BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (http://bmjopen.bmj.com).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Assessment of patients' expectations: development and validation of the Expectation for Treatment Scale (ETS)

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-026712
Article Type:	Research
Date Submitted by the Author:	15-Sep-2018
Complete List of Authors:	Barth, Juergen; University Hospital Zurich, Institute for Complementary and Integrative Medicine Kern, Alexandra; Institute for Complementary and Integrative Medicine, University Hospital Zurich and University Zurich, Zurich Switzerland Lüthi, Sebastian; Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich, Zurich, Switzerland Witt, Claudia M.; Institute for Complemtary and Integrative Medicine, University Hospital Zurich and University of Zurich, Zurich, Switzerland; Institute for Social Medicine, Epidemiology and Health Economics, Charité
Keywords:	expectations, placebo, assessment, pain, acupuncture

SCHOLARONE™
Manuscripts

For peer review only

Assessment of patients' expectations: development and validation of the Expectation for Treatment Scale (ETS)

Running title: Expectation for Treatment Scale (ETS)

Submission to British Medical Journal

Jürgen Barth*, Alexandra Kern, Sebastian Lüthi, Claudia M. Witt

Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich, Sonneggstrasse 6, 8006 Zurich, Switzerland Dr. Jürgen Barth Research Director, M.A. Alexandra Kern PhD student, M.A. Sebastian Lüthi Master student, Professor Dr. Claudia M. Witt, Chair and Director

Institute for Social Medicine, Epidemiology and Health Economics, Charité – Universitätsmedizin Berlin, Luisenstrasse 57, Berlin, Germany, Professor Dr. Claudia M. Witt, research group leader

Center for Integrative Medicine, University of Maryland, School of Medicine, 520 W Lombard Street, Baltimore, MD, USA Professor Dr. Claudia M. Witt, senior scientist

*Correspondence:

Dr. Jürgen Barth, University Hospital Zurich and University of Zurich, Switzerland

Fon: +41 (0)44 255 4896

mail@juergen-barth.de

www.iki.usz.ch

Word count: 4298

For peer review only

Abstract

Objective: Development of a short instrument for the assessment of expectations (Expectation for Treatment Scale, ETS), using acupuncture as case example.

Design: Cross sectional assessment with retest after one week.

Setting: Web survey with patients suffering from pain

Methods: In a three-step approach, we reduced the initially collected number of items from 17 to 9 and to 5, including expectations about coping ability, vitality, physical health and reduction of patient complaints. Items were selected according to internal consistency (Cronbach's alpha), convergent and divergent validity with related constructs (optimism, pessimism, resilience, perceived sensitivity to medicines, depression, and others), one-week retest reliability (intraclass correlation coefficient, ICC) and exploratory and confirmatory factor analysis.

Results: A total of 102 pain patients were included, and 54 of these patients completed the retest assessment. The final version of the ETS consisted of 5 items and had an excellent Cronbach's alpha (.90) with 72.33% variance on one single factor. Depression, pessimism and perceived sensitivity to medicines showed positive correlations with our expectation measure (r = 0.23, r = 0.20, r = 0.34, respectively), the correlation between the ETS and optimism was low (r = -0.07), and no correlation between the ETS and resilience was found (r = -0.07). The retest ICC was .86, which showed high stability over one week. A confirmatory factor analysis (N = 439) with data from patients with low back pain confirmed the single factor structure of the instrument.

Conclusion: The ETS showed strong psychometric properties and covered a distinct construct. As next step it will be used for other disorders and treatments in which expectations are also an important predictor of treatment effects.

For peer review only

Keywords: Expectations, placebo, assessment, pain, optimism, acupuncture

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Strengths and limitations of this study

- It is the first project to systematically develop a general measure to assess patients' expectations across medical conditions and treatments with the involvement of patients, earlier scale and empirical data.
- The Expectation for Treatment Scale (ETS) is a short and reliable measure which capture outcome related expectation of patients.
- The association between ETS and related constructs was explored.
- The development of the ETS was done by an online survey with pain patients and the initial findings were confirmed by a patient sample with low back pain
- The transfer of the ETS in different clinical conditions and settings should be pre-tested in order to explore if patients can evaluate the respective treatment on this level.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Introduction

Patients' expectations alter responses to interventions in placebo research as well as in clinical studies. In experimental research the instruction to receive an active intervention goes along with large effects on pain reduction (effect size = 0.75) [1]. Expectations can be regarded as a major driver for changes in symptoms and other health related outcomes [2]. Clinical studies have revealed that patients' positive expectations are related to reduced pain after a medical treatment [3-5] and this phenomenon has also been observed for other medical conditions [6]. For acupuncture, this association was investigated in several studies with heterogeneous findings: Patients with chronic pain had higher odds (odds ratio 2.11 [95%CI 1.32-3.34]) of benefitting from acupuncture treatment if they expected a better treatment response at baseline (controlled for other clinical baseline characteristics) [7]. However, there are also studies that found no influence of expectations on the outcome [8 9].

These differences in the association of expectations and outcomes might be partly explained by the fact that each study used a specific measure and differences between measures might hide or exaggerate associations between expectations and intervention outcomes. Additionally, the match between expectations of patients and treatment providers might be relevant for the success of a specific treatment [10].

Expectation is a well-known and often-used term. Nevertheless a clear definition and a sharp distinction from associated constructs is important for the development of a measure [11]. A basic definition suggests that ‘‘patient expectations’’ cover treatment-related outcome expectations [12]. Related constructs such as optimism, self-efficacy, and hope have some shared facets with expectations, such as positive cognition, but differ on the level of the construct [13 14]. Related constructs are sometimes included in expectation measures, and in other cases, only cognitions about treatment outcome relationships are included [9 15 16].

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Along the definition by Bowling et al. (2012), we therefore designed our measure to assess expectations related to a clinical intervention with a clinically relevant outcome from a patient's perspective.

For further research in the field of expectations a strong measure with high acceptance across clinical fields would be needed for several reasons. First, a reliable measure with high internal consistency at a specific time point is a prerequisite to use expectations as a robust predictor. Second, ceiling effects were a common problem in the measurement of expectations, since patients who are seeking help from a specific treatment often expect a large benefits and one established expectation measure (Acupuncture Expectancy Scale (AES)) [17]. Therefore, some authors have used the term "realistically expect" to capture expectations in a recent study with a comprehensive 10-item assessment of expectations, hope and beliefs (EXPECT) [18] with limmited internal consistency. Third, a measurement of expectation should be stable in a reasonable timeframe: The EXPECT scale showed only a moderate retest reliability with an intra-class correlation of ICC = 0.75. Most importantly, the association of EXPECT with the established AES, a benchmark measure for the field of acupuncture, was unexpectedly low (r = .54).

This study aimed to develop a measure to capture patient expectations with a short scale in a reliable way. The newly developed scale (Expectation for Treatment Scale, ETS) ideally has a strong association with available measures of expectations, moderate associations with related constructs (e.g., optimism, pessimism) and no association with personality (e.g., neuroticism). Furthermore, it should be constructed in a way that it can be adapted to other treatments and used in a universal way in different clinical fields independent of patients' complaints.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Methods

To address the problem of the abovementioned ceiling effects, we developed the items of our scale (ETS) in a pilot study with pain patients in our outpatient clinic to measure expected benefits before treatment from the patient's perspective. Based on existing questionnaires on patient expectations [9 17 19-22]. We created a list of 17 items covering different facets of expectations. After testing the items, we selected items with low skewness, a large range of responses (i.e., variation), and high correlations between items.

Participants

For the main validation, study subjects of the convenience sample were recruited in different ways. Three regional patient organizations distributed the information, and we used several email distribution lists (including the University of Zurich, Switzerland and the Charité University Hospital Berlin, Germany). Furthermore, patients from the previous year seeking treatment for a musculoskeletal condition at the Institute for Complementary and Integrative Medicine of the University Hospital Zurich were contacted by email. All participants had to be at least at age 18 and according to their own evaluation sufficient knowledge to understand German language.

The study link was accessed by 522 subjects and 244 started with the survey. In total 142 subjects were excluded due to insufficient pain (N = 113), missing data in the nine items version of the ETS (N = 14) and more than 20% missing across all items (N = 13). Data from patients suffering from pain (N = 102) with retest data from 54 patients were available.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Procedures

Subjects who decided to participate activated an access link to the online survey provided via email. The survey was hosted by https://www.soscisurvey.de/, and all data were collected electronically and anonymously. Participants were able to complete the survey within 15 minutes. All questions were created in a forced choice format, meaning that subjects were not able to continue to the next set of questions without completion of the previous ones. Upon starting the survey, a short address of welcome, followed by an introduction to the study, was provided. Subjects who agreed to be contacted again for the retest assessment provided their email address at the end of the survey. The retest sample received a consecutive ETS one week after completion of the first survey. The study was granted ethics approval by the local ethics committee (Kantonale Ethikkommission Zürich, No. 48-2015).

Measures

The survey included demographic information (age, gender, education, employment status, and country of residence), questions about any previous experience with acupuncture (current or earlier), the degree to which the acupuncture treatment has been successful (numeric rating scale from 1 to 10) and the reason for the treatment (pain, mental health or non-specific). Patients who suffered from pain (binary variable with yes vs. no options) were asked to give information about their pain. We assessed average pain intensity (numeric rating scale from 1 to 10) if a physician was consulted or if any other kind of therapy was used and the number of days of restriction due to pain. Subsequently,

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

subjects filled in the 9-item ETS. Below we provide a detailed description of all measures in this study.

Expectation for Treatment Scale

The first version of the ETS consisted of nine items (e.g., "I expect the treatment [acupuncture] will help me to cope with my complaints."). Each item was to be rated on a four-point scale ranging from 1 to 4 (partially disagree, partially agree, agree, definitely agree). We decided to use a four-point scale instead of a five-point scale for two reasons. First, the chosen format forces the patient to provide an answer with a direction (higher or lower than the middle answer choice) instead of opting for the middle category, which can often be the case in situations of insecurity (such as the estimation of future events here) [23]. Second, the lowest answer on the scale of "definitely disagree" can be considered unexpected for patients motivated to undergo an acupuncture treatment. After the statistical analyses, the nine-item version was reduced to a final 5-item version of the ETS. A detailed description of these five items is provided in Table 1.

Insert Table 1 about here

Optimism and pessimism

Both concepts were assessed using the German version of the Life Orientation Test—Revised [24]. The questionnaire consists of six self-report items (plus four filler items), each rated on a five-point Likert scale ranging from 0 (strongly disagree) to 4 (strongly agree). The data were separated into optimism and pessimism scores, as recommended by Glaesmer et al. (2008). Each possible score ranged from 0 to 12, indicating that higher

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

values display either a higher peculiarity for optimism or pessimism. In the present study, the Cronbach's alpha was acceptable for optimism (0.72) and questionable for pessimism (0.60).

Depression

The Patient Health Questionnaire (PHQ-9) is widely used as a standard instrument for diagnosing depression in primary care and is considered to be well validated [25]. The PHQ-9 consists of nine questions operating according to the modified Diagnostic and Statistical Manual, Fourth Edition criteria. Patients indicated their answers on a scale ranging from 0 (not at all) to 3 (almost every day) [26]. A higher score indicates a higher severity of depressive symptoms (ranging between 0 and 27). In the present study, the internal consistency of the scale was excellent (Cronbach's alpha = 0.90).

Resilience

As a protective personality factor, resilience was appraised using the Resilience Scale [27]. In the current study, we used the German Version Resilienzskala (RS-11) [28]. The instrument consists of eleven questions with seven response alternatives ranging from 1 (disagree) to 7 (fully agree). Higher scores indicate higher resilience. In the present study, the internal consistency was excellent (Cronbach's alpha = 0.90).

Body-Efficacy Expectation

To measure beliefs concerning bodily coping capabilities, the Body-Efficacy Expectation (BEE) questionnaire [29] was used. The instrument consists of five items using a four-point Likert scale ranging from 1 (not true at all) to 4 (exactly true). A higher score indicates stronger beliefs in one's bodily coping capabilities. In the present study, the Cronbach's alpha was 0.75, indicating an acceptable internal consistency.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Perceived Sensitivity to Medicines

To assess the individual's sensitivity to medicines, we used the Perceived Sensitivity to Medicines scale (PSM) [30]. The PSM is considered a reliable and valid measure composed of five self-report questions to assess perceived sensitivity to the potential adverse effects of medicines. Responses are scored on a five-point Likert scale, and patients' item scores are summed to provide a total score ranging between 5 and 25. Higher scores point toward a high perceived sensitivity to the potential adverse effects of medicines. In the current study, the scale showed excellent internal consistency (Cronbach's alpha = 0.94).

Neuroticism and openness to experience

To assess the distinctness of neuroticism and openness to experience among patients, we used the corresponding subscales of the NEO Five Factor Inventory (NEO-FFI) [31]. To that end, 23 questions were extracted from the 60-item NEO-FFI. The instrument utilizes a five-point Likert response format, from 1 (strongly refuse) to 5 (strongly agree). A higher score indicates a higher value for neuroticism and openness to experience. In the present study, the Cronbach's alpha was 0.85 for neuroticism and 0.72 for openness to experience, which indicates good to acceptable internal consistency.

Acupuncture Expectancy Scale

Mao and colleagues developed the AES aiming to measure patients' expected response from acupuncture [17], which was developed for only one clinical intervention (i.e. acupuncture). The instrument consists of four items. The answers are given using a five-point Likert scale ranging from 1 (not at all agree) to 5 (completely agree). A higher score points toward higher expectancies. In the present study, the internal consistency of the scale was considered good (Cronbach's alpha = 0.88).

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Statistical Analyses

All data analyses were executed using SPSS (version 22, SPSS Inc., Chicago, Illinois, USA). As a first step, we used the nine items of the first ETS version to explore homogeneity and diversity between items. We conducted a descriptive analysis of the data with the aim of detecting possible floor or ceiling effects and assessing the distribution of the data. Internal consistency was examined through reliability analyses (Cronbach's alpha), with the item-total correlation corrected and the Cronbach's alpha if the item is deleted. Retest reliability for the nine items was assessed by the intraclass correlation coefficient (ICC). We set for each item a minimal acceptable ICC of .60 and for the total score an ICC of .80. Three criteria guided the decision to keep an item for the final version: 1) it should have a high item-total correlation corrected and no low internal consistency if deleted, 2) the item should not overlap too strongly in content with another item that might be included, and 3) the item should not contribute to ceiling effects in the final version of the scale, which means items with lower values were preferred.

In a second step, we used the reduced version of the scale to generate a total sum score of the five items (ranging from 5 to 20). To examine divergent and convergent validity, correlations between the ETS sum score and the other measures were calculated. We assumed a very high correlation between the ETS and the most strongly related construct (r > .70; AES), moderate correlations with strongly related constructs (r > .30; LOT-R optimism, inverse with LOT-R pessimism), small correlations with less related constructs (r > .20; PHQ-9; RS-11, BEE) and no correlation with unrelated constructs (PSM, NEO-FFI neuroticism, NEO-FFI openness to experience).

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

To test our assumption of one general factor an explorative factor analysis using a varimax rotation, an Eigenvalue of more than one and the Scree test was used with the 5 ETS items to determine the number of underlying factors. In addition we conducted a confirmatory factor analysis with an independent sample of chronic low back pain patients, which had been included in a randomized controlled trial (registration number DRKS-ID: DRKS00010191). These patients filled in the ETS. The data has been used to test the single factor structure. The best model used correlated error terms between item 2 and item 5 and is shown in Figure 1. Multiple goodness of fit tests [32] were used to evaluate model, including the comparative fit index (CFI; [33]), the normed fit index (NFI; [32]) the goodness of fit index (GFI; [34]) and the root mean squared error of approximation (RMSEA). A CFI greater than 0.90 indicates a good fit to the data [35]. A NFI and GFI greater than 0.90 indicates a good fit to the data [36]. A RMSEA with values of less than 0.08 indicates a good fit to the data [37], while values greater than 0.10 suggest strongly that the model fit is unsatisfactory. The Chi-square goodness of fit test can be regarded as a general test for the acceptability of the model: a statistically significant χ^2 indicates that a significant proportion of variance remains unexplained by the model [32]. Confirmatory factor analysis was conducted with the Analysis of Moment Structures (AMOS) software version 25 which is part of the SPSS package.

Results

Sample characteristics

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Three-quarters of the patients were female, one-third had a high school degree or higher vocational training and approximately 70% were currently employed. Approximately two-thirds of the surveys were conducted with patients in Switzerland, and one-third came from Germany. The characteristics of patients with current pain are displayed in Table 2. Descriptive information about the scales used in the study is presented in the appendix.

Insert Table 2 about here

First version of the ETS

In an initial analysis, we included nine items of the ETS. Descriptive results and results from scale analyses are presented in the appendix. In general, the standard deviations showed large variance, with the exception of item 6 ("The treatment [acupuncture] is in general effective") and item 7 ("I myself have positive expectations about the treatment [acupuncture]"). The mean value of item 6 was relatively high, which can be seen as an indicator of a ceiling effect. Therefore, we decided to exclude item 6 (i.e., general expectations about the effectiveness of the treatment [acupuncture]) and item 7 (i.e., personal expectation of a treatment) from the final ETS.

The item-total correlation corrected and the Cronbach's alpha if an item is deleted are indicators for the coherence of the single item meaning with the total scale (see appendix). The item-total correlation corrected did not give any strong reason for the exclusion of items. However, there had been some items with very high total correlations between items (item 2 and item 9). Item 2 ("I expect my complaints will be much better because of the treatment [acupuncture]") and item 9 ("I expect that after the treatment [acupuncture] my complaints will be considerably better") were very similar in terms of their

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

wording, and we decided to keep only one of the items for the final version. The reason behind this decision was that the ICC value for item 2 was lower than that of item 9. In addition, the mean value for item 9 was lower than the mean of item 2, which consequently can contribute to a larger variation in the final scale since ceiling effects can be avoided. The decision to exclude item 1 was based on content. Based on pre-study feedback, a misunderstanding of “clearly reduced” became evident. Furthermore, item 1 (“I expect the treatment [acupuncture] will clearly reduce my complaints”) and item 4 (“I expect the treatment [acupuncture] will make my complaints disappear”) covered similar topics. The factor loading confirmed our earlier decision to eliminate item 6 (“The treatment [acupuncture] is in general effective”) since the factor loading was relatively low. All items of the first version contributed to one single factor with 65.51% of explained variance. Additional information concerning the normal distribution for the 9-item ETS version, the 5-item ETS version as well as for the AES items can be found in the appendix. It becomes apparent that the 9-item ETS version total score is normally distributed. For the 5-items ETS version and the AES items there had been substantial number of subjects with either the lowest or the highest score on the scale.

The final version of the ETS consists of five items: the previous item 3 (“I expect the treatment [acupuncture] will help me to cope with my complaints”) covers coping ability, item 4 (“I expect the treatment [acupuncture] will make my complaints disappear”) covers total absence of complaints, item 5 (“I expect the treatment [acupuncture] will improve my energy”) covers an energy increase, item 8 (“I expect the treatment [acupuncture] will improve my physical performance”) covers an improvement of physical functioning and item 9 (“I expect that after the treatment [acupuncture] my complaints will be considerably better”) covers a considerable decrease in symptoms. To summarize, the ETS captures an

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

expected decrease in symptoms, an expected increase in energy and expected overall wellbeing after a treatment.

Insert Table 3 about here

The Cronbach's alpha of the final version was 0.90, which is an excellent value for a very short scale (Table 3). Retest reliability was excellent for the five items. All items contributed substantially to the final scale, and the factor loading indicated a single factor structure with 72.33% of explained variance.

