QUALITY INSPECTION METHODOLOGY IN DORSOPATHY’S TREATMENT

Chebykin A.V., Supilnikov A.A., Pavlova O.N.

LEU Institution «Institute of Medicine «REAVIZ», Samara, e-mail: casiopeya13@mail.ru

Standardized questionnaires as a method of treatment’s results assessment became routine in clinical trials practice, including vertebrology practice. There were developed series of expert approaches for patients with back pain, which are definitely important for analysis of administered treatment. In our work we implemented in a consistent manner the most often used scales, questionnaires and inventories: SF-36, SF-12, ODI, RDQ, QBPDQ, BPFS, VAS, NRS, CPQG, MPQ, WQL, Macnab, Prolo, LPOS. In order to get an objective estimation of clinical and performance status in patients with vertebrological disorders we carried out general and instrumental methods of analysis. However, none of them allow getting a clear view of the fact to what extent the disease and diagnosis «vertebral osteochondrosis» limit live activity of the real patient, his ability to participate in social, occupational activity and activities of daily living. In accordance with principles of evidence-based medicine we decided to carry out treatment’s results assessment by using the quality of life index. New quality of life evaluation technic and author’s recommendations for data assessment were established.

Keywords: vertebrology, dorsopathy, quality of life control, quality inspection

Treatment quality inspection by using standardized quantitative rating scales and questionnaires became routine in clinical trials practice, including vertebrological and neurologic practice. They are used for matched patient groups’ selection, response to the treatment’s comparison, treatment outcomes’ prognosis and risk groups’ identification [1, 2, 4, 5, 8, 11, 16, 21, 22, 28, 31, 32]. Nowadays there is no unified approach of these methods using in Russia. In this paper we give the ways to systematization of the most commonly used scales and questionnaires, and provide recommendations concerning data estimation.

1. Quality of life estimation

Usually it is used an integrated specification – quality of life (QOL), that comply with CONSORT recommendations (Consolidated Standards of Reporting Trials) [8]. QOL is particularly important for patients with comorbidities, as these conditions can influence on treatment efficacy. It is also important for different trials’ results comparison, carrying-out an economic analysis and clear understanding of the problem in the context of health care service modernization.

Quality of Life questionnaire SF-36 (Short Form) [6, 27, 30] was developed by RAND (Research and Development corporation) as part of Medical Outcome Study, MOS. Later research team published the shipping version of RAND-36™ questionnaire. Questionnaires SF-36 and RAND-36 consists of the same set of questions, but have differences in evaluating «general health» and «pain». It should be taken into account while comparing study results obtained by using questionnaire’s different modifications [27]. Questionnaire SF-36 is not specific for treatment’s results assessment in patients with vertebrogenous disorders. But it is important for QOL evaluation in patients expecting vertebral surgical operations, that’s proved by a number of studies.

In general, SF-36 corresponds with specificity, accuracy, sensitivity and number of questions. There is wide experience of using it among big patients groups. Also SF-36 has advantages in results distribution (mean and standard deviation) in large and varied samplings. This questionnaire has been translated into more than 40 languages. There are also it short versions – SF-12 and SF-8 [30]. Using questionnaire SF-12 in large population-based cohort study where QOL assessment isn’t a primarily endpoint could be a good compromise between quality of the study and time needed for filling and data handling [6].

Oswestry Disability Questionnaire (Oswestry Disability Index (ODI)) was developed in 1980 [10]. Nowadays it is commonly used for disability status assessment in patients with spinal disorders [7, 15]. Version 2.1a of Oswestry Disability Questionnaire is available now; it consists of 10 sections. The maximum score for each section is 5. Oswestry Disability Index is calculated as follows: (total score of the patient/total possible raw score) X 100.

Roland-Morris Disability Questionnaire (RDQ) was published by M. Roland and R. Morris in 1983 [24]. We used this questionnaire for assessment of low back pain influence on disability. RDQ was used for assessment patients with acute and subacute back pain syndrome [9]. Questionnaire consists of 24 questions. Doctor adds up the number of items checked by the patient; the score can therefore vary from 0 to 24. The more the sum is, the more level of disability is. Clinical improvement over time can be graded based on the analysis of serial questionnaire scores; the improvement express as a percentage.

Quebek Back Pain Disability Scale, QB-PDQ [16], was developed by authors’ team in...
1995. It measures the difficulty in performing 20 daily activities by 5-point scale. The item scores are summed for a total score between 0 and 100, with higher numbers representing lower levels of QOL. The set of questions for Quebec Back Pain Disability Scale came out of vast number of signs as a result of factorial analysis, confidence estimation and correlation with regard to sensitivity. The scale authors supposed that this method represents the most accurate changes in patients QOL.

