The Blatt-Kupperman menopausal index: a critique

Elizabeth Alder

Department of Epidemiology and Public Health, University of Dundee, Ninewells Hospital and Medical School, Dundee DD1 9SY, Scotland, UK

Accepted 5 January 1998

Abstract

The Blatt-Kupperman menopausal index has been used widely in studies of climacteric symptoms, but it now needs to be reassessed. The original index was derived from clinical experience in New York in the 1950s. The index was a combination of self report and physician ratings; it omitted measures of vaginal dryness and loss of libido; no demographic data of the sample were given; weighting was used without statistical justification; terms were ill defined; categories included overlapping scores, and most importantly scores were summed without being based on independent factors. Modern psychometrics has led to the publication of reliable and valid scales. These should be used in climacteric research in preference to the Blatt-Kupperman index. © 1998 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Menopause; Blatt-Kupperman index; Measurement; Psychometrics

1. Introduction

Many clinical studies of the climacteric have attempted to measure varied patterns of symptoms and complaints, and in the last 40 years a number of research studies have used an index of climacteric symptoms known as the Kupperman index or the Blatt-Kupperman index.

In 1952, in a paper on estradiol implants, Kupperman described a numerical summation of eleven menopausal complaints and termed it the menopausal index [1]. This was based on Dr Kupperman’s experience in treating women with menopausal complaints and arose from his desire to provide a numerical score for complaints which could then be used to evaluate treatments. The following year he described a modification to the menopausal index, which allowed for some symptoms to be weighted more than others [2]. This index was then used in subsequent papers on menopausal symptoms and was widely quoted [3].

Ballinger [4] published a major review on the psychiatric aspects of the menopause and pointed out that most of the general population surveys of the 1960s and 1970s used symptom checklists based on the Blatt-Kupperman index, but failed
to comment on its limitations. Bungay et al. [5] used it in a study of 1120 women aged between 30 and 64. They found a peak of ‘minor mental illnesses’ just before the menopause and a peak of complaints of flushes and sweats at the time of the menopause. Cittadini et al. [6] in Argentina and Crona et al. [7] in Sweden used their own version of the Blatt-Kupperman index to assess the effect of ORG OD 14 in climacteric symptoms, but much of the early work on climacteric symptoms was methodologically flawed [8].

Marslew et al. [9] carried out a randomised control trial of combined hormone replacement therapy in order to investigate the effects of progesterone. They used the Kupperman index to assess changes over time and differences between the therapies. Again the index is used uncritically. They note: ‘The Kupperman index is a well documented method of recording the severity of climacteric symptoms. It provides an integral assessment of the commonest symptoms and the questionnaire is easy for the patient to handle and to complete accurately after receiving brief instructions’ p. 14. In this study, they used the weightings suggested by Kupperman although in an earlier study [10] they only weighted hot flushes.

More recently Wiklund et al. [11] used a modified Blatt-Kupperman index in a study of the effect of transdermal therapy on quality of life. Two hundred and twenty three women were randomly allocated to either transdermal therapy (20 mg:24 h) or placebo, given as patches twice a week. The symptoms that make up the Index were assessed by the same gynaecologist before and after twelve months treatment. It is not stated whether the symptoms were weighted. A number of other quality of life assessments were based on self report questionnaires. The climacteric symptoms used in the Kupperman index were also self rated by the women on visual analogue scales and combined into three factors (psychologic, vaso-motor and other). The Kupperman index totals were compared before and after treatment and there were significant differences ($P<0.001$, Wilcoxon test) with transdermal oestradiol therapy being superior to placebo ($P=0.0001$). The self rated symptom scores also showed significant differences in the same direction as did the other quality of life measures. No correlations were made between measures and no comments were made on any differences in discrimination between the clinical and self assessments.

In 1996, a further twelve papers or posters which were presented at the 8th International Congress of the Menopause in Sydney, Australia, reported using the Kupperman scale [12–22].

