Appraisal and evidence synthesis of two reviews

Select two reviews of your choice. One should be a Systematic Review and the other a Traditional Selective Review. Critically discuss the quality of your reviews with respect to a standard assessment tool suitable for reviews.

The aim of a review is to give an overview of the primary studies of a particular subject and analyse them in a way, which is thorough, unbiased and reproducible, should it be repeated.

about the gathering of the primary data that is being reviewed. A systematic review is a method of synthesising primary research and studies. It is based on having a clearly formulated question that uses systematic and explicit methods to identify, select and critically appraise relevant research, and to collect and analyse data from the studies that are included in the review (Cochrane Collaboration Glosasry, 2010). It also uses an objective and transparent approach for research synthesis, with the aim of minimizing bias. Statistical methods known as meta-analysis may or may not be used to analyse and summarise the results of the included studies (Wiesler & McGauran, 2010).

Systematic reviews differ from narrative reviews. A narrative review differs in that the question being answered is usually less specific and has a wider focus of interest (Cook et al, 1997). A narrative review may look more at the interaction of variables within a problem or the development and management of a problem, while a systematic review will usually focus on a specific and answerable question; therefore have a clear defined outcome that it is gathering evidence about. A narrative review summarizes different primary studies from which conclusions may be drawn in to holistic

interpretation contributed by the reviewer’s own experience, existing theories and models with results having a qualitative rather than a quantitative meaning (Campbell Collaborative, 2001; Kirkevold, 1997).

As with all types of research reviews both systematic and narrative need to be appraised for their relevance and value to the subject it relates. A number of tools have been developed to help with the process of critically appraising of research. For the purpose of this assignment, the tool that will be used was produced by the Critical Appraisal Skills Programme (CASP), defining questions and prompts to assess the review, based on questions developed by Oxman et al (1994). This is an appraisal tool that assesses three main issues; is the study valid? What

the results are and whether those results help locally. The aim of this tool has been designed to answer these questions in a detailed and systematic manner. Furthermore, this tool has been selected as it has been specifically designed to support evidence-based practice in health and social care (Public Health Resources Unit, 2006) and therefore it forces the reviewer to also consider the perspective of the patient and it is of relevance to the articles that have been selected here.

The systematic review to be used for this assignment is Gava et al (2009) ‘ Psychological Treatments Versus Treatment as Usual for Obsessive Compulsive Disorder’. This review was chosen from The Cochrane Library, as the systematic reviews here have all been carried out using precise methodology, are updated in line with new research and are specifically

intended to help anyone involved in healthcare, including patients (The Cochrane Collaboration, 2010).

The narrative review to be considered will be Abramowitz et al (2001) ‘ Cognitive Behavioural Therapy for Obsessive Compulsive Disorder: A review of the treatment literature. This review was found in a search on the CINAHL database. The CASP tool to be used is specific to systematic reviews (PHRU, 2006); however it will be adapted here to also appraise the narrative review. The ten questions asked in the CASP tool will now be considered.

Did the review ask a clearly focused question?

Liberati et al (2009) stated that authors should always identify their report as a systematic review or meta-analysis. Although sensitive search strategies have been developed to identify reviews, inclusion of the terms systematic review or meta-analysis in the title may improve indexing and identification (Montori et al, 2005). Furthermore, the title of a systematic review should be informative making key information easily accessible to the reader. This should include reflecting PICOS approach (participants, interventions, comparators, outcomes and study design) providing key information about the scope of the review (Liberati et al, 2009). The systematic review used for the purpose of this assignment was

‘ Psychological Treatments versus Treatment as Usual for Obsessive Compulsive Disorder (OCD) (Review)’. This title failed to provide details of any participants or settings. The intervention and comparator are given as psychological treatments and treatment as usual, however, these had not been clearly specified and details of the outcome measure had not been

provided either. Furthermore, the title simply stated ‘ review’. Such terms do not enlighten the reader whether the review was systematic or whether a meta-analysis was performed (Liberati et al, 2009). Therefore, it has to be noted that the review failed to ask a clearly focused question.

