European Journal of Integrative Medicine 7 (2015) 423–428

Contents lists available at ScienceDirect

European Journal of Integrative Medicine

journal homepage: www.elsevier.com/eujim

Research paper

Delphi methodology in health research: how to do it?☆

Esmé G Trevelyan*, Prof Nicola Robinson

School of Health and Social Care, London South Bank University, 103 Borough Road, London SE1 0AA. UK

ARTICLE INFO

Article history:
Received 15 April 2015
Received in revised form 18 June 2015
Accepted 5 July 2015

Keywords:
Delphi
Methodology
Consensus
Expert opinion
Guidelines

ABSTRACT

Introduction: Delphi technique is widely used to develop consensus on group opinion. However, no strict guidelines exist and various methods are often employed. The aim of this article was to reflect on Delphi methodology and provide guidance useful to researchers in integrative medicine.
Methods: Two parallel Delphi studies were undertaken to achieve consensus on how to treat phantom limb pain with acupuncture. Whilst completing these studies methodological issues relating to Delphi technique were identified which may be of use to other researchers.
Results: Ten areas were identified;, use of the term 'expert', sample size and sample heterogeneity/homogeneity, iteration, structure of round one, optimal number of response categories, inclusion/exclusion of data in subsequent rounds, participant feedback, defining consensus, stability of response and agreement, attrition.
Conclusions: Defining and using the term 'expert' is problematic. Three rounds are optimal. Round one data collection and analysis need structuring to avoid generation of unmanageable amounts of data. Subsequent rounds should consider using Likert Scales with four to seven categories, with even number of categories eliminating the problems associated with midpoints. To ensure rigour, data should not be excluded from round three. Participant feedback should include both central tendency and a measure of dispersion and be presented graphically. Consensus should be clearly defined and not confused with stability of response or agreement. Attrition can be minimised by ensuring participants are well informed and through a short time frame between rounds. It is intended that this guidance may help future researchers.

1. Introduction

The Delphi technique is now a widely used methodology which has the advantage over other consensus methods of not requiring face to face contact [1] whilst still guiding group opinion [2]. It can facilitate wider group participation than other consensus techniques (such as nominal group technique and consensus development conference) so avoiding recruitment bias due to participants geographical location.

Delphi technique evolved due to limitations of traditional methods used to gain group opinion for policy development [3] and was founded on the premise that unstructured, face to face group predictions were weaker than individual statistical predictions [4] The original Delphi method was developed in the 1950's by Olaf Helmer, Norman Dalkey and Nicholas Rescher of the Rand Corporation to forecast the impact of technology on warfare [4]. It has subsequently been used in healthcare, marketing, education, information systems, transportation engineering [5] and complementary and alternative medicine (CAM) to establish guidelines and establish key components of an intervention [6,7,8].

Delphi technique is defined by its use of 'experts', and its use of a series of questionnaires interspersed with controlled feedback and provides information on group opinion [5]. It is an appropriate methodology when there is lack of agreement, incomplete knowledge, uncertainty or lack of evidence [9]. This technique does not intend to challenge statistical or model based procedures but instead is intended for use in situations where statistical methods are not practical or possible [5].

Delphi technique has four main characteristics; anonymity between participants, iteration with controlled feedback of group opinion, statistical aggregation of group response and expert input [3]. Anonymity allows views to be expressed and changed privately [5]. Iteration with controlled feedback allows 'communication' between participants and perspectives to be shared [9] and allows participants to change views [10]. Statistical aggregation of group responses allows for data to be analysed and interpreted [10].

☆ This article belongs to the Special Issue on Diagnostic Techniques and Outcome Measures for Integrated Health.

* Corresponding author at: School of Health and Social Care, London South Bank University, 103 Borough Road, London SE1 0AA, UK.
E-mail addresses: trevelye@lsbu.ac.uk (E.G. Trevelyan), nicky.robinson@lsbu.ac.uk (P.N. Robinson).

http://dx.doi.org/10.1016/j.eujim.2015.07.002
1876-3820/

Expert input means each participant is informed in the area under study [11].