Convergent and divergent validity

We expected a moderate correlation between the ETS and the LOT-R optimism and an inverse correlation with pessimism. However, these assumptions were not confirmed by our findings (Table 4). Interestingly, higher pessimism was moderately associated with higher expectations, both for the ETS and the AES. The ETS showed a small but significant correlation with the PHQ-9, which is in line with our assumption. The correlation between the ETS and the RS-11 or the BEE was close to zero, which is unexpected according to our assumptions. However, we found a moderate correlation between the ETS and the PSM, for which higher expectations were associated with a higher sensitivity to medication. The shared aspect of this association is the responsiveness to a medical treatment. Personality traits (NEO-FFI) were not associated with the ETS, which is in line with our assumptions. The ETS score has a very high correlation (> .90) with the AES which can be regarded as benchmark measure for acupuncture expectations.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Insert Table 4 about here

Confirmatory factor analysis

In addition we conducted a CFA with the data from the 439 ETS questionnaires mentioned above with the model presented in Figure 1. The factor loadings were between 0.609 (Item 2) and 0.796 (Item 4). We found an acceptable model fit in the Chi-square statistics with (Chi-square = 5.859, d.f. = 4, p = .210) indicating that the model is able to explain the data structure in general. The CFI was .998 in our model and indicates a very good model fit. Similarly the NFI (.993) and GFI (.995) confirms the excellent model fit. The RMSEA (.033) suggests also that the model with one single factor explains the data very well. The good internal consistency of the ETS in this sample (Cronbach alpha = .836) confirmed the findings of the scale development.

Insert Figure 1 about here

Discussion

The ETS is a well-validated and brief 5-item scale for measuring patient's expectations with excellent test-retest properties. We had been able to overcome ceiling effects, which had been a limitation of earlier measures [17].This improvement might be the consequence of using only four response categories instead of a larger scale with rather inadequate categories, (namely, "definitely disagree"). The ETS has excellent measurement properties concerning Cronbach's alpha, retest reliability over one week and single factor

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

structure was replicated in a clinical sample. The ETS can be used for research and clinical purposes equally.

The ETS could be adapted to other clinical situations and treatments beyond acupuncture, which is an important step toward implementing treatment expectations as working mechanisms among a variety of patient populations in a standardized way. Pain disorders, mental disorders and functional symptoms might be the most appropriate fields, since clinical research indicates the high relevance of expectations for treatment outcomes. The ETS can be used by patients with lower levels of health literacy, since items are kept very descriptive and intuitive. We avoided the use of percentages, technical terms and probabilities to make the ETS as easy to administer in patients as possible. Compared to the EXPECT scale [18], we did not ask for specific improvement according to symptoms, which can be challenging task for patients.

Earlier findings about the expectation outcome association in clinical studies had been limited by the diversity of measures. A strong measure is a prerequisite to accurately predict treatment responders based on pre-treatment expectations. After further investigation of the results from a systematic review by Prady and colleagues [15], we found that out of ten trials, only five provided their exact item wording for measuring expectations. Since the assessment instruments are not available, it is difficult to replicate the studies. Three out of the five mentioned studies used only one item in order to assess expectations. There is no reporting at all of Cronbach's alpha in the two remaining studies.

As mentioned earlier, the ETS showed a substantial correlation with the AES. Nevertheless, there are some differences worthy of mention between the two measures. In contrast to Jun Mao and his colleagues, we included measures of convergent and divergent validity in our study in order to enhance information about the unique aspects of expectations compared to related constructs. Moreover, the ETS can be used for a

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

multitude of interventions and can easily be adapted to other clinical situations, since no illness-specific symptoms are mentioned in the scale. Furthermore, there are differences concerning item wording and scaling (four vs. five response categories) between the ETS and the AES.

Another notable issue that deserves to be discussed is the moderate correlation between high pessimism and high expectations for both the ETS and the AES. Our initial assumption was to find a negative association between both expectation measures and pessimism, which was not confirmed. Furthermore, we found no correlation between optimism and the ETS in our study. The findings about negative cognitions (pessimism) are in line with our results of the small but significant correlation of the ETS and depression (PHQ-9). A possible explanation for both findings might be that in pain patients, depression might be associated with more severe medical symptoms, which might lower patients' expectations in some cases depending on treatment history.

Expectations are used in clinical populations as observational data, but experimental studies suggest that expectations can be changed by verbal interventions themselves [1 38]. A study of pain patients receiving acupuncture showed larger treatment effects if the patients received a briefing about expectations in advance [39]. However, the study only used observer ratings to check for the adequate manipulation of expectations by therapists. The real change of expectations from a patient's perspective was not measured, and therefore, other working mechanisms (like empathy or adherence to recommendations) might have influenced the results.

Limitations

There are some limitations to our study. First, even though collection of data via a web survey holds many advantages, there are also some downsides to this method since

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

recruitment is done in an unstructured way. Since the ETS questionnaire was also validated in a paper-based version in a clinical population this limitation is of minor importance. Second, in order to enhance the external validity of our study other clinical populations should be investigated in order to prove the validity of our scale across clinical conditions. So far the application seems feasible and valid for patients with pain. Third, there might be clinical intervention where patients might have a very vague idea about procedures and outcomes. we did not collect data from a clinical pain population *prior* to their acupuncture treatment. The patients from our sample provided information about their previous experience with acupuncture, but such an assessment could be done more comprehensively in order to capture the underlying beliefs and experiences of patients and their relevance for expectations.

Conclusion

The ETS is a short and validated measure that can be used in the German and English languages and can contribute to the understanding of patient expectations for treatment outcomes. The field of acupuncture served as example to develop the ETS, but it could be easily adapted for other treatments and clinical contexts. The ETS fills a gap by providing a strong and flexible measure that can serve as a basis for upcoming predictor analyses of treatment expectations in clinical studies.

Funding

This study was funded by the Swiss National Science Foundation (Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung), grant No. 105319_159833.

Declaration of competing interests

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

The present research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authorship and contributorship

JB was involved in the conception and execution of the scale and survey, the analysis and interpretation of data, the writing of the first draft and the revision of the manuscript. AK was involved in the analysis and interpretation of data and the writing of the manuscript. SL was involved in programming and conducting the survey, managing and analysing data, and revising the manuscript. CW was involved in the conception of the scale and survey, the interpretation of the data and the revision of the manuscript.

License for publication statement

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above.

Data sharing statement

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Date are available from the first author on request.

Acknowledgement

We would like to thank Dr. George Lewith, who made important suggestions to the phrasing of questions in our scale.

References

1. Peerdeman KJ, van Laarhoven, Antoinette IM, Keij, Sascha M, Vase, Lene, Rovers, Maroeska M, Peters, Madelon L, Evers, Andrea WM. Relieving patients' pain with expectation interventions: a meta-analysis. *Pain* 2016;157(6):1179-91.
2. Kirsch IE. How expectancies shape experience: American Psychological Association 1999.
3. Bialosky JE, Bishop, Mark D, Cleland, Joshua A. Individual expectation: an overlooked, but pertinent, factor in the treatment of individuals experiencing musculoskeletal pain. *Physical therapy* 2016;90(9):1345-55.
4. Bishop FL, Yardley, Lucy, Prescott, Philip, Cooper, Cyrus, Little, Paul, Lewith, George T. Psychological covariates of longitudinal changes in back-related disability in patients undergoing acupuncture. *The Clinical journal of pain* 2015;31(3):254.
5. Iles RA, Davidson, Megan, Taylor, Nicholas F, O'Halloran, Paul. Systematic review of the ability of recovery expectations to predict outcomes in non-chronic non-specific low back pain. *Journal of occupational rehabilitation* 2009;19(1):25-40.
6. Mondloch MV, Cole DC, Frank JW. Does how you do depend on how you think you'll do? A systematic review of the evidence for a relation between patients' recovery expectations and health outcomes. *Canadian Medical Association Journal* 2001;165(2):174-79.
7. Linde K, Witt, C. M., Streng, A., Weidenhammer, W., Wagenpfeil, S., Brinkhaus, B., Willich, S. N., Melchart, D. The impact of patient expectations on outcomes in four randomized controlled trials of acupuncture in patients with chronic pain. *Pain* 2007;128(3):264-71. doi: 10.1016/j.pain.2006.12.006
8. Foster NE, Thomas, E., Hill, J. C., Hay, E. M. The relationship between patient and practitioner expectations and preferences and clinical outcomes in a trial of exercise and acupuncture for knee osteoarthritis. *European journal of pain* 2010;14(4):402-09. doi: 10.1016/j.ejpain.2009.06.010
9. Sherman KJ, Cherkin, D. C., Ichikawa, L., Avins, A. L., Delaney, K., Barlow, W. E., Khalsa, P. S., Deyo, R. A. Treatment expectations and preferences as predictors of outcome of acupuncture for chronic back pain. *Spine* 2010;35(15):1471-7. doi: 10.1097/BRS.0b013e3181c2a8d3
10. Barth J, Schafroth, Larissa, Witt, Claudia M. Overlap and differences between patient and provider expectations for treatment outcomes: the case of acupuncture. *The Journal of Pain* 2016;17(6):685-93.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

11. Haanstra TM, Tilbury C, Kamper SJ, et al. Can optimism, pessimism, hope, treatment credibility and treatment expectancy be distinguished in patients undergoing total hip and total knee arthroplasty? *PloS one* 2015;10(7):e0133730.
12. Bowling A, Rowe, G., Lambert, N., Waddington, M., Mahtani, K. R., Kenten, C., Howe, A., Francis, S. A. The measurement of patients' expectations for health care: a review and psychometric testing of a measure of patients' expectations. *Health technology assessment* 2012;16(30):1-509. doi: 10.3310/hta16300
13. Bandura A. Self-efficacy mechanism in human agency. *Am Psychol* 1982;37(2):122.
14. Carver CS, Scheier MF, Segerstrom SC. Optimism. *Clin Psychol Rev* 2010;30(7):879-89. doi: 10.1016/j.cpr.2010.01.006 [published Online First: 2010/02/23]
15. Prady SL, Burch, J., Vanderbloemen, L., Crouch, S., MacPherson, H. Measuring expectations of benefit from treatment in acupuncture trials: A systematic review. *Complementary therapies in medicine* 2015;23(2):185-99. doi: 10.1016/j.ctim.2015.01.007
16. Laferton JA, Kube, Tobias, Salzmann, Stefan, Auer, Charlotte J, Shedden-Mora, Meike C. Patients' expectations regarding medical treatment: A critical review of concepts and their assessment. *Frontiers in Psychology* 2017;8
17. Mao JJ, Armstrong, Katrina, Farrar, John T, Bowman, Marjorie A. Acupuncture expectancy scale: development and preliminary validation in China. *Explore: The Journal of Science and Healing* 2007;3(4):372-77.
18. Jones SM, Lange, Jane, Turner, Judith, Cherkin, Dan, Ritenbaugh, Cheryl, Hsu, Clarissa, Berthoud, Heidi, Sherman, Karen. Development and validation of the EXPECT questionnaire: assessing patient expectations of outcomes of complementary and alternative medicine treatments for chronic pain. *The Journal of Alternative and Complementary Medicine* 2016;22(11):936-46.
19. Constantino MJ, Arnkoff DB, Glass CR, et al. Expectations. *Journal of clinical psychology* 2011;67(2):184-92. doi: 10.1002/jclp.20754 [published Online First: 2010/12/04]
20. Faller H, Vogel H, Bosch B. Erwartungen von Rehabilitanden hinsichtlich der Methoden und Ergebnisse ihrer Rehabilitation-Eine kontrollierte Studie mit Rückenschmerz-und onkologischen Patienten. *Die Rehabilitation* 2000;39(04):205-14.
21. Kravitz RL. Measuring patients' expectations and requests. *Annals of internal medicine* 2001;134(9 Pt 2):881-8.
22. Schulte D. Messung der Therapieerwartung und Therapieevaluation von Patienten (PATHEV). *Zeitschrift für Klinische Psychologie und Psychotherapie* 2005;34(3):176-87.
23. Eibner F, Barth J, Helmes A, et al. Variations in subjective breast cancer risk estimations when using different measurements for assessing breast cancer risk perception. *Health, Risk & Society* 2006;8(2):197-210.
24. Glaesmer H, Hoyer, Jürgen, Klotsche, Jens, Herzberg, Philipp Yorck. Die deutsche version des Life-Orientation-Tests (LOT-R) zum dispositionellen Optimismus und Pessimismus. *Zeitschrift für Gesundheitspsychologie* 2008;16(1):26-31.
25. Löwe B, Unützer, Jürgen, Callahan, Christopher M, Perkins, Anthony J, Kroenke, Kurt. Monitoring depression treatment outcomes with the patient health questionnaire-9. *Medical care* 2004;42(12):1194-201.
26. Ell K, Unützer, Jurgen, Aranda, Maria, Sanchez, Kathleen, Lee, Pey-Jiuan. Routine PHQ-9 depression screening in home health care: Depression prevalence, clinical and treatment characteristics, and screening implementation. *Home health care services quarterly* 2006;24(4):1-19.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

27. Wagnild GM, Young, Heather M. Development and psychometric evaluation of the Resilience Scale. *Journal of Nursing Measurement* 1993
28. Schumacher J, Leppert, Karena, Gunzelmann, Thomas, Strauß, Bernhard, Brähler, Elmar. Die Resilienzskala–Ein Fragebogen zur Erfassung der psychischen Widerstandsfähigkeit als Personmerkmal. *Zeitschrift für Klinische Psychologie, Psychiatrie und Psychotherapie* 2005;53:16-39.
29. Schutzler L, Witt, Claudia M. Body-Efficacy Expectation: Assessment of Beliefs concerning Bodily Coping Capabilities with a Five-Item Scale. *Evidence-based complementary and alternative medicine : eCAM* 2013;2013:152727. doi: 10.1155/2013/152727
30. Horne R, Faasse, Kate, Cooper, Vanessa, Diefenbach, Michael A, Leventhal, Howard, Leventhal, Elaine, Petrie, Keith J. The perceived sensitivity to medicines (PSM) scale: an evaluation of validity and reliability. *British journal of health psychology* 2013;18(1):18-30.
31. Borkenau P, Ostendorf, Fritz. NEO-Fünf-Faktoren Inventar:(NEO-FFI); nach Costa und McCrae: Göttingen: Hogrefe 1993.
32. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. *Psychological bulletin* 1980;88(3):588.
33. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull* 1990;107(2):238-46. [published Online First: 1990/03/01]
34. Jöreskog KG, Sörbom D. LISREL 8: Structural equation modeling with the SIMPLIS command language: Scientific Software International 1993.
35. Kline RB. Principles and practice of structural equation modeling: Guilford publications 2015.
36. Marsh HW, Balla JR, McDonald RP. Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological bulletin* 1988;103(3):391.
37. Browne MW, Cudeck R. Alternative ways of assessing model fit. In. KA, Bollen, JS Long,(Eds.), Testing structural equation models (pp. 136-162): Newbury Park, CA: Sage, 1993.
38. Colagiuri B, Smith, Caroline A. A systematic review of the effect of expectancy on treatment responses to acupuncture. *Evidence-based complementary and alternative medicine : eCAM* 2012;2012:857804. doi: 10.1155/2012/857804
39. Suarez-Almazor ME, Looney, Carol, Liu, Yanfang, Cox, Vanessa, Pietz, Kenneth, Marcus, Donald M, Street, Richard L. A randomized controlled trial of acupuncture for osteoarthritis of the knee: Effects of patient-provider communication. *Arthritis care & research* 2010;62(9):1229-36.
40. Derogatis LR. SCL-90-R: Symptom Checklist-90-R: administration, scoring, and procedures manual: NCS Pearson 1996.

Tables and Figures

Table 1. Description of the 5 final items in English and German

Table 2. Sample characteristics of subjects with current pain (N = 102)

Table 3. Scale and factor analysis of the 5 item ETS for subjects with current pain (N = 102)

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 4. Convergent and divergent validity of the 5 item version of the ETS for subjects with current pain (N = 102)

Figure 1. Confirmatory factor analysis with factor loadings (N = 439) of the final 5 item ETS version in patients with chronic low back pain

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 1. Description of the 5 final items in English and German

Description: The overarching construct of ETS is treatment-related outcome expectation of patients. Outcomes of interest for patients are coping (item 1), vitality (item 3) and physical health (item 4). Two items (items 2 and 5) capture the reduction of patient complaints, either an absence or a considerable reduction thereof. The scale can be adapted to specific treatments of interest. Instead of the term “treatment” a specific term can be used (i.e. acupuncture, counseling, physiotherapy). We used the term complaints [german Beschwerden] as already used by other authors like in the symptom checklist (SCL-90-R) [40]. Therefore, any kind of complaints can be captured and allow the use of this instrument across patient groups and settings. The term complaints can be adapted if needed to the clinical situation if one specific target symptom is predominant (i.e. pain, depression, nausea). The scale should be used in a clinical setting by patients with sufficient German or English language skills.

For peer review only

English version

There are several statements below that capture your expectations about the [acupuncture] treatment. Please indicate to what extent these statements apply to you personally. **There are no right or wrong answers. We are only interested in your current personal thoughts.**

*Please select **one** option.*

1	2	3	4
partially disagree	partially agree	agree	definitely agree

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Expectation for Treatment Scale (ETS)

1	I expect the treatment [acupuncture] will help me to cope with my complaints.	1	2	3	4
2	I expect the treatment [acupuncture] will make my complaints disappear.	1	2	3	4
3	I expect the treatment [acupuncture] will improve my energy.	1	2	3	4
4	I expect the treatment [acupuncture] will improve my physical performance.	1	2	3	4
5	I expect that after the treatment [acupuncture] my complaints will be considerably better.	1	2	3	4

*We used the term acupuncture in our study. The general version of ETS uses the term treatment.

567

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

German version

Im Folgenden finden Sie einige Aussagen darüber, was man von einer Behandlung [Akupunktur] erwarten kann. Geben Sie bitte an, inwieweit die einzelnen Aussagen für Sie persönlich zutreffen. **Es gibt keine richtigen und falschen Antworten. Uns interessiert nur, was Sie persönlich denken.**

*Bitte wählen Sie **eine** Antwort aus.*

1	2	3	4
trifft eher nicht zu	trifft eher zu	trifft sehr zu	trifft völlig zu

1	Ich erwarte, dass ich durch die Behandlung [Akupunktur] mit meinen Beschwerden besser umgehen kann.	1	2	3	4	
2	Ich erwarte, dass meine Beschwerden durch die Behandlung [Akupunktur] verschwinden.	1	2	3	4	
3	Ich erwarte, dass sich meine Energie durch die Behandlung [Akupunktur] verbessert.	1	2	3	4	
4	Ich erwarte durch die Behandlung [Akupunktur] eine verbesserte körperliche Leistungsfähigkeit.	1	2	3	4	
5	Ich erwarte, dass sich nach der Behandlung [Akupunktur] meine Beschwerden deutlich verbessern.	1	2	3	4	

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

*Wir haben den Begriff Akupunktur in unserer Studie verwendet. Die allgemeine Version des ETS verwendet den Begriff Behandlung.

Scoring of the questionnaire

Add up the score of the individual items and use the number of filled in questions as denominator.

568

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 2. Sample characteristics of subjects with current pain (N = 102)

	Frequencies // Mean (SD)	Percentage
Age	49.16 (14.04)	
Gender		
Female	76	75%
Male	26	25%
Education		
Vocational school	39	38%
Higher school certificate or Higher vocational training	29	28%
University	34	33%
Employment status		
Working	72	71%
Retired	10	10%
Other	20	20%
Country of residence		
Switzerland	68	67%
Germany	34	33%
Acupuncture experience		
Current	13	13%
Earlier	51	50%

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

None	38	37%
Success of acupuncture [range 1-10]*	6.5 (2.80)	
Reason for treatment**		
Pain	55	86%
Mental health	13	20%
Other	24	38%
General Health***	3.21 (.92)	
Pain characteristics		
Intensity [range 0-10]****	6.5 (1.92)	
Physician consulted*****	68	67%
Any treatment*****	66	65%
Days of restriction due to pain*****	46.76 (63.74)	
Location******		
Back	58	57%
Neck	44	43%
Knee	21	21%
Head	20	20%
Hip	16	16%
Other	33	32%

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

*Range from no success (1) to very successful (10) treatment outcome of the last acupuncture treatment

**Multiple answers were possible

***Range from 1 = excellent to 5 = bad

****0 indicating no pain and 10 indicating maximum pain

******In the last 6 month period

******Multiple answers were possible

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 3. Scale and factor analysis of the 5 item ETS for subjects with current pain (N = 102)

Item	Mean (SD)	Item-total correlation corrected	Cronbach`s alpha if item deleted	Factor loading	Retest ICC [95%-CI]
Item 3	2.38 (1.03)	.727	.887	.826	.697 [.517,.818]
Item 4	1.97 (.96)	.736	.883	.834	.724 [.556,.835]
Item 5	2.41 (1.0)	.776	.875	.865	.840 [.732,.907]
Item 8	2.33 (.83)	.817	.869	.891	.749 [.592,.851]
Item 9	2.52 (.90)	.737	.883	.834	.836 [.725,.904]
			Cronbach's α	% of variance	
ETS 5			.901	72.325	.856 [.757,.917]

Item 3: I expect the treatment [acupuncture] will help me to cope with my complaints.