The Back Pain Function Scale of Stratford, BPFS [25], was developed by P. Stratford and L. Riddle in 2000 to evaluation functional ability in patients with back pain. It measures the ability in performing the most common activities (12) by 5-point scale: any of usual housework, recreational or sporting activities, performing heavy activities around home, hobbies, putting shoes or socks, bending, lifting things from the floor, sleeping, standing or sitting for 1 hour, going up 2 stairs, driving for 1 hour. The score strongly correlates with the abovementioned Roland-Morris questionnaire. By comparison with QBPDQ the ODI has the advantages in evaluating patients with low back pain [12]. Roland-Morris Disability Questionnaire and Oswestry Disability Questionnaire are specific to vertebrologists, so they are easy operating and fail-safe [3, 6].

2. Evaluation of present pain intensity

Pain is the subjective symptom that is under study of vertebrology [2]. Most vertebrologists agree that pain relief is the main parameter of benign treatment outcome. Moreover, many patients expect significant or full pain relief after appropriate treatment [14]. Pain’s severity evaluation differs from pain’s influence evaluation on general well-being. Pain’s severity is characterized by the degree of patient distress, whereas pain’s influence is complex term reflecting changes in mental status caused by pain and pain influence on patient’s QOL. Pain’s severity evaluation is enough advanced, while there are a lot of open questions in pain’s influence assessment. So, it’s impossible to divide questionnaires and scales in two groups, such as scales to evaluate only pain’s severity or only QOL.

The simplest, most convenient and commonly used scale for pain’s severity evaluation is visual analog scale – VAS. VAS is usually a horizontal line, 10 cm in length [29]. The patient should mark the point on the line that corresponding to the pain’s severity he experienced. One end of the line is marked «0» that means «no pain», the other end is marked «10» that means «worst possible pain». The VAS score is determined by measuring in millimeters from the left hand end of the line to the point that the patient marks.

Numerical rating scale (NRS) is also widely used for pain’s severity evaluation. It consists of 11 points from 0 (no pain) to 10 (worst possible pain). Its advantages are independence from good eyesight, availability of writing materials and possibility to use them. It can be used even during phone conversation with patient. Scales with pictures of happy and unhappy faces are used for children. VAS and NRS are used for subjective patient pain’s evaluation during examination. In the number of scales pain and QOL are evaluated simultaneously (some of them pay more attention to pain’s influence, whereas others concentrate on QOL). VAS and NRS usability is based on the fact that they could be used for time course pain’s evaluation within 24 hours or a week. Retrospective analysis is not preferable, so pain’s memories could be inaccurate or even aberrant. It should be taken into account that pain’s severity evaluation by using one of the scales (e.g. VAS) is subjective and couldn’t reflect real patient’s condition, especially in terms of anaesthetics influence. So, it is reasonable to use scales with different assessment principles.

While evaluating chronic and recurrent pain syndrome it’s important to assess pain’s severity during definite time interval instead of definite moment as at the clinical visit.

Chronic pain grade questionnaire, CPGQ [29], was developed in 1992 by Von Korff and J. Ormel [29]. Its distinctive feature is measurement of pain’s duration, intensity and pain’s influence on daily activities, rest and work during last month.

McGill Pain Questionnaire (MPQ) [7, 19, 20] was developed in 1975 by R. Melzack at Canadian university. It was translated into several languages. It helps to measure the sensory, affective and other aspects of chronic pain. The questionnaire consists of 11 sensory and 4 affective verbal characteristics: 78 adjectives describing pain are classified into 20 subclasses according to semantic meaning increasing in quantitative terms. After analyzing the questionnaire three pain’s characteristics are determined: sensory, affective and general. MPQ could be used for evaluation of pain’s characteristics changes before and after treatment. The 2 major measures are: the Pain Rating Index (the sum of the scale values of each word chosen or their arithmetic mean) and the number of words chosen. Obtained results could be used not only for pain’s evaluation, but also for patient’s emotional state evaluation. Obtained data are not parametric, but could be used in statistical processing. Nevertheless, MPQ isn’t used very often in vertebrogenic pain syndrome’s studies because of extensive effort and absence of necessity in such detail pain’s
characterization. CPGQ, SF-12 and ODI are found in between scales mainly evaluating QOL and scales evaluating only pain syndrome.

3. Disability examination

Not much attention at literary sources is paid to outcomes’ assessment in the context of professional suitability and possibility of employment [2]. However, these criteria are very important for economic analysis of the health care industry and also for estimation their influence on QOL and treatment satisfaction of patient, employer and doctor. In our opinion, occupational status should be evaluated at the first visit to doctor and after the rehabilitation programme. It is recommended to check the time of disability appeared, rehabilitation period’s duration and disability status (if applicable). For example, SF-36 has questions about limitation of work capability in the social role functioning section. However, the questionnaire doesn’t describe disability status whereas evaluate capabilities to different kinds of activities.

The Work Limitations Questionnaire was published by D. Lerner et al. in 2001 to estimate disability status in patients with chronic pain syndromes [2, 18]. It consists of 24 items combined into 4 subscales:

1. «Time management» contains 5 items that address difficulty handling time and scheduling demands.
2. «Physical demands» includes 6 items that covers a person’s ability to perform job tasks that involve bodily strength, movement, endurance, coordination and flexibility.
3. «Mental-interpersonal demands» includes 9 items addressing cognitive job tasks, and on-the-job social interactions.
4. «Output demands» includes 5 items concerning diminished work quantity and quality. Subscale scores range from 0 (limited none of the time) to 100 (limited all of the time) and represent the reported amount of time in the prior two weeks respondents were limited on-the-job.

It should be noted that disability status is evaluated not only by specifically developed scales like WLQ, but also by most QOL questionnaires as previously mentioned.

4. Disease outcome measures

The important outcome criterion is treatment satisfaction of patient. There are a lot of approaches to quantitative assessment of this value. Some of them contain only several general questions, whereas others are very specialized [2].

Subjective Macnab’s scale is the most mentioned and simplest in use scale. By this scale patient estimate the result of the treatment as «excellent>, «good>, «fair» or «poor».

The Patient Satisfaction Scale was developed in 2002 by T. Morita. It contains questions connected with awareness of treatment, emotional support and treatment efficacy. In general it helps to estimate patient’s satisfaction with medical care at the hospital. According to Byval’tsev V.A. et al., (2011) patient’s satisfaction with treatment consists of many components, so it is impossible to do complete evaluation by one scale. For example, some patients give priority to communication with doctor rather than equipment used during surgery. So, this fact should be noted while using scales described in the paper.

Prolo scale [23] is used for evaluation of patients’ economic and functional status. It was developed by neurosurgeon D. Prolo in 1986 especially for patients who have undergone spine surgeries. Two aspects could be estimated by Prolo scale: economic outcome (with due regard to disability status) and functional outcome (with due regard to patient’s physical activity). The final score is calculated by summing up of two criteria scores: economic and functional status. Score of 9–10 are considered excellent, 7–8 – good, 5–6 – fair and < 4 – poor [5, 23]. It’s not necessary to evaluate economic status of spine surgeries as a part of routine medical practice, but it could be useful for healthcare managers. However, general treatment cost’s calculation also could be one of the surgery’s outcome criteria.

The Low-Back Outcome Scale (LBOS) [11, 25] was published in 1992 for measuring functional treatment outcome in patients with low back pain [13, 26]. Treatment outcomes are estimated as «excellent), «good), «fair» or «poor» according to answers to 13 questions about pain’s intensity, working capacity, capability to active physical and daily activity. So, LBOS helps to evaluate outcomes by taking into account many aspects of patient’s everyday activities. It could be recommended for routine use.

During the course of trials conducting we developed specialized questionnaire «QOL of patient with spine disorder». It was designed with accordance to following general requirements: universality, reliability, repeatability, usability, laconicism, standardization, correspondence with main QOL criteria recommended by World Health Organization (1992), scientific-production association «Medsofteconominform» (2000), Ju.P. Lisiñ’s guidance (2011).

It was decided to develop specialized software program to automatic data processing based on specialized questionnaire «QOL of patient with spine disorder» (Fig. 1).

Algorithms and software program were designed for automatic data processing [4].
Fig. 1. Main window for input, processing, analyzing and information storing of patients with spine disorders

Calculating of relative values including values of abovementioned expert method before and after treatment are represented in Table 1 and 2.

