The Internet Journal of Allied Health Sciences and Practice

<http://ijahsp.nova.edu>

A Peer Reviewed Publication of the College of Allied Health & Nursing at Nova Southeastern University

Dedicated to allied health professional practice and education

<http://ijahsp.nova.edu> Vol. 3 No. 1 ISSN 1540-580X

Development of a Generic Critical Appraisal Tool by Consensus: Presentation of First Round Delphi Survey Results

Jeannie Burnett, M.App.Sc

Karen Grimmer, PhD

Saravana Kumar, M.App.Sc

Center for Allied Health Evidence

Citation:

Burnett, J., Grimmer, K., Kumar, S. Development of a Generic Critical Appraisal Tool by Consensus: Presentation of First Round Delphi Survey Results. *The Internet Journal of Allied Health Sciences and Practice*. January 2005. Volume 3 Number 1.

ABSTRACT

The growing importance of evidence based practice is necessitating academics and clinicians to be able to make judgments about the quality of the body of research evidence pertaining to clinical questions. There are numerous critical appraisal tools to assist this process. These are mostly designed for specific research designs, and tend not to reflect the particular concerns of allied health professionals, such as accuracy of diagnosis, adequate description of intervention, and sensitivity and utility of outcome measures. This paper reports the findings of a study which sought expert opinion on the essential criteria for critical appraisal, and whether a generic critical appraisal tool could be developed for allied health use. A modified Delphi technique was used to identify experts, and determine key criteria. Fifteen Australian allied health professionals participated, and identified key criteria as clinical relevance, methodological robustness, statistical robustness, aims that are clearly stated and conclusions that are reasonable considering the results. In terms of the development of a generic critical appraisal tool for all research designs, the opinion was that to adequately deal with critical appraisal of qualitative and quantitative research designs within a generic tool would be challenging.

INTRODUCTION

Critical appraisal forms the basis of uptake of evidence in clinical practice. It is through the application of critical appraisal that researchers, clinicians and other stakeholders in health care can evaluate the strength of available evidence.¹⁻³ This process enables stakeholders to make informed judgments about the effectiveness of therapies.⁴⁻⁷ Historically, evidence based medicine began in medical disciplines, with a recent adoption into of allied health.⁴ A recent systematic review of critical appraisal tools found one hundred and ninety-three different published critical appraisal tools.⁸ The 108 papers that were included in the review were in the most part, specific to quantitative research designs with very few being developed specifically with allied health requirements in mind. This review found no "Gold Standard" critical appraisal tool, and identified the need to further investigate the needs of allied health.

Allied health interventions differ from medical interventions in the following ways:

- The use of clinically reasoned diagnosis as opposed to the applicability of diagnostic tests and imaging that is available in medical diagnosis.⁹
- Often multiple interventions are provided in one treatment session. This requires research to provide reproducible descriptions of interventions in terms of their relevance to the diagnosis, their intensity and frequency, their order of administration and instructions given to the patient.¹⁰
- Clients are often seen over a period of time (an episode of care). This necessitates regular follow-up to ascertain short and long term effectiveness of interventions.^{11, 12}

- The need to demonstrate the use of appropriate outcome measures that can sensitively and reliably detect a change in impairment, function and participation status.¹³ Such outcome measures should reflect the needs of relevant stakeholders (clinician, patient, insurer etc).^{9, 10, 12}

For these reasons critical appraisal tools that do not reflect the perspectives of allied health may not provide sufficiently sensitive or appropriate information about the quality of the body of research evidence for therapies. While there are critical tools developed by Allied Health, the systematic review by Katrak et al identified that consensus as to the appropriate criteria in critical appraisal tools is lacking.^{8, 14-16}

The other feature of existing critical appraisal tools is that they are predominantly design- specific. The question therefore is whether a generic critical appraisal tool that can be applied across quantitative and qualitative research designs can be constructed. The abovementioned review of critical appraisal tools identified five papers that either presented a generic critical appraisal tool, or alluded to potential criteria relevant to such a tool.^{8, 17-21}

Common themes were:

- Sufficient number of people included in the study (power calculations for experimental designs).
- Was the study methodology described in sufficient detail to allow replication?
- Was subject compliance dealt with? eg was the intervention acceptable to subjects.
- Do the conclusions make sense biologically, sociologically and economically?
- Are the results of the study applicable in a clinical setting and was clinical relevance of the findings dealt with.