Item 4: I expect the treatment [acupuncture] will make my complaints disappear.

Item 5: I expect the treatment [acupuncture] will improve my energy.

Item 8: I expect the treatment [acupuncture] will improve my physical performance.

Item 9: I expect that after the treatment [acupuncture] my complaints will be considerably better.

For peer review only

Table 4. Convergent and divergent validity of the 5 item version of the ETS for subjects with current pain (N = 102). Pearson Correlation, Significance level (2-tailed), N of patients.

	ETS	LOT-R Optimism	LOT-R Pessimism	PHQ-9	RS-11	BEE	PSM	NEO-FFI Neuroticism	NEO-FFI Openness to experience
ETS	1 102								
LOT-R Optimism	-.066 .517 99	1 99							
LOT-R Pessimism	.204* .043 99	-.247* .014 99	1 99						
PHQ-9	.233* .020 100	-.567** .000 99	.312** .002 99	1 100					
RS-11	-.073 .474 99	.546** .000 99	-.460** .000 99	-.572** .000 99	1 99				

For peer review only

Expectation for Treatment Scale (ETS)

BEE	.032	.259**	-.074	-.176	.407**	1			
	.754	.010	.468	.081	.000				
	99	99	99	99	99	99			
PSM	.344**	-.078	.201*	.306**	-.049	.011	1		
	.000	.445	.046	.002	.631	.912			
	100	99	99	100	99	99	100		
NEO-FFI	.104	-.578**	.467**	.631**	-.682**	-.284**	.238*	1	
Neuroticism	.316	.000	.000	.000	.000	.005	.020		
	95	95	95	95	95	95	95	95	
NEO-FFI	-.072	.297**	-.167	-.090	.302**	.122	.132	-.130	1
Openness to	.492	.004	109	.386	.003	.241	.205	.211	
experience	94	94	94	94	94	94	94	94	94
Cronbach Alpha	.901	.727	.608	.900	.900	.757	.940	.854	.726

**Correlation is significant at the 0.01 level (2-tailed)

*Correlation is significant at the 0.05 level (2-tailed)

ETS = Expectation for Treatment Scale; LOT-R = Life-Orientation Test; PHQ-9 = Patients Health Questionnaire-9; RS-11 = Die Resilienzskala - Ein Fragebogen zur Erfassung der psychischen Widerstandsfähigeit als Personenmerkmal; BEE = Body-Efficacy Expectation; PSM = The Perceived Sensitivity to Medicines Scale; NEO-FFI = NEO Five-Factor Inventory

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Figure 1. Confirmatory factor analysis with factor loadings (N = 439) of the final 5 item ETS version in patients with chronic low back pain

For peer review only

Expectation for Treatment Scale (ETS)

617 Appendix (Barth et al.)

APPENDIX Table 1: Sample characteristics in scales in subjects with current pain (N = 102)

	Median Mean (SD)
ETS (9 items) (range 9-36)	22 22.63 (.64)
ETS (5 items) (range 5-20)	11 11.62 (4.01)
AES (range 4-20)	12.0 11.82 (4.10)
PHQ-9 (range 0-27)	7 8.46 (6.06)
RS-11 (range 11-77)	60 58.36 (11.13)
BEE (range 5-20)	14 13.71 (2.90)

For peer review only

PSM (range 5-25)	11 12.77 (6.13)
LOT-R Optimism (range 0-12)	8 7.93 (2.77)
LOT-R Pessimism (range 0-12)	8 7.93 (2.38)
NEO-FFI Neuroticism (range 0-48)	21 21.08 (8.19)
NEO-FFI Openness to experience (range 0-44)	30 29.84 (5.77)

Expectation for Treatment Scale; AES= Acupuncture Expectancy Scale; LOT-R= Life-Orientation Test; PHQ-9= Patients Health Questionnaire-9; RS-11= Die Resilienzskala - Ein Fragebogen zur Erfassung der psychischen Widerstandsfähigeit als Personenmerkmal; BEE= Body-Efficacy Expectation; PSM= The Perceived Sensitivity to Medicines Scale; NEO-FFI= NEO Five-Factor Inventory

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Expectation for Treatment Scale (ETS)

Appendix Table 2. Between item correlations for sample with current pain (N = 102). Pearson Correlation and significance (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Expectation for Treatment Scale (ETS)

Item 1	1								639
Item 2	.840**	1							640
	.000								641
Item 3	.636**	.605**	1						642
	.000	.000							643
Item 4	.568**	.627**	.580**	1					644
	.000	.000	.000						645
Item 5	.533**	.611**	.664**	.637**	1				646
	.000	.000	.000	.000					647
Item 6	.540**	.556**	.442**	.437**	.505**	1			648
	.000	.000	.000	.000	.000				649
Item 7	.692**	.679**	.541**	.575**	.509**	.521**	1		650
	.000	.000	.000	.000	.000	.000			651
Item 8	.528**	.571**	.673**	.664**	.762**	.470**	.631**	1	652
	.000	.000	.000	.000	.000	.000	.000		653
Item 9	.733**	.754**	.605**	.673**	.603**	.592**	.701**	.673**	1
	.000	.000	.000	.000	.000	.000	.000	.000	654

655

Item 1: I expect the treatment [acupuncture] will clearly

656 reduce my complaints.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Expectation for Treatment Scale (ETS)

Item 2: I expect my complaints will be much better because of the treatment [acupuncture].

Item 3: I expect the treatment [acupuncture] will help me to cope with my complaints.

Item 4: I expect the treatment [acupuncture] will make my complaints disappear.

Item 5: I expect the treatment [acupuncture] will improve my energy.

Item 6: The treatment [acupuncture] is in general effective.

Item 7: I myself have positive expectations about the treatment [acupuncture].

Item 8: I expect the treatment [acupuncture] will improve my physical performance.

Item 9: I expect that after the treatment [acupuncture] my complaints will be considerably better.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Appendix Table 3. Scale and factor analysis of the 9 item ETS for subjects with current pain (N = 102)

Item	Mean (SD)	Item-total correlation corrected	Cronbach`s alpha if item deleted	Factor loading	Retest ICC [95%-CI]
Item 1	2.71 (1.0)	.782	.920	.839	.684 [.499,.810]
Item 2	2.66 (.93)	.816	.918	.863	.725[.557,.836]
Item 3**	2.38 (1.03)	.736	.924	.790	.697 [.517,.818]
Item 4**	1.97 (.96)	.735	.923	.793	.724 [.556,.835]
Item 5**	2.41 (1.0)	.745	.923	.799	.840 [.732,.907]
Item 6	2.97 (.59)	.612	.931	.685	.649 [.450,.787]
Item 7	2.68 (.65)	.744	.924	.806	.658 [.462,.793]
Item 8**	2.33 (.83)	.772	.921	.821	.749 [.592,.851]
Item 9**	2.52 (.90)	.826	.917	.874	.836 [.725,.904]
			Cronbach's α	% of variance	
ETS 9			.930	65.512	.879 [.794,.930]

**marked items are included in the final version of its scale

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Expectation for Treatment Scale (ETS)

Item 1: I expect the treatment [acupuncture] will clearly reduce my complaints.

Item 2: I expect my complaints will be much better because of the treatment [acupuncture].

Item 3: I expect the treatment [acupuncture] will help me to cope with my complaints.

Item 4: I expect the treatment [acupuncture] will make my complaints disappear.

Item 5: I expect the treatment [acupuncture] will improve my energy.

Item 6: The treatment [acupuncture] is in general effective.

Item 7: I myself have positive expectations about the treatment [acupuncture].

Item 8: I expect the treatment [acupuncture] will improve my physical performance.

Item 9: I expect that after the treatment [acupuncture] my complaints will be considerably better.

For Peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

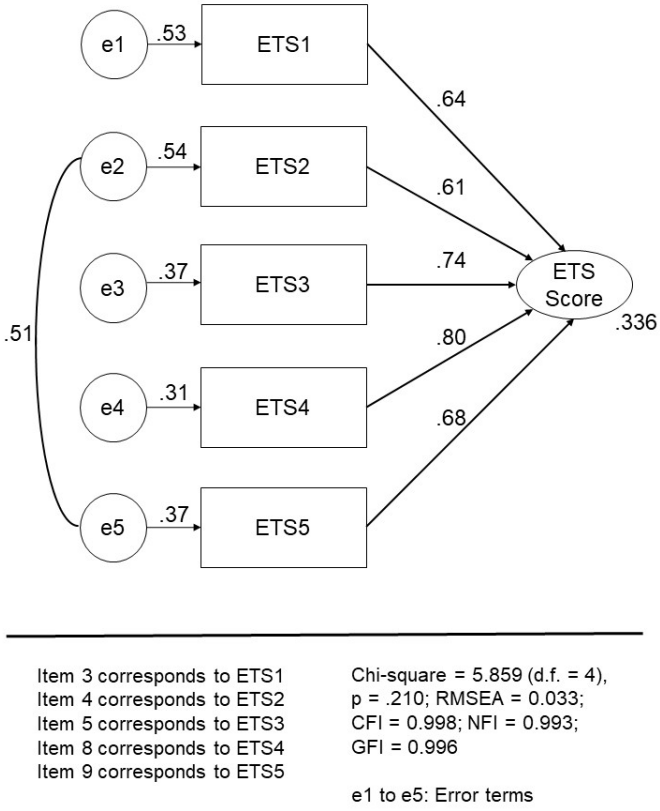


Figure 1. Confirmatory factor analysis with factor loadings (N = 439) of the final 5 item ETS version in patients with chronic low back pain

190x338mm (96 x 96 DPI)

BMJ Open

Assessment of patients' expectations: development and validation of the Expectation for Treatment Scale (ETS)

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-026712.R1
Article Type:	Research
Date Submitted by the Author:	10-Feb-2019
Complete List of Authors:	Barth, Jürgen; University Hospital Zurich, Institute for Complementary and Integrative Medicine Kern, Alexandra; Institute for Complementary and Integrative Medicine, University Hospital Zurich and University Zurich, Zurich Switzerland Lüthi, Sebastian; Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich, Zurich, Switzerland Witt, Claudia M.; Institute for Complemtary and Integrative Medicine, University Hospital Zurich and University of Zurich, Zurich, Switzerland; Institute for Social Medicine, Epidemiology and Health Economics, Charité
Primary Subject Heading:	Complementary medicine
Secondary Subject Heading:	Patient-centred medicine, Research methods
Keywords:	expectations, placebo, assessment, pain, acupuncture
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
ETS_STROBE_checklist_cross-sectional_20190118.doc	

SCHOLARONE™
Manuscripts

Assessment of patients' expectations: development and validation of the Expectation for Treatment Scale (ETS)

Running title: Expectation for Treatment Scale (ETS)

Submission to British Medical Journal

Jürgen Barth*[1], Alexandra Kern [1], Sebastian Lüthi [1], Claudia M. Witt [1, 2, 3]

[1] Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich, Sonneggstrasse 6, 8006 Zurich, Switzerland Dr. Jürgen Barth Research Director, M.A. Alexandra Kern PhD student, M.A. Sebastian Lüthi Master student, Professor Dr. Claudia M. Witt, Chair and Director

[2] Institute for Social Medicine, Epidemiology and Health Economics, Charité – Universitätsmedizin Berlin, Luisenstrasse 57, Berlin, Germany, Professor Dr. Claudia M. Witt, research group leader

[3] Center for Integrative Medicine, University of Maryland, School of Medicine, 520 W Lombard Street, Baltimore, MD, USA Professor Dr. Claudia M. Witt, senior scientist

*Correspondence:

Dr. Jürgen Barth, University Hospital Zurich and University of Zurich, Switzerland

Phone: +41 (0)44 255 4896

mail@juergen-barth.de

www.iki.usz.ch

Word count: 4526

For peer review only

Abstract

Objective: The development of a short self-report instrument for the assessment of expectations (Expectation for Treatment Scale, ETS), using acupuncture as a case example.

Design: A cross-sectional assessment with retest after one week.

Setting: A web-based survey with patients suffering from pain.

Methods: In a three-step approach, we reduced the initially collected number of items from 17 to 9 and to 5, including expectations about coping ability, vitality, physical health and reduction of patient complaints. Items were selected according to internal consistency (Cronbach's alpha), convergent and divergent validity with related constructs (optimism, pessimism, resilience, perceived sensitivity to medicines, depression, and others), one-week retest reliability (intraclass correlation coefficient, ICC) and exploratory and confirmatory factor analysis.

Results: A total of 102 pain patients were included, and 54 of these patients completed the retest assessment. The final version of the ETS consisted of 5 items and had an excellent Cronbach's alpha (.90), with 72.33% variance on one single factor. Depression, pessimism and perceived sensitivity to medicines showed positive correlations with our expectation measure (r = 0.23, r = 0.20, r = 0.34, respectively), the correlation between the ETS and optimism was low (r = -0.07), and no correlation between the ETS and resilience was found (r = -0.07). Convergent validity was confirmed with a high correlation (r > .90) between ETS and a treatment-specific measure of expectations. The retest ICC was .86, which showed high stability over one week. A confirmatory factor analysis (N = 439) with data from patients with low back pain confirmed the single-factor structure of the instrument.

For peer review only

Conclusion: The ETS showed strong psychometric properties and covered a distinct construct. As a next step, the ETS might be implemented in different clinical conditions and settings to investigate psychometrics and its predictive power for treatment outcomes.

Keywords: Expectations, placebo, assessment, pain, optimism, acupuncture

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Strengths and limitations of this study

- It is the first project to systematically develop a general measure to assess patients' expectations across medical conditions and treatments with the involvement of patients, earlier scales and empirical data.
- The Expectation for Treatment Scale (ETS) is a short and reliable measure that captures outcome-related expectations of patients.
- The association between the ETS and related constructs was explored.
- The development of the ETS was done via an online survey with pain patients, and the initial findings were confirmed using confirmatory factor analysis in a patient sample with low back pain.
- The transfer of the ETS in different clinical conditions and settings should be pre-tested to explore whether patients can evaluate the respective treatment at this level.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Introduction

Patients' expectations alter responses to interventions in placebo research and in clinical studies. In experimental research, the instruction to receive an active intervention accompanies large effects on pain reduction (effect size = 0.75)[1]. Expectations can be considered a major driver of changes in symptoms and other health-related outcomes[2]. Clinical studies have revealed that patients' positive expectations are related to reduced pain after a medical treatment[3–5], and this phenomenon has also been observed for other medical conditions[6]. For acupuncture, this association was investigated in several studies with heterogeneous findings: Patients with chronic pain had higher odds (odds ratio 2.11 [95% CI 1.32-3.34]) of benefitting from acupuncture treatment if they expected a better treatment response at baseline (controlled for other clinical baseline characteristics)[7]. However, there are also studies that found no influence of expectations on the outcome[8 9].

These differences in the association of expectations and outcomes might be partly explained by the fact that each study used a newly invented measure, and differences between measures might hide or exaggerate associations between expectations and intervention outcomes. Additionally, the match between expectations of patients and treatment providers might be relevant for the success of a specific treatment[10].

Expectation is a well-known and oft-used term. A clear definition and a sharp distinction from associated constructs is important for the development of a measure[11]. In the context of medical treatments, the term "expectations" describes cognitions about treatment-related health outcomes in the future after a specific intervention[12 13]. Patients can consider a treatment more or less beneficial for their complaints or disease at a specific time-point (i.e., outcome expectations)[13]. Role expectations also capture the role of a

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

patient and the therapist during the treatment. In other words, a patient might consider himself rather inactive during treatment in defining treatment goals and expects an active therapist to achieve a good treatment outcome. However, our purpose was to develop a scale on ‘‘patient expectations’’ that covers treatment-related outcome expectations. Bowling et al. (2012) provide an insightful summary about the theoretical underpinning of expectations[14]. Related constructs such as optimism, self-efficacy, and hope share some facets with expectations but differ on the level of the construct[15 16]. Optimism can be viewed as a trait characteristic of a person with high stability over time and situations. Optimism is defined as “the extent to which people hold generalized factorable expectancies for their future” (Carver et al, 2010, p. 879)[16]. Self-efficacy is also a construct at a general level (i.e., “Perceived self-represents an optimistic sense of personal competence […]”; Scholz et al., 2002; p. 342)[17]. If self-efficacy is related to a specific behaviour or problem, it captures the strength of a belief to cope in a situation successfully (for example Pain Self-Efficacy Questionnaire; PSEQ)[18]. Hope should also be considered conceptually different: “Expectations and hopes are very different concepts. Hopes tend to be based more upon emotions or wishes, things that individuals want reality to be, whereas expectations tend to rely more heavily upon rational thought and logical reasoning” Woolhead et al., 2003 p. 1656)[19]. Related constructs are sometimes included in expectation measures, whereas in other cases, only cognitions about treatment outcome relationships are included[9 20 21]. Following the definition by Bowling et al. (2012), we therefore designed our measure to assess expectations related to a clinical intervention with a clinically relevant outcome from a patient’s perspective.

Earlier findings about the expectation outcome association in clinical studies have been limited by the diversity of measures. Several authors claimed diversity in covered

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

concepts, time-point of assessment and problems to evaluate the validity of the measures[22–24]. A strong measure is a prerequisite to accurately predict treatment responses based on pretreatment expectations. A closer investigation of the results from a systematic review about acupuncture expectation measures by Prady and colleagues[20] showed that of ten trials, only five provided their exact item wording for measuring expectations. Because many of the assessment instruments are not publicly available, it is difficult to replicate the studies. Three of the five mentioned studies used only one item to assess expectations. There is no reporting at all of Cronbach's alpha in the two remaining studies.

For further research in the field of expectations, a strong measure with high acceptance across clinical fields would be needed for several reasons. First, a reliable measure with high internal consistency at a specific time-point is a prerequisite to use expectations as a robust predictor. Second, ceiling effects were a common problem in the measurement of expectations, because patients who are seeking help from a specific treatment often expect large benefits; otherwise, they would not be attracted by this treatment. This problem was apparent in an established expectation measure for acupuncture treatment that served as a benchmark measure for our scale (Acupuncture Expectancy Scale (AES))[25]. Ceiling effects are particularly problematic because the predictive power of such skewed variables is low. In the case of expectations, many research questions address the prediction of treatment outcomes; therefore, a measure with sufficient variation between patients is needed. Some authors have used the term "realistically expect" to capture expectations in a recent study with a comprehensive 10-item assessment of expectations, hope and beliefs (EXPECT)[26] with limited internal consistency. Third, a measurement of expectation should be stable in a reasonable timeframe: The EXPECT scale showed only a moderate retest reliability, with an intra-class correlation of

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

ICC = 0.75. Most importantly, the association of EXPECT with the established AES, a benchmark measure for the field of acupuncture, was unexpectedly low (r = .54).

This study aimed to develop a measure to reliably capture patient expectations with a short scale. The newly developed scale (Expectation for Treatment Scale, ETS) ideally has a strong association with available measures of expectations, moderate associations with related constructs (e.g., optimism and pessimism) and no association with personality (e.g., neuroticism). Furthermore, it should be constructed such that it can be adapted to other treatments and used universally in different clinical fields independent of patients' complaints.

Methods

To address the problem of the abovementioned ceiling effects, we developed the items of our scale (ETS) in a pilot study with pain patients in our outpatient clinic to measure expected benefits before treatment from the patient's perspective. Based on existing questionnaires on patient expectations[9 25 27–30], we created a list of 17 items covering different facets of expectations that fit with our aim to develop a measure for treatment-related outcome expectations.