The narrative review of the Abramowitz et al (2001) study was ‘ Cognitive-Behavioural Therapy for Obsessive Compulsive Disorder: A review of the treatment literature’. As with the systematic review this paper also failed to use the PICOS approach, nevertheless, as a narrative review summaries the research more generally, these details may not even be appropriate (Cook et al, 1997). In this review the intervention was specified as Cognitive-Behavioural Therapy (CBT) and it had been clearly indicated that this was a traditional review. Even though the question was not clearly focused, the author does provide the reader with an insight as to what the article contains.

Did the review include the right type of study?

The studies discussed in the systematic review included looking at any psychological treatment compared to any treatment as usual for adults with OCD, which appeared to be appropriate for answering the research question. All of the included studies were randomised controlled trials (RCTs) which appeared to be an appropriate study design with the focus being on the effectiveness of interventions (Hill & Spittlehouse, 2003). However, by including RCT’s only and excluding other forms of research, this may result on limiting the conclusions drawn in the review. This appears of particular relevance to healthcare where qualitative forms of data, such as patient’s

experience of an intervention, should be considered of great importance (Petticrew, 2001).

The studies discussed in the narrative review also seem to address the question as research carried out on CBT for OCD was drawn upon. It also stated that the review focused particularly on exposure and response prevention treatment. This decision appeared to reflect the author's own interest rather than a thorough search of the available literature, which may lead to some form of bias being introduced. The review draws largely upon RCT's and meta-analyses, which may be appropriate as the focus of interest was effectiveness (Hill & Spittlehouse, 2003). However, limiting to the inclusion of only this research design and in a narrative review, it would seem that a range of research could have been incorporated easily as this review appeared limited from not having done so. As explicit details of each study used were not provided, other methodology could have been included but not clearly described.

Did the reviewers try to identify all relevant studies?

To help identify all potentially appropriate research, a thorough literature search must be conducted in order to carry out a systematic review. This involves not only searching electronic databases, but also checking the reference lists of any selected articles in order to identify further research (Greenhalgh, 1997b), searching for non-English language research and unpublished research (Centre for Reviews & Disseminations, 2008). This helps to reduce any bias in research that is published in non-English languages, as research showing significant results tend to be published in

English journals. Furthermore, by having English publications only may have a greater impact on the results than there actually was.

By searching a number of various electronic journal databases, reference lists, unpublished and ongoing research and also research in non-English language indicated that Gava et al (2009) attempted to conduct a thorough literature research.

There were no details or information provided to indicate how the studies were obtained or how the search was conducted for the narrative review. Even though researches from other countries have been referenced, no indication has been given whether this was taken from non-English language publications. Emphasis has been placed on research into exposure and

response prevention therapy and the author’s own work has also been cited. This however places some aspect of bias on the review as the reviewer has a clear interest in the topic currently being reviewed.

Did the reviewers assess the quality of the included studies?

The Gava et al (2009) systematic review describes rigorous quality assessment. A pre-determined scoring system was used which should help to reduce bias by making quality criteria objective and unrelated to the outcomes of the research (Greenhalgh, 1997b). To help reduce the chances of any bias emerging, quality assessments were carried out by two authors independently. However, it must be noted that all articles were included regardless of their quality standard. Furthermore, it has been suggested that

not all the authors made the method of randomisation used explicit and they were not always blind to the allocation.

Abramowitz et al's (2001) review does not mention any formal quality assessment. It must be noted that some informal judgements have been made. Furthermore, it has been noted that some of the research discussed in the review failed to be conducted to a high standard as reference was made to the use of control groups where comparison treatments used was known to be of no benefit to sufferers of OCD. This is unethical as participants allocated to the control groups would have received ineffective treatment which would have had an impact on the results as participants in these groups would have experienced higher anxiety levels and frustration making gains in the experimental group appear much higher. As narrative reviews do not normally employ pre-defined quality assessment tools, research included is likely to be of inconsistent quality and is more at risk of bias (Cook et al, 1997).

If the results of the studies have been combined, was it reasonable to do so?

Results of each study used can be combined in a systematic review to increase the sensitivity to significant results (Crombie, 1996). However, this should only be done when the studies

and their results are similar enough to make the combination of findings meaningful (Crombie, 1996).