2. The index

The Blatt-Kupperman menopausal index (Table 1) included 11 symptoms which were rated on a four point scale from 0 = none to 3 = severe. These ratings were then summed across the 11 items and the score was categorised as follows: none 0–5, mild 5–10, moderate 10–15, severe 15+. In later papers [3], weightings were introduced and the scores were categorised as follows: mild 15–20, moderate 20–35, severe 35+.

3. The critique

The Index was pioneering in its attempt to quantify symptoms, which was needed to test the efficacy of menopausal treatments, but this was 40 years ago. The Index does not meet with accepted standards of psychometrics today [23], and for the following reasons it can be argued that it should now be abandoned.

3.1. Biased assessment

The Blatt-Kupperman index was originally based on ‘the severity of her presenting symptoms as estimated by the examining physician, and also according to the concept the patient had with respect to the severity of her own complaints’ ([1], p. 326).

The scale combined physician and patient assessments, but the relative contributions of each are unclear. The patient’s reporting of her symptoms could have been influenced by her relationship with her physician making it more likely that
Table 1
The Blatt-Kupperman menopausal index [2]

<table>
<thead>
<tr>
<th>Symptoms</th>
<th>Example</th>
<th>Factor</th>
<th>Severity</th>
<th>Numerical conversion factor × severity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vasomotor</td>
<td></td>
<td>4</td>
<td>M</td>
<td>8</td>
</tr>
<tr>
<td>Paraesthesia</td>
<td></td>
<td>2</td>
<td>+</td>
<td>6</td>
</tr>
<tr>
<td>Insomnia</td>
<td></td>
<td>2</td>
<td>+</td>
<td>6</td>
</tr>
<tr>
<td>Nervousness</td>
<td></td>
<td>2</td>
<td>M</td>
<td>4</td>
</tr>
<tr>
<td>Melancholia</td>
<td></td>
<td>1</td>
<td>+</td>
<td>3</td>
</tr>
<tr>
<td>Vertigo</td>
<td></td>
<td>1</td>
<td>S</td>
<td>1</td>
</tr>
<tr>
<td>Weakness (fatigue)</td>
<td></td>
<td>1</td>
<td>S</td>
<td>1</td>
</tr>
<tr>
<td>Arthralgia and myalgia</td>
<td></td>
<td>1</td>
<td>M</td>
<td>2</td>
</tr>
<tr>
<td>Headaches</td>
<td></td>
<td>1</td>
<td>S</td>
<td>1</td>
</tr>
<tr>
<td>Palpitations</td>
<td></td>
<td>1</td>
<td>+</td>
<td>+3</td>
</tr>
<tr>
<td>Formication</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Menopausal index (sum) 35

Coding of symptoms: 0 (none) = 0, S (slight) = 1, M (moderate) = 2, and + (severe) = 3.

she would report symptoms if she had a good relationship with him or less likely if he was perceived as unsympathetic. The authors of Kupperman’s papers are all male, and the women may have reported what they thought the physician expected to hear. There is a body of evidence [24] that shows that symptom reporting is influenced by many factors other than physiological change. The expectations of physicians may also influence their ratings of the complaints. In a clinical context, this is accepted practice, but Kupperman and his colleagues proposed the index for use as a tool that allowed ‘statistical comparison between the various compounds that were employed’ ([2], p. 796). They therefore proposed the scale for use in research, and not only in clinical practice.

3.2. Sexuality

The Index does not include two frequently described climacteric symptoms: vaginal dryness and loss of libido. Along with vasomotor symptoms, vaginal dryness is the only symptom that has been shown to be clearly oestrogen related [25,26]. These symptoms may have been excluded by Kupperman and colleagues because of lack of other evidence of hypo-oestrogenic effects at that time. However, low oestrogen levels in lactating women are associated with vaginal dryness [27], and it seems unlikely that this was unknown to obstetricians and gynaecologists. It is more likely that sexuality was not considered an issue to be discussed or that it was not perceived as a medical problem.

3.3. Social context

The selection of symptoms and their weighting was based on clinical experience in New York in the 1950s and Kupperman gives no data on the social background of his patients. We do not know how typical these women were of menopausal women in general or even of menopausal women attending menopause clinics. Those who reported symptoms in the 1950s (possibly born last century) may be very different from those women who report menopausal symptoms today.

3.4. Weighting of factors

In the second paper, Blatt et al. [2] weighted vasomotor symptoms by a factor of four, and nervousness, insomnia, and paraesthesia by a factor of two. They did this because ‘we believe that they are the most important symptoms comprising the menopausal syndrome and are responsible in great part for the patient’s presenting herself
for therapy' ([2], p. 795). Weighting is a mathematical procedure, but in this scale it is not based on any justification. The advances in statistical analysis in the last 20 years have enabled psychometricians to calculate weighting objectively [23].

3.5. Definition of terms

The terms included in the Index are not defined and some are no longer in common use. Formication means a sensation ‘as of ants crawling up the skin’, which is presumably different from paraesthesia which is ‘tingling’. While these terms may be recognisable, it may be difficult to interpret them, and it is unlikely that they are familiar to patients. If the index is used by a physician there may be different interpretations of these terms, and if used by the patients themselves they may be misunderstood. The Flesch reading index for this list of symptoms is 35.6, which means that it would be understood by less than 24% of the population [28].

3.6. Scoring

The categories include overlapping scores. If a patient has a score of 20 she could be categorised as either moderate or as severe.

3.7. Summing scores

The most important limitation of the Blatt-Kupperman index is that it is not statistically valid to sum scores of diverse symptoms, unless they have been shown by factor analysis to be independent. There have been a number of factor analytic studies of menopausal complaints that have identified a very different set of important symptoms [29–32]. These findings have then been incorporated into scales that are statistically valid [32–34].

3.8. Modified scales

The Blatt Kupperman index is still being used today. In many cases the original scale has been changed and bears little resemblance to the scale proposed by Kupperman. The selection of symptoms varies and some so-called ‘Kupperman scales’ include extra items such as urinary symptoms. In some cases the scale is then called a ‘modified Kupperman Index’. The number of items also varies and ranges from nine [35] to 18 [36]. We often do not know in what way the items are weighted. In most cases the Index is used as a self report scale, unlike the original combination of physician and self report. In these variations of the Blatt-Kupperman index, it is effectively only the principle of assigning a numerical value to reported symptoms that is being used. The description of a measure by name and reference to a journal publication may give it spurious credibility. The conclusions drawn from research based on invalid measurement may be unsound.

4. Modern psychometrics

Since Kupperman first introduced his index over 40 years ago, there have been improvements in measurement of psychological symptoms or climacteric complaints, based on sound psychological knowledge [23]. In terms of the requirements of modern scales we can compare the Blatt-Kupperman Index with three published scales (Table 2). The two most important properties of psychometric scales are reliability and validity. Reliability is the extent to which the scale...
is consistent, or effective, and no data on this is given for the Blatt-Kupperman index. It probably has face validity, in that it appears to measure what it attempts to measure. However, there is no concurrent validity showing how it agrees with other scales.

Scores on established scales are usually standardised and we assume a normal distribution. No means or standard deviations were given for scores on the Blatt-Kupperman Index. The list of symptoms is derived from clinical experience and not subjected to factor analysis.

5. Conclusion

The attempt by New York clinicians in the 1950s to measure climacteric symptoms when statistics and psychology were still young, is to be applauded. Now we approach a new century with more powerful statistical analysis using new technology, sophistication in scientific method and international research. There is a move for evidence based medicine and for this we need sound measurement, and the Blatt-Kupperman Index has severe limitations. If we combine clinical sensitivity and understanding of women with menopausal problems, with research using properly designed tools, we increase our scientific rigour.

References