Patient and Public Involvement

Twenty patients completed the questionnaires, provided sociodemographic information and were asked for written comments about the accessibility of the questions. In addition, two patients were interviewed by a qualitative researcher. Two health professionals (one

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

acupuncturist and a doctor assistant) also verbally commented about the appropriateness of the questions. Based on these data, we selected items with low skewness and a large range of responses (i.e., variation). High correlations between items and the findings from the qualitative feedback were also considered.

Participants

For the main validation, study subjects of the convenience sample were recruited using different methods. Three regional patient organizations distributed the information, and we used several email distribution lists (including the University of Zurich, Switzerland and the Charité University Hospital Berlin, Germany). Furthermore, patients from the previous year seeking treatment for a musculoskeletal condition at the Institute for Complementary and Integrative Medicine of the University Hospital Zurich were contacted by email. All participants were required to be at least age 18 and, according to their own evaluation, have sufficient knowledge to understand German. The included pain patients had to suffer from at least some pain at the day of the assessment (> 0 on a numeric rating scale from 0 (no pain) to 10 (worst pain)).

The study link was accessed by 522 subjects, and 244 started the survey. In total, 142 subjects were excluded due to insufficient pain (N = 113), missing data in the nine-item version of the ETS (N = 14) and more than 20% missing responses across all items (N = 13). Data from patients suffering from pain (N = 102) with retest data from 54 patients were available.

Procedures

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Subjects who decided to participate activated an access link to the online survey provided via email. The survey was hosted by https://www.soscisurvey.de/, and all data were collected electronically and anonymously. Participants were able to complete the survey within 15 minutes. All questions were created in a forced choice format, meaning that subjects were not able to continue to the next set of questions without completing the previous set. Upon starting the survey, a short welcome message, followed by an introduction to the study, was provided. Subjects who agreed to be contacted again for the retest assessment provided their email address at the end of the survey. The retest sample received a follow-up ETS one week after completion of the first survey. The study was granted ethics approval by the local ethics committee (Kantonale Ethikkommission Zürich, No. 48-2015).

Measures

The survey included demographic information (age, gender, education, employment status, and country of residence), questions about any previous experience with acupuncture (current or earlier), the degree to which the acupuncture treatment was successful (numeric rating scale from 1 to 10, indicating no success to much success) and the reason for the treatment (pain, mental health or non-specific). Patients who suffered from pain (binary variable with yes vs. no option) were asked to give information about their pain. We assessed average pain intensity (numeric rating scale from 1 to 10, indicating no pain to worse pain), whether a physician was consulted, whether any other type of therapy was used and the number of days of restriction due to pain. Subsequently, subjects completed the 9-item ETS. Below, we provide a detailed description of all measures in this study.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Expectation for Treatment Scale

The first version of the ETS consisted of nine items (e.g., "I expect the treatment [acupuncture] will help me to cope with my complaints."). Each item was to be rated on a four-point scale ranging from 1 to 4 (partially disagree, partially agree, agree, definitely agree). We decided to use a four-point scale instead of a five-point scale for two reasons. First, the chosen format forces the patient to provide an answer with a direction (higher or lower than the middle answer choice) instead of opting for the middle category, which often occurs in situations of insecurity (such as the present estimation of future events)[31]. Second, the lowest answer on the scale of "definitely disagree" can be considered unexpected for patients motivated to undergo an acupuncture treatment. After the statistical analyses, the nine-item version was reduced to a final 5-item version of the ETS. A detailed description of these five items is provided in the appendix Table 1. These five items were translated into English by two bilingual researchers and translated back into German by two other bilingual researchers. The wording was improved based on feedback from Dr. George Lewith. The final English version is presented in the appendix Table 1.

Optimism and pessimism

Both concepts were assessed using the German version of the Life Orientation Test – Revised[32]. The questionnaire consists of six self-report items (plus four filler items), each rated on a five-point Likert scale ranging from 0 (strongly disagree) to 4 (strongly agree). The data were separated into optimism and pessimism scores, as recommended by Glaesmer et al. (2008). Each score can range from 0 to 12, with higher values indicating

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

either higher optimism or pessimism. In the present study, the Cronbach's alpha was acceptable for optimism (0.72) and questionable for pessimism (0.60).

Depression

The Patient Health Questionnaire (PHQ-9) is widely used as a standard instrument for diagnosing depression in primary care and is considered well validated[33]. The PHQ-9 consists of nine questions operating according to the modified Diagnostic and Statistical Manual, Fourth Edition criteria. Patients indicated their answers on a scale ranging from 0 (not at all) to 3 (almost every day)[34]. A higher score indicates a higher severity of depressive symptoms (ranging between 0 and 27). In the present study, the internal consistency of the scale was excellent (Cronbach's alpha = 0.90).

Resilience

As a protective personality factor, resilience was appraised using the Resilience Scale[35]. In the current study, we used the German Version Resilienzskala (RS-11)[36]. The instrument consists of eleven questions, with seven response alternatives ranging from 1 (disagree) to 7 (fully agree). Higher scores indicate higher resilience. In the present study, the internal consistency was excellent (Cronbach's alpha = 0.90).

Body-Efficacy Expectation

To measure beliefs concerning bodily coping capabilities, the Body-Efficacy Expectation (BEE) questionnaire[37] was used. The instrument consists of five items using a four-point Likert scale ranging from 1 (not true at all) to 4 (exactly true). A higher score indicates stronger beliefs in one's bodily coping capabilities. In the present study, the Cronbach's alpha was 0.75, indicating acceptable internal consistency.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Perceived Sensitivity to Medicines

To assess the individual's sensitivity to medicines, we used the Perceived Sensitivity to Medicines scale (PSM)[38]. The PSM is considered a reliable and valid measure composed of five self-report questions to assess perceived sensitivity to the potential adverse effects of medicines. Responses are scored on a five-point Likert scale, and patients' item scores are summed to provide a total score ranging between 5 and 25. Higher scores point towards a high perceived sensitivity to the potential adverse effects of medicines. In the current study, the scale showed excellent internal consistency (Cronbach's alpha = 0.94).

Neuroticism and openness to experience

To assess the distinctiveness of neuroticism and openness to experience among patients, we used the corresponding subscales of the NEO Five-Factor Inventory (NEO-FFI)[39]. To that end, 23 questions were extracted from the 60-item NEO-FFI. The instrument utilizes a five-point Likert response format, from 1 (strongly refuse) to 5 (strongly agree). A higher score indicates a higher value for neuroticism and openness to experience. In the present study, the Cronbach's alpha was 0.85 for neuroticism and 0.72 for openness to experience, which indicates good to acceptable internal consistency.

Acupuncture Expectancy Scale

Mao and colleagues developed the AES to measure patients' expected response from acupuncture[25]. The scale was developed for only one clinical intervention (i.e., acupuncture). The instrument consists of four items. The answers are given using a five-point Likert scale ranging from 1 (not at all agree) to 5 (completely agree). A higher score points towards higher expectancies. In the present study, the internal consistency of the scale was considered good (Cronbach's alpha = 0.88).

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Statistical Analyses

All data analyses were executed using SPSS (version 22, SPSS Inc., Chicago, Illinois, USA). As a first step, we used the nine items of the first ETS version to explore homogeneity and diversity between items. We conducted a descriptive analysis of the data to detect possible floor or ceiling effects and assess the distribution of the data. Internal consistency was examined through reliability analyses (Cronbach's alpha), with the corrected item-total correlation and the Cronbach's alpha if the item is deleted. Retest reliability for the nine items was assessed by the intraclass correlation coefficient (ICC). We set for each item a minimal acceptable ICC of .60 and for the total score an ICC of .80. Three criteria guided the decision to keep an item for the final version: 1) it should have a high corrected item-total correlation and no low internal consistency if deleted, 2) the item should not overlap too strongly in content with another item that might be included, and 3) the item should not contribute to ceiling effects in the final version of the scale, which means items with lower values were preferred.

In a second step, we used the reduced version of the scale to generate a total sum score of the five items (ranging from 5 to 20). To examine divergent and convergent validity, correlations between the ETS sum score and the other measures were calculated. We assumed a very high correlation between the ETS and the most strongly related construct ($r > .70$; AES), moderate correlations with strongly related constructs ($r > .30$; LOT-R optimism, inverse with LOT-R pessimism), small correlations with less related constructs ($r > .20$; PHQ-9; RS-11, BEE) and no correlation with unrelated constructs (PSM, NEO-FFI neuroticism, NEO-FFI openness to experience).

To test our assumption of one general factor, an explorative factor analysis using a varimax rotation, an Eigenvalue of more than one and the Scree test was used with the 5

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

ETS items to determine the number of underlying factors. In addition, we conducted a confirmatory factor analysis with an independent sample of chronic low back pain patients, which had been included in a randomized controlled trial (registration number DRKS-ID: DRKS00010191). These patients completed the ETS. The data were used to test the single-factor structure. The best model used correlated error terms between item 2 and item 5 and is shown in Figure 1. Multiple goodness of fit tests[40] were used to evaluate the model, including the comparative fit index (CFI;[41]), the normed fit index (NFI;[40]), the goodness of fit index (GFI;[42]) and the root mean squared error of approximation (RMSEA). A CFI greater than 0.90 indicates a good fit to the data[43]. An NFI and GFI greater than 0.90 indicates a good fit to the data[44]. A RMSEA with values of less than 0.08 indicates a good fit to the data[45], whereas values greater than 0.10 suggest strongly that the model fit is unsatisfactory. The Chi-square goodness of fit test can be considered a general test for the acceptability of the model; a statistically significant χ^2 indicates that a significant proportion of variance remains unexplained by the model[40]. Confirmatory factor analysis was conducted with the Analysis of Moment Structures (AMOS) software version 25 which is part of the SPSS package.

Results

Sample characteristics

Three-quarters of the patients were female, one-third had a high school degree or higher vocational training, and approximately 70% were currently employed. Approximately two-thirds of the surveys were conducted with patients in Switzerland, and one-third came from Germany. The characteristics of patients with current pain are displayed in Table 1.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Descriptive information about the scales used in the study is presented in the appendix Table 2.

Insert Table 1 about here

First version of the ETS

In an initial analysis, we included nine items of the ETS. Descriptive results and results from scale analyses are presented in the appendix Table 3. In general, the standard deviations showed large variances, with the exception of item 6 (“The treatment [acupuncture] is in general effective”) and item 7 (“I myself have positive expectations about the treatment [acupuncture]”). The mean value of item 6 was relatively high, which can be considered an indicator of a ceiling effect. Therefore, we decided to exclude item 6 (i.e., general expectations about the effectiveness of the treatment [acupuncture]) and item 7 (i.e., personal expectation of a treatment) from the final ETS.

The item-total correlation and the Cronbach’s alpha if an item is deleted are indicators for the coherence of the single item meaning with the total scale (see appendix Table 4). The corrected item-total correlation did not give any strong reason for the exclusion of items. However, there were items with very high total correlations between items (item 2 and item 9). Item 2 (“I expect my complaints will be much better because of the treatment [acupuncture]”) and item 9 (“I expect that after the treatment [acupuncture], my complaints will be considerably better”) were very similar in terms of their wording; therefore, we decided to keep only one of the items for the final version. The reason for this decision was that the ICC value for item 2 was lower than that of item 9. In addition, the

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

mean value for item 9 was lower than the mean of item 2, which consequently can contribute to a larger variation in the final scale because ceiling effects can be avoided. The decision to exclude item 1 was based on content. Our pilot study indicated problems with understanding the meaning of "clearly reduced". Furthermore, item 1 ("I expect the treatment [acupuncture] will clearly reduce my complaints") and item 4 ("I expect the treatment [acupuncture] will make my complaints disappear") covered similar topics. The factor loading confirmed our earlier decision to eliminate item 6 ("The treatment [acupuncture] is in general effective") because the factor loading was relatively low. All items of the first version contributed to one single factor with 65.51% of explained variance. The 9-item ETS version total score is normally distributed. For the 5-item ETS version and the AES items, there were a substantial number of subjects with either the lowest or the highest score on the scale.

The final version of the ETS consists of five items: the previous item 3 ("I expect the treatment [acupuncture] will help me to cope with my complaints") covers coping ability; item 4 ("I expect the treatment [acupuncture] will make my complaints disappear") covers total absence of complaints; item 5 ("I expect the treatment [acupuncture] will improve my energy") covers an energy increase; item 8 ("I expect the treatment [acupuncture] will improve my physical performance") covers an improvement of physical functioning; and item 9 ("I expect that after the treatment [acupuncture] my complaints will be considerably better") covers a considerable decrease in symptoms. To summarize, the ETS captures an expected decrease in symptoms, an expected increase in energy and expected overall wellbeing after a treatment.

Insert Table 2 about here

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

The Cronbach's alpha of the final version was 0.90, which is an excellent value for a very short scale (Table 3). Retest reliability was excellent for the five items. All items contributed substantially to the final scale, and the factor loading indicated a single factor structure with 72.33% of explained variance.

Convergent and divergent validity

We expected a moderate correlation between the ETS and the LOT-R optimism and an inverse correlation with pessimism. However, these assumptions were not confirmed by our findings (Table 3). Interestingly, higher pessimism was moderately associated with higher expectations, both for the ETS and the AES. The ETS showed a small but significant correlation with the PHQ-9, which is consistent with our assumption. The correlation between the ETS and the RS-11 or the BEE was close to zero, which is unexpected according to our assumptions. However, we found a moderate correlation between the ETS and the PSM, for which higher expectations were associated with a higher sensitivity to medication. The shared aspect of this association is the responsiveness to a medical treatment. Personality traits (NEO-FFI) were not associated with the ETS, which is consistent with our assumptions. The ETS score has a very high correlation (> .90) with the AES, which can be considered a benchmark measure for acupuncture expectations.

Insert Table 3 about here

Confirmatory factor analysis

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

In addition, we conducted a CFA with the data from the 439 ETS questionnaires mentioned above with the model presented in Figure 1. The factor loadings were between 0.609 (Item 2) and 0.796 (Item 4). We found an acceptable model fit in the Chi-square statistics (Chi-square = 5.859, d.f. = 4, p = .210), indicating that the model is able to explain the data structure in general. The CFI was .998 in our model, indicating a very good model fit. Similarly, the NFI (.993) and GFI (.995) confirm the excellent model fit. The RMSEA (.033) also suggests that the model with one single factor explains the data very well. The good internal consistency of the ETS in this sample (Cronbach's alpha = .836) confirmed the findings of the scale development.

Insert Figure 1 about here

Discussion

The ETS is a well-validated and brief 5-item scale for measuring patient expectations, with excellent test-retest properties. We were able to overcome ceiling effects, which had been a limitation of earlier measures[25]. This improvement might result from using only four response categories instead of a larger scale with rather inadequate categories (namely, "definitely disagree"). The ETS has excellent measurement properties concerning Cronbach's alpha; retest reliability over one week and single factor structure were replicated in a clinical sample. The ETS can be used for research and clinical purposes equally.

The ETS could be adapted to other clinical situations and treatments beyond acupuncture, which would be an important step towards implementing treatment

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

expectations as standardized working mechanisms among a variety of patient populations. Pain disorders, mental disorders and functional symptoms might be the most appropriate fields, because clinical research indicates the high relevance of expectations for treatment outcomes.

As mentioned earlier, the ETS showed a substantial correlation with the AES. Nevertheless, there are differences worthy of mention between the two measures. In contrast to Jun Mao and his colleagues, we included measures of convergent and divergent validity in our study to enhance information about the unique aspects of expectations compared with related constructs. Moreover, the ETS can be used for a multitude of interventions and can easily be adapted to other clinical situations because no illness-specific symptoms are mentioned in the scale. Furthermore, there are differences concerning item wording and scaling (four vs. five response categories) between the ETS and the AES.

Another notable issue that deserves discussion is the moderate correlation between high pessimism and high expectations for both the ETS and the AES. Our initial assumption was that we would find a negative association between both expectation measures and pessimism. This assumption was not confirmed. Furthermore, we found no correlation between optimism and the ETS in our study. The findings about negative cognitions (pessimism) are consistent with our result of a small but significant correlation of the ETS and depression (PHQ-9). A possible explanation for both findings might be that in pain patients, depression might be associated with more-severe medical symptoms that might lower patients' expectations in some cases depending upon treatment history.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Limitations

There are limitations to our study. First, although collection of data via a web-based survey holds many advantages, there are also downsides to this method because recruitment is done using an unstructured approach. Because the ETS questionnaire was also validated in a paper-based version in a clinical population, this limitation is of minor importance. Second, to enhance the external validity of our study, other clinical populations should be investigated to prove the validity of our scale across clinical conditions. Thus far, the application appears feasible and valid for patients with pain. Third, there might be clinical intervention in situations in which patients might have a very vague idea about procedures and outcomes. We did not collect data from a clinical pain population *prior* to their acupuncture treatment. The patients from our sample provided information about their previous experience with acupuncture, but such an assessment could be done more comprehensively to capture the underlying beliefs and experiences of patients and their relevance for expectations.

Conclusion

The ETS is a short and validated measure that can contribute to the understanding of patient expectations for treatment outcomes. The field of acupuncture served as example to develop the ETS, but it could be easily adapted for other treatments and clinical contexts. The ETS fills a gap by providing a strong and flexible measure that can serve as a basis for upcoming predictor analyses of treatment expectations in clinical studies.

Funding

This study was funded by the Swiss National Science Foundation (Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung), grant No. 105319_159833.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Declaration of competing interests

No, there are no competing interests for any author.

Authorship and contributorship

JB was involved in the conception and execution of the scale and survey, the analysis and interpretation of data, the writing of the first draft and the revision of the manuscript. AK was involved in the analysis and interpretation of data and the writing of the manuscript. SL was involved in programming and conducting the survey, managing and analysing the data, and revising the manuscript. CW was involved in the conception of the scale and survey, the interpretation of the data and the revision of the manuscript.

License for publication statement

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above.

Data-sharing statement

Date are available from the first author on request.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Acknowledgement

We would like to thank Dr. George Lewith, who made important suggestions on the phrasing of questions in our scale.

