patient. Rehabilitation is implemented at the level appropriate to the patient, and progressed sequentially as power and endurance increase; progression is guided by continuous re-assessment. Recently, it is believed that there is co-contraction of transversus abdominis during a PFM contraction (Sapsford *et al*, 2001) and it is recommended that this is now incorporated into treatment protocols.

The recommendations outlined above have not been tested clinically, and may need modification to cater for the personality and lifestyle of individual subjects. The number and type of daily repetitions will depend on the initial assessment, and several sessions per day are suggested. This is contrary to general muscle training programmes (2 to 5 sessions per week for athletes) but is considered necessary when dealing with a flat muscle with a small cross-sectional area, as such a muscle is easily fatigued. Fitness relates to the amount of work a muscle can produce and the time it takes to recover (Reilly, 1981) and so careful planning to ensure overload, but avoiding excess fatigue, is essential. A poor performance due to inadequate rest can be demoralising for both an athlete and an incontinent woman; furthermore, PFM fatigue may lead to a temporary aggravation of urinary symptoms.

Muscle weakness leads to reduction in mitochondrial oxidative capacity as indicated by a decrease in the succinate dehydrogenase (SDH) activity (Eriksson and Haggmark, 1979). Consequently, these same authors maintain that it is necessary to gradually increase the SDH reserves by repeated contractions which will deplete levels and stimulate greater production. This reinforces the theory of overload and supports the need for a programme of regular daily contractions advanced in this study.

The proposed PERFECT assessment scheme has the disadvantage of appearing complicated and some therapists may well prefer a simpler technique. On the other hand, many patients appreciate the rationale behind the assessment and individual exercise programme, and respond accordingly. Due to the subjectivity of the tests, especially evaluation of strength, standardised training is required to enable this scheme to be used in any multi-centre studies, to ensure accurate data collection. However, individual clinicians should be able to detect a change in any of the four parameters from the information contained in this report.

**References**

**Benvenuti, F, Caputo, G M, Bandenelli, S *et al* (1987)**. 'Re-educative treatment of female genuine stress incontinence', *American Journal of Obstetrics and Gynecology*, **154**, 1, 58-64.

**Bland, J M and Altman, D G (1999)**. 'Measuring agreement in method comparison studies', *Statistical Methods in Medical Research*, **8**, 135-160.

**Bø, K, Hagen, R, Jorgensen, J *et al* (1989)**. 'The effect of two different pelvic floor muscle exercise programs in the treatment of urinary stress incontinence in women', *Neurourology and Urodynamics*, **8**, 4, 355-356.

**Brink, C A, Sampselle, C M, Wells, T J *et al* (1989)**. 'A digital test for pelvic muscle strength in older women with urinary incontinence', *Nursing Research*, **38**, 196-199.

**Chiarelli, P (1989)**. 'Women's waterworks: Curing incontinence', *Century Magazines*, New South Wales.

**Edwards, R H T (1978)**. 'Physiological analysis of skeletal muscle weakness and fatigue', *Clinical Science and Molecular Medicine*, **54**, 463-470.

**Eriksson, E and Haggmark, T (1979)**. 'Comparison of isometric muscle training and electrical stimulation supplementing isometric muscle training in the recovery after major ligament surgery', *American Journal of Sports Medicine*, **7**, 169-171.

**Gilpin, S A, Gosling, J A, Smith, A R B *et al* (1989)**. 'The pathogenesis of genito-urinary prolapse and stress incontinence of urine: A histological and histochemical study', *British Journal of Obstetrics and Gynaecology*, **96**, 31-38.

**Gosling, J A, Dixon, J S and Humpherson, J R (1981)**. 'A comparative study of the human external sphincter and periurethral levator ani muscles', *Journal of Urology*, **53**, 35-41.

**Graber, B, Kline-Graber, G and Golden, C J (1981)**. 'A circumvaginal muscle nomogram: A new diagnostic tool for evaluation of sexual dysfunction', *Journal of Psychiatry*, **42**, 157-161.

**Haslam, J (2002)**. 'Biofeedback' in: Laycock, J and Haslam, J (eds) *Therapeutic Management of Incontinence and Pelvic Pain*, Springer-Verlag, London, in press.

**Hendrickson, L S (1981)**. 'The frequency of stress incontinence in women before and after the implementation of an exercise program', *Issues in Health Care of Women*, **3**, 81-92.

**Isherwood, P J and Rane, A (2000)**. 'Comparative assessment of pelvic floor strength using a perineometer and digital examination', *British Journal of Obstetrics and Gynaecology*, **107**, 1007-11.

**Jahnke, M T, Proske, U and Struppler, A (1989)**. 'Measurements of muscle stiffness, the electromyogram and activity in single muscle spindles of human muscles following conditioning by passive stretch or contraction', *Brain Research*, **493**, 103-112.

**Jones, D A and Rutherford, O M (1987)**. 'Human muscle strength training: The effects of three different regimes and the nature of the resultant changes', *Journal of Physiology*, **391**, 1-11.

**Kegel, A H (1948)**. 'Progressive resistance exercise in the functional restoration of the perineal muscles', *American Journal of Obstetrics and Gynecology*, **56**, 238.

**Kegel, A H (1956)**. 'Early genital relaxation: New technique of diagnosis and non-surgical treatment', *Obstetrics and Gynecology*, **8**, 545-550.

**Khullar, V (2002)**. 'Investigations' in: Laycock, J and Haslam, J (eds) *Therapeutic Management of Incontinence and Pelvic Pain*, Springer-Verlag, London, in press.

**Laycock, J and Jerwood, D (1994)**. 'Development of the Bradford perineometer', *Physiotherapy*, **80**, 139-142.

**Millard, R J (1987)**. *Overcoming Incontinence*, Thorsons, Wellingborough, page 21.

**Miller, J, Ashton-Miller, J and DeLancey, J O L (1996)**. 'The knack: Use of precisely-timed pelvic muscle contraction can reduce leakage in SUI', *Neurourology and Urodynamics*, **15**, 4, 392-393.

**Reilly, T (1981)**. *Sports Fitness and Sports Injuries*, Faber and Faber, London.

**Sapsford, R R, Hodges, P W, Richardson, C A *et al* (2001)**. 'Co-activation of the abdominal and pelvic floor muscles during voluntary exercises', *Neurourology and Urodynamics*, **20**, 31-42.

**Vierhout, M E and Jansen, H (1989)**. 'Supine and sitting transrectal ultrasonographic evaluation of the bladder neck during relaxation, straining and squeezing', *Neurourology and Urodynamics*, **8**, 301-302.

**Worth, A M, Dougherty, M C and McKey, P L (1986)**. 'Development and testing of the circumvaginal muscles rating scale', *Nursing Research*, **35**, 3, 166-168.

## Key Messages

- Digital assessment of the contractility of PFM can be carried out during vaginal examination.
- Assessment should include evaluation of PFM strength and endurance, and reflex activity (during cough).
- The PERFECT assessment is easy to perform, reliable and reproducible.
- The assessment scheme described provides information for a patient-specific exercise programme.