These common generic criteria provide a reasonable basis upon which to develop an allied health generic tool. The aim of this study was to determine consensus amongst content experts as to essential criteria for a generic tool that could be applied across research designs and which was applicable to allied health requirements.

METHOD

A Modified Delphi Technique was used to investigate this question.²² Delphi surveying technique is designed to turn opinion into consensus via asking content experts questions which are then coded into key issues. These issues are re-presented to the respondents for further consideration and comment.²³⁻²⁵ Delphi and other consensus gaining techniques have been previously been used in the development of critical appraisal tools.^{1, 26-33}

Subject Recruitment:

For the purposes of this study an expert was defined as someone who has a known or stated interest in the topic.³⁴ In the first step, heads of Schools of Physiotherapy, Podiatry, Speech Pathology, Occupational Therapy, Nursing and Medical Radiations around Australia were contacted. They were asked to identify colleagues who had an interest in Evidence Based Practice. This list was supplemented by names of colleagues known to the authors as having an interest in evidence based practice. Those on this list were contacted via e-mail and invited to participate in the survey. They were advised that they would show their consent by completing and returning the questionnaire provided as an attachment to the email.

In step 2 a list of Allied Health professionals employed by university faculties was derived by performing an internet search of all Australian Allied Health faculties. Those who stated an interest in Evidence based practice in their staff biography. This list was mutually exclusive to the list derived in step.¹

Inclusion and Exclusion Criteria:

Inclusion was based on a known or reported interest in EBP. Response to the invitation to participate was considered to fill the inclusion criteria, and implied consent to participate in the study. There were no exclusion criteria, as all individuals approached were considered to have a significant interest in evidence-based practice.

Information Provision

Those on the list were e-mailed a letter inviting them to fill out a questionnaire that was sent as an attachment. Consent was obtained through the completion of the questionnaire (see appendix 1 -consent paragraph).

The questionnaire recorded the following information:

- Demographic data: profession and position currently held.

- How many times participants had used a critical appraisal tool in the past month.
- Which critical appraisal tool(s) they used most frequently.

Data Collection and Synthesis:

Respondents to the questionnaire were asked to list the core elements of a critical appraisal tool (see Figure 1). These elements were divided into those pertaining to internal validity of a study, and those to external validity. Respondents were initially asked to rank the elements in order of importance using numbers (1 being the most important). Finally respondents were asked to identify up to three colleagues who also had an interest in EBP. Those named would also receive an invitation to participate by completing the questionnaire.

Figure 1: Questionnaire emailed to expert list.

FIRST ROUND DELPHI QUESTIONNAIRE

Development of a critical appraisal tool for use with allied health research.

Thank you for agreeing to fill out the following questionnaire.

PROFESSION:

CURRENT PLACE OF EMPLOYMENT AND POSITION HELD:

1. How many times in the past month have you used a critical appraisal tool (approx).
2. Please indicate the tool(s) that you use most frequently (optional).
3. Please list criteria relevant to internal validity.

In the right hand column rank numerically those that you consider essential.

CRITERIA	RATING

4. Please list criteria relevant to external validity.

In the right hand column rank numerically those that you consider essential

CRITERIA	RATING

4. Finally we ask that you identify up to 3 colleagues who share an interest in evidence based practice. They will be asked to fill out section 1 and 2 of this form. This section is optional.

1. ...
2. ...
3. ...

Thank you for finding the time to fill out this questionnaire.

The interview delivered questionnaire was conducted in a semi-structured manner, hence rating of criteria was not called for. The interviews were conducted face to face or by telephone. The interviewer (JB) transcribed the interviews, which were then forwarded to the interviewee for verification. All interviewees were invited to alter the transcript to ensure that it represented their view correctly. This validated transcript was then used in data analysis. The interviewees were also asked for their opinion as to whether a generic tool would be useful in the process of collation of evidence for Allied Health therapies.

Due to the two methods of questionnaire delivery, a snowballing approach was taken when combining the findings to ensure validity of findings.

Data Collation and Analysis:

The responses were collated by one person (JB) and the list was cross-checked by a second independent person (KG) to address potential bias in the inclusion of appraisal elements.

Data collation was undertaken using an Excel spread sheet. The frequencies with which items were mentioned were tabulated. The frequency with which a criterion was entered in order of preference (as first in importance, second in importance) was determined in order to determine the relative importance of each criterion. A summary list that was reflective of responses was then developed.

RESULTS

Responses to Surveying

Figure 2 provides a flow chart of responses to contacts via the two steps in the emailed survey

Figure 1: Flowchart of Response to Email Rounds

The respondent sample characteristics are presented in Tables 1 and 2 in terms of profession and employment type.

Table 1: Proportion of respondent sample academic versus clinician.

	Questionnaire (n=8)	Interview (n=7)	Total
Academic only	6	5	11(73%)
Clinical only	1		1(7%)
Clinical and Academic	1	2	3(20%)
Totals	8	7	15 (100%)

Table 2: Breakdown of respondent sample by profession.

Profession	Questionnaire (n=8)	Interview (n=7)	Total
Physiotherapist	7	2	9 (60%)
Occupational Therapist	1	3	4(26%)
Podiatrist	1	1	2(13%)
Social Worker			
Speech Therapist		1	1(1%)

As is evident from the critical appraisal usage responses from the sample, detailed in Table 3; 10 of the respondents stated that they had not formally used a critical appraisal tool in the past month. Of these, one indicated that in the past 6 months she had read 10 articles at least per week for her doctoral studies and had used a mental checklist to rate the quality of the article. Four are involved in actively teaching evidence-based practice, one held a senior clinical position and the other two held academic positions. Thus despite the lack of recent use of a critical appraisal we were satisfied that those responding fitted our definition of an expert.³⁴

Table 3: Frequency of use of critical appraisal tools.

Times in the past month used CAT	
Daily	1
2-3 daily	1
2 times past month	1
4	1
5	1
0	10

Tables 4 and 5 present the frequencies of criterion as mentioned in the questionnaire and interviews.

Table 4: Criteria code for internal validity and occurrence within the responses.

Criteria	Questionnaire Frequency	Interview Frequency	TOTAL
Internal validity			
Outcome measures psychometry (ie reliable and valid).	5	4	9
Baseline equivalence in characteristics (including outcome measures) and inclusion/exclusion criteria stated.	5	1	6
Appropriate study question and design.	3	4	7
Randomisation	4	1	5
Assessor blinding	4		4
Sampling (technique, power, size)	3	12	15
Intention to treat, follow-up of dropouts, dropout <85%	5	1	6
Therapist blinding	1		1
Subject blind and compliant	1		1
Presence of bias.	1		1
Triangulation (Qual)	2		2
Quotes provided (qual)	1		1
Ethics approvals noted	1		1
History (consideration of natural history as effect or maturation)	2		2
Appropriate statistical tests	5	1	6
Confounders and effect modifiers dealt with	1	1	2
Methodological robustness		3	3

Table 5: Occurrence of criterion for external validity within the responses.

External validity:	Frequency	Frequency	Total
Clinical relevance (includes applicability)	9	9	18
Statistical significance (including effect size)	6		6
Sample representative of larger population	3	4	7
Effect modifiers identified.	1		1
Aims clear and contextualized	1	2	3
Conclusions appropriate and relevant to results.	2	3	5
Drop outs reported	1		1
Time study undertaken	1		1
Patient compliance reported.	1		1
Intervention described	1	1	2
Applicability to population (culturally sensitive)		1	1
Divergent findings		1	1
Presentation of stream to coding.		1	1
Issues re bias in publication.		1	1

Coding the 'Clinical Relevance' Responses

The most complex responses pertained to clinical relevance of the research. Examples are provided below of quotations that were coded as 'clinical relevance'.

- "generalisability 2.extent to which theory derived from the research can be applied in other settings 3. Reproducibility of methods/findings."
- "Discussion of clinical importance"
- "Implication for field research. Clinical relevance"
- "Specification of the intervention (dosage-response effect and timing; and disease outcomes, disease severity, co-interventions and patient characteristics."
- "...intervention appropriate to clinical setting (number of treatments. type of treatment etc)."

Coding the responses relevant to the other themes was straightforward and thus pertinent quotations were not considered necessary for this section.

In the interview round of surveying only five new criteria were mentioned. These were;

1. Methodological Robustness
2. Applicability to Population (in particular taking into account the cultural mix of the target population).
3. Divergent findings are presented.
4. Presentation of stream to coding.
5. Issues concerning bias in publication stated..

It illustrates that snowballing occurred with this second sample. It is apparent that some measure of how well the study was carried out (within the boundaries of whatever design has been chosen) was considered important. Therefore in reality only four new criteria were mentioned as "methodological robustness" really encompasses the other criteria stated that deal with design specific requirements (eg sampling technique and randomisation for certain experimental designs).

The common themes of criteria from the sample were:

- Having a justified size, well described sample was important in making a judgment about the internal validity as well as how generalisable (external validity) the results were.
- That an appropriate study design for the research question was important.
- Statistical tests that are appropriate to the design and research question.
- Design specific indicators of robustness are important. Twenty nine of the responses were related to design specific features. These came from categories relating to blinding, randomisation, and statistical appropriateness for quantitative, and triangulation and provision of streaming to coding for qualitative.

- Clinical relevance and applicability is vital in giving research results meaning, particularly when investigating clinical interventions. It is of note that one respondent indicated that clinical relevance was not an important part of critical appraisal. It was her opinion that this was the responsibility of the reader, not the author, to determine this construct. Clinical relevance can encompass the consideration of whether there is a clinically significant effect size, whether an intervention is applicable to the clinical setting, and whether the intervention or phenomena under investigation is important to the various stakeholders.
- Outliers in the data set need to be investigated, using intention to treat analysis where appropriate or presentation of outlying data or opinions. Also important here was that dropouts were at an acceptable level.

Opinion Surrounding the Development of a Generic Critical Appraisal Tool for Allied Health.

In comments made as part of the questionnaire all eight of the questionnaire respondents indicated that design specific tools are perhaps the most valid. For example "I believe that the most appropriate items in a critical review form will depend on the type of evidence that is evaluated. For example, the qualitative and quantitative paradigms are completely different and therefore different items would be appropriate in each situation."

The range of responses in the interview to the question of the usefulness and plausibility of developing a generic critical appraisal tool are quoted below;

- "Not useful at all. Critical Appraisal Tools need to be design specific. Can't see the call for it. As a profession physiotherapy deal with multiple outcome measures, so can deal with interpreting systematic reviews based on different critical appraisal tools." (PT)
- "Did mention that a tool for Allied Health would be beneficial as tools (eg Sackett criteria) are based on medical model." (OT)
- "Challenge is to capture the issues of internal validity that are design specific" (ST)
- "Yes.....Problem with PEDro (14) is that it only applies to certain designs...a tool that applied to cohorts, case control as well for example would be helpful. Interested to see design features pertaining to internal validity included." (PT)
- "Not sure that Generic tool would be useful" (OT)
- "...Qualitative and quantitative studies have a different place in evidence. They ask quite different questions. Qualitative asks for opinion, experience of the subjects....quantitative looks for an effect. Therefore potentially a generic tool not helpful in determining levels of evidence."

The last quotation came from a participant who had a self professed bias towards qualitative research.

Usefulness of a Generic Critical Appraisal Tool

Table 6 outlines an even spread of opinion (for and against) the development of a generic critical appraisal tool. The themes relevant to the comments from those respondents who were not in favour of a generic tool were around the challenge of including sufficient sensitivity regarding design specific features. Those that were for the development of a generic critical appraisal tool highlighted the need for consensus to develop such a critical appraisal tool. This group also pointed out the challenge of dealing with design specific criterion within a generic tool.

DISCUSSION

This is the first known Australian study which has attempted to develop a general critical appraisal tool relevant to allied health. The overall findings from this study were the importance of clinical relevance, sample size and characteristics and design specific robustness as features of a critical appraisal tool (CAT). The sample in this study was biased towards academics (73% employed in academic, 23% in combined clinical and academic roles). Only one respondent came from a clinical background. It is reasonable to expect that a sample of clinicians may provide a different set of criteria. The sampling process used in this study attempted to apply a systematic approach to inviting both clinicians and academics to respond. It is perhaps indicative of where the interest and work in evidence based practice is done that such an academic sample was attained. Despite the loading of academics in the sample it is noteworthy that criteria pertaining to clinical relevance were a prominent feature of the criteria list.

The question of whether this sample is representative of the broader Allied health professions remains. The sample size was small (n=15), and the major proportion of respondents came from the Physiotherapy and Occupational professions. No Social workers responded to our invitation to participate. Two speech therapists agreed and one was interviewed; the other for practical reasons could not be contacted within the time frame of the study. Future work into developing consensus into any critical

appraisal tool for Allied Health should seek to access a wider scope of professionals. Despite this limitation, the fact that snowballing occurred supports the validity of the spread of opinions gained.

One methodological limitation was the exclusive use of e-mail to contact potential collaborators. This mode of contact is inherently fallible. Many academics delete e-mails that are not directly related to their work. This is necessary as they are faced with a lot of e-mails each day. It does pose an issue in terms of low response rates. This is a problem inherent in using email for recruitment. As there is no ideal sample size for Delphi techniques (22-25), it was decided that once we had attained fifteen and discovered that snowballing was occurring, that recruitment could cease. Future studies may use mail-out and e-mail to attempt to improve the hit rate.

When considering the validity of this data it is pertinent to compare the list generated from the systematic review of critical appraisal tools (8) with the themes generated from this study sample. Those themes were;

- Was the study methodology described in sufficient detail to allow replication?
- Was subject compliance dealt with? eg was the intervention acceptable to subjects.
- Are the results of the study applicable in a clinical setting and was clinical relevance of the findings dealt with.
- Do the conclusions make sense biologically, sociologically and economically?

Comparing to the consensus list generated by this survey, all except the last point occurs. This means that triangulation can be considered to have occurred, thus providing a measure of validity to the results of this study. This supports the statement made in the introduction that these themes form a reasonable the basis for further development of a generic critical appraisal tool.

One of common features of the responses was the opinion that clinical relevance is not dealt with well in existing tools. Those issues pertaining to clinical relevance are not given enough weight. This is consistent with the findings of the abovementioned systematic review. It is in this area that any tool being developed would need to focus on.

Development of a Generic Critical Appraisal Tool

Participants' opinions around the development of a generic tool generic tool ranged from at best challenging, and at worst impossible. The main reason given for this was that it was felt that inclusion of design specific features is vital in determining the quality of an article. It is seen that how a study is carried out that can determine the value of the results. There are different design features that apply to quantitative and qualitative designs. For a generic tool to be useful both paradigms need to be represented so as to gain a sensitive measure. There has been considerable debate in evidence based practice writing as to how much importance is placed on internal validity compared to clinical applicability. It has been suggested that more importance be placed upon how applicable findings are, how homogenous samples are across studies, and how similar to a clinical group they are (18,20,21) . Other authors state that internal validity and design specific features need to feature heavily in quality appraisal tools (19,35). The conclusion here seems that a generic tool needs to include a means to rate the quality of the way that particular design was carried out. This may be done with a feeder question to another tool, or with a question that links to explanatory notes about how to rate the internal validity of different designs within the tool. This is the challenge for the next phase of this project.

Opinion as to whether a generic critical appraisal tool would be helpful in collating evidence for Allied Health was not consistent amongst our respondents. Negative opinion centred on the feeling that a generic tool may potentially be not critical enough. That key features of internal validity that are specific to the research design could be understated in such a tool. Were opinion was positive, this was tempered by the feeling that to develop such a tool would be a "huge ask" (ST). The theme of the opinions can be summarised by considering an analogy. Is it better to have two different size screwdrivers that serve specific purposes, or one that can do both jobs? The risk is that the medium size screwdriver may do neither job well enough. The challenge has been thus laid out to develop a tool that is generic, but does deal with issues of internal validity adequately.

CONCLUSION

The results of this survey identify six main themes that content experts in the sample consider essential in a critical appraisal tool. These are:

- Clinical relevance and applicability.
- Sample characteristics (size, techniques and inclusion/exclusion criteria).

- Statistical and methodological robustness.
- Conclusions that are appropriate given the results ("not making a mountain nor a molehill out of the results").
- Study design that is appropriate to the question.
- Aims that are clear and contextualised.

These are criteria that can be applied across quantitative and qualitative research paradigms, so can serve as a basis for the development of a generic critical appraisal tool.

The question of whether it is better to have two different size screw drivers for the job, or one medium size one to do both jobs, could be resolved through the process of developing a draft tool and testing it for validity and reliability. This would comprise a second phase of this Delphi study and further consensus gaining studies targeting Allied Health Practitioners.

REFERENCES

1. Oxman, AD, Guyatt GH. *Validation of an index of the quality of review articles*. Journal of Clinical epidemiology 1991; 44: 1271-1278
2. Dickersin K, Berlin JA. *Meta-analysis: state of the science*. Epidemiological review 1992; 14:154-176
3. Mulrow CD. *Rationale for systematic reviews*. BMJ: 1994;309: 597-599.
4. Thomson-O'Brien MA, Moreland, J. *Evidence-based Information Circle*. Physiotherapy Canada 1998;50:171-205.
5. Sackett DL, Straus SE, Richardson WS, Rosenberg W and Haynes RB. *Evidence-based Medicine. How to practice and teach EBM*. 2nd Edition. Churchill Livingstone; London; 2000
6. Oxman AD, Sackett DL, Guyatt GH et al. *User's guide to the medical literature. How to get started*. JAMA 1993;270:2093-2095.
7. National Health and Medical Research Council: *How to Review the Evidence: Systematic Identification and Review of the Scientific Literature*. Canberra; 2000.
8. Katrak P, Bialocerkowski AE, Massy-Westropp N, Kumar VS, Grimmer KA. *A systematic review of the content of critical appraisal tools*. BMC Medical Research Methodology 2004; 4:22 (16 September 2004)
9. Bialocerkowski AE, Grimmer KA, Milanese SF, Kumar S: *Application of current research evidence to clinical physiotherapy practice*. Journal of Allied Health Research [in press].
10. Grimmer K, Milanese S, Bialocerkowski A *Clinical guidelines for low back pain: A physiotherapy perspective*. Physiotherapy Canada 2003; 55(4): 1-9.
11. Grimmer K, Bowman P, Roper J. *Episodes of allied health outpatient care: an investigation of service delivery in acute public hospital settings*. Disability and Rehabilitation 2000;22(1/2): 80-87.
12. Grimmer KA, Milanese S, Bialocerkowski AE, Kumar S. *Producing and implementing evidence in clinical practice: the therapies' dilemma*. Physiotherapy 2004.[in press]
13. www.who.int/classifications/icf/en/
14. Physiotherapy Evidence Database (PEDro). <http://ptwww.cchs.usyd.edu.au/pedro/> October 2003.
15. De Vet HCW, De Bie RA, Van der Heijden GJ, Verhagen AP, Sijpkens P, Kipschild PG. *Systematic reviews on the basis of methodological criteria*. Physiotherapy 1997;83: 284-289.
16. Law M, Stewart D, Pollock N et al. *Occupational Therapy Evidence-Based Research Group*: McMaster University. <http://www-fls.mcmaster.ca/rehab/ebp/>
17. Forrest JL, Miller SA. *Evidence-based decision making in action: Part 2-Evaluating and applying the clinical evidence*. The Journal of Contemporary Dental Practice 2002; 4:42-52
18. Lohr KN and Carey TS. *Assessing 'best evidence': issues in grading the quality of studies for systematic reviews*. Joint Commission J Qual Improvement 1999; 25:470-479
19. Harris RP, Helfand M, Woolf SH, et al. *Current methods of the U.S. Preventive Services Task Force: A review of the process*. American Journal of Preventive Medicine 2001;20:21-35
20. Evans D. (2003) *Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions*. Journal of Clinical Nursing; 12: 77-84.
21. Chestnut RM, Carney N, Maynard H, Patterson P, Mann NC, Helfand M. *Rehabilitation for Traumatic Brain Injury*. Evidence Report/Technology Assessment No.
22. McKenna, H.. *The Delphi technique: a worthwhile approach for nursing?* Journal of advanced Nursing 1994; 19: 1221-1225.
23. Williams, PL, and Webb, C. *The Delphi Technique: a methodological discussion*. Journal of Advanced Nursing 1994; 19: 180-86.
24. Linstrome, H. and, Turoff, M. *The Delphi Method; techniques and applications*. London; Wesley Publishing;1975.

25. Hasson, F., Keeney, S. and McKenna, H (2000). *Research guidelines for the Delphi survey technique*. Journal of advanced Nursing; 32(4) : 1008-1015.
26. Oxman AD, Guyatt GH, Singer J, Goldsmith CH, Hutchinson BG, Milner RA, Streiner DL: *Agreement among reviewers of review articles*. J Clin Epidemiol 1991;44:91-98.
27. Guyatt GH, Sackett DL, Sinclair JC, Hoyward R, Cook DJ, Cook RJ: *User's guide to the medical literature. IX. A method for grading health care recommendations*. J Am Med Assoc 1995;274:1800-1804.
28. Cranney A, Tugwell P, Shea B, Wells G: *Implications of OMERACT outcomes in arthritis and osteoporosis for Cochrane metaanalysis*. J Rheumatol 1997;24:1206-1207.
29. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, DeVet HCW: *Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative*. Clin Chem 2003;49:1-6.
30. Downs SH, Black N: *The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions*. J Epidemiol Comm Health 1998;52:377-384.
31. Sindhu F, Carpenter L, Seers K: *Development of a tool to rate the quality assessment of randomized controlled trials using a Delphi technique*. J Advanced Nurs 1997;25:1262-1268.
32. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. *Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of observational studies in epidemiology (MOOSE) group*. J Am Med Assoc 2000;283:2008-2012.
33. Verhagen AP, de Vet HC, de Bie RA, Kessels AGH, Boers M, Bouter LM et al. *The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus*. J Clin Epidemiol 1998;51:1235-1241.
34. Sumison, T .*The Delphi Technique: an adaptive research tool*. British Journal of Occupational Therapy: 1998; 61(4): 153-156.
35. Greer N, Mosser G, Logan G, Halaas GW *A practical approach to evidence grading*. Joint Commission Journal of Quality Improvement 2000; 26:700-712.

Appendix 1: Introduction and invitation to participate- letter emailed to list of content experts.

28th April.

To whom it may concern,

We are writing from the Centre for Allied Health Evidence (University of South Australia) to ask for your collaboration in identifying and testing criteria for a critical appraisal tool for use with allied health literature.

In a systematic review it was identified that there are over 190 published tools, containing a wide variety of criteria on which studies are critically appraised. The lack of a gold standard means that researchers face the decision of which tool to use, and indeed the results of reviews often need to be considered in the context of the chosen tool. Our aim is therefore to gain consensus on the criteria that are important for the appraisal of allied health research of quantitative and qualitative research designs. It is hoped that this will provide the Allied Health professions with a tool that is applicable across quantitative and qualitative research designs.

This phase (phase one) involves asking for your participation by:

- Ø Part 1: Identifying up to 3 Allied Health professionals with a similar interest in evidence based practice. These people will receive a similar invitation to participate.
- Ø Part 2: Identify the core elements of a critical appraisal tool for allied health evidence, that could cross quantitative and qualitative research boundaries

Ethics approval has been gained through the University of South Australia Ethics Committee.

ATTACHED IS A BRIEF QUESTIONARE THAT YOU CAN FILL OUT, SAVE, THEN E-MAIL BACK TO THE ADDRESS AT THE BOTTOM OF THIS LETTER.

The responses to part 2 will be collated and rated as to the level of importance each criterion is given. The Delphi technique of surveys will be employed. This means that the collated data will be submitted to you for further comment. In this way consensus as to the core criteria for a critical appraisal tool for Allied Health will be attained. This means that you will receive two e-mails for this part of this project.

The result will be that a critical appraisal tool will be developed. In the next phase we will undertake reliability and validity testing of the said tool. This will involve sending the tool out to you and asking you to apply it to a relevant paper in your field. It would be appreciated if you would indicate in your response your availability for this testing.

If you do not wish to participate could you reply and indicate so.

Thank you for considering this collaborative project,

We hope that you agree to participate and look forward to receiving your responses to the following address:

Jeannie Burnett-
Centre for Allied Health Research- University of South Australia
jeannie.burnett@unisa.edu.au
Phone: 08-83022086 (Wednesdays only)
We ask that you respond by 31st May 2004.
Yours sincerely,

Ms Jeannie Burnett- M.App.Sc (Physiotherapy).

Associate Professor Karen Grimmer- Director CAHE.

CONSENT

Project Title: **Development of a critical appraisal tool for use with allied health research.**

Primary Researcher's name: Karen Grimmer

I have received information about this research project.

I understand the purpose of the research project and my involvement in it.

I understand that I may withdraw from the research project at any stage.

I understand that while information gained during the study may be published, I will not be identified and my personal results will remain confidential unless I wish to be named as a collaborator.

I have provided information about the research to the research participant and believe that he/she understands what is involved.

I understand that by completing and returning the questionnaire attached, I am giving consent.