References

1. Peerdeman KJ, van Laarhoven, Antoinette IM, Keij, Sascha M, Vase, Lene, Rovers, Maroeska M, Peters, Madelon L, Evers, Andrea WM. Relieving patients' pain with expectation interventions: a meta-analysis. *Pain* 2016;157(6):1179-91.
2. Kirsch IE. How expectancies shape experience: American Psychological Association 1999.
3. Bialosky JE, Bishop, Mark D, Cleland, Joshua A. Individual expectation: an overlooked, but pertinent, factor in the treatment of individuals experiencing musculoskeletal pain. *Physical therapy* 2016;90(9):1345-55.
4. Bishop FL, Yardley, Lucy, Prescott, Philip, Cooper, Cyrus, Little, Paul, Lewith, George T. Psychological covariates of longitudinal changes in back-related disability in patients undergoing acupuncture. *The Clinical journal of pain* 2015;31(3):254.
5. Iles RA, Davidson, Megan, Taylor, Nicholas F, O'Halloran, Paul. Systematic review of the ability of recovery expectations to predict outcomes in non-chronic non-specific low back pain. *Journal of occupational rehabilitation* 2009;19(1):25-40.
6. Mondloch MV, Cole DC, Frank JW. Does how you do depend on how you think you'll do? A systematic review of the evidence for a relation between patients' recovery expectations and health outcomes. *Canadian Medical Association Journal* 2001;165(2):174-79.
7. Linde K, Witt, C. M., Streng, A., Weidenhammer, W., Wagenpfeil, S., Brinkhaus, B., Willich, S. N., Melchart, D. The impact of patient expectations on outcomes in four randomized controlled trials of acupuncture in patients with chronic pain. *Pain* 2007;128(3):264-71. doi: 10.1016/j.pain.2006.12.006
8. Foster NE, Thomas, E., Hill, J. C., Hay, E. M. The relationship between patient and practitioner expectations and preferences and clinical outcomes in a trial of exercise and acupuncture for knee osteoarthritis. *European journal of pain* 2010;14(4):402-09. doi: 10.1016/j.ejpain.2009.06.010
9. Sherman KJ, Cherkin, D. C., Ichikawa, L., Avins, A. L., Delaney, K., Barlow, W. E., Khalsa, P. S., Deyo, R. A. Treatment expectations and preferences as predictors of outcome of acupuncture for chronic back pain. *Spine* 2010;35(15):1471-7. doi: 10.1097/BRS.0b013e3181c2a8d3
10. Barth J, Schafroth, Larissa, Witt, Claudia M. Overlap and differences between patient and provider expectations for treatment outcomes: the case of acupuncture. *The Journal of Pain* 2016;17(6):685-93.
11. Haanstra TM, Tilbury C, Kamper SJ, et al. Can optimism, pessimism, hope, treatment credibility and treatment expectancy be distinguished in patients undergoing total hip and total knee arthroplasty? *PloS one* 2015;10(7):e0133730.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

12. Crow R, Gage H, Hampson S, et al. The role of expectancies in the placebo effect and their use in the delivery of health care: A systematic review. *Health Technology Assessment* 1999;3(3):1-96. doi: https://doi.org/10.3310/hta3030
13. Constantino MJ, Ametrano RM, Greenberg RP. Clinician interventions and participant characteristics that foster adaptive patient expectations for psychotherapy and psychotherapeutic change. *Psychotherapy* 2012;49(4):557-69. doi: 10.1037/a0029440
14. Bowling A, Rowe, G., Lambert, N., Waddington, M., Mahtani, K. R., Kenten, C., Howe, A., Francis, S. A. The measurement of patients' expectations for health care: a review and psychometric testing of a measure of patients' expectations. *Health technology assessment* 2012;16(30):1-509. doi: 10.3310/hta16300
15. Bandura A. Self-efficacy mechanism in human agency. *Am Psychol* 1982;37(2):122.
16. Carver CS, Scheier MF, Segerstrom SC. Optimism. *Clin Psychol Rev* 2010;30(7):879-89. doi: 10.1016/j.cpr.2010.01.006 [published Online First: 2010/02/23]
17. Scholz U, Doña BG, Sud S, et al. Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment* 2002;18(3):242.
18. Nicholas MK. The pain self-efficacy questionnaire: taking pain into account. *European Journal of Pain* 2007;11(2):153-63.
19. Woolhead G, Donovan J, Dieppe P. Patient expectations and total joint arthroplasty. *The Journal of Rheumatology* 2003;30(7):1656-57.
20. Prady SL, Burch, J., Vanderbloemen, L., Crouch, S., MacPherson, H. Measuring expectations of benefit from treatment in acupuncture trials: A systematic review. *Complementary therapies in medicine* 2015;23(2):185-99. doi: 10.1016/j.ctim.2015.01.007
21. Laferton JA, Kube, Tobias, Salzmann, Stefan, Auer, Charlotte J, Shedden-Mora, Meike C. Patients' expectations regarding medical treatment: A critical review of concepts and their assessment. *Frontiers in Psychology* 2017;8
22. Haanstra TM, van den Berg T, Ostelo RW, et al. Systematic review: do patient expectations influence treatment outcomes in total knee and total hip arthroplasty? *Health and Quality of Life Outcomes* 2012;10(1):152.
23. van Hartingsveld F, Ostelo RW, Cuijpers P, et al. Treatment-related and patient-related expectations of patients with musculoskeletal disorders: a systematic review of published measurement tools. *The Clinical Journal of Pain* 2010;26(6):470-88.
24. Zywiel MG, Mahomed A, Gandhi R, et al. Measuring expectations in orthopaedic surgery: a systematic review. *Clinical Orthopaedics and Related Research®* 2013;471(11):3446-56.
25. Mao JJ, Armstrong, Katrina, Farrar, John T, Bowman, Marjorie A. Acupuncture expectancy scale: development and preliminary validation in China. *Explore: The Journal of Science and Healing* 2007;3(4):372-77.
26. Jones SM, Lange, Jane, Turner, Judith, Cherkin, Dan, Ritenbaugh, Cheryl, Hsu, Clarissa, Berthoud, Heidi, Sherman, Karen. Development and validation of the EXPECT questionnaire: assessing patient expectations of outcomes of complementary and alternative medicine treatments for chronic pain. *The Journal of Alternative and Complementary Medicine* 2016;22(11):936-46.
27. Constantino MJ, Arnkoff DB, Glass CR, et al. Expectations. *Journal of clinical psychology* 2011;67(2):184-92. doi: 10.1002/jclp.20754 [published Online First: 2010/12/04]

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

28. Faller H, Vogel H, Bosch B. Erwartungen von Rehabilitanden hinsichtlich der Methoden und Ergebnisse ihrer Rehabilitation-Eine kontrollierte Studie mit Rückenschmerz-und onkologischen Patienten. *Die Rehabilitation* 2000;39(04):205-14.
29. Kravitz RL. Measuring patients' expectations and requests. *Annals of internal medicine* 2001;134(9 Pt 2):881-8.
30. Schulte D. Messung der Therapieerwartung und Therapieevaluation von Patienten (PATHEV). *Zeitschrift für Klinische Psychologie und Psychotherapie* 2005;34(3):176-87.
31. Eibner F, Barth J, Helmes A, et al. Variations in subjective breast cancer risk estimations when using different measurements for assessing breast cancer risk perception. *Health, Risk & Society* 2006;8(2):197-210.
32. Glaesmer H, Hoyer, Jürgen, Klotsche, Jens, Herzberg, Philipp Yorck. Die deutsche version des Life-Orientation-Tests (LOT-R) zum dispositionellen Optimismus und Pessimismus. *Zeitschrift für Gesundheitspsychologie* 2008;16(1):26-31.
33. Löwe B, Unützer, Jürgen, Callahan, Christopher M, Perkins, Anthony J, Kroenke, Kurt. Monitoring depression treatment outcomes with the patient health questionnaire-9. *Medical care* 2004;42(12):1194-201.
34. Ell K, Unützer, Jurgen, Aranda, Maria, Sanchez, Kathleen, Lee, Pey-Jiuan. Routine PHQ-9 depression screening in home health care: Depression prevalence, clinical and treatment characteristics, and screening implementation. *Home health care services quarterly* 2006;24(4):1-19.
35. Wagnild GM, Young, Heather M. Development and psychometric evaluation of the Resilience Scale. *Journal of Nursing Measurement* 1993
36. Schumacher J, Leppert, Karena, Gunzelmann, Thomas, Strauß, Bernhard, Brähler, Elmar. Die Resilienzskala–Ein Fragebogen zur Erfassung der psychischen Widerstandsfähigkeit als Personmerkmal. *Zeitschrift für Klinische Psychologie, Psychiatrie und Psychotherapie* 2005;53:16-39.
37. Schutzler L, Witt, Claudia M. Body-Efficacy Expectation: Assessment of Beliefs concerning Bodily Coping Capabilities with a Five-Item Scale. *Evidence-based complementary and alternative medicine : eCAM* 2013;2013:152727. doi: 10.1155/2013/152727
38. Horne R, Faasse, Kate, Cooper, Vanessa, Diefenbach, Michael A, Leventhal, Howard, Leventhal, Elaine, Petrie, Keith J. The perceived sensitivity to medicines (PSM) scale: an evaluation of validity and reliability. *British journal of health psychology* 2013;18(1):18-30.
39. Borkenau P, Ostendorf, Fritz. NEO-Fünf-Faktoren Inventar:(NEO-FFI); nach Costa und McCrae: Göttingen: Hogrefe 1993.
40. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. *Psychological bulletin* 1980;88(3):588.
41. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull* 1990;107(2):238-46. [published Online First: 1990/03/01]
42. Jöreskog KG, Sörbom D. LISREL 8: Structural equation modeling with the SIMPLIS command language: Scientific Software International 1993.
43. Kline RB. Principles and practice of structural equation modeling: Guilford publications 2015.
44. Marsh HW, Balla JR, McDonald RP. Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological bulletin* 1988;103(3):391.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

45. Browne MW, Cudeck R. Alternative ways of assessing model fit. In. KA, Bollen, JS Long,(Eds.), Testing structural equation models (pp. 136-162): Newbury Park, CA: Sage, 1993.

Tables and Figures

Table 1. Sample characteristics of subjects with current pain (N = 102)

Table 2. Scale and factor analysis of the 5-item ETS for subjects with current pain (N = 102)

Table 3. Convergent and divergent validity of the 5-item version of the ETS for subjects with current pain (N = 102). Pearson Correlation, Significance level (2-tailed), N of patients

Figure 1. Confirmatory factor analysis with factor loadings (N = 439) of the final 5-item ETS version in patients with chronic low back pain

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 1. Sample characteristics of subjects with current pain (N = 102)

	Frequencies // Mean (SD)	Percentage
Age	49.16 (14.04)	
Gender		
Female	76	75%
Male	26	25%
Education		
Vocational school	39	38%
Higher school certificate or Higher vocational training	29	28%
University	34	33%
Employment status		
Working	72	71%
Retired	10	10%
Other	20	20%
Country of residence		
Switzerland	68	67%
Germany	34	33%
Acupuncture experience		
Current	13	13%

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Earlier	51	50%
None	38	37%
Success of acupuncture [range 1–10]*	6.5 (2.80)	
Reason for treatment**		
Pain	55	86%
Mental health	13	20%
Other	24	38%
General Health***	3.21 (.92)	
Pain characteristics		
Intensity [range 0–10]****	6.5 (1.92)	
Physician consulted*****	68	67%
Any treatment*****	66	65%
Days of restriction due to pain*****	46.76 (63.74)	
Location******		
Back	58	57%
Neck	44	43%
Knee	21	21%
Head	20	20%
Hip	16	16%
Other	33	32%

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

627

628 *Range from no success (1) to very successful (10) treatment outcome of the last

629 acupuncture treatment

630 **Multiple answers were possible.

631 ***Range from 1 = excellent to 5 = bad

632 ****0 indicating no pain and 10 indicating maximum pain

633 ******In the last 6-month period

634 ******Multiple answers were possible.

635

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 2. Scale and factor analysis of the 5-item ETS for subjects with current pain (N = 102)

Item	Mean (SD)	Item-total correlation corrected	Cronbach`s alpha if item deleted	Factor loading	Retest ICC [95%-CI]
Item 3	2.38 (1.03)	.727	.887	.826	.697 [.517, .818]
Item 4	1.97 (.96)	.736	.883	.834	.724 [.556, .835]
Item 5	2.41 (1.0)	.776	.875	.865	.840 [.732, .907]
Item 8	2.33 (.83)	.817	.869	.891	.749 [.592, .851]
Item 9	2.52 (.90)	.737	.883	.834	.836 [.725, .904]
			Cronbach's α	% of variance	
ETS 5			.901	72.325	.856 [.757, .917]

Item 3: I expect the treatment [acupuncture] will help me to cope with my complaints.

Item 4: I expect the treatment [acupuncture] will make my complaints disappear.

Item 5: I expect the treatment [acupuncture] will improve my energy.

For peer review only

Item 8: I expect the treatment [acupuncture] will improve my physical performance.

Item 9: I expect that after the treatment [acupuncture], my complaints will be considerably better.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 3. Convergent and divergent validity of the 5-item version of the ETS for subjects with current pain (N = 102). Pearson Correlation, Significance level (2-tailed), N of patients

Correlation Significance level Number of patients	ETS	LOT-R Optimism	LOT-R Pessimism	PHQ-9	RS-11	BEE	PSM	NEO-FFI Neuroticism	NEO-FFI Openness to experience
ETS	1 102								
LOT-R Optimism	-.066 .517 99	1 99							
LOT-R Pessimism	.204* .043 99	-.247* .014 99	1 99						
PHQ-9	.233* .020 100	-.567** .000 99	.312** .002 99	1 100					
RS-11	-.073 .474 99	.546** .000 99	-.460** .000 99	-.572** .000 99	1 99				

For peer review only

Expectation for Treatment Scale (ETS)

BEE	.032	.259**	-.074	-.176	.407**	1			
	.754	.010	.468	.081	.000				
	99	99	99	99	99	99			
PSM	.344**	-.078	.201*	.306**	-.049	.011	1		
	.000	.445	.046	.002	.631	.912			
	100	99	99	100	99	99	100		
NEO-FFI	.104	-.578**	.467**	.631**	-.682**	-.284**	.238*	1	
Neuroticism	.316	.000	.000	.000	.000	.005	.020		
	95	95	95	95	95	95	95	95	
NEO-FFI	-.072	.297**	-.167	-.090	.302**	.122	.132	-.130	1
Openness to	.492	.004	109	.386	.003	.241	.205	.211	
experience	94	94	94	94	94	94	94	94	94
Cronbach's Alpha	.901	.727	.608	.900	.900	.757	.940	.854	.726

**Correlation is significant at the 0.01 level (2-tailed).

*Correlation is significant at the 0.05 level (2-tailed).

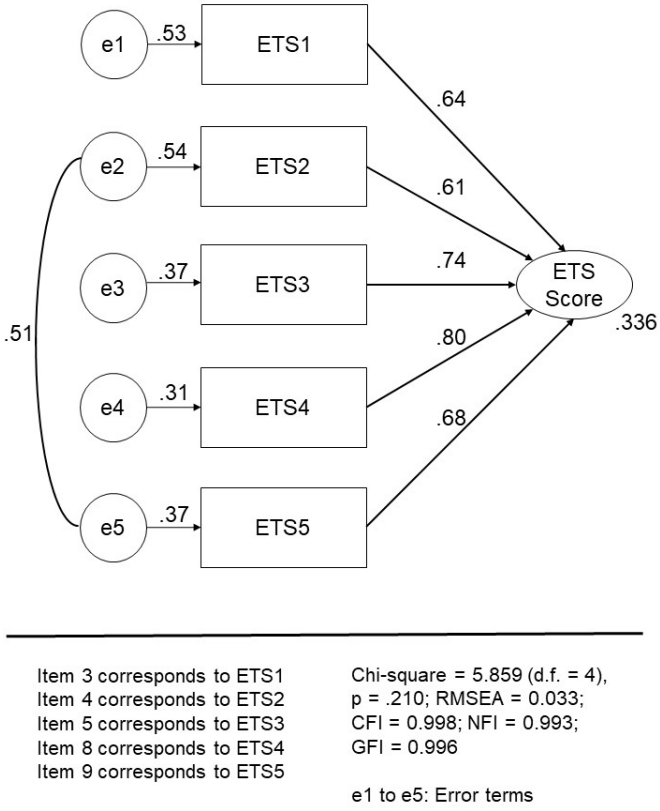
ETS = Expectation for Treatment Scale; LOT-R = Life-Orientation Test; PHQ-9 = Patient Health Questionnaire-9; RS-11 = Die Resilienzskala - Ein Fragebogen zur Erfassung der psychischen Widerstandsfähigeit als Personenmerkmal; BEE = Body-Efficacy Expectation; PSM = Perceived Sensitivity to Medicines Scale; NEO-FFI = NEO Five-Factor Inventory

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Figure 1. Confirmatory factor analysis with factor loadings (N = 439) of the final 5-item ETS version in patients with chronic low back pain

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml



Confirmatory factor analysis with factor loadings (N = 439) of the final 5-item ETS version in patients with chronic low back pain.

Appendix (Barth et al.; ETS)

APPENDIX Table 1. Description of the 5 final items in English and German

Table 1. Description of the 5 final items in English and German

Description: The overarching construct of ETS is treatment-related outcome expectation of patients. Outcomes of interest for patients are coping (item 1), vitality (item 3) and physical health (item 4). Two items (items 2 and 5) capture the reduction of patient complaints, either an absence or a considerable reduction thereof. The scale can be adapted to specific treatments of interest. Instead of the term “treatment”, a specific term can be used (e.g., acupuncture, counselling, or physiotherapy). We used the term complaints [German Beschwerden] as previously used by other authors as in the symptom checklist (SCL-90-R)[46]. Therefore, any type of complaint can be captured and allow the use of this instrument across patient groups and settings. The term complaints can be adapted if needed to the clinical situation if one specific target symptom is predominant (e.g., pain, depression, or nausea). The scale should be used in a clinical setting by patients with sufficient German or English language skills.

For peer review only

English version

There are several statements below that capture your expectations about the [acupuncture] treatment. Please indicate to what extent these statements apply to you personally. **There are no right or wrong answers. We are only interested in your current personal thoughts.**

*Please select for each statement **one** response.*

1. I expect the treatment [acupuncture] will help me to cope with my complaints.

partially disagree	partially agree	agree	definitely agree
❑	❑	❑	❑

2. I expect the treatment [acupuncture] will make my complaints disappear.

partially disagree	partially agree	agree	definitely agree
❑	❑	❑	❑

3. I expect the treatment [acupuncture] will improve my energy.

partially disagree	partially agree	agree	definitely agree
❑	❑	❑	❑

4. I expect the treatment [acupuncture] will improve my physical performance.

partially disagree	partially agree	agree	definitely agree
❑	❑	❑	❑

5. I expect that after the treatment [acupuncture], my complaints will be considerably better.

partially disagree	partially agree	agree	definitely agree
❑	❑	❑	❑

*We used the term acupuncture in our study. The general version of ETS uses the term treatment.

For peer review only

German version

Im Folgenden finden Sie einige Aussagen darüber, was man von einer Behandlung [Akupunktur] erwarten kann. Geben Sie bitte an, inwieweit die einzelnen Aussagen für Sie persönlich zutreffen. **Es gibt keine richtigen und falschen Antworten. Uns interessiert nur, was Sie persönlich denken.**

*Bitte wählen Sie für jede Aussage **eine** Antwort aus.*

1. Ich erwarte, dass ich durch die Behandlung [Akupunktur] mit meinen Beschwerden besser umgehen kann.

trifft eher nicht zu	trifft eher zu	trifft sehr zu	trifft völlig zu
❑	❑	❑	❑

2. Ich erwarte, dass meine Beschwerden durch die Behandlung [Akupunktur] verschwinden.

trifft eher nicht zu	trifft eher zu	trifft sehr zu	trifft völlig zu
❑	❑	❑	❑

3. Ich erwarte, dass sich meine Energie durch die Behandlung [Akupunktur] verbessert.

trifft eher nicht zu	trifft eher zu	trifft sehr zu	trifft völlig zu
❑	❑	❑	❑

4. Ich erwarte durch die Behandlung [Akupunktur] eine verbesserte körperliche Leistungsfähigkeit.

trifft eher nicht zu	trifft eher zu	trifft sehr zu	trifft völlig zu
❑	❑	❑	❑

5. Ich erwarte, dass sich nach der Behandlung [Akupunktur] meine Beschwerden deutlich verbessern.

trifft eher nicht zu	trifft eher zu	trifft sehr zu	trifft völlig zu
❑	❑	❑	❑

*Wir haben den Begriff Akupunktur in unserer Studie verwendet. Die allgemeine Version des ETS verwendet den Begriff Behandlung.

For peer review only

Scoring of the questionnaire

The values from the 5 single items are summed to build the ETS sum score (min 5, max 20). If one item has a missing value, the sum score can be calculated. Therefore, the values of the remaining 4 items are summed, divided by 4 and multiplied by 5. However, multiple imputation procedures to impute the missing value should be preferred over this re-calculation. In the case of more than one missing value, imputation procedures are needed, and no manual recalculation should be considered.