### Table 1

<table>
<thead>
<tr>
<th>Sign</th>
<th>Weighted average</th>
<th>Degree of impact</th>
<th>Relative difference</th>
<th>Mean error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tonomiometriya (before treatment)</td>
<td>-0.0171</td>
<td>3034.5029</td>
<td>-0.0171</td>
<td>0.0004</td>
</tr>
<tr>
<td>Pain’s severity measured by VAS before treatment</td>
<td>-0.0088</td>
<td>47.7401</td>
<td>-0.0088</td>
<td>0.0002</td>
</tr>
<tr>
<td>Muscle strength (by Haribov) before treatment in affected segment</td>
<td>0.0182</td>
<td>3034.5534</td>
<td>0.0182</td>
<td>0.0004</td>
</tr>
<tr>
<td>Integrated QOL before treatment by author’s questionnaire</td>
<td>-0.0039</td>
<td>3034.3997</td>
<td>-0.0039</td>
<td>0.0001</td>
</tr>
<tr>
<td>Duration of disease recurrence before visiting a doctor</td>
<td>-0.0507</td>
<td>66.227</td>
<td>-0.0507</td>
<td>0.0012</td>
</tr>
<tr>
<td>Numeric rating scale before treatment</td>
<td>-0.0165</td>
<td>3034.5621</td>
<td>-0.0165</td>
<td>0.0004</td>
</tr>
<tr>
<td>CPS before treatment</td>
<td>0.0096</td>
<td>3034.5629</td>
<td>0.0096</td>
<td>0.0002</td>
</tr>
<tr>
<td>Oswestry Disability Questionnaire before treatment</td>
<td>0.0103</td>
<td>39.3806</td>
<td>0.0103</td>
<td>0.0003</td>
</tr>
<tr>
<td>Roland-Morris questionnaire before treatment</td>
<td>0.0053</td>
<td>3034.5615</td>
<td>0.0053</td>
<td>0.0001</td>
</tr>
<tr>
<td>Neck Pain and Disability Index (Vernon-Mior) before treatment</td>
<td>0</td>
<td>3034.563</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>R.Watkins score before treatment</td>
<td>-0.0057</td>
<td>3034.5572</td>
<td>-0.0057</td>
<td>0.0001</td>
</tr>
<tr>
<td>McGill Pain Questionnaire (short form) before treatment</td>
<td>0.0011</td>
<td>70.0629</td>
<td>0.0011</td>
<td>0</td>
</tr>
<tr>
<td>Waddell Disability Index before treatment</td>
<td>-0.0127</td>
<td>3034.5629</td>
<td>-0.0127</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

It should be noted that many questionnaires and scales consist of Likert-type questions (american psychologist’s scale) and the respondent is asked to evaluate the level of agreement or disagreement by five levels:

1) strongly disagree;
2) disagree;
3) neither agree nor disagree;
4) agree;
5) strongly agree.
Central tendency and variance could be calculated while processing data obtained by using Likert scale. These values should be considered as median or mode with interquartile range, in other words non-parametric tests should be used. Central limit theorem helps to carry out a parametric analysis [2, 3, 4]. In connection with these facts, it could be recommended to use non-parametric tests for processing data obtained by scales and questionnaires. It could be explained by the fact that many scales describe nominal data, so probability distribution is not always Gaussian distribution. Nowadays there are different software programs that help to perform statistical analysis (i.e. Statistica, StatSoft, Inc).

We used the software package «STATISTICA 6.0» in our work. It helps to figure the data in accordance with Gaussian probability law. Mathematical analysis’ results of obtained data are represented in Fig. 2 for describing basic tendencies. The x-axis represents patients groups before and after treatment, the y-axis represents values of integral QOL index measured in scores.

Represented data show that some patients had high integral QOL index (about 26 scores) before treatment, whereas half of patients (50% percentile rank) had integral QOL index within 10–15 scores which is equivalent of poor quality of life.

As a result of appropriate combination treatment we observed marked increase in integral QOL index among most patients. 50% percentile rank includes values within 17–27 scores which is equivalent of high or fair level of this index.

Scales for evaluating QOL, occupational disability and capability are designed for between-group analysis. Many experts consider that they could be used for individual clinical decision-making. It is necessary to take into account significant variation in scores for each scale. There are two types of significance: statistical and clinical. In statistics, a result is called statistically significant if there is statistical evidence that there is a rather large difference. If there is a statistical significance in evaluating by different scales it doesn’t mean that there is appropriate clinical significance. So, it’s important to determine minimal important change – least significant change for patient. Knowing the minimal important change helps to evaluate results before and after treatment and draw the conclusion concerning the importance of health gain for the patient. So, some experts think that minimal important change is the main value for making personal opinion and clinical decision. Moreover, minimal important change is useful for determining sample size for clinical trial.

Generally, SF-36 is used as a standard for detection minimal important change. According to clinical trial results as a part of VIII International Forum on Primary Care Research on Low Back Pain (Amsterdam, 2006) were determined following minimal important changes: 15 mm for the VAS, 2 scores for the NRS, 5 scores for the Roland Disability Questionnaire, 10 for the Oswestry Disability Index, and 20 for the QBDQ. It was also mentioned that a 30% improvement was considered a useful threshold for identifying clinically meaningful improvement on each of these measures [21].

So, criteria universalization of observation results by using above-described methods in vertebrology helps to objectify and compare treatment outcomes in different clinics and centres. It could simplify professional communication and improve clinical trials quality in Russia.
References