The Gava et al (2009) review did combine the results of all the studies reviewed. The results of each study were presented in a forest plot showing

the mean, standard deviation and confidence interval for each. Heterogeneity was also considered both by eyeballing the data and also by statistical analysis, which found no significant differences in the main analyses, but did find significant differences on some of the secondary analyses. General reasons for heterogeneity were discussed at the start of the paper, but the specific heterogeneity found was not discussed in great lengths. Furthermore, a weighting system was used to combine the results. This means that in the analyses stages studies displaying lower level quality were then given a lower weight making this an effective method for including such studies.

The Abramowitz et al (2001) review did not combine the results of all the studies reviewed. In general terms the CRD (2008) state that the intentions of a narrative review is generally to provide a summary of relevant research rather than to synthesis or re-analyse. Furthermore, they also state that this could lead to further biased conclusions based on the reviewers own opinions rather than on stringent analyses that could be recreated by other researchers.

How the results presented and what are the main result?

In the systematic review, the results have been presented in the body of the text and also as in forest plots. CRD (2008) state that results should be expressed in formats that are easily understood. The most commonly used graphic is the forest plot as it provides a simple representation of the precision of individual and overall results and of the variation between – study results (CRD, 2008). The results indicated that psychological

treatments led to more improvements in OCB symptoms than did treatments as usual. Improvements in dropout rates, quality of life, anxiety and depression levels in both psychological and treatment were also taken into consideration. Psychological therapy was broken down further to look at the variables being expressed, the mean differences being calculated for continuous variables, but

it also took into consideration and reported on the individual effects of CBT, Cognitive Therapy and Behaviour Therapy.

Abramowitz et al (2001) review also discussed results for each study reviewed. This indicated that exposure and response prevention were both effective therapies for OCD. However, since no statistics had been used then it was not possible to assess how large and significant the results were without referring to the original research. The research presented tends to be interpreted in favour of exposure and response prevention therapy, even with the results appear to suggest little difference (Jüni et al, 2001). The conclusions drawn here are based on the author’s own opinions and could have been subjected to biased interpretation of results or detection bias (Jüni et al, 2001).

How precise are the results?

The use of confidence intervals in systematic reviews gives the reader an idea of how precise the results can be considered to be (Hopkins, 2001). Confidence intervals describe the range within which a result for the whole population would occur for a specified proportion of times a survey or test was repeated among a sample of the population. Confidence intervals are a

standard way of expressing the statistical accuracy of a survey-based estimate (Young & Bolton, 2009). The confidence interval expresses the uncertainty around the point estimate, describing a range of values within which it is reasonably certain that the true effect lies; wider confidence interval reflects greater uncertainty (CRD, 2008). Where a 95% confidence interval is reported then we can be reasonably confidence that the range included the ‘ true’ value for the population as a whole. Formally we would expect it to contain the ‘ true’ value 95% of the time (Young & Bolton, 2009). Although intervals can be reported for any level of confidence, in most systematic reviews of health interventions, the 95% confidence interval is used (CRD, 2008).

Eight studies (11 study comparisons) were identified, all of which compared cognitive and/or behavioural treatments versus treatment as usual control groups. Seven studies (ten comparisons) had usable data for meta-analyses. These studies demonstrated that patients

receiving any variant of cognitive behavioural treatment exhibited significantly fewer symptoms post-treatment than those receiving treatment as usual (SMD -1. 24, 95%, CI -1. 61 to -0. 87, 1^2 test for heterogeneity was not significant at 33. 4%) (Gava et al, 2009). Different types of cognitive and / or behavioural treatments showed similar differences in effect when compared with treatment as usual. Results obtained for CBT on OCD symptoms exhibited that the overall mean difference (fixed effects) was in favour of psychological treatments (WMD -7. 73, 95%, CI -9. 92, -5. 55). The 1^2 test of heterogeneity was not significant at 33. 4%. Results obtained for cognitive therapy on OCD symptoms exhibited that the overall standardised

mean difference (random effects) were slightly in favour of psychological treatments (SMD -1. 20, 95%, CI -2. 66, 0. 25). The 1^2 test of heterogeneity was not significant at 74. 2%. The overall treatment effect appeared to be influenced by differences in baseline severity (Gava et al, 2009).

Within the Abramowitz et al (2001) paper, there does not seem to be any confidence intervals apparent within, thus leaving us unable to analyse how precise the results, which are described within the paper truly are. In comparison to the systematic review, this approach appears to be highly less open; the author is able to narrate the results in however way they please for example they could suggest that they are highly significant or interpret them as equal to their own theoretical standing but without the need of referral to the original case, however there is no way a certain conclusion as to whether this has occurred or not be drawn up.