For peer review only

APPENDIX Table 2: Sample characteristics in scales in subjects with current pain (N = 102)

	Median Mean (SD)
ETS (9 items) (range 9–36)	22 22.63 (.64)
ETS (5 items) (range 5–20)	11 11.62 (4.01)
AES (range 4–20)	12.0 11.82 (4.10)
PHQ-9 (range 0–27)	7 8.46 (6.06)
RS-11 (range 11–77)	60 58.36 (11.13)
BEE (range 5–20)	14 13.71 (2.90)
PSM (range 5–25)	11 12.77 (6.13)
LOT-R Optimism (range 0–12)	8 7.93 (2.77)
LOT-R Pessimism (range 0–12)	8 7.93 (2.38)
NEO-FFI Neuroticism (range 0–48)	21 21.08 (8.19)
NEO-FFI Openness to experience (range 0–44)	30 29.84 (5.77)

Expectation for Treatment Scale; AES= Acupuncture Expectancy Scale; LOT-R= Life-Orientation Test; PHQ-9= Patient Health Questionnaire-9; RS-11= Die Resilienzskala - Ein Fragebogen zur Erfassung der psychischen Widerstandsfähigeit als Personenmerkmal; BEE= Body-Efficacy Expectation; PSM= Perceived Sensitivity to Medicines Scale; NEO-FFI= NEO Five-Factor Inventory

For peer review only

Appendix Table 3. Scale and factor analysis of the 9-item ETS for subjects with current pain (N = 102)

Item	Mean (SD)	Item-total correlation corrected	Cronbach`s alpha if item deleted	Factor loading	Retest ICC [95%–CI]
Item 1	2.71 (1.0)	.782	.920	.839	.684 [.499, .810]
Item 2	2.66 (.93)	.816	.918	.863	.725[.557, .836]
Item 3**	2.38 (1.03)	.736	.924	.790	.697 [.517, .818]
Item 4**	1.97 (.96)	.735	.923	.793	.724 [.556, .835]
Item 5**	2.41 (1.0)	.745	.923	.799	.840 [.732, .907]
Item 6	2.97 (.59)	.612	.931	.685	.649 [.450, .787]
Item 7	2.68 (.65)	.744	.924	.806	.658 [.462, .793]
Item 8**	2.33 (.83)	.772	.921	.821	.749 [.592, .851]
Item 9**	2.52 (.90)	.826	.917	.874	.836 [.725, .904]
			Cronbach's α	% of variance	
ETS 9			.930	65.512	.879 [.794, .930]

**marked items are included in the final version of its scale.

Item 1: I expect the treatment [acupuncture] will clearly reduce my complaints.

Item 2: I expect my complaints will be much better because of the treatment [acupuncture].

Item 3: I expect the treatment [acupuncture] will help me to cope with my complaints.

Item 4: I expect the treatment [acupuncture] will make my complaints disappear.

For peer review only

Item 5: I expect the treatment [acupuncture] will improve my energy.

Item 6: The treatment [acupuncture] is in general effective.

Item 7: I myself have positive expectations about the treatment [acupuncture].

Item 8: I expect the treatment [acupuncture] will improve my physical performance.

Item 9: I expect that after the treatment [acupuncture], my complaints will be considerably better.

For peer review only

Appendix Table 4. Between-item correlations for sample with current pain (N = 102). Pearson Correlation and significance (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9
Item 1	1								
Item 2	.840** .000	1							
Item 3	.636** .000	.605** .000	1						
Item 4	.568** .000	.627** .000	.580** .000	1					
Item 5	.533** .000	.611** .000	.664** .000	.637** .000	1				
Item 6	.540** .000	.556** .000	.442** .000	.437** .000	.505** .000	1			
Item 7	.692** .000	.679** .000	.541** .000	.575** .000	.509** .000	.521** .000	1		
Item 8	.528** .000	.571** .000	.673** .000	.664** .000	.762** .000	.470** .000	.631** .000	1	
Item 9	.733** .000	.754** .000	.605** .000	.673** .000	.603** .000	.592** .000	.701** .000	.673** .000	1

Item 1: I expect the treatment [acupuncture] will clearly reduce my complaints.

Item 2: I expect my complaints will be much better because of the treatment [acupuncture].

For peer review only

Item 3: I expect the treatment [acupuncture] will help me to cope with my complaints.

Item 4: I expect the treatment [acupuncture] will make my complaints disappear.

Item 5: I expect the treatment [acupuncture] will improve my energy.

Item 6: The treatment [acupuncture] is in general effective.

Item 7: I myself have positive expectations about the treatment [acupuncture].

Item 8: I expect the treatment [acupuncture] will improve my physical performance.

Item 9: I expect that after the treatment [acupuncture], my complaints will be considerably better.

For peer review only

TROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

	Item No	Recommendation	page
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	8
Methods			
Study design	4	Present key elements of study design early in the paper	9
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	9
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	9
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	n.a.
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	10-13
Bias	9	Describe any efforts to address potential sources of bias	n.a.
Study size	10	Explain how the study size was arrived at	n.a.
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	n.a.
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	14-15
		(*b*) Describe any methods used to examine subgroups and interactions	n.a.
		(*c*) Explain how missing data were addressed	n.a.
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	n.a.
		(*e*) Describe any sensitivity analyses	n.a.
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	15 and Table 2
		(b) Give reasons for non-participation at each stage	n.a.
		(c) Consider use of a flow diagram	n.a.
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	n.a.
		(b) Indicate number of participants with missing data for each variable of interest	n.a.
Outcome data	15*	Report numbers of outcome events or summary measures	n.a.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	n.a.
		(*b*) Report category boundaries when continuous variables were categorized	n.a.
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n.a.
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	19
Discussion			
Key results	18	Summarise key results with reference to study objectives	19
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	21
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	21
Generalisability	21	Discuss the generalisability (external validity) of the study results	21
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	21

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

BMJ Open

Assessment of patients' expectations: development and validation of the Expectation for Treatment Scale (ETS)

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-026712.R2
Article Type:	Research
Date Submitted by the Author:	14-Apr-2019
Complete List of Authors:	Barth, Jürgen; University Hospital Zurich and University of Zurich, Institute for Complementary and Integrative Medicine Kern, Alexandra; University Hospital Zurich and University of Zurich, Institute for Complementary and Integrative Medicine Lüthi, Sebastian; University Hospital Zurich and University Zurich, Institute for Complementary and Integrative Medicine Witt, Claudia M.; University Hospital Zurich and University of Zurich , Institute for Complementary and Integrative Medicine; Institute for Social Medicine, Epidemiology and Health Economics, Charité
Primary Subject Heading:	Complementary medicine
Secondary Subject Heading:	Patient-centred medicine, Research methods
Keywords:	expectations, placebo, assessment, pain, acupuncture
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
ETS_STROBE_checklist_cross-sectional_20190118.doc	

SCHOLARONE™
Manuscripts

Assessment of patients' expectations: development and validation of the Expectation for Treatment Scale (ETS)

Running title: Expectation for Treatment Scale (ETS)

Submission to British Medical Journal

Jürgen Barth*[1], Alexandra Kern [1], Sebastian Lüthi [1], Claudia M. Witt [1, 2, 3]

[1] Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich, Sonneggstrasse 6, 8006 Zurich, Switzerland Dr. Jürgen Barth Research Director, M.A. Alexandra Kern PhD student, M.A. Sebastian Lüthi Master student, Professor Dr. Claudia M. Witt, Chair and Director

[2] Institute for Social Medicine, Epidemiology and Health Economics, Charité – Universitätsmedizin Berlin, Luisenstrasse 57, Berlin, Germany, Professor Dr. Claudia M. Witt, research group leader

[3] Center for Integrative Medicine, University of Maryland, School of Medicine, 520 W Lombard Street, Baltimore, MD, USA Professor Dr. Claudia M. Witt, senior scientist

*Correspondence:

Dr. Jürgen Barth, University Hospital Zurich and University of Zurich, Switzerland

Phone: +41 (0)44 255 4896

mail@juergen-barth.de

www.iki.usz.ch

Word count: 4526

For peer review only

Abstract

Objective: The development of a short self-report instrument for the assessment of expectations (Expectation for Treatment Scale, ETS), using acupuncture as a case example.

Design: A cross-sectional assessment with retest after one week.

Setting: A web-based survey with patients suffering from pain.

Methods: In a three-step approach, we reduced the initially collected number of items from 17 to 9 and to 5, including expectations about coping ability, vitality, physical health and reduction of patient complaints. Items were selected according to internal consistency (Cronbach's alpha), convergent and divergent validity with related constructs (optimism, pessimism, resilience, perceived sensitivity to medicines, depression, and others), one-week retest reliability (intraclass correlation coefficient, ICC) and exploratory and confirmatory factor analysis.

Results: A total of 102 pain patients were included, and 54 of these patients completed the retest assessment. The final version of the ETS consisted of 5 items and had an excellent Cronbach's alpha (.90), with 72.33% variance on one single factor. Depression, pessimism and perceived sensitivity to medicines showed positive correlations with our expectation measure (r = 0.23, r = 0.20, r = 0.34, respectively), the correlation between the ETS and optimism was low (r = -0.07), and no correlation between the ETS and resilience was found (r = -0.07). Convergent validity was confirmed with a high correlation (r > .90) between ETS and a treatment-specific measure of expectations. The retest ICC was .86, which showed high stability over one week. A confirmatory factor analysis (N = 439) with data from patients with low back pain confirmed the single-factor structure of the instrument.

For peer review only

Conclusion: The ETS showed strong psychometric properties and covered a distinct construct. As a next step, the ETS might be implemented in different clinical conditions and settings to investigate psychometrics and its predictive power for treatment outcomes.

Keywords: Expectations, placebo, assessment, pain, optimism, acupuncture

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Strengths and limitations of this study

- It is the first project to systematically develop a general measure to assess patients' expectations across medical conditions and treatments with the involvement of patients, earlier scales and empirical data.
- The Expectation for Treatment Scale (ETS) is a short and reliable measure that captures outcome-related expectations of patients.
- The association between the ETS and related constructs was explored.
- The development of the ETS was done via an online survey with pain patients, and the initial findings were confirmed using confirmatory factor analysis in a patient sample with low back pain.
- The transfer of the ETS in different clinical conditions and settings should be pre-tested to explore whether patients can evaluate the respective treatment at this level.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Introduction

Patients' expectations alter responses to interventions in placebo research and in clinical studies. In experimental research, the instruction to receive an active intervention accompanies large effects on pain reduction (effect size = 0.75)[1]. Expectations can be considered a major driver of changes in symptoms and other health-related outcomes[2]. Clinical studies have revealed that patients' positive expectations are related to reduced pain after a medical treatment[3-5], and this phenomenon has also been observed for other medical conditions[6]. For acupuncture, this association was investigated in several studies with heterogeneous findings: Patients with chronic pain had higher odds (odds ratio 2.11 [95% CI 1.32-3.34]) of benefitting from acupuncture treatment if they expected a better treatment response at baseline (controlled for other clinical baseline characteristics)[7]. However, there are also studies that found no influence of expectations on the outcome[8 9].

These differences in the association of expectations and outcomes might be partly explained by the fact that each study used a newly invented measure, and differences between measures might hide or exaggerate associations between expectations and intervention outcomes. Additionally, the match between expectations of patients and treatment providers might be relevant for the success of a specific treatment[10].

Expectation is a well-known and oft-used term. A clear definition and a sharp distinction from associated constructs is important for the development of a measure[11]. In the context of medical treatments, the term "expectations" describes cognitions about treatment-related health outcomes in the future after a specific intervention[12 13]. Patients can consider a treatment more or less beneficial for their complaints or disease at a specific time-point (i.e., outcome expectations)[13]. Role expectations also capture the role of a

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

patient and the therapist during the treatment. In other words, a patient might consider himself rather inactive during treatment in defining treatment goals and expects an active therapist to achieve a good treatment outcome. However, our purpose was to develop a scale on ‘‘patient expectations’’ that covers treatment-related outcome expectations. Bowling et al. (2012) provide an insightful summary about the theoretical underpinning of expectations[14]. Following the definition by Bowling et al. (2012), we therefore designed our measure to assess expectations related to a clinical intervention with a clinically relevant outcome from a patient’s perspective.

Earlier findings about the expectation outcome association in clinical studies have been limited by the diversity of measures. Several authors claimed diversity in covered concepts, time-point of assessment and problems to evaluate the validity of the measures[15-17]. A strong measure is a prerequisite to accurately predict treatment responses based on pretreatment expectations. A closer investigation of the results from a systematic review about acupuncture expectation measures by Prady and colleagues[18] showed that of ten trials, only five provided their exact item wording for measuring expectations. Because many of the assessment instruments are not publicly available, it is difficult to replicate the studies. Three of the five mentioned studies used only one item to assess expectations. There is no reporting at all of Cronbach’s alpha in the two remaining studies.

For further research in the field of expectations, a strong measure with high acceptance across clinical fields would be needed for several reasons. First, a reliable measure with high internal consistency at a specific time-point is a prerequisite to use expectations as a robust predictor. Second, ceiling effects were a common problem in the measurement of expectations, because patients who are seeking help from a specific treatment often expect large benefits; otherwise, they would not be attracted by this

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

treatment. This problem was apparent in an established expectation measure for acupuncture treatment that served as a benchmark measure for our scale (Acupuncture Expectancy Scale (AES))[19]. Ceiling effects are particularly problematic because the predictive power of such skewed variables is low. In the case of expectations, many research questions address the prediction of treatment outcomes; therefore, a measure with sufficient variation between patients is needed. Some authors have used the term "realistically expect" to capture expectations in a recent study with a comprehensive 10-item assessment of expectations, hope and beliefs (EXPECT)[20] with limited internal consistency. Third, a measurement of expectation should be stable in a reasonable timeframe: The EXPECT scale showed only a moderate retest reliability, with an intra-class correlation of ICC = 0.75. Most importantly, the association of EXPECT with the established AES, a benchmark measure for the field of acupuncture, was unexpectedly low (r = .54).

This study aimed to develop a measure to reliably capture patient expectations with a short scale. The newly developed scale (Expectation for Treatment Scale, ETS) ideally has a strong association with available measures of expectations, moderate associations with related constructs (e.g., optimism and pessimism) and no association with personality (e.g., neuroticism). Furthermore, it should be constructed such that it can be adapted to other treatments and used universally in different clinical fields independent of patients' complaints.

Methods

To address the problem of the abovementioned ceiling effects, we developed the items of our scale (ETS) in a pilot study with pain patients in our outpatient clinic to measure

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

expected benefits before treatment from the patient's perspective. Based on existing questionnaires on patient expectations[9 19 21-24], we created a list of 17 items covering different facets of expectations that fit with our aim to develop a measure for treatment-related outcome expectations. We had been particularly interested in questionnaires from the field of acupuncture research and other non-pharmacological interventions, and the questions should be applicable in an applied context. The first and senior author were responsible for the selection of these items.

Patient and Public Involvement

Twenty patients completed the questionnaires, provided sociodemographic information and were asked for written comments about the accessibility of the questions. In addition, two patients were interviewed by a qualitative researcher. Two health professionals (one acupuncturist and a doctor assistant) also verbally commented about the appropriateness of the questions. Based on these data, we selected items with low skewness and a large range of responses (i.e., variation). High correlations between items and the findings from the qualitative feedback were also considered.

Participants

For the main validation, study subjects of the convenience sample were recruited using different methods. Three regional patient organizations distributed the information, and we used several email distribution lists (including the University of Zurich, Switzerland and the Charité University Hospital Berlin, Germany). Furthermore, patients from the previous year seeking treatment for a musculoskeletal condition at the Institute for

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Complementary and Integrative Medicine of the University Hospital Zurich were contacted by email. All participants were required to be at least age 18 and, according to their own evaluation, have sufficient knowledge to understand German. The included pain patients had to suffer from at least some pain at the day of the assessment (> 0 on a numeric rating scale from 0 (no pain) to 10 (worst pain)).

The study link was accessed by 522 subjects, and 244 started the survey. In total, 142 subjects were excluded due to insufficient pain (N = 113), missing data in the nine-item version of the ETS (N = 14) and more than 20% missing responses across all items (N = 13). Data from patients suffering from pain (N = 102) with retest data from 54 patients were available.

Procedures

Subjects who decided to participate activated an access link to the online survey provided via email. The survey was hosted by https://www.soscisurvey.de/, and all data were collected electronically and anonymously. Participants were able to complete the survey within 15 minutes. All questions were created in a forced choice format, meaning that subjects were not able to continue to the next set of questions without completing the previous set. Upon starting the survey, a short welcome message, followed by an introduction to the study, was provided. Subjects who agreed to be contacted again for the retest assessment provided their email address at the end of the survey. The retest sample received a follow-up ETS one week after completion of the first survey. The study was granted ethics approval by the local ethics committee (Kantonale Ethikkommission Zürich, No. 48-2015).

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Measures

The survey included demographic information (age, gender, education, employment status, and country of residence), questions about any previous experience with acupuncture (current or earlier), the degree to which the acupuncture treatment was successful (numeric rating scale from 1 to 10, indicating no success to much success) and the reason for the treatment (pain, mental health or non-specific). Patients who suffered from pain (binary variable with yes vs. no option) were asked to give information about their pain. We assessed average pain intensity (numeric rating scale from 1 to 10, indicating no pain to worse pain), whether a physician was consulted, whether any other type of therapy was used and the number of days of restriction due to pain. Subsequently, subjects completed the 9-item ETS. Below, we provide a detailed description of all measures in this study.

Expectation for Treatment Scale

The first version of the ETS consisted of nine items (e.g., "I expect the treatment [acupuncture] will help me to cope with my complaints."). Each item was to be rated on a four-point scale ranging from 1 to 4 (partially disagree, partially agree, agree, definitely agree). We decided to use a four-point scale instead of a five-point scale for two reasons. First, the chosen format forces the patient to provide an answer with a direction (higher or lower than the middle answer choice) instead of opting for the middle category, which often occurs in situations of insecurity (such as the present estimation of future events)[25]. Second, the lowest answer on the scale of "definitely disagree" can be considered unexpected for patients motivated to undergo an acupuncture treatment. After the statistical

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

analyses, the nine-item version was reduced to a final 5-item version of the ETS. A detailed description of these five items is provided in the appendix Table 1. These five items were translated into English by two bilingual researchers and translated back into German by two other bilingual researchers. The wording was improved based on feedback from Dr. George Lewith. The final English version is presented in the appendix Table 1.

Optimism and pessimism

Both concepts were assessed using the German version of the Life Orientation Test – Revised[26]. The questionnaire consists of six self-report items (plus four filler items), each rated on a five-point Likert scale ranging from 0 (strongly disagree) to 4 (strongly agree). The data were separated into optimism and pessimism scores, as recommended by Glaesmer et al. (2008). Each score can range from 0 to 12, with higher values indicating either higher optimism or pessimism. In the present study, the Cronbach's alpha was acceptable for optimism (0.72) and questionable for pessimism (0.60).

Depression

The Patient Health Questionnaire (PHQ-9) is widely used as a standard instrument for diagnosing depression in primary care and is considered well validated[27]. The PHQ-9 consists of nine questions operating according to the modified Diagnostic and Statistical Manual, Fourth Edition criteria. Patients indicated their answers on a scale ranging from 0 (not at all) to 3 (almost every day)[28]. A higher score indicates a higher severity of depressive symptoms (ranging between 0 and 27). In the present study, the internal consistency of the scale was excellent (Cronbach's alpha = 0.90).

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Resilience

As a protective personality factor, resilience was appraised using the Resilience Scale[29]. In the current study, we used the German Version Resilienzskala (RS-11)[30]. The instrument consists of eleven questions, with seven response alternatives ranging from 1 (disagree) to 7 (fully agree). Higher scores indicate higher resilience. In the present study, the internal consistency was excellent (Cronbach's alpha = 0.90).

Body-Efficacy Expectation

To measure beliefs concerning bodily coping capabilities, the Body-Efficacy Expectation (BEE) questionnaire[31] was used. The instrument consists of five items using a four-point Likert scale ranging from 1 (not true at all) to 4 (exactly true). A higher score indicates stronger beliefs in one's bodily coping capabilities. In the present study, the Cronbach's alpha was 0.75, indicating acceptable internal consistency.

Perceived Sensitivity to Medicines

To assess the individual's sensitivity to medicines, we used the Perceived Sensitivity to Medicines scale (PSM)[32]. The PSM is considered a reliable and valid measure composed of five self-report questions to assess perceived sensitivity to the potential adverse effects of medicines. Responses are scored on a five-point Likert scale, and patients' item scores are summed to provide a total score ranging between 5 and 25. Higher scores point towards a high perceived sensitivity to the potential adverse effects of medicines. In the current study, the scale showed excellent internal consistency (Cronbach's alpha = 0.94).

