Outcome Measures Associated with Perceived Stress

9

Deborah Ann Hall

Abbreviations

ACT   Acceptance and commitment therapy
BAI   Beck Anxiety Inventory
BDI   Beck Depression Inventory
CBT   Cognitive behaviour therapy
DSM   Diagnostic and Statistical Manual for Mental Disorders
HADS   Hospital Anxiety and Depression Scale
PROM   Patient-Reported Outcome Measure
PSQ   Perceived Stress Questionnaire
PSS   Perceived Stress Scale
STAI   State-Trait Anxiety Inventory
TRQ   Tinnitus Reaction Questionnaire
TRT   Tinnitus retraining therapy

9.1 Clinical Trial Designs

This chapter describes a number of outcome measures associated with the perceived stress experienced by people with tinnitus and that are used in clinical research. Such measures are used to determine whether an intervention aimed at alleviating tinnitus leads to a meaningful patient benefit. The choice of outcome is one of the

D.A. Hall

National Institute for Health Research Nottingham Biomedical Research Centre, Nottingham, UK

Otology and Hearing Group, Division of Clinical Neuroscience, School of Medicine, University of Nottingham, Nottingham, UK
e-mail: deborah.hall@nottingham.ac.uk

most important fundamental aspects of clinical trial design, but that choice is often driven by the type of clinical trial, in particular whether it is an explanatory or a pragmatic trial (see Tutorial 9.1).

> *The choice of outcome is one of the most important fundamental aspects of clinical trial design.*

Types of clinical trial designs are broadly outlined in Tutorial 9.1. Feasibility and pilot trials typically answer questions about the process and procedures of how a trial might be run, but they don't directly test patient benefit. Explanatory and pragmatic trials do. Explanatory trials are most common in the tinnitus field. They typically answer the question about whether a particular treatment for tinnitus works 'under ideal conditions'. For these trials, outcomes are often condition-specific. A good example of this would be the measurement of the perceived stress associated with tinnitus. Pragmatic trials are broader in scope since they typically answer the question about whether a particular treatment for tinnitus works in everyday clinical practice. For these trials, outcomes tend not to be restricted to any single health condition. Good examples of this would be the measurement of general perceived stress, generalised anxiety or depressive symptoms, or even health-related quality of life.

Tutorial 9.1 Clinical Trial Designs

Clinical trial designs typically fall into four broad levels or categories (Arain et al. 2010; Williams et al. 2015):

Feasibility. A feasibility study tries out pieces of an explanatory clinical trial in order to answer the question about whether that main study can be done. Feasibility is used to test important parameters that are needed to design the main study: (1) scientific basis, (2) process, (3) resources and (4) management.

Pilot. A pilot study is a 'miniature' version of the main study that is run to test whether the components of the main study can all work together. Unlike a feasibility study, a pilot resembles the main study in many respects, including an assessment of the primary outcome. However, hypothesis testing is considered inappropriate because a pilot tends to be underpowered.

Other terms for feasibility and pilot trials are 'proof of concept' or 'Phase I' trials.

Explanatory. An explanatory trial typically answers the question 'can this treatment work under ideal conditions?' It is deliberately designed to give the maximum chance of showing an effect, if one is present. For example, the sample is often a highly selected and homogenous group exhibiting good compliance, the intervention is tightly defined, the comparator may be a

placebo, outcomes are often condition-specific and may include biomarkers as well as condition-specific questionnaires, and the study end point tends to be short term (e.g. 6 weeks).

Pragmatic. A pragmatic trial typically answers the question 'we now know the intervention can work, but how well does it work in everyday clinical practice?' The design aims to test an intervention in a study environment that is closer to real life in terms of sample, intervention, active comparator and outcomes. Outcomes are generic rather than specific and often quality of life questionnaires. Another aim of a pragmatic trial is to ensure that the intervention can be implemented in routine healthcare settings and that the primary outcome is clinically important and easily understood by a range of users, including clinicians, patients, policy makers and health commissioners. The study end point tends to be longer term (e.g. 6 months).

9.2   Definition of Perceived Stress

General situations in life are appraised as stressful when they are unpredictable, uncontrollable or overloading. In the context of this chapter, perceived stress reflects the experienced level of stress as a function of objective stressful events, coping processes and personality factors (Cohen et al. 1983). This perspective is based on Lazarus's (1966) transactional model of stress which argues that the experience of a stressor is influenced by evaluations on the part of the person as to how well they can manage a stressor given their coping resources.

Stress therefore specifically refers to the subjective components of stress. There is an overlap in the operational definitions of perceived stress and the symptomatology of anxiety and depression. Stress can be manifest in a variety of patient-reported complaints about tinnitus and include the person's emotional state, physical state, their performance and behaviours in everyday life, relationships with others and overall quality of life. **Stress-related outcome** directly refers to one of these domains of complaint. It is important to acknowledge that there are physiological characteristics of stress, such as biomarkers like pupillary dilation and salivary cortisol and amylase. These have rarely been used in explanatory or pragmatic trials of tinnitus interventions and so are not considered further. Another term used throughout this chapter is **stress-related outcome instrument**. This refers to the way in which a stress-related outcome is measured or quantified. Instruments relating to perceived stress are typically questionnaires made up of a number of items. On each item, the person is asked to rate the frequency or severity of a complaint, or the degree to which he/she agrees or disagrees with a statement. Such instruments are sometimes called Patient-Reported Outcome Measures (PROMS). PROMS are important for measuring and improving the quality of patient care.

9.3 Patient-Reported Complaints About Tinnitus

For some people with tinnitus, their experience way exceeds the perception of a sound inside the head or ears and causes problems in daily life such as sleep disturbance, difficulties concentrating, and poor psychological well-being (Tyler and Baker 1983; Jakes et al. 1985), ultimately impairing overall quality of life. For counselling purposes during clinical management, it may be informative to identify perceptual attributes of the tinnitus (e.g. pitch and loudness). However, it is more clinically meaningful to identify the symptoms and functional impacts for each individual patient since these are most likely to determine joint decision-making about preferred management options (Henry et al. 2005).

Patient-reported complaints about tinnitus are many and varied. Some of the most influential studies to identify patient-reported complaints of tinnitus were those published before the advent of any tinnitus-related questionnaire measures. Two studies (Tyler and Baker 1983; Jakes et al. 1985) are worthy of note because they have informed decisions about the construction of many of the tinnitus-related questionnaires that have followed. Decisions informed by these data have been the choice of questionnaire subscales, empirical support for their validity to patients (content validity), as well as the selection and wording of individual questionnaire items. I also summarise findings from a third study which is more contemporary, but nevertheless addresses the same issue. Coincidentally, all three studies just so happen to have been conducted in the UK.
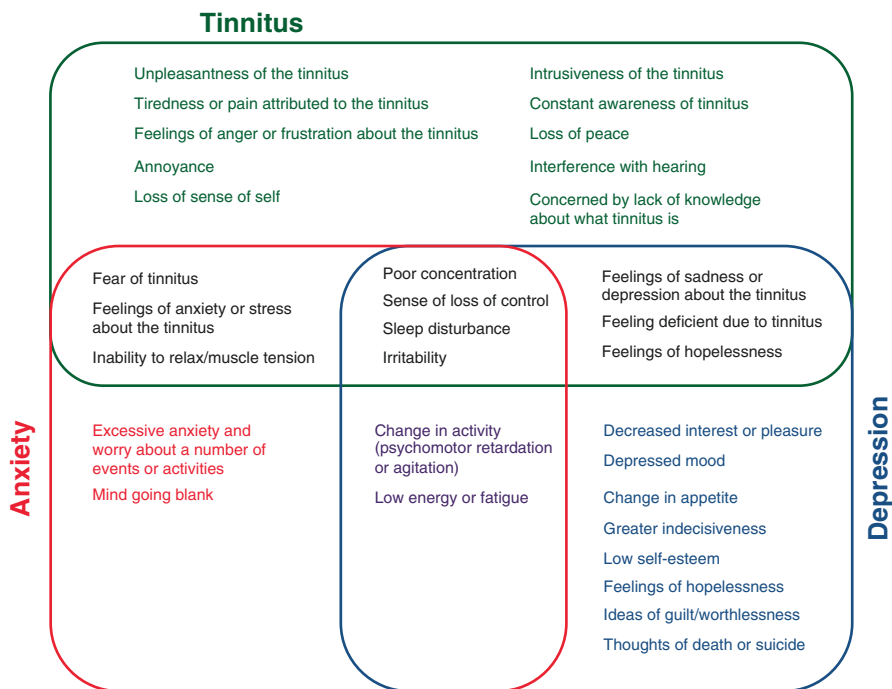
The first is a patient-centred study published in 1983 by Tyler and Baker. Data were collected from 72 people who were members of a tinnitus self-help group. People were asked to list the difficulties that they had as a result of their tinnitus. Instructions were 'Please make a list of the difficulties which you have as a result of your tinnitus. List them in order of importance, starting with the biggest difficulties. Write down as many of them as you can' p. 150. Respondents reported between one and 13 complaints. Subjective evaluation of the free-text response data demonstrated the diversity of tinnitus complaints. Four thematic categories were described: (1) effects of hearing (e.g. understanding speech, appreciation of music and localising sounds), (2) effects on lifestyle (e.g. getting to sleep, family problems and avoiding quiet situations), (3) effects on general health/healthcare (e.g. pain/headaches, tiredness and giddiness/imbalance/fuzzy head) and (4) emotional problems (e.g. concentration/confusion, despair/frustration/depression and annoyance/irritation/inability to relax).

The second is a patient-centred study published in 1985 by Jakes and colleagues. Data were collected from 82 patients who attended a neuro-otology clinic and whose main presenting symptom was tinnitus. Patients were asked to

complete a questionnaire concerning 19 different features of tinnitus and other symptoms. These were presented as closed questions requiring a rating of the frequency or severity, according to predefined descriptors given by the authors. Illustrative examples are (1) 'I find the noises are now/bearable/unbearable' and (2) 'The noises are affecting me to the extent that I am now/not depressed/somewhat depressed/extremely depressed'. The authors also collected audiometric data about the hearing status of all patients. Statistical analysis of the quantitative response data using factor analysis confirmed the multifactorial nature of tinnitus complaints. Eleven factors explained 86% of the variance. These categories were (1) hearing loss, (2) tinnitus-related distress, (3) intrusiveness of tinnitus, (4) interference on music and TV, (5) tinnitus loudness, (6) sleep disturbance, (7) vertigo, (8) use of medication, (9) impact on work, (10) bilaterality of tinnitus and (11) auditory thresholds, in descending order of percentage variance explained.

Our team in Nottingham has conducted a contemporary assessment of the same issue, but using a much larger clinical sample (Watts et al. 2016). Through collaboration with Jacqueline Sheldrake, we had the good fortune to obtain anonymised clinical interview data from 988 patients whom attended the Tinnitus and Hyperacusis Centre (London, UK) between 1989 and 2014. All patients answered the open-ended question 'Why is tinnitus a problem?' Thematic analysis was used to code and collate individual responses into groups or themes according to the domain of the patient-reported problem. Complaints covered the person's emotional state, physical state, their performance and behaviours in everyday life, relationships with others and overall quality of life. Each domain included only those responses that were judged to relate to the same theoretical construct. We can think of these domains as stress-related outcomes that are relevant to tinnitus. Overall, the free-text response data indicated 18 distinct domains of tinnitus-associated complaints. The domains 'tinnitus-related fear' and 'constant awareness' had the highest number of mentions by individual participants indicating that it was a very frequent complaint. The top five also included 'loss of quiet', 'annoyance' and 'effects on quality of life'. Many of the 18 domains were those identified in the two previous studies. However, four particular domains were not highlighted by the two prior studies. These were (1) 'feeling deficient because of tinnitus', (2) 'sense of loss of control', (3) 'concerned by lack of knowledge about what tinnitus is' and (4) 'loss of sense of self'.
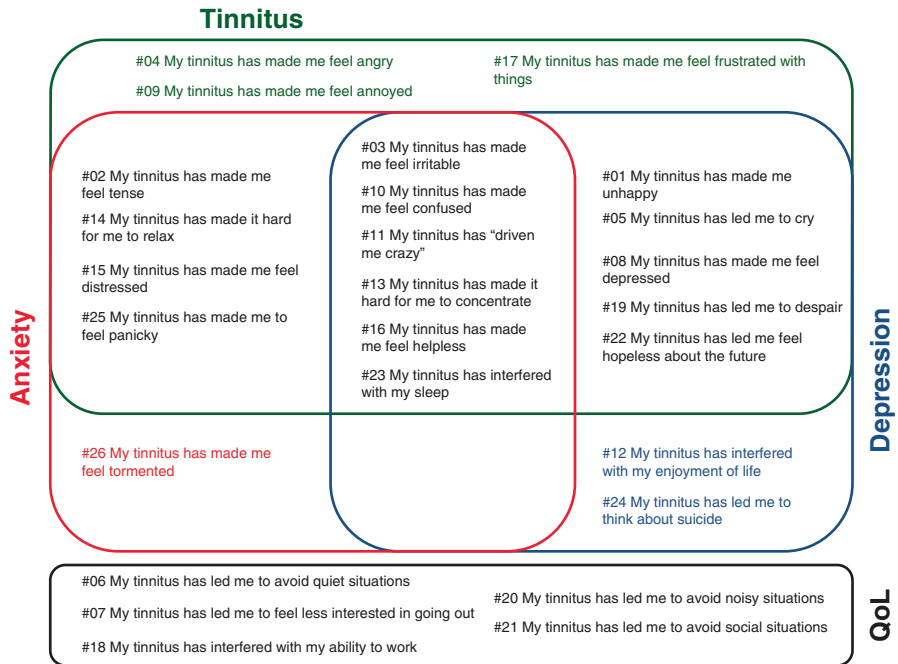
The number and diversity of patient-reported complaints about tinnitus are illustrated in Fig. 9.1, encircled within the green box. These examples have been collated from those patient responses reported by the above three independent pieces of research.

Tinnitus

Unpleasantness of the tinnitus	Intrusiveness of the tinnitus
Tiredness or pain attributed to the tinnitus	Constant awareness of tinnitus
Feelings of anger or frustration about the tinnitus	Loss of peace
Annoyance	Interference with hearing
Loss of sense of self	Concerned by lack of knowledge about what tinnitus is

Fear of tinnitus	Poor concentration	Feelings of sadness or depression about the tinnitus
Feelings of anxiety or stress about the tinnitus	Sense of loss of control	Feeling deficient due to tinnitus
Inability to relax/muscle tension	Sleep disturbance	Feelings of hopelessness
	Irritability	

Anxiety

Depression

Excessive anxiety and worry about a number of events or activities	Change in activity (psychomotor retardation or agitation)	Decreased interest or pleasure
Mind going blank	Low energy or fatigue	Depressed mood
		Change in appetite
		Greater indecisiveness
		Low self-esteem
		Feelings of hopelessness
		Ideas of guilt/worthlessness
		Thoughts of death or suicide

Fig. 9.1 Correspondence between patient-reported symptoms of tinnitus and the diagnostic symptoms of anxiety and depression. Symptoms of tinnitus are based on a qualitative analysis of patient-reported complaints collated by Watts et al. (2016) and Tyler and Baker (1983). Mental health symptoms are based on DSM-5 criteria for generalised anxiety and persistent and major depression (American Psychiatric Association 2013). Functional impacts that impair quality of life are not included here, but are common to all conditions. These impacts include but are not restricted to avoidance behaviours, reduced social participation and negative effect on work

9.4   Measuring Perceived Stress Associated with Tinnitus

The Tinnitus Reaction Questionnaire (TRQ) was constructed in Australia in 1991. It was designed specifically to measure a single attribute of tinnitus: the perceived stress related to tinnitus (Wilson et al. 1991). The TRQ selectively measures perceived stress associated with tinnitus. It has 26 items which all start off with the phrase 'My tinnitus has…' (see Fig. 9.2 for examples). Ratings for each item are made on a 5-point Likert scale (scored 0–4) with the category labels (not at all/a little of the time/a good deal of the time/almost all of the time). Scoring involves the simple addition of the category score selected by the respondent. This gives a range of scores from 0 to 104, with a high score representing greater distress. There is very little literature on the psychometric properties of the TRQ. Although Wilson et al. (1991) describe four factors resulting from their factor analysis (general distress, interference, severe distress, avoidance behaviours), the statistical outputs from the

Tinnitus

#04 My tinnitus has made me feel angry

#09 My tinnitus has made me feel annoyed

#17 My tinnitus has made me feel frustrated with things

Anxiety

#02 My tinnitus has made me feel tense

#14 My tinnitus has made it hard for me to relax

#15 My tinnitus has made me feel distressed

#25 My tinnitus has made me to feel panicky

#03 My tinnitus has made me feel irritable

#10 My tinnitus has made me feel confused

#11 My tinnitus has "driven me crazy"

#13 My tinnitus has made it hard for me to concentrate

#16 My tinnitus has made me feel helpless

#23 My tinnitus has interfered with my sleep

#01 My tinnitus has made me unhappy

#05 My tinnitus has led me to cry

#08 My tinnitus has made me feel depressed

#19 My tinnitus has led me to despair

#22 My tinnitus has led me feel hopeless about the future

Depression

#26 My tinnitus has made me feel tormented

#12 My tinnitus has interfered with my enjoyment of life

#24 My tinnitus has led me to think about suicide

#06 My tinnitus has led me to avoid quiet situations

#07 My tinnitus has led me to feel less interested in going out

#18 My tinnitus has interfered with my ability to work

#20 My tinnitus has led me to avoid noisy situations

#21 My tinnitus has led me to avoid social situations

QoL

Fig. 9.2 Items taken from the Tinnitus Reaction Questionnaire (Wilson et al. 1991) and mapped onto the symptoms of tinnitus and the symptoms of anxiety and depression. Symptoms of tinnitus are based on Watts et al. (2016) and Tyler and Baker (1983). Mental health symptoms are based on DSM-5 criteria for generalised anxiety and persistent and major depression (American Psychiatric Association 2013). Functional impacts are not included here, but are common to all conditions. These impacts include but are not restricted to avoidance behaviours, reduced social participation, negative effect on work and impaired quality of life

factor analysis show that most of these items are very closely related to one another. So there is little value in treating this questionnaire as if it has multiple subscales. A single global score is adequate.

If a questionnaire looks like it is going to measure what it is supposed to measure, then it has what is termed 'face validity'. To assess the face validity of the TRQ, Fig. 9.2 illustrates how the 26 items map onto the framework of stress-related complaints already presented in Fig. 9.1. The wording of each item was carefully evaluated for its meaning, and it was ascertained whether it would fit within one of the symptom domains. Only 3 of the 26 items appear to be restricted to tinnitus-related stress (#04 My tinnitus has made me feel angry, #09 My tinnitus has made me feel annoyed, #17 My tinnitus has made me feel frustrated with things). Section 8.5 considers how many of the remaining tinnitus-related complaints share characteristics with impaired psychological well-being.

The impetus for the construction of the TRQ came from a need for a reliable measurement instrument for evaluating the effects of psychological interventions on the ability of people to cope with tinnitus (Ireland et al. 1985). Up until this point,

tinnitus-related questionnaires were purposefully broad in scope, and they included items that asked patients about a wide variety of complaints. The Tinnitus Questionnaire (Hallam et al. 1988) and the Tinnitus Handicap Questionnaire (Kuk et al. 1990) are both good examples of broad-ranging multi-attribute questionnaire instruments that were constructed before the TRQ, but they measure much more than simply perceived stress.

The Tinnitus Questionnaire has 52 items. Ratings for each item are made on a 3-point Likert scale (scored 0–2) with the category labels (true/partly true/not true). Only 41 items are scored, and the total score is scaled so that the global score ranges from 0 to 82, with higher score indicating greater severity of tinnitus symptoms. The first assessment of the psychometric (statistical) properties of the Tinnitus Questionnaire, using factor analysis techniques, identified three orthogonal factors covering (1) emotional distress, (2) auditory difficulties and (3) sleep disturbance (Hallam et al. 1988). A later reinvestigation by Hallam in 1996 using data from a different sample of tinnitus patients identified five orthogonal factors covering (1) emotional and cognitive distress, (2) intrusiveness, (3) auditory perceptual difficulties, (4) sleep disturbance, and (5) somatic complaints. Hence, there is some uncertainty about what domains of tinnitus-related complaints are measured by the Tinnitus Questionnaire.

The Tinnitus Handicap Questionnaire has 27 items. Ratings for each item are made on a 100-point numerical scale (from 0 = strongly disagree to 100 = strongly agree). The total score is scaled so that the global score ranges from 0 to 100, with higher score indicating greater severity of tinnitus symptoms. The Tinnitus Handicap Questionnaire has three subscales covering (1) social, emotional and physical effects of tinnitus, (2) hearing ability and unease and (3) the individual's perception of tinnitus. It has been pointed out that items on the first two subscales are very closely related to one another, both in terms of the semantic content (i.e. meaning) of the items and the statistical outputs from the factor analysis (see Kennedy et al. 2004; Fackrell et al. 2014). These observations indicate that the Tinnitus Handicap Questionnaire is particularly sensitive to the social, emotional and physical functioning aspects of tinnitus-related distress, but arguably this subscale actually covers three different discrete domains.

Although this summary of tinnitus-related questionnaires is not exhaustive, it serves to highlight the general emphasis on questionnaires that measure a broad range of dimensions of tinnitus complaint. Later questionnaires are little different in this respect (e.g. Tinnitus Handicap Inventory, Newman et al. 1996). Unlike these broad-scope questionnaires, the developers of the TRQ were explicit in their aim to assess a narrow range of tinnitus characteristics. In other words, their aim was to create a single-attribute questionnaire instrument that focused on perceived stress associated with tinnitus.

Given the overlap in patient-reported complaints for tinnitus, anxiety and depressive symptoms, one should expect a high degree of association between tinnitus-related questionnaire scores and questionnaire scores for anxiety and/or

depression. Convergent validity is a term that describes the extent to which the underlying construct of one questionnaire corresponds to other questionnaire constructs that are theoretically similar. It is measured by calculating the correlation coefficients between the questionnaire scores and assessing the strength of the association. Convergent validity is indicated by a strong Pearson correlation coefficient ($r > 0.60$) (Andresen 2000). The TRQ has been examined by correlating with scores for depression and anxiety questionnaires. TRQ has strong convergent validity with the Beck Depression Inventory (BDI) (Beck et al. 1961). Wilson et al. (1991) reported correlations of $r = 0.63$ and $r = 0.87$ for two independent samples of participants, while Robinson et al. (2003) reported a correlation of 0.66. With respect to anxiety, TRQ also correlates well. Correlation coefficients reported by Wilson et al. (1991) were 0.60 and 0.74 for state anxiety and 0.58 and 0.71 for trait anxiety, as measured by the State-Trait Anxiety Inventory (STAI) (Spielberger et al. 1970). Overall, the TRQ seems to be measuring similar theoretical constructs associated with general perceived stress. These findings raise an important question about whether the TRQ measures any sufficiently distinct aspect of tinnitus-related stress that is not captured by measures of general psychological well-being.

These results set the TRQ apart from the Tinnitus Questionnaire (Hallam et al. 1988) and the Tinnitus Handicap Questionnaire (Kuk et al. 1990). For comparison, the Tinnitus Questionnaire and Tinnitus Handicap Questionnaire generally had weaker correlations: BDI ($r = 0.51$ and $r = 0.62$, respectively) and Hamilton Rating Scale for Depression (Hamilton 1960) ($r = 0.48$ and $r = 0.57$, respectively) (Robinson et al. 2003). Overall, the Tinnitus Questionnaire and Tinnitus Handicap Questionnaire may therefore be measuring different theoretical constructs. One might speculate that this difference reflects the other (non-stress) outcome domains contained within these broad-scope multi-attribute tinnitus instruments.

9.5 Associations Between Tinnitus and Psychological Well-Being

From the patient-reported complaints (Sect. 9.3), it is clear that tinnitus is associated with considerable perceived stress manifest as feelings of anxiety, sadness or depression, irritability, inability to relax, etc. These symptoms are not restricted to tinnitus. Symptom overlap in tinnitus, depression and anxiety can act as a confounder in estimating the severity of either condition. Figure 9.1 illustrates this point by mapping out the correspondence between patient-reported symptoms of tinnitus and the diagnostic symptoms of generalised anxiety and depression according to the Diagnostic and Statistical Manual for Mental Disorders (DSM) edition 5 (American Psychiatric Association 2013). Four of the patient-reported complaints reported by people with tinnitus seem common to all three conditions (poor concentration, sense of loss of control, sleep disturbance and irritability).

Three further complaints are common to tinnitus and anxiety (fear, feelings of anxiety or stress and inability to relax), and three more are common to tinnitus and depression (feelings of sadness or depression, feeling imperfect and feelings of hopelessness). This high degree of association is also seen in the construction of the TRQ. Turning to Fig. 9.2, one can see how 18 of the 26 items from the TRQ appear to map onto domains relating to general anxiety or depression (or the intersections thereof). Associations between tinnitus and psychological well-being have important implications when we turn to discuss how perceived stress is measured.

9.6   Measuring General Perceived Stress

One of the most widely used measures of stress is the Perceived Stress Scale (PSS), developed in the USA (Cohen et al. 1983). This questionnaire was designed specifically to measure global perceived stress. Up until this point, measurements of stress typically focused on objective indicators (e.g. frequencies) of specific stressors such as chronic illness, bereavement, and retirement. But this focus on external life event stressors and the cumulative minor stressors of everyday life overlooked the influence on individual's subjective interpretation of that stressor.

The PSS therefore asks questions about whether a person feels under pressure from specific worries. It has 14 items which ask individuals to rate how often they experienced particular feelings and thoughts *in the past month*. Items were designed to tap into how unpredictable, uncontrollable and overloaded people find their lives. An example item is 'In the last month, how often have you felt that you were unable to control the important things in your life?' Ratings on each item are made on a 5-point Likert scale (scored 0–4) with the category labels/never/almost never/sometimes/fairly often/very often. Seven of the items are positively worded and seven are negatively worded. The positive items are reverse scored, and then the global score is the sum across all 14 items. A high score therefore reflects a high degree of perceived stress with the global score ranging from 0 to 56.

The first major assessment of the psychometric properties of the PSS, using factor analysis techniques, identified two orthogonal factors covering (1) the negatively worded items (e.g. been upset, unable to control things, felt nervous and stressed) and (2) the positively worded items (e.g. dealt successfully with hassles, effectively coping, felt confident) (Cohen and Williamson 1988). Informed by this dataset, a shorter 10-item version was produced, and again this had the same two-factor structure.

Results for convergent validity have been usefully summarised as part of a systematic review of the psychometric properties of the PSS (Lee 2012). Overall findings support the conclusion that the questionnaire score is either moderately or strongly correlated with scores for depression and anxiety questionnaires, as measured using the BDI, Hospital Anxiety and Depression Scale (HADS) (Zigmond

and Snaith 1983), STAI, and Depression, Anxiety and Stress Scale (DASS) (Lovibond and Lovibond 1995). These findings indicate that the PSS seems to be measuring similar theoretical constructs associated with stress.

The Perceived Stress Questionnaire (PSQ) is also concerned with the cognitive appraisal about aspects of everyday life and the emotional reaction to them (Levenstein et al. 1993). Many of the questions have the format 'you feel…'. For example, 'You feel that too many demands are being made on you/You feel frustrated'. The original PSQ comprised 30 items, spanning seven factors (harassment, irritability, lack of joy, fatigue, worries, tension and overload). Ratings for each item are made on a 4-point Likert scale (scored 1–4) with the category labels/almost never/sometimes/often/usually. Raw scores are transformed into a *stress index* from 0 (lowest possible level of stress) to 1 (highest possible level of stress). Just as in the PSS, one version of the PSQ asks individuals to rate how often they experienced particular feelings and thoughts *in the past month*. A second version of the PSQ asks about events *in the past 2 years*.

The 'past month' version of the PSQ demonstrated acceptable convergent validity with the PSS ($r = 0.73$) and trait anxiety measured using the STAI ($r = 0.75$), but weaker correlations with depression ($r = 0.56$).

In 2001, the PSQ was translated into German and re-evaluated on a broad sample of participants (Fliege et al. 2001). The resulting German version has a reduced set of 20 items, covering four factors (joy, worries, tension and demands). Although the labels given to three of the factors are equivalent across languages, it is important to note that the items that correspond to the factors are different. Thus, any subscale scores should not be directly compared across the English and German versions.

The DASS (Lovibond and Lovibond 1995) is another widely used questionnaire that includes a measure of perceived stress. This questionnaire comprises 42 items covering three separate scales of stress, anxiety and depression over the past week. Each scale has 14 items. For example, one of the stress scale items is 'I found myself getting upset by quite trivial things'. Ratings for each item are made on a 4-point (scored 0–3) with the following category labels: did not apply to me at all/ applied to me to some degree, or some of the time/applied to me to a considerable degree, or a good part of time/applied to me very much, or most of the time. Scores of depression, anxiety and stress are calculated by summing the scores for the relevant items and are interpreted according to five symptom severities (normal, mild, moderate, severe and extremely severe).

There is relatively little data on the DASS in tinnitus. However, one article does report questionnaire findings in a sample of 100 patients with tinnitus attending an out-patient otorhinolaryngology clinic (Gomaa et al. 2014). Severe to extremely severe stress was observed in 33% of patients. The proportion of patients with severe to extremely severe depression and anxiety was somewhat greater (51% and 54%, respectively). Figure 9.3 shows the pattern of stress, anxiety and depressive comorbidities in this sample, plotted as a function of tinnitus 'severity'. Tinnitus severity was measured using a Visual Analogue Scale.
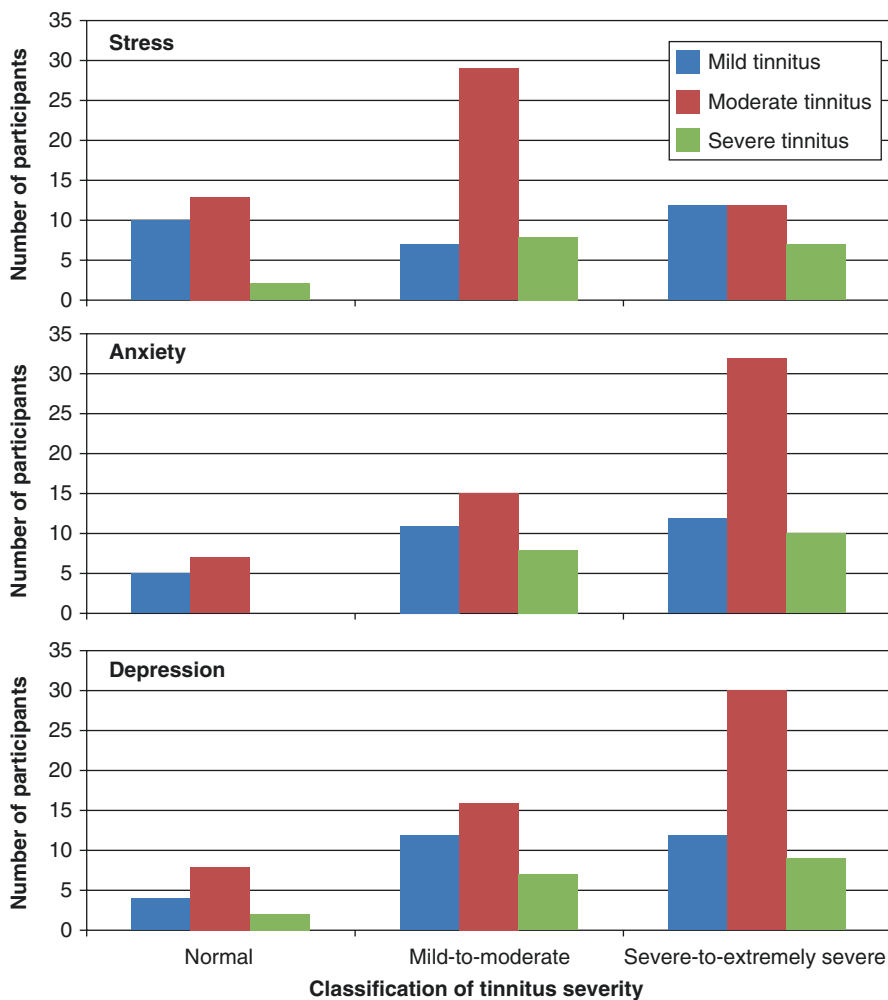
Fig. 9.3 Comparisons between level of stress, anxiety and depressive symptoms assessed in a sample of 100 tinnitus patients, using the DASS (Gomaa et al. 2014). Score classifications for the DASS are stress (normal, 0–14; mild-to-moderate, 15–25; severe-to-extremely severe, 26+), anxiety (normal, 0–7; mild-to-moderate, 8–14; severe-to-extremely severe, 15+) and depression (normal, 0–9; mild-to-moderate, 10–20; severe-to-extremely severe, 21+)

9.7   Applications of Stress-Related Questionnaire Instruments

The discussion so far has shown how patient-reported complaints have informed the construction of questionnaire instruments and has demonstrated the commonalities between complaints of tinnitus, stress, anxiety and depression. None of the issues so far concerning questionnaire construction are necessarily restricted to outcomes

used to evaluate treatment-related change. They are equally applicable to the purposes of screening, diagnosis and prognosis. However, the way that a questionnaire is constructed should be informed by the purpose for which it is intended. Tutorial 9.2 explains more about these different applications. This section considers how the intended application of each stress-related questionnaire defines what statistical properties of the instrument are most important during its creation.

Tutorial 9.2 Purpose of Questionnaires

Potential applications of such questionnaires typically fall into three broad categories (Kirschner and Guyatt 1985):

Discrimination. A discriminative tool is used to distinguish between individuals or groups, generally as part of a screening or diagnostic procedure. For example, to quantify the burden of stress on individual tinnitus patients so that healthcare provision can be tailored more effectively.

Prediction. A predictive tool is used to classify individuals into predefined categories generally as part of a screening or diagnostic procedure. For example, to identify clues to a prognosis.

Evaluation. An evaluative tool is used to measure the magnitude of change over time in an individual or group on the complaint of interest. For example, to quantify treatment benefits in clinical trials and for measuring quality-adjusted life years in cost-utility analysis.

It is not unusual for developers of tinnitus-related questionnaires to claim that theirs is a multipurpose instrument. For example, on the TRQ, Wilson et al. (1991) claimed 'such a scale may provide a useful assessment device in clinical practice and in further research on psychological aspects of tinnitus. It may be useful as a screening instrument in the selection of distressed samples, as a means to distinguish tinnitus sufferers who cope with the problem from those who do not cope well, and as a measure of psychological distress before and after treatment' p. 198.

For questionnaires assessing how a person feels and functions in day-to-day activities, the psychometric (statistical) requirements to maximise the discriminative, predictive or evaluative properties of the questionnaire are often at odds with one another. Table 9.1 describes key issues to be considered when devising a strategy for constructing a questionnaire for discrimination or for evaluation. Prediction is not discussed further because it is not an issue that has been widely investigated in the tinnitus field.

For those readers particularly interested in the measurement properties of patient-reported outcome instruments, the COSMIN checklist (http://www.cosmin.nl/) is a useful generic tool for evaluating the methodological quality of studies reporting the construction of an instrument.

While a discriminative strategy places an emphasis on attempting to sample all important, relatively stable aspects of functional status common to most members of each functional class, an evaluative strategy places an emphasis on restricting

Table 9.1 Major issues for consideration in the construction and evaluation of outcome instruments (informed by Kirschner and Guyatt 1985)

Issue to consider	Discriminative strategy	Evaluative strategy
Selecting the questionnaire items	Complaints are: • Important to patients with tinnitus • Universally applicable to people with tinnitus • Stable over time	Complaints: • Are likely to change • Will be responsive to a clinically significant change, as a result of the intervention of interest
Choosing the format of the response options available to patients	• Short response sets which facilitate the same interpretation from person to person	• Response sets have sufficient gradations to register change
Reducing the total item pool based on performance in the relevant setting	• Remove items where variability between-subjects is not related to tinnitus • Consider time and effort needed by the subject	• Remove unresponsive items
Ensuring measurement of true differences relative to the overall variance	• Variation between subjects is large and remains stable across testing intervals	• Variation within-individuals remains stable across testing intervals
Validity	• Include all characteristics of tinnitus that are common to most people • Relationship between the instrument score and external measures at a single point in time	• Include only characteristics that are salient with respect to clinically important treatment-related change. • Relationship between changes in the instrument score and external measures over time
Responsiveness	• N/A	• Known power of the test to detect a minimal clinically important difference (i.e. all parameters for computing the sample size required to observe a predefined change in the population)

measurement only to those salient activities and feelings that are subject to clinically important treatment-related change. Kirschner and Guyatt (1985) point out that this distinction has often been neglected in the health status measurement literature. An outcome instrument for measuring stress-related symptoms in people with tinnitus before and after (psychological) treatment should certainly *not* be seeking to assess all of the 20 distinct domains of tinnitus-associated complaints that are given in Fig. 9.1. Criticism of tinnitus questionnaires that 'measure a limited number of constructs' (p. 144) (Newman et al. 1996) is not a valid criticism for questionnaires that are primarily to be used for an evaluative (outcome) purpose.

Suffice it to say that few tinnitus-related questionnaires have been developed specifically according to an evaluative strategy (Fackrell et al. 2014). And the TRQ, PSS, PSQ and DASS are no exceptions. Their psychometric properties as outcome instruments in the tinnitus population are not yet established.

9.7.1   Use of Questionnaires for Diagnosing Perceived Stress

Just like other chronic conditions, tinnitus in the general population is associated with perceived stress and its associated symptoms of anxiety and depression. Table 9.2 lists some of the instruments that have been used for assessing stress,

Table 9.2   List of instruments used for assessing stress, anxiety and depression in clinical research

Instrument for assessing stress, anxiety and depression	Reference	Instrument used for diagnosis	Instrument used as primary outcome	Instrument used as secondary outcome
		Number of studies out of 16	Number of studies out of 228	
Perceived stress associated with tinnitus				
Tinnitus Reaction Questionnaire (TRQ)	Wilson et al. (1991)	2 (13%)	11 (5%)	2 (<1%)
General perceived stress				
Perceived Stress Scale (PSS)	Cohen et al. (1983, 1988)	0	1 (<1%)	0
Perceived Stress Questionnaire (PSQ)	Levenstein et al. (1993)	0	3 (1%)	0
Depression, Anxiety and Stress Scale (DASS)	Gomaa et al. (2014)	0	0	1 (<1%)
Anxiety and depression				
Anxiety Sensitivity Index	Reiss et al. (1986)	1 (6%)	0	1 (<1%)
Beck Anxiety Inventory (BAI)	Beck et al. (1988)	1 (6%)	0	1 (<1%)
Beck Depression Inventory (BDI)	Beck et al. (1961)	5 (31%)	7 (3%)	13 (6%)
Composite International Diagnostic Interview Short Form	World Health Organization	1 (6%)	0	0
Comprehensive Psychopathological Rating Scale	Asberg et al. (1978)	1 (6%)	0	0
Hopkins Symptoms Checklist (HSCL)	Derogatis et al. (1974)	1 (6%)	0	0
Hospital Anxiety and Depression Scale (HADS)	Zigmond et al. (1983)	5 (31%)	7 (3%)	27 (12%)

(continued)

Table 9.2 (continued)

Instrument for assessing stress, anxiety and depression	Reference	Instrument used for diagnosis	Instrument used as primary outcome	Instrument used as secondary outcome
		Number of studies out of 16	Number of studies out of 228	
M.I.N.I.—International Neuropsychiatric Interview	www.medical-outcomes.com/index/mini	2 (13%)	0	0
Major Depression Inventory	World Health Organization	0	0	5 (2%)
State-Trait Anxiety Inventory (STAI)	Spielberger et al. (1970)	2 (13%)	3 (1%)	3 (1%)
Structured Clinical Interview for the Diagnostic and Statistical Manual for Mental Disorders (DSM)	American Psychiatric Association	5 (31%)	0	0
Symptom Checklist 90 Revised	www.pearsonclinical.com	1 (6%)	1 (<1%)	0

Instruments used for diagnosis of stress, anxiety and depression comorbid with tinnitus are taken from Pinto et al. (2014). Instruments used for evaluation of treatment-related outcome are taken from an unpublished systematic review (Hall et al. 2015)

Note that only the 'most typical' are reported here. Occasional use of other questionnaires is reported, but these constitute <1% when combined across diagnosis and treatment outcome studies

anxiety and depression in tinnitus research. Consistent with the previous descriptions, these are classified according to measures of perceived stress associated with tinnitus, general perceived stress, anxiety and depression.

Those questionnaires used for diagnosis of tinnitus-related comorbidities are given in column 3. It is not an exhaustive list, but instead reflects data reported by Pinto et al. (2014) in a systematic review focusing on the *diagnosis* of mental disorders associated with tinnitus. Using such measures, it was confirmed that the prevalence of anxiety and depression is high in patients with tinnitus (Pinto et al. 2014). Indeed, this pattern was reported in 15 of the 16 included studies. Seven studies reported a significant positive correlation between the presence and severity of depression and the severity and annoyance of tinnitus, while four reported a similar pattern for anxiety. On the basis of this evidence, the authors conclude that the presence of a comorbid depression or anxiety worsens the prognosis of tinnitus-related stress. It is important to note that only five out of the 16 included studies used a psychiatric diagnosis of a mental health disorder according to the Structured Clinical Interview for the DSM edition 3 or 4 (Table 9.1). Patient-reported questionnaires, namely, the HADS and the BDI, were equally popular. But detection of depressive symptoms by self-report scales does not automatically mean the diagnostic criteria for a psychiatric disorder are fulfilled (Langguth et al. 2011).

9.7.2 Use of Questionnaires for Evaluating Treatment-Related Change in People with Tinnitus

A number of co-workers and I have recently completed a systematic review of clinical trials assessing the treatment of adults with tinnitus (Hall et al. 2016). Only trials with ≥20 participants were eligible, thus excluding the majority of feasibility and pilot studies. The objective was to identify and evaluate the current reported outcome domains and instruments. There was no restriction on inclusion criteria for participants in the trial, for the type of intervention or for the type of outcome evaluation. From 1574 articles and trial registrations published since July 2006 to March 2015, 228 met our inclusion criteria for the review. Overall for the primary evaluation of treatment benefit (i.e. the primary outcome), 78 different instruments were used. For the secondary evaluation of treatment benefit (i.e. the secondary outcome), 108 different instruments were used. Perhaps, unsurprisingly, the most popular instruments were condition-specific, i.e. they were assessing multiple attributes related to the impact of tinnitus. The Tinnitus Handicap Inventory (Newman et al. 1996) was most frequently used.

The other interesting questionnaire instruments with respect to the aim of this review are the TRQ which was used in 13 studies, the PSS which was used in one study, the PSQ which was used in three studies and the DASS which was used in one study and was scored separately for each of the three subscales (Table 9.2).

Table 9.3 reports some of the characteristics of these studies including the trial design, type of intervention(s), description of the outcome domain, end points, definition of the minimal clinically important difference, details of the sample size calculation and the sample size itself. For controlled trials, Table 9.3 also reports whether a statistically significant difference between groups was detected ($p < 0.05$).

As expected from the description of clinical trial designs given in Sect. 9.1, explanatory trials formed the majority of studies using the TRQ as a condition-specific stress outcome. Eight of the 13 studies using the TRQ were of this design, with another being a pilot trial for a later planned explanatory trial. Only one study was described as a pragmatic trial, and this used a general perceived stress outcome, as would be expected.

In general, definitions of the minimal clinically important difference and details of the sample size calculation were poorly reported (Table 9.3). This may reflect a lack of awareness on the part of the investigators or (for the change measure) a paucity of knowledge about what these parameters should be for the target population. These pieces of information are crucial aspects of good trial design for explanatory and pragmatic trials. The interested reader is directed to Tutorial 9.3 for more details. Whatever the reason, there is a risk that many of these trial designs are underpowered.

Many (but not all) of the interventions being assessed using the TRQ, PSS, PSQ and DASS were psychological interventions. This is consistent with the purpose of these questionnaires to measure complaints associated with perceived stress. In other words, investigators tended to describe an interest in reducing patient distress or stress and most therefore tended to choose a stress-related questionnaire as a

Table 9.3 Table reporting details of clinical trials conducted and/or reported between 2006 and 2015, using stress-related outcomes

Reference	Clinical trial design	Type of intervention(s)	Author-defined outcome domain	Primary or secondary outcome?	Sample size	Minimal clinically important difference	Sample size calculation	End points	Group diff.
Tinnitus reaction questionnaire									
ClinicalTrials. gov ID: NCT00724152	Pilot—randomised controlled trial	1-Tinnitus education, 2-Tinnitus education plus CBT, 3-Standard care (not defined)	Not reported	Secondary	66	Not reported	Not reported	6, 14 and 30 weeks	Protocol only
Abbott et al. (2009)	Explanatory—cluster randomised trial	1-Internet-based CBT 2-Internet-based information only	Tinnitus distress	Primary	1–32 2–24	Not reported	Not reported	6 weeks	n.s
Davis et al. (2007)	Explanatory—randomised controlled trial	1-Neuromonics one-stage protocol 2-Neuromonics two-stage protocol	Tinnitus distress	Primary	1–16, 2–19	A threshold for a clinically significant improvement was set at 40% of pretreatment TRQ	Not reported	2, 4, 6 and 12 months	n.s.
Kaldo et al. (2007)	Explanatory—randomised controlled trial	1-Tinnitus self-help book and telephone calls 2-Delayed tinnitus self-help book	Tinnitus annoyance and distress	Primary	1–34 2–38	50% reduction in TRQ mean score = clinically significant improvement	Not reported	6 weeks	sig.
Kaldo et al. (2008)	Explanatory—randomised controlled trial	1-Internet-based CBT 2-Group-based CBT	Tinnitus annoyance and distress	Primary	1–26 2–25	Not reported	Not reported	Not clearly defined	n.s.
Malouff et al. (2010)	Explanatory—randomised controlled trial	1-Tinnitus self-help book (Henry and Wilson) 2-Delayed tinnitus self-help book	Psychological distress specifically associated with tinnitus	Primary	1–84 2–78	Not reported	Not reported	2 months	n.s.

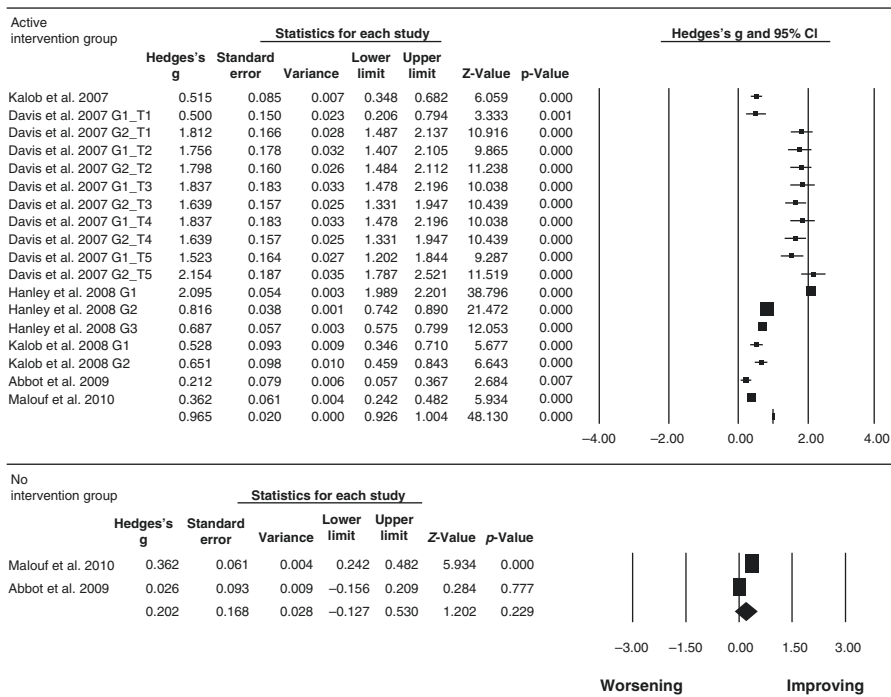| Robinson et al. (2008)	Explanatory—randomised controlled trial	1-Immediate CBT 2-Delayed CBT	Psychological distress related to tinnitus	Primary	1–38 2–27	Not reported	'We conducted a power analysis and identified a goal of recruiting 60 patients total, 30 per group, to be able to detect a moderate effect size at a power of 0.8 and a p value of 0.05 for our primary outcome'	8 weeks	sig. (but groups not matched on baseline TRQ)
Heijneman et al. (2012)	Explanatory—randomised crossover trial	1-Phase shift sound therapy 2-Placebo	Tinnitus-related distress	Primary	25	Not reported	'To determine the sample size, the 1.5-decrease in tinnitus loudness (from 6.4 to 4.9, standard deviation 2.2) on a 10-point visual analog scale was used, which was described by Vermeire et al. While testing 2-sided with $a = 0.05$ and 80% power, it was found that 19 patients would be needed to obtain significance, in a crossover randomized control study design'	1 week after the end of each treatment	n.s.

(continued)

Table 9.3 (continued)

Reference	Clinical trial design	Type of intervention(s)	Author-defined outcome domain	Primary or secondary outcome?	Sample size	Minimal clinically important difference	Sample size calculation	End points	Group diff.
ISRCTN ID: ISRCTN17631678	Explanatory—randomised crossover trial	1-Phase out auditory stimulation 2-Pacebo	Not reported	Secondary	60	Not reported	Not reported	5 and 9 weeks	Protocol only
Hanley et al. (2008)	Not applicable—observational study	customised acoustic stimulation coupled with counselling (Neuromonics)	Tinnitus distress	Primary	470	A threshold for a clinically significant improvement was set at 40% of pretreatment TRQ, in line with prior clinical studies of this technique	Not reported	Approx. 37 weeks	N/A
ClinicalTrials. gov ID: NCT00730834	Not applicable—observational study	Customised acoustic stimulation coupled with counselling (Neuromonics)	Not reported	Primary	53	Not reported	Not reported	6, 12, 24, 36 months	Protocol only
Távora-Vieira et al. (2011)	Not applicable—retrospective data analysis	1-Neuromonics standard protocol 2-Neuromonics variation of standard protocol	Not reported	Primary	1–13 2–13	The success rate of the treatment is defined by Neuromonics protocols as a reduction of at least 40% in TRQ scores.	Not reported	2, 4 and 6 months	Sig (4 months only)
McNeill et al. (2012)	Not applicable—retrospective data analysis	Hearing aids	Tinnitus distress	Primary	70	Not reported	Not reported	3 months	N/A

Perceived stress scale

Hesser et al. (2012)	Explanatory—randomised controlled trial	1-Internet ACT 2-Internet CBT 3-Monitored online discussion forum	Perceived stress	Secondary	1–35 2–32 3–32	Not reported	Not reported	Not reported	3 and 8 weeks and 1 year	n.s

Perceived stress questionnaire

Seydel et al. (2010)	Explanatory—randomised controlled trial	1-Modified TRT 2-Waiting list	Stress	Primary	1–192 2–45	Not reported	Not reported	Not reported	3 months	n.s.
ISRCTN ID: ISRCTN38408464	Pragmatic—randomised controlled trial	1-TRT with *Ginkgo biloba* extract (tablet) 2-TRT alone	Not reported	Primary	300	Not reported	Not reported	Not reported	6, 12 and 24 weeks	Protocol only
Olze et al. (2012)	Not applicable—semi-retrospective study	Cochlear implantation	Stress	Primary	40	Not reported	Not reported	Not reported	After implantation (not defined)	N/A

Depression Anxiety and Stress Scale (DASS)

Abbott et al. (2009)	Explanatory—cluster randomised trial	1-Internet-based CBT 2-Internet-based information only	Tinnitus distress	Secondary	1–32 2–24	Not reported	Not reported	Not reported	6 weeks	n.s

The TRQ is the most popular and is available in English, French, German and Swedish. The PSS has been translated from English into other languages including German, Swedish, Japanese and Arabic; PSQ is available in (at least) English and German, while DASS is available in over 40 languages

primary measure of treatment efficacy. For example, with respect to 13 studies using the TRQ as an outcome measure, nine trials assessed cognitive behaviour therapy (CBT) or an equivalent counselling approach, and one was an evaluation of a self-help book. Comparison of findings across studies is limited by the different study designs, choice of controls (active or waiting list), small sample sizes and various study end points. But some general statements can be made. First, the questionnaire instruments were generally responsive to detecting a reduction in scores after treatment, compared to before treatment (Fig. 9.4). However, for controlled trials, the most important result is the comparison of treatment-related change between groups, and here, the ability of the questionnaire instruments to detect these more subtle effects appeared less successful. When you simply look at the findings as reported by the study investigators, the TRQ appears to be rather mixed in its ability to detect significant changes in patient-reported stress between groups. Only three studies reported significant differences between groups (Kaldo et al. 2007; Robinson et al. 2008; Távora-Vieira et al. 2011). Four studies used general

Active intervention group	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Kalob et al. 2007	0.515	0.085	0.007	0.348	0.682	6.059	0.000
Davis et al. 2007 G1_T1	0.500	0.150	0.023	0.206	0.794	3.333	0.001
Davis et al. 2007 G2_T1	1.812	0.166	0.028	1.487	2.137	10.916	0.000
Davis et al. 2007 G1_T2	1.756	0.178	0.032	1.407	2.105	9.865	0.000
Davis et al. 2007 G2_T2	1.798	0.160	0.026	1.484	2.112	11.238	0.000
Davis et al. 2007 G1_T3	1.837	0.183	0.033	1.478	2.196	10.038	0.000
Davis et al. 2007 G2_T3	1.639	0.157	0.025	1.331	1.947	10.439	0.000
Davis et al. 2007 G1_T4	1.837	0.183	0.033	1.478	2.196	10.038	0.000
Davis et al. 2007 G2_T4	1.639	0.157	0.025	1.331	1.947	10.439	0.000
Davis et al. 2007 G1_T5	1.523	0.164	0.027	1.202	1.844	9.287	0.000
Davis et al. 2007 G2_T5	2.154	0.187	0.035	1.787	2.521	11.519	0.000
Hanley et al. 2008 G1	2.095	0.054	0.003	1.989	2.201	38.796	0.000
Hanley et al. 2008 G2	0.816	0.038	0.001	0.742	0.890	21.472	0.000
Hanley et al. 2008 G3	0.687	0.057	0.003	0.575	0.799	12.053	0.000
Kalob et al. 2008 G1	0.528	0.093	0.009	0.346	0.710	5.677	0.000
Kalob et al. 2008 G2	0.651	0.098	0.010	0.459	0.843	6.643	0.000
Abbot et al. 2009	0.212	0.079	0.006	0.057	0.367	2.684	0.007
Malouf et al. 2010	0.362	0.061	0.004	0.242	0.482	5.934	0.000
	0.965	0.020	0.000	0.926	1.004	48.130	0.000

No intervention group	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Malouf et al. 2010	0.362	0.061	0.004	0.242	0.482	5.934	0.000
Abbot et al. 2009	0.026	0.093	0.009	−0.156	0.209	0.284	0.777
	0.202	0.168	0.028	−0.127	0.530	1.202	0.229

Worsening   Improving

Fig. 9.4 Effect sizes for pre- versus post-intervention, within-group comparisons are shown for the Tinnitus Reaction Questionnaire. To be eligible for inclusion, group mean, standard deviation and sample size had to be reported. In some cases, this information was read from a graphical figure. Several studies report more than one active intervention group (G1–G3) and/or more than one end point (T1–T5). These details are reported in Table 9.3. The meta-analysis must be interpreted with caution due to the heterogeneity of the clinical trial design, but there seems to be a general responsiveness to active intervention over time

stress measures (PSS, PSQ and DASS) to assess the efficacy of CBT or an equivalent counselling approach, but none of these detected a significant between-group effect where findings were reported. The lack of statistical significance is likely to be another marker of underpowered clinical trial designs.

From the same systematic review of clinical trials assessing the treatment of adults with tinnitus, it can be seen that anxiety and depression questionnaires have also been used to determine treatment-related change in people with tinnitus. Table 9.2 illustrates their distribution. Just as was the case for diagnosis of mental health comorbidities with tinnitus (Pinto et al. 2014), the BDI, HADS and STAI were the preferred tools for assessing treatment-related changes in anxiety and depression in people with tinnitus. These questionnaire instruments have also been widely used outside the tinnitus field, in those clinical trials of interventions which are targeted at the treatment of anxiety and depression disorders (Churchill et al. 2013; Hunot et al. 2007, 2013; Joyce and Herbison 2015; Mayo-Wilson and Montgomery 2013; Ori et al. 2015). And the Beck Anxiety Inventory (BAI) (Beck 1988) has been used too.

Tutorial 9.3 Sample Size Calculation

It is never practical to study the whole population. Instead studies must select a subset of participants, which is smaller in size, but adequately represents the population from which it is drawn. This means that true inferences about the population can be made from the results obtained. This subset of participants is known as the **sample**, and the number of participants is known as the **sample size**.

The calculation of an adequate sample size is a crucial in the design of explanatory and pragmatic clinical trials. It is the process by which we calculate the optimum number of participants required to be able to arrive at ethically and scientifically valid results.

Generally, the sample size depends on:

- **Outcome instrument.** A single instrument should be predefined so that scores on this measure will be used to determine whether the treatment is beneficial or not. This is called the primary outcome.
- **Pooled standard deviation.** Standard deviation is a measure of variability in the scores measured by the primary outcome instrument within the population. It is usually estimated from previously reported studies, including pilot work.
- **Acceptable level of significance.** Statistical significance is denoted by the 'p' value and convention is a p value of 5%. A $p = 0.05$ means that the investigators accept the erroneous detection of a difference 5 out of 100 times, when actually no difference exists. This is called the Type I error ('false positive').
- **Side of hypothesis testing.** For $p = 0.05$, a two-tailed test allocates 0.025 to testing the statistical significance in one direction (i.e. improving) and 0.025 to testing statistical significance in the other direction (i.e. worsening). In contrast, a one-tailed test allocates all 0.05 to one tail of the distri-

bution of the test statistic. Convention is for two-sided testing because there is rarely sufficient prior knowledge about the intervention effects at the end point of interest.

- **Power.** Statistical power relates to the probability of failing to detect a difference when actually there *is* a difference. Power is denoted as a percentage and convention is 80%. A power of 80% means that investigators accept that one in five times (i.e. 20%) a real difference will be missed. This is called the Type II error ('false negative').
- **Expected difference.** Just because a treatment-related change is statistically significant at $p < 0.05$, it does not necessarily mean that it is worth implementing in clinical practice. Any treatment-related benefit should also be meaningful to patients.

A challenge is thus to define the difference between the treatment and control groups in the scores measured by the primary outcome instrument that can be considered clinically meaningful. This is called the 'minimal clinically important difference'.

Additional factors can be taken into account when calculating the final sample size, and these include the expected drop-out rate, an unequal allocation ratio and the objective and design of the study.

9.8   Take-Home Messages

This chapter has defined and discussed perceived stress in the context of clinical research. The following key points have been discussed and are worth highlighting again in this concluding section.

Measuring Perceived Stress Associated with Tinnitus
- The Tinnitus Reaction Questionnaire (TRQ) is the *only* questionnaire that selectively measures perceived stress related to tinnitus for use in quantifying patient benefit from psychological interventions.
- The TRQ seems to be closely associated with patient-reported outcome measures for depression and anxiety. Indeed it can be argued that they are to some extent measuring the same underlying constructs.
- The TRQ holds promise as a responsive instrument for detecting improvements over time within a group of patients receiving a psychological intervention.

> **Measuring General Perceived Stress**
> - The Perceived Stress Scale (PSS), Perceived Stress Questionnaire (PSQ) and the Depression, Anxiety and Stress Scale (DASS) all measure general aspects of perceived stress.
> - There is considerably less experience in using them for tinnitus research.

Stress measures are most likely to be suited for evaluating the effectiveness of psychological interventions, than non-pharmacological interventions such as sound therapy and electrophysiology. However, further research is needed to understand how well-suited these questionnaires are for use as an outcome measure in clinical trials for tinnitus. In particular, their responsiveness properties should be better characterised before any recommendations are made for clinical trial design. In the absence of such knowledge, informed decisions about sample size and what difference should be expected in order to interpret that the treatment can be noticed by patients. Until this point, there is a potential risk that null findings are not indicative of an ineffective intervention, but simply the wrong choice of outcome measure.

The relationship between certain personality traits, depressive mood, anxiety and tinnitus is highly relevant for understanding the degree to which tinnitus is amenable to intervention. Personality traits, such as emotional resilience and adaptive coping strategies, may enable one individual to be much less affected by stressors that would otherwise have negative health impacts. These modulatory factors are rarely taken into account when evaluating the impact of an intervention on perceived stress, yet they are of fundamental importance to understanding treatment efficacy and may explain some of the individual variability that is often seen in clinical practice and in clinical trials.

Disclaimer   The views expressed in this publication are those of the author and not necessarily of the NHS, the National Institute for Health Research, or the Department of Health.

References

Abbott JA, Kaldo V, Klein B et al (2009) A cluster randomised trial of an internet-based intervention program for tinnitus distress in an industrial setting. Cogn Behav Ther 38(3):162–173

American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders, 5th edn. American Psychiatric Association Publishing, Arlington

Andresen EM (2000) Criteria for assessing the tools of disability outcomes research. Arch Phys Med Rehabil 81(12 Suppl 2):S15–S20

Arain M, Campbell MJ, Cooper CL et al (2010) What is a pilot or feasibility study? A review of current practice and editorial policy. BMC Med Res Methodol 10(1):67

Åsberg M, Montgomery SA, Perris C et al (1978) A comprehensive psychopathological rating scale. Acta Psychiatrica Scandinavica 57(S271):5–27

Beck AT, Epstein N, Brown G et al (1988) An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 56(6):893–897

Beck AT, Ward CH, Mendelson M et al (1961) An inventory for measuring depression. Arch Gen Psychiatry 4(6):561–571

Churchill R, Moore THM, Furukawa TA et al (2013) 'Third wave' cognitive and behavioural therapies versus treatment as usual for depression. Cochrane Database Syst Rev. doi:10.1002/14651858.CD008705.pub2

ClinicalTrials.gov Identifier: NCT00724152

ClinicalTrials.gov Identifier: NCT00730834

Cohen S, Williamson G (1988) Perceived stress in a probability sample of the United States. In: Spacapam S, Oskamp S (eds) The social psychology of health: Claremont symposium on applied social psychology. Sage, Newbury Park, CA

Cohen S, Kamarck T, Mermelstein R (1983) A global measure of perceived stress. J Health Soc Behav 24(4):385–396

Davis PB, Paki B, Hanley PJ (2007) Neuromonics tinnitus treatment: third clinical trial. Ear Hear 28(2):242–259

Derogatis LR, Lipman RS, Rickels K et al (1974) The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. Behav Sci 19(1):1–15

Fackrell K, Hall DA, Barry J et al (2014) Tools for tinnitus measurement: development and validity of questionnaires to assess handicap and treatment effects. In: Signorelli F, Turjman F (eds) Tinnitus: causes, treatment and short & long-term health effects. Nova Science Publishers Inc, New York, pp 13–60

Fliege H, Rose M, Arck P et al (2001) Validierung des "Perceived Stress Questionnaire" (PSQ) an einer deutschen Stichprobe. Diagnostica 47:142–152

Gomaa MAM, Elmagd MHA, Elbadry MM et al (2014) Depression, anxiety and stress scale in patients with tinnitus and hearing loss. Eur Arch Otorhinolaryngol 271(8):2177–2184

Hall DA, Szczepek A, Kennedy V et al (2015) Current reported outcome domains in studies of adults with a focus on the treatment of tinnitus: protocol for a systematic review. BMJ Open 11;5(11):e009091.

Hall DA, Haider H, Szczepek AJ et al (2016) Systematic review of outcome domains and instruments used in clinical trials of tinnitus treatments in adults. Trials 17(1):270.

Hallam RS, Jakes SC, Hinchcliffe R (1988) Cognitive variables in tinnitus annoyance. Br J Clin Psychol 27(3):231–222

Hamilton M (1960) A rating scale for depression. J Neurol Neurosurg Psychiatry 23(1):52–62

Hanley PJ, Davis PB, Paki B et al (2008) Treatment of tinnitus with a customized, dynamic acoustic neural stimulus: clinical outcomes in general private practice. Ann Otol Rhinol Laryngol 117(11):791–799

Heijneman KM, de Kleine E, van Dijk P (2012) A randomized double-blind crossover study of phase-shift sound therapy for tinnitus. Otolaryngol Head Neck Surg 147(2):308–315

Henry JA, Dennis KC, Schechter MA (2005) General review of tinnitus prevalence, mechanisms, effects and management. J Speech Lang Hear Res 48(5):1204–1235

Hesser H, Gustafsson T, Lundén C et al (2012) A randomized controlled trial of internet-delivered cognitive behavior therapy and acceptance and commitment therapy in the treatment of tinnitus. J Consult Clin Psychol 80(4):649–661

Hunot V, Churchill R, Teixeira V et al (2007) Psychological therapies for generalised anxiety disorder. Cochrane Database Syst Rev. doi:10.1002/14651858.CD001848.pub4

Hunot V, Moore THM, Caldwell DM et al (2013) 'Third wave' cognitive and behavioural therapies versus other psychological therapies for depression. Cochrane Database Syst Rev. doi:10.1002/14651858.CD008704.pub2

Ireland CE, Wilson PH, Tonkin JP et al (1985) An evaluation of relaxation training in the treatment of tinnitus. Behav Res Ther 23:423–430

ISRCTN registry: ISRCTN17631678

ISRCTN registry: ISRCTN38408464

Jakes SC, Hallam RS, Chambers C et al (1985) A factor analytical study of tinnitus complaint behaviour. Audiology 24:195–206

Joyce J, Herbison GP (2015) Reiki for depression and anxiety. Cochrane Database Syst Rev. doi:10.1002/14651858.CD006833.pub2

Kaldo V, Cars S, Rahnert M et al (2007) Use of a self-help book with weekly therapist contact to reduce tinnitus distress: a randomized controlled trial. J Psychosom Res 63(2):195–202

Kaldo V, Levin S, Widarsson J et al (2008) Internet versus group cognitive-behavioral treatment of distress associated with tinnitus: a randomized controlled trial. Behav Ther 39(4):348–359

Kennedy V, Wilson C, Stephens D (2004) Quality of lie and tinnitus. Audiol Med 2(1):29–40

Kirshner B, Guyatt G (1985) A methodological framework for assessing health indices. J Chronic Dis 38(1):27–36

Kuk FK, Tyler RS, Russell D et al (1990) The psychometric properties of a tinnitus handicap questionnaire. Ear Hearing 11(6):434–445

Langguth B, Landgrebe M, Kleinjung T et al (2011) Tinnitus and depression. World J Biol Psychiatry 12(7):489–500

Lazarus RS (1966) Psychological stress and the coping process. McGraw-Hill, New York

Lee E-H (2012) Review of the psychometric evidence of the perceived stress scale. Asian Nurs Res 6(4):121–127

Levenstein S, Prantera C, Varvo V et al (1993) Development of the perceived stress questionnaire: a new tool for psychosomatic research. J Psychosom Res 37(1):19–32

Lovibond PF, Lovibond SH (1995) Manual for depression, anxiety and stress scales. Psychology Foundation, Sydney

Malouff JM, Noble W, Schutte NS et al (2010) The effectiveness of bibliotherapy in alleviating tinnitus-related distress. J Psychosom Res 68(3):245–251

Mayo-Wilson E, Montgomery P (2013) Media-delivered cognitive behavioural therapy and behavioural therapy (self-help) for anxiety disorders in adults. Cochrane Database Syst Rev. doi:10.1002/14651858.CD005330.pub4

McNeill C, Távora-Vieira D, Alnafjan F (2012) Tinnitus pitch, masking, and the effectiveness of hearing aids for tinnitus therapy. Int J Audiol 51(12):914–919

Newman CW, Jacobson GP, Spitzer JB (1996) Development of the tinnitus handicap inventory. Arch Otolaryngol Head Neck Surg 122:143–148

Olze H, Szczepek AJ, Haupt H et al (2012) The impact of cochlear implantation on tinnitus, stress and quality of life in postlingually deafened patients. Audiol Neurootol 17(1):2–11

Ori R, Amos T, Bergman H et al (2015) Augmentation of cognitive and behavioural therapies (CBT) with d-cycloserine for anxiety and related disorders. Cochrane Database Syst Rev. doi:10.1002/14651858.CD007803.pub2

Pinto PCL, Marcelos CM, Mezzasalma MA et al (2014) Tinnitus and its association with psychiatric disorders: systematic review. J Laryngol Otol 128(8):660–664

Reiss S, Peterson RA, Gursky DM, McNally RJ (1986) Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. Behav Res Ther 24(1):1–8

Robinson SK, McQuaid JR, Viire ES et al (2003) Relationship of tinnitus questionnaires to depressive symptoms, quality of well-being and internal focus. Int Tinnitus J 9(2):97–103

Robinson SK, Viirre ES, Bailey KA et al (2008) A randomized controlled trial of cognitive-behavior therapy for tinnitus. Int Tinnitus J 14(5):119–126

Seydel C, Haupt H, Szczepek AJ et al (2010) Long-term improvement in tinnitus after modified tinnitus retraining therapy enhanced by a variety of psychological approaches. Audiol Neurootol 15(2):69–80

Spielberger CD, Gorsuch RL, Lushene RE (1970) STAI manual. Consulting Psychologists Press, Palo Alto, CA

Távora-Vieira D, Eikelboom RH, Miller S (2011) Neuromonics tinnitus treatment for patients with significant level of hearing loss: an adaptation of the protocol. Int J Audiol 50(12):881–886

Tyler RS, Baker LJ (1983) Difficulties experienced by tinnitus sufferers. J Speech Hear Disord 48(2):150–154

Watts E, Fackrell K, Smith S et al (2016) Why is tinnitus a problem? A qualitative analysis of problems reported by tinnitus patients. 10th International Tinnitus Research Initiative Conference, Nottingham, UK. http://www.tinnitusresearch.org/images/files/meetings/files2016/TRI2016_Abstract_Book.pdf. Accessed 26 May 2017

Williams HC, Burden-Teh E, Nunn AJ (2015) What is a pragmatic clinical trial? J Invest Dermatol 135. doi:10.1038/jid.2015.134

Wilson PH, Henry J, Bowen M et al (1991) Tinnitus reaction questionnaire: psychometric properties of a measure of distress associated with tinnitus. J Speech Lang Hear Res 34(1):197–201

Zigmond AS, Snaith RP (1983) The hospital anxiety and depression scale. Acta Psychiatr Scand 67(6):361–370