Can the results be applied to the local population?

Within the Gava et al (2009) paper, population details and setting for each study do not seem to be clearly apparent. Although a high percentage are noted as outpatient, a number of settings were not highlighted as to their location or setting. Due to this lack of knowledge, and that some of the studies could have been carried out in a inpatient setting, it would not be possible to ascertain a generalised result to the local population with information found in the ‘ inpatient studies’. It was stated that the statistical data, or demographics, of the participants

were of a similar and consistent nature however there was no in depth analysis or description of them. Due to this small fact, therefore it would not

have been possible to state that these were similar to the local population. The duration of treatments described when provided locally was usually shorter thus it can assume some of the methods could not be provided within the local settings. Therefore, the generalisation of the results requires some degree of accuracy and precision for there are often vast and clearly significant differences between the sample and the settings used, here and the local population.

Once again, within the Abramowitz et al (2001) review, there was no description of the population meaning that it was not possible to ascertain a generalised collaboration of results. Furthermore, as one of the studies had been carried out in an inpatient setting, then it would not have been possible to generalise these results within such settings. The duration of these studies are described as being of 60-90 minutes on a daily basis which frankly could not have been possible locally. Therefore it is vital that care is taken when attempting to create a generalised result from this review as things that appear to be insignificant, as such small details can be inevitably crucial elements.

Were all the important outcomes considered?

The Gava et al (2009) review discussed various outcomes as well as improvements in OCD symptoms. Also considered were the outcomes for anxiety symptoms, depressive symptoms, dropout rates and quality of life. These elements were not only likely to be of a high degree of importance to the service providers but also to the service users themselves. Service leads, mental health commissioners and policy makers / government officials may

also be interested in dropout rates and reasons for dropout i. e. at initial assessment or at follow up appointment. Details of each therapy session missed, cancelled, failed to attend would be important as well.

The narrative review also discussed various improvements in OCD symptoms with each intervention. Consideration was given to the distress linked with exposure therapy, as it would be an important factor for all parties concerned i. e. patient and their families, and also taking into account the supporting role of the carer while patient receives therapy.

Abramowitz et al (2001) review state that a relatively large number of clients refuse participation in exposure therapy because of anxiety evoking elements of treatment (i. e. confronting feared stimuli). This places emphasis on the importance of understanding the rationale for exposure therapy, demonstrating mastery of case conceptualization, and providing the client with a convincing explanation of why confronting feared situations will result in long-term abatement of obsessions and compulsions.

Should policy or practice change as a result of the evidence contained in the review?

The current use of CBT for OCD is supported by the Gava et al (2009) review. Despite this, any RCT or other psychological therapy was unidentifiable by the author whilst other types of therapy did not fit the required format needed in order to carry out a RCT to the same quality that a CBT does thus meaning it has more controlled research in comparison to other forms of therapy. One criticism of systematic reviews is their bias towards certain methodologies for important research may not be included thus leading to

biased conclusions. Further research is therefore needed to make comparisons between CBT and other types of therapy; it may be that a mixed methods review would have been carried out in order to do this effectively (Dixon-Woods et al, 2004)

A biased review is also present in Abramowtiz et al (2001). In this case it is towards studies in support of exposure and response prevention therapy. No decision can be reached from this review for a thorough search to identify all research in the area -quality assessment and the clear display of results – was not completed. Conclusions are more likely to be biased and not be representative of all the evidence in the field within narrative views for they do not follow a predefined and set methodology thus, narrative reviews are less likely to influence policy and practise.

Conclusion

It is evident that both methodologies consist of advantages and disadvantages. Systematic reviews use pre-defined methodologies with the intention of reducing bias making the results to appear more robust. Nevertheless a critique appears to be the evident over emphasis of RCT which is argued to find less generalisable results. On the other hand, narrative review results appear to be more generalisable and tend to be more flexible in the incorporation of other methodologies. However within these reviews, the scientific, pre-defined strategies are not employed, thus considering them to have a higher risk of leading to bias. To conclude, it appears that both of these reviews can be of significant help and are easily

justifiable for use, depending on the question that the researcher is trying to answer and the point at which the evidence base is at in its development.