Neuroticism and openness to experience

To assess the distinctiveness of neuroticism and openness to experience among patients, we used the corresponding subscales of the NEO Five-Factor Inventory (NEO-FFI)[33]. To that end, 23 questions were extracted from the 60-item NEO-FFI. The instrument

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

utilizes a five-point Likert response format, from 1 (strongly refuse) to 5 (strongly agree). A higher score indicates a higher value for neuroticism and openness to experience. In the present study, the Cronbach's alpha was 0.85 for neuroticism and 0.72 for openness to experience, which indicates good to acceptable internal consistency.

Acupuncture Expectancy Scale

Mao and colleagues developed the AES to measure patients' expected response from acupuncture[19]. The scale was developed for only one clinical intervention (i.e., acupuncture). The instrument consists of four items. The answers are given using a five-point Likert scale ranging from 1 (not at all agree) to 5 (completely agree). A higher score points towards higher expectancies. In the present study, the internal consistency of the scale was considered good (Cronbach's alpha = 0.88).

Statistical Analyses

All data analyses were executed using SPSS (version 22, SPSS Inc., Chicago, Illinois, USA). As a first step, we used the nine items of the first ETS version to explore homogeneity and diversity between items. We conducted a descriptive analysis of the data to detect possible floor or ceiling effects and assess the distribution of the data. Internal consistency was examined through reliability analyses (Cronbach's alpha), with the corrected item-total correlation and the Cronbach's alpha if the item is deleted. Retest reliability for the nine items was assessed by the intraclass correlation coefficient (ICC). We set for each item a minimal acceptable ICC of .60 and for the total score an ICC of .80. Three criteria guided the decision to keep an item for the final version: 1) it should have a high corrected item-total correlation and no low internal consistency if deleted, 2) the item should not overlap too strongly in content with another item that might be included, and 3)

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

the item should not contribute to ceiling effects in the final version of the scale, which means items with lower values were preferred.

In a second step, we used the reduced version of the scale to generate a total sum score of the five items (ranging from 5 to 20). To examine divergent and convergent validity, correlations between the ETS sum score and the other measures were calculated. The selection of measures was based on theoretical assumptions: With another measure of expectations (AES) we hypothesised very high correlations, since three of the five items of the ETS cover similar topics as the AES (coping, disappearance of complaints, energy) even though instruction and response options differ. We assumed a very high correlation between the ETS and the most strongly related construct (correlation about .70; AES) and a moderate correlation with strongly related constructs (correlation about .30; LOT-R optimism, inverse with LOT-R pessimism). Optimism can be viewed as a trait characteristic of a person with high stability over time and situations. Optimism is defined as "the extent to which people hold generalized favourable expectancies for their future" (Carver et al, 2010, p. 879)[34]. We included optimisms and the counterpart pessimism to assess the overlap between expectation and this personality trait.

Smaller correlations with less related constructs (correlation about .20; PHQ-9; RS-11, BEE) were assumed. Explanatory styles (i.e. expectation about future events) are associated with depressed mood with similar correlations [35]. Self-efficacy is also a construct at a general level (i.e., "Perceived self-represents an optimistic sense of personal competence […]"; Scholz et al., 2002; p. 342)[36]. If self-efficacy is related to a specific behaviour or problem, it captures the strength of a belief to cope in a situation successfully (for example Pain Self-Efficacy Questionnaire; PSEQ)[37]. In our study we used the Body Efficacy Expectation (BEE) since this scale assess the "conviction that one's body is able to

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

heal and take care of itself by dealing with pathogens and other health-threatening demands on its own." (Schützler & Witt, 2013; p. 2). Resources for health also capture positive beliefs on how to deal with a difficult situation in life. We hypothesized for both dimension only low correlations since such resources are rather general and neither related to a specific disorder nor to a specific time-point in life. Finally we assumed no correlation with unrelated constructs (PSM, NEO-FFI neuroticism, NEO-FFI openness to experience), however we wanted to assess these dimension since the ETS might be used in upcoming placebo / nocebo research and several studies have shown that these dimensions are possibly related to placebo / nocebo responses [38 39].

To test our assumption of one general factor, an explorative factor analysis using a varimax rotation, an Eigenvalue of more than one and the Scree test was used with the 5 ETS items to determine the number of underlying factors. In addition, we conducted a confirmatory factor analysis with an independent sample of chronic low back pain patients, which had been included in a randomized controlled trial (registration number DRKS-ID: DRKS00010191). These patients completed the ETS. The data were used to test the single-factor structure. The best model used correlated error terms between item 2 and item 5 and is shown in Figure 1. Multiple goodness of fit tests[40] were used to evaluate the model, including the comparative fit index (CFI;[41]), the normed fit index (NFI;[40]), the goodness of fit index (GFI;[42]) and the root mean squared error of approximation (RMSEA). A CFI greater than 0.90 indicates a good fit to the data[43]. An NFI and GFI greater than 0.90 indicates a good fit to the data[44]. A RMSEA with values of less than 0.08 indicates a good fit to the data[45], whereas values greater than 0.10 suggest strongly that the model fit is unsatisfactory. The Chi-square goodness of fit test can be considered a general test for the acceptability of the model; a statistically significant χ^2 indicates that a significant proportion

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

of variance remains unexplained by the model[40]. Confirmatory factor analysis was conducted with the Analysis of Moment Structures (AMOS) software version 25 which is part of the SPSS package.

Results

Sample characteristics

Three-quarters of the patients were female, one-third had a high school degree or higher vocational training, and approximately 70% were currently employed. Approximately two-thirds of the surveys were conducted with patients in Switzerland, and one-third came from Germany. The characteristics of patients with current pain are displayed in Table 1. Descriptive information about the scales used in the study is presented in the appendix Table 2.

Insert Table 1 about here

First version of the ETS

In an initial analysis, we included nine items of the ETS. Descriptive results and results from scale analyses are presented in the appendix Table 3. In general, the standard deviations showed large variances, with the exception of item 6 (“The treatment [acupuncture] is in general effective”) and item 7 (“I myself have positive expectations about the treatment [acupuncture]”). The mean value of item 6 was relatively high, which can be

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

considered an indicator of a ceiling effect. Therefore, we decided to exclude item 6 (i.e., general expectations about the effectiveness of the treatment [acupuncture]) and item 7 (i.e., personal expectation of a treatment) from the final ETS.

The item-total correlation and the Cronbach's alpha if an item is deleted are indicators for the coherence of the single item meaning with the total scale (see appendix Table 4). The corrected item-total correlation did not give any strong reason for the exclusion of items. However, there were items with very high total correlations between items (item 2 and item 9). Item 2 ("I expect my complaints will be much better because of the treatment [acupuncture]") and item 9 ("I expect that after the treatment [acupuncture], my complaints will be considerably better") were very similar in terms of their wording; therefore, we decided to keep only one of the items for the final version. The reason for this decision was that the ICC value for item 2 was lower than that of item 9. In addition, the mean value for item 9 was lower than the mean of item 2, which consequently can contribute to a larger variation in the final scale because ceiling effects can be avoided. The decision to exclude item 1 was based on content. Our pilot study indicated problems with understanding the meaning of "clearly reduced". Furthermore, item 1 ("I expect the treatment [acupuncture] will clearly reduce my complaints") and item 4 ("I expect the treatment [acupuncture] will make my complaints disappear") covered similar topics. The factor loading confirmed our earlier decision to eliminate item 6 ("The treatment [acupuncture] is in general effective") because the factor loading was relatively low. All items of the first version contributed to one single factor with 65.51% of explained variance. The 9-item ETS version total score is normally distributed. For the 5-item ETS version and the AES items, there were a substantial number of subjects with either the lowest or the highest score on the scale.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

The final version of the ETS consists of five items (see Table 2): the previous item 3 ("I expect the treatment [acupuncture] will help me to cope with my complaints") covers coping ability; item 4 ("I expect the treatment [acupuncture] will make my complaints disappear") covers total absence of complaints; item 5 ("I expect the treatment [acupuncture] will improve my energy") covers an energy increase; item 8 ("I expect the treatment [acupuncture] will improve my physical performance") covers an improvement of physical functioning; and item 9 ("I expect that after the treatment [acupuncture] my complaints will be considerably better") covers a considerable decrease in symptoms. To summarize, the ETS captures an expected decrease in symptoms, an expected increase in energy and expected overall wellbeing after a treatment.

Insert Table 2 about here

The Cronbach's alpha of the final version was 0.90, which is an excellent value for a very short scale (Table 2). Retest reliability was excellent for the five items. All items contributed substantially to the final scale, and the factor loading indicated a single factor structure with 72.33% of explained variance.

Convergent and divergent validity

We expected a moderate correlation between the ETS and the LOT-R optimism and an inverse correlation with pessimism. However, these assumptions were not confirmed by our findings (Table 3). Interestingly, higher pessimism was moderately associated with higher expectations, both for the ETS and the AES. The ETS showed a small but significant

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

correlation with the PHQ-9, which is consistent with our assumption. The correlation between the ETS and the RS-11 or the BEE was close to zero, which is unexpected according to our assumptions. However, we found a moderate correlation between the ETS and the PSM, for which higher expectations were associated with a higher sensitivity to medication. The shared aspect of this association is the responsiveness to a medical treatment. Personality traits (NEO-FFI) were not associated with the ETS, which is consistent with our assumptions. The ETS score has a very high correlation (> .90) with the AES, which can be considered a benchmark measure for acupuncture expectations.

Insert Table 3 about here

Confirmatory factor analysis

In addition, we conducted a CFA with the data from the 439 ETS questionnaires mentioned above with the model presented in Figure 1. The factor loadings were between 0.609 (Item 2) and 0.796 (Item 4). We found an acceptable model fit in the Chi-square statistics (Chi-square = 5.859, d.f. = 4, p = .210), indicating that the model is able to explain the data structure in general. The CFI was .998 in our model, indicating a very good model fit. Similarly, the NFI (.993) and GFI (.995) confirm the excellent model fit. The RMSEA (.033) also suggests that the model with one single factor explains the data very well. The good internal consistency of the ETS in this sample (Cronbach's alpha = .836) confirmed the findings of the scale development.

Insert Figure 1 about here

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Discussion

The ETS is a well-validated and brief 5-item scale for measuring patient expectations, with excellent test-retest properties. We were able to overcome ceiling effects, which had been a limitation of earlier measures[19]. This improvement might result from using only four response categories instead of a larger scale with rather inadequate categories (namely, “definitely disagree”). The ETS has excellent measurement properties concerning Cronbach’s alpha; retest reliability over one week and single factor structure were replicated in a clinical sample. The ETS can be used for research and clinical purposes equally.

The ETS could be adapted to other clinical situations and treatments beyond acupuncture, which would be an important step towards implementing treatment expectations as standardized working mechanisms among a variety of patient populations. Pain disorders, mental disorders and functional symptoms might be the most appropriate fields, because clinical research indicates the high relevance of expectations for treatment outcomes.

Another notable issue that deserves discussion is the moderate correlation between high pessimism and high expectations for both the ETS and the AES. Our initial assumption was that we would find a negative association between both expectation measures and pessimism. This assumption was not confirmed. Furthermore, we found no correlation between optimism and the ETS in our study, which is contrary to the underlying theoretical framework of Bowling et al. in this study. The findings about negative cognitions (pessimism) are consistent with our result of a small but significant correlation of the ETS and depression (PHQ-9). A possible explanation for both findings might be that in pain

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

patients, depression might be associated with more-severe medical symptoms that might lower patients' expectations in some cases depending upon treatment history.

Limitations

There are limitations to our study. First, although collection of data via a web-based survey holds many advantages, there are also downsides to this method because recruitment is done using an unstructured approach. Because the ETS questionnaire was also validated in a paper-based version in a clinical population, this limitation is of minor importance. Second, to enhance the external validity of our study, other clinical populations should be investigated to prove the validity of our scale across clinical conditions. Thus far, the application appears feasible and valid for patients with pain. Third, there might be clinical intervention in situations in which patients might have a very vague idea about procedures and outcomes. We did not collect data from a clinical pain population *prior* to their acupuncture treatment. The patients from our sample provided information about their previous experience with acupuncture, but such an assessment could be done more comprehensively to capture the underlying beliefs and experiences of patients and their relevance for expectations.

Conclusion

The ETS is a short and validated measure that can contribute to the understanding of patient expectations for treatment outcomes. The field of acupuncture served as example to develop the ETS, but it could be easily adapted for other treatments and clinical contexts. The ETS fills a gap by providing a strong and flexible measure that can serve as a basis for upcoming predictor analyses of treatment expectations in clinical studies.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Funding

This study was funded by the Swiss National Science Foundation (Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung), grant No. 105319_159833.

Declaration of competing interests

The present study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authorship and contributorship

JB was involved in the conception and execution of the scale and survey, the analysis and interpretation of data, the writing of the first draft and the revision of the manuscript. AK was involved in the analysis and interpretation of data and the writing of the manuscript. SL was involved in programming and conducting the survey, managing and analysing the data, and revising the manuscript. CW was involved in the conception of the scale and survey, the interpretation of the data and the revision of the manuscript.

License for publication statement

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v)

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above.

Data-sharing statement

Date are available from the first author on request.

Acknowledgement

We would like to thank Dr. George Lewith, who made important suggestions on the phrasing of questions in our scale.

References

1. Peerdeman KJ, van Laarhoven, Antoinette IM, Keij, Sascha M, Vase, Lene, Rovers, Maroeska M, Peters, Madelon L, Evers, Andrea WM. Relieving patients' pain with expectation interventions: a meta-analysis. *Pain* 2016;157(6):1179-91.
2. Kirsch IE. How expectancies shape experience: American Psychological Association 1999.
3. Bialosky JE, Bishop, Mark D, Cleland, Joshua A. Individual expectation: an overlooked, but pertinent, factor in the treatment of individuals experiencing musculoskeletal pain. *Physical therapy* 2016;90(9):1345-55.
4. Bishop FL, Yardley, Lucy, Prescott, Philip, Cooper, Cyrus, Little, Paul, Lewith, George T. Psychological covariates of longitudinal changes in back-related disability in patients undergoing acupuncture. *The Clinical journal of pain* 2015;31(3):254.
5. Iles RA, Davidson, Megan, Taylor, Nicholas F, O'Halloran, Paul. Systematic review of the ability of recovery expectations to predict outcomes in non-chronic non-specific low back pain. *Journal of occupational rehabilitation* 2009;19(1):25-40.
6. Mondloch MV, Cole DC, Frank JW. Does how you do depend on how you think you'll do? A systematic review of the evidence for a relation between patients' recovery expectations and health outcomes. *Canadian Medical Association Journal* 2001;165(2):174-79.
7. Linde K, Witt, C. M., Streng, A., Weidenhammer, W., Wagenpfeil, S., Brinkhaus, B., Willich, S. N., Melchart, D. The impact of patient expectations on outcomes in four randomized controlled trials of acupuncture in patients with chronic pain. *Pain* 2007;128(3):264-71. doi: 10.1016/j.pain.2006.12.006
8. Foster NE, Thomas, E., Hill, J. C., Hay, E. M. The relationship between patient and practitioner expectations and preferences and clinical outcomes in a trial of exercise

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

and acupuncture for knee osteoarthritis. *European journal of pain* 2010;14(4):402-09. doi: 10.1016/j.ejpain.2009.06.010
9. Sherman KJ, Cherkin, D. C., Ichikawa, L., Avins, A. L., Delaney, K., Barlow, W. E., Khalsa, P. S., Deyo, R. A. Treatment expectations and preferences as predictors of outcome of acupuncture for chronic back pain. *Spine* 2010;35(15):1471-7. doi: 10.1097/BRS.0b013e3181c2a8d3
10. Barth J, Schafroth, Larissa, Witt, Claudia M. Overlap and differences between patient and provider expectations for treatment outcomes: the case of acupuncture. *The Journal of Pain* 2016;17(6):685-93.
11. Haanstra TM, Tilbury C, Kamper SJ, et al. Can optimism, pessimism, hope, treatment credibility and treatment expectancy be distinguished in patients undergoing total hip and total knee arthroplasty? *PloS one* 2015;10(7):e0133730.
12. Crow R, Gage H, Hampson S, et al. The role of expectancies in the placebo effect and their use in the delivery of health care: A systematic review. *Health Technology Assessment* 1999;3(3):1-96. doi: https://doi.org/10.3310/hta3030
13. Constantino MJ, Ametrano RM, Greenberg RP. Clinician interventions and participant characteristics that foster adaptive patient expectations for psychotherapy and psychotherapeutic change. *Psychotherapy* 2012;49(4):557-69. doi: 10.1037/a0029440
14. Bowling A, Rowe, G., Lambert, N., Waddington, M., Mahtani, K. R., Kenten, C., Howe, A., Francis, S. A. The measurement of patients' expectations for health care: a review and psychometric testing of a measure of patients' expectations. *Health technology assessment* 2012;16(30):1-509. doi: 10.3310/hta16300
15. Haanstra TM, van den Berg T, Ostelo RW, et al. Systematic review: do patient expectations influence treatment outcomes in total knee and total hip arthroplasty? *Health and Quality of Life Outcomes* 2012;10(1):152.
16. van Hartingsveld F, Ostelo RW, Cuijpers P, et al. Treatment-related and patient-related expectations of patients with musculoskeletal disorders: a systematic review of published measurement tools. *The Clinical Journal of Pain* 2010;26(6):470-88.
17. Zywiel MG, Mahomed A, Gandhi R, et al. Measuring expectations in orthopaedic surgery: a systematic review. *Clinical Orthopaedics and Related Research®* 2013;471(11):3446-56.
18. Prady SL, Burch, J., Vanderbloemen, L., Crouch, S., MacPherson, H. Measuring expectations of benefit from treatment in acupuncture trials: A systematic review. *Complementary therapies in medicine* 2015;23(2):185-99. doi: 10.1016/j.ctim.2015.01.007
19. Mao JJ, Armstrong, Katrina, Farrar, John T, Bowman, Marjorie A. Acupuncture expectancy scale: development and preliminary validation in China. *Explore: The Journal of Science and Healing* 2007;3(4):372-77.
20. Jones SM, Lange, Jane, Turner, Judith, Cherkin, Dan, Ritenbaugh, Cheryl, Hsu, Clarissa, Berthoud, Heidi, Sherman, Karen. Development and validation of the EXPECT questionnaire: assessing patient expectations of outcomes of complementary and alternative medicine treatments for chronic pain. *The Journal of Alternative and Complementary Medicine* 2016;22(11):936-46.
21. Constantino MJ, Arnkoff DB, Glass CR, et al. Expectations. *Journal of clinical psychology* 2011;67(2):184-92. doi: 10.1002/jclp.20754 [published Online First: 2010/12/04]