The aim of this article was to reflect on the Delphi technique as a research tool in healthcare. It draws upon a practical example from two studies carried out by the authors and highlights areas for consideration when using this technique [12]. The paper identifies a number of areas which need careful consideration, which could be useful for researchers when deciding whether to use this methodology and provides a list of the main issues (drawn from the above) to consider when using this technique (Table 1). The paper attempts to address areas not commonly considered in other reviews, such as what to consider when designing round one, the size of Likert scale, items to include in subsequent rounds and differentiating between consensus, agreement and reliability. The paper identifies areas needing consideration for the future use of the technique to inform integrative medicine research and provides a 'how to' guide.

2. Methods

In trying to achieve consensus from practitioners on how to treat phantom limb pain with acupuncture, two parallel quasi-anonymous studies were undertaken with a group of acupuncture practitioners both with and without previous experience of treating phantom limb pain [12]. Both studies used a Classical Delphi approach, (using an open first round to facilitate generation of ideas) administered online. One study included seven participants with past experience of treating phantom limb pain (PLP) (six acupuncturists ± physiotherapists and one physiotherapist) and the other 16 participants with no past experience of treating PLP (eight acupuncturists and eight physiotherapists). Two hypothetical case studies were provided and participants were asked twelve open-ended questions about the pathology and treatment of this condition. Subsequent rounds included statements generated from round one which participants ranked on a six point Likert scale. An a priori criterion of three rounds was set and participants were also asked to rate the protocol, developed from the Delphi study. Participants were asked to respond to each round within 7–10 days. First round data was qualitative and was analysed using qualitative content analysis. In subsequent rounds the data was quantitative and a group median of 5–6 was considered to mean agreement and an inter-quartile range (IQR) $\leq$1.75 was considered indicative of consensus. Wilcoxon matched pairs signed rank test was used to determine stability of results using a p value of $\leq$0.05.

3. Results

The data collected as part of this study and the resulting acupuncture protocol is reported elsewhere [12]. However, in summary, combining data from the two studies, nineteen participants completed all Delphi rounds (four dropped out). A

Table 1
Summary of areas to consider when designing a Delphi study.

Area of consideration	Recommendation
Expert panel	• Avoid labelling participants as 'experts' without consideration of the controversies of this label and the difficulties in defining 'experts'. • Be explicit about participant's expertise. • Recognise the limitations of Delphi studies (group consensus is not synonymous with 'best' or 'correct' results). • Consider the homogeneity or heterogeneity of the sample when deciding sample size.
Iteration	• Consider setting number of rounds *a priori.* • Three rounds of Delphi are optimal. • If developing a protocol, consider asking participants to rate the protocol post completion of the Delphi study.
Round 1	• Use few well-structured open-ended questions or a modified approach to develop initial statements. • Beware of generating large amounts of data for subsequent round.
Round 2	• Optimal number of response categories of the Likert scale lie between four to seven. • Consider the potential pitfalls of using a Likert scale with a midpoint. • Consider providing a 'no comment option' if participants have different backgrounds and knowledge.
Subsequent rounds	• Ideally recirculate all data from round two. • Be aware that omitting data can introduce bias and prevent full analysis of results.
Participant feedback	• Use both central tendency and a measure of dispersion to aggregate data (median and IQR). • Consider providing visual feedback (bar charts) to provide information on the distribution of data.
Consensus	• Define if consensus is being used to determine if agreement exists or as a stopping guideline. • Differentiate between stability, agreement and consensus. • A measurement of variance in response is appropriate for determining consensus.
Stability of response / internal validity	• Stability of response should not be confused with consensus. • If analysed, theWilcoxon matched-pairs signed rank test or ICC are appropriate methods to inferentially determine stability of response. Stability could also be determined though providing data on the median and IQR across rounds or through graphical representation.
Attrition	• Consider only recruiting participants with an interest in the topic. • Ensure participants are fully informed of the commitments of a Delphi study