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

22. Faller H, Vogel H, Bosch B. Erwartungen von Rehabilitanden hinsichtlich der Methoden und Ergebnisse ihrer Rehabilitation-Eine kontrollierte Studie mit Rückenschmerz-und onkologischen Patienten. *Die Rehabilitation* 2000;39(04):205-14.
23. Kravitz RL. Measuring patients' expectations and requests. *Annals of internal medicine* 2001;134(9 Pt 2):881-8.
24. Schulte D. Messung der Therapieerwartung und Therapieevaluation von Patienten (PATHEV). *Zeitschrift für Klinische Psychologie und Psychotherapie* 2005;34(3):176-87.
25. Eibner F, Barth J, Helmes A, et al. Variations in subjective breast cancer risk estimations when using different measurements for assessing breast cancer risk perception. *Health, Risk & Society* 2006;8(2):197-210.
26. Glaesmer H, Hoyer, Jürgen, Klotsche, Jens, Herzberg, Philipp Yorck. Die deutsche version des Life-Orientation-Tests (LOT-R) zum dispositionellen Optimismus und Pessimismus. *Zeitschrift für Gesundheitspsychologie* 2008;16(1):26-31.
27. Löwe B, Unützer, Jürgen, Callahan, Christopher M, Perkins, Anthony J, Kroenke, Kurt. Monitoring depression treatment outcomes with the patient health questionnaire-9. *Medical care* 2004;42(12):1194-201.
28. Ell K, Unützer, Jurgen, Aranda, Maria, Sanchez, Kathleen, Lee, Pey-Jiuan. Routine PHQ-9 depression screening in home health care: Depression prevalence, clinical and treatment characteristics, and screening implementation. *Home health care services quarterly* 2006;24(4):1-19.
29. Wagnild GM, Young, Heather M. Development and psychometric evaluation of the Resilience Scale. *Journal of Nursing Measurement* 1993
30. Schumacher J, Leppert, Karena, Gunzelmann, Thomas, Strauß, Bernhard, Brähler, Elmar. Die Resilienzskala–Ein Fragebogen zur Erfassung der psychischen Widerstandsfähigkeit als Personmerkmal. *Zeitschrift für Klinische Psychologie, Psychiatrie und Psychotherapie* 2005;53:16-39.
31. Schutzler L, Witt, Claudia M. Body-Efficacy Expectation: Assessment of Beliefs concerning Bodily Coping Capabilities with a Five-Item Scale. *Evidence-based complementary and alternative medicine : eCAM* 2013;2013:152727. doi: 10.1155/2013/152727
32. Horne R, Faasse, Kate, Cooper, Vanessa, Diefenbach, Michael A, Leventhal, Howard, Leventhal, Elaine, Petrie, Keith J. The perceived sensitivity to medicines (PSM) scale: an evaluation of validity and reliability. *British journal of health psychology* 2013;18(1):18-30.
33. Borkenau P, Ostendorf, Fritz. NEO-Fünf-Faktoren Inventar:(NEO-FFI); nach Costa und McCrae: Göttingen: Hogrefe 1993.
34. Carver CS, Scheier MF, Segerstrom SC. Optimism. *Clin Psychol Rev* 2010;30(7):879-89. doi: 10.1016/j.cpr.2010.01.006 [published Online First: 2010/02/23]
35. Peterson C, Vaidya RS. Explanatory style, expectations, and depressive symptoms. *Personality and individual differences* 2001;31(7):1217-23.
36. Scholz U, Doña BG, Sud S, et al. Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment* 2002;18(3):242.
37. Nicholas MK. The pain self-efficacy questionnaire: taking pain into account. *European Journal of Pain* 2007;11(2):153-63.
38. Rosen A, Yi J, Kirsch I, et al. Effects of subtle cognitive manipulations on placebo analgesia - An implicit priming study. *European Journal of Pain* 2017;21(4):594-604. doi: 10.1002/ejp.961 [published Online First: 2016/10/18]

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

39. Kelley JM, Lembo AJ, Ablon JS, et al. Patient and practitioner influences on the placebo effect in irritable bowel syndrome. *Psychosom Med* 2009;71(7):789-97. doi: 10.1097/PSY.0b013e3181acee12 [published Online First: 2009/08/08]
40. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. *Psychological bulletin* 1980;88(3):588.
41. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull* 1990;107(2):238-46. [published Online First: 1990/03/01]
42. Jöreskog KG, Sörbom D. LISREL 8: Structural equation modeling with the SIMPLIS command language: Scientific Software International 1993.
43. Kline RB. Principles and practice of structural equation modeling: Guilford publications 2015.
44. Marsh HW, Balla JR, McDonald RP. Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological bulletin* 1988;103(3):391.
45. Browne MW, Cudeck R. Alternative ways of assessing model fit. In. KA, Bollen, JS Long,(Eds.), Testing structural equation models (pp. 136-162): Newbury Park, CA: Sage, 1993.

Tables and Figures

Table 1. Sample characteristics of subjects with current pain (N = 102)

Table 2. Scale and factor analysis of the 5-item ETS for subjects with current pain (N = 102)

Table 3. Convergent and divergent validity of the 5-item version of the ETS for subjects with current pain (N = 102). Pearson Correlation, Significance level (2-tailed), N of patients

Figure 1. Confirmatory factor analysis with factor loadings (N = 439) of the final 5-item ETS version in patients with chronic low back pain

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 1. Sample characteristics of subjects with current pain (N = 102)

	Frequencies // Mean (SD)	Percentage
Age	49.16 (14.04)	
Gender		
Female	76	75%
Male	26	25%
Education		
Vocational school	39	38%
Higher school certificate or Higher vocational training	29	28%
University	34	33%
Employment status		
Working	72	71%
Retired	10	10%
Other	20	20%
Country of residence		
Switzerland	68	67%
Germany	34	33%
Acupuncture experience		
Current	13	13%

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Earlier	51	50%
None	38	37%
Success of acupuncture [range 1–10]*	6.5 (2.80)	
Reason for treatment**		
Pain	55	86%
Mental health	13	20%
Other	24	38%
General Health***	3.21 (.92)	
Pain characteristics		
Intensity [range 0–10]****	6.5 (1.92)	
Physician consulted*****	68	67%
Any treatment*****	66	65%
Days of restriction due to pain*****	46.76 (63.74)	
Location******		
Back	58	57%
Neck	44	43%
Knee	21	21%
Head	20	20%
Hip	16	16%
Other	33	32%

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

633

634 *Range from no success (1) to very successful (10) treatment outcome of the last

635 acupuncture treatment

636 **Multiple answers were possible.

637 ***Range from 1 = excellent to 5 = bad

638 ****0 indicating no pain and 10 indicating maximum pain

639 ******In the last 6-month period

640 ******Multiple answers were possible.

641

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 2. Scale and factor analysis of the 5-item ETS for subjects with current pain (N = 102)

Item	Mean (SD)	Item-total correlation corrected	Cronbach`s alpha if item deleted	Factor loading	Retest ICC [95%-CI]
Item 3	2.38 (1.03)	.727	.887	.826	.697 [.517, .818]
Item 4	1.97 (.96)	.736	.883	.834	.724 [.556, .835]
Item 5	2.41 (1.0)	.776	.875	.865	.840 [.732, .907]
Item 8	2.33 (.83)	.817	.869	.891	.749 [.592, .851]
Item 9	2.52 (.90)	.737	.883	.834	.836 [.725, .904]
			Cronbach's α	% of variance	
ETS 5			.901	72.325	.856 [.757, .917]

Item 3: I expect the treatment [acupuncture] will help me to cope with my complaints.

Item 4: I expect the treatment [acupuncture] will make my complaints disappear.

Item 5: I expect the treatment [acupuncture] will improve my energy.

For peer review only

Item 8: I expect the treatment [acupuncture] will improve my physical performance.

Item 9: I expect that after the treatment [acupuncture], my complaints will be considerably better.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 3. Convergent and divergent validity of the 5-item version of the ETS for subjects with current pain (N = 102). Pearson Correlation, Significance level (2-tailed), N of patients

Correlation Significance level Number of patients	ETS	LOT-R Optimism	LOT-R Pessimism	PHQ-9	RS-11	BEE	PSM	NEO-FFI Neuroticism	NEO-FFI Openness to experience
ETS	1 102								
LOT-R Optimism	-.066 .517 99	1 99							
LOT-R Pessimism	.204* .043 99	-.247* .014 99	1 99						
PHQ-9	.233* .020 100	-.567** .000 99	.312** .002 99	1 100					
RS-11	-.073 .474 99	.546** .000 99	-.460** .000 99	-.572** .000 99	1 99				

For peer review only

Expectation for Treatment Scale (ETS)

BEE	.032	.259**	-.074	-.176	.407**	1			
	.754	.010	.468	.081	.000				
	99	99	99	99	99	99			
PSM	.344**	-.078	.201*	.306**	-.049	.011	1		
	.000	.445	.046	.002	.631	.912			
	100	99	99	100	99	99	100		
NEO-FFI	.104	-.578**	.467**	.631**	-.682**	-.284**	.238*	1	
Neuroticism	.316	.000	.000	.000	.000	.005	.020		
	95	95	95	95	95	95	95	95	
NEO-FFI	-.072	.297**	-.167	-.090	.302**	.122	.132	-.130	1
Openness to	.492	.004	109	.386	.003	.241	.205	.211	
experience	94	94	94	94	94	94	94	94	94
Cronbach's Alpha	.901	.727	.608	.900	.900	.757	.940	.854	.726

**Correlation is significant at the 0.01 level (2-tailed).

*Correlation is significant at the 0.05 level (2-tailed).

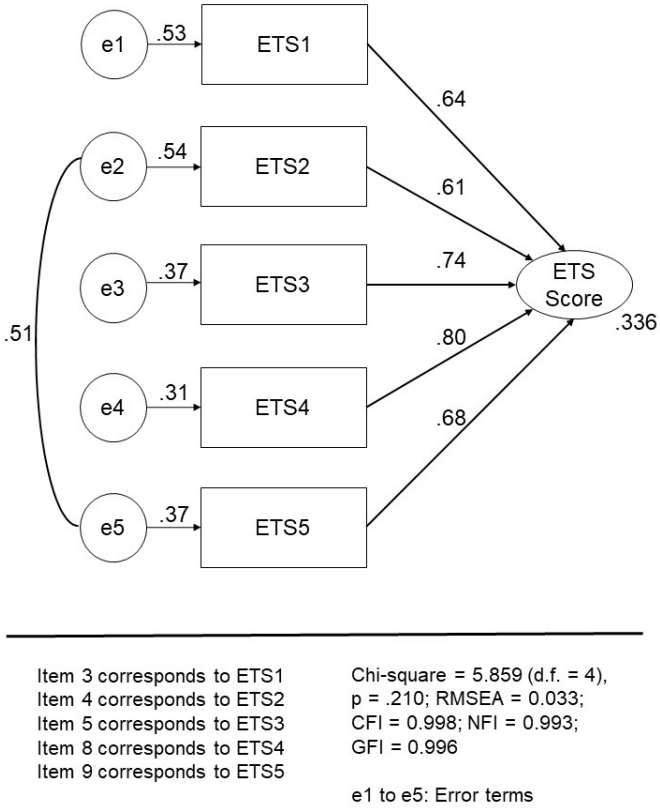
ETS = Expectation for Treatment Scale; LOT-R = Life-Orientation Test; PHQ-9 = Patient Health Questionnaire-9; RS-11 = Die Resilienzskala - Ein Fragebogen zur Erfassung der psychischen Widerstandsfähigeit als Personenmerkmal; BEE = Body-Efficacy Expectation; PSM = Perceived Sensitivity to Medicines Scale; NEO-FFI = NEO Five-Factor Inventory

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Figure 1. Confirmatory factor analysis with factor loadings (N = 439) of the final 5-item ETS version in patients with chronic low back pain

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml



Confirmatory factor analysis with factor loadings (N = 439) of the final 5-item ETS version in patients with chronic low back pain.

Appendix (Barth et al.; ETS)

APPENDIX Table 1. Description of the 5 final items in English and German

Table 1. Description of the 5 final items in English and German

Description: The overarching construct of ETS is treatment-related outcome expectation of patients. Outcomes of interest for patients are coping (item 1), vitality (item 3) and physical health (item 4). Two items (items 2 and 5) capture the reduction of patient complaints, either an absence or a considerable reduction thereof. The scale can be adapted to specific treatments of interest. Instead of the term “treatment”, a specific term can be used (e.g., acupuncture, counselling, or physiotherapy). We used the term complaints [German Beschwerden] as previously used by other authors as in the symptom checklist (SCL-90-R)[46]. Therefore, any type of complaint can be captured and allow the use of this instrument across patient groups and settings. The term complaints can be adapted if needed to the clinical situation if one specific target symptom is predominant (e.g., pain, depression, or nausea). The scale should be used in a clinical setting by patients with sufficient German or English language skills.

For peer review only

English version

There are several statements below that capture your expectations about the [acupuncture] treatment. Please indicate to what extent these statements apply to you personally. **There are no right or wrong answers. We are only interested in your current personal thoughts.**

*Please select for each statement **one** response.*

1. I expect the treatment [acupuncture] will help me to cope with my complaints.

partially disagree	partially agree	agree	definitely agree
❑	❑	❑	❑

2. I expect the treatment [acupuncture] will make my complaints disappear.

partially disagree	partially agree	agree	definitely agree
❑	❑	❑	❑

3. I expect the treatment [acupuncture] will improve my energy.

partially disagree	partially agree	agree	definitely agree
❑	❑	❑	❑

4. I expect the treatment [acupuncture] will improve my physical performance.

partially disagree	partially agree	agree	definitely agree
❑	❑	❑	❑

5. I expect that after the treatment [acupuncture], my complaints will be considerably better.

partially disagree	partially agree	agree	definitely agree
❑	❑	❑	❑

*We used the term acupuncture in our study. The general version of ETS uses the term treatment.

For peer review only

German version

Im Folgenden finden Sie einige Aussagen darüber, was man von einer Behandlung [Akupunktur] erwarten kann. Geben Sie bitte an, inwieweit die einzelnen Aussagen für Sie persönlich zutreffen. **Es gibt keine richtigen und falschen Antworten. Uns interessiert nur, was Sie persönlich denken.**

*Bitte wählen Sie für jede Aussage **eine** Antwort aus.*

1. Ich erwarte, dass ich durch die Behandlung [Akupunktur] mit meinen Beschwerden besser umgehen kann.

trifft eher nicht zu	trifft eher zu	trifft sehr zu	trifft völlig zu
❑	❑	❑	❑

2. Ich erwarte, dass meine Beschwerden durch die Behandlung [Akupunktur] verschwinden.

trifft eher nicht zu	trifft eher zu	trifft sehr zu	trifft völlig zu
❑	❑	❑	❑

3. Ich erwarte, dass sich meine Energie durch die Behandlung [Akupunktur] verbessert.

trifft eher nicht zu	trifft eher zu	trifft sehr zu	trifft völlig zu
❑	❑	❑	❑

4. Ich erwarte durch die Behandlung [Akupunktur] eine verbesserte körperliche Leistungsfähigkeit.

trifft eher nicht zu	trifft eher zu	trifft sehr zu	trifft völlig zu
❑	❑	❑	❑

5. Ich erwarte, dass sich nach der Behandlung [Akupunktur] meine Beschwerden deutlich verbessern.

trifft eher nicht zu	trifft eher zu	trifft sehr zu	trifft völlig zu
❑	❑	❑	❑

*Wir haben den Begriff Akupunktur in unserer Studie verwendet. Die allgemeine Version des ETS verwendet den Begriff Behandlung.

For peer review only

Scoring of the questionnaire

The values from the 5 single items are summed to build the ETS sum score (min 5, max 20). If one item has a missing value, the sum score can be calculated. Therefore, the values of the remaining 4 items are summed, divided by 4 and multiplied by 5. However, multiple imputation procedures to impute the missing value should be preferred over this re-calculation. In the case of more than one missing value, imputation procedures are needed, and no manual recalculation should be considered.

For peer review only

APPENDIX Table 2: Sample characteristics in scales in subjects with current pain (N = 102)

	Median Mean (SD)
ETS (9 items) (range 9–36)	22 22.63 (.64)
ETS (5 items) (range 5–20)	11 11.62 (4.01)
AES (range 4–20)	12.0 11.82 (4.10)
PHQ-9 (range 0–27)	7 8.46 (6.06)
RS-11 (range 11–77)	60 58.36 (11.13)
BEE (range 5–20)	14 13.71 (2.90)
PSM (range 5–25)	11 12.77 (6.13)
LOT-R Optimism (range 0–12)	8 7.93 (2.77)
LOT-R Pessimism (range 0–12)	8 7.93 (2.38)
NEO-FFI Neuroticism (range 0–48)	21 21.08 (8.19)
NEO-FFI Openness to experience (range 0–44)	30 29.84 (5.77)

Expectation for Treatment Scale; AES= Acupuncture Expectancy Scale; LOT-R= Life-Orientation Test; PHQ-9= Patient Health Questionnaire-9; RS-11= Die Resilienzskala - Ein Fragebogen zur Erfassung der psychischen Widerstandsfähigeit als Personenmerkmal; BEE= Body-Efficacy Expectation; PSM= Perceived Sensitivity to Medicines Scale; NEO-FFI= NEO Five-Factor Inventory

For peer review only

Appendix Table 3. Scale and factor analysis of the 9-item ETS for subjects with current pain (N = 102)

Item	Mean (SD)	Item-total correlation corrected	Cronbach`s alpha if item deleted	Factor loading	Retest ICC [95%–CI]
Item 1	2.71 (1.0)	.782	.920	.839	.684 [.499, .810]
Item 2	2.66 (.93)	.816	.918	.863	.725[.557, .836]
Item 3**	2.38 (1.03)	.736	.924	.790	.697 [.517, .818]
Item 4**	1.97 (.96)	.735	.923	.793	.724 [.556, .835]
Item 5**	2.41 (1.0)	.745	.923	.799	.840 [.732, .907]
Item 6	2.97 (.59)	.612	.931	.685	.649 [.450, .787]
Item 7	2.68 (.65)	.744	.924	.806	.658 [.462, .793]
Item 8**	2.33 (.83)	.772	.921	.821	.749 [.592, .851]
Item 9**	2.52 (.90)	.826	.917	.874	.836 [.725, .904]
			Cronbach's α	% of variance	
ETS 9			.930	65.512	.879 [.794, .930]

**marked items are included in the final version of its scale.

Item 1: I expect the treatment [acupuncture] will clearly reduce my complaints.

Item 2: I expect my complaints will be much better because of the treatment [acupuncture].

Item 3: I expect the treatment [acupuncture] will help me to cope with my complaints.

Item 4: I expect the treatment [acupuncture] will make my complaints disappear.

For peer review only

Item 5: I expect the treatment [acupuncture] will improve my energy.

Item 6: The treatment [acupuncture] is in general effective.

Item 7: I myself have positive expectations about the treatment [acupuncture].

Item 8: I expect the treatment [acupuncture] will improve my physical performance.

Item 9: I expect that after the treatment [acupuncture], my complaints will be considerably better.

For peer review only

Appendix Table 4. Between-item correlations for sample with current pain (N = 102). Pearson Correlation and significance (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9
Item 1	1								
Item 2	.840** .000	1							
Item 3	.636** .000	.605** .000	1						
Item 4	.568** .000	.627** .000	.580** .000	1					
Item 5	.533** .000	.611** .000	.664** .000	.637** .000	1				
Item 6	.540** .000	.556** .000	.442** .000	.437** .000	.505** .000	1			
Item 7	.692** .000	.679** .000	.541** .000	.575** .000	.509** .000	.521** .000	1		
Item 8	.528** .000	.571** .000	.673** .000	.664** .000	.762** .000	.470** .000	.631** .000	1	
Item 9	.733** .000	.754** .000	.605** .000	.673** .000	.603** .000	.592** .000	.701** .000	.673** .000	1

Item 1: I expect the treatment [acupuncture] will clearly reduce my complaints.

Item 2: I expect my complaints will be much better because of the treatment [acupuncture].

For peer review only

Item 3: I expect the treatment [acupuncture] will help me to cope with my complaints.

Item 4: I expect the treatment [acupuncture] will make my complaints disappear.

Item 5: I expect the treatment [acupuncture] will improve my energy.

Item 6: The treatment [acupuncture] is in general effective.

Item 7: I myself have positive expectations about the treatment [acupuncture].

Item 8: I expect the treatment [acupuncture] will improve my physical performance.

Item 9: I expect that after the treatment [acupuncture], my complaints will be considerably better.

For peer review only

TROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

	Item No	Recommendation	page
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	8
Methods			
Study design	4	Present key elements of study design early in the paper	9
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	9
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	9
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	n.a.
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	10-13
Bias	9	Describe any efforts to address potential sources of bias	n.a.
Study size	10	Explain how the study size was arrived at	n.a.
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	n.a.
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	14-15
		(*b*) Describe any methods used to examine subgroups and interactions	n.a.
		(*c*) Explain how missing data were addressed	n.a.
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	n.a.
		(*e*) Describe any sensitivity analyses	n.a.
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	15 and Table 2
		(b) Give reasons for non-participation at each stage	n.a.
		(c) Consider use of a flow diagram	n.a.
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	n.a.
		(b) Indicate number of participants with missing data for each variable of interest	n.a.
Outcome data	15*	Report numbers of outcome events or summary measures	n.a.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	n.a.
		(*b*) Report category boundaries when continuous variables were categorized	n.a.
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n.a.
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	19
Discussion			
Key results	18	Summarise key results with reference to study objectives	19
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	21
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	21
Generalisability	21	Discuss the generalisability (external validity) of the study results	21
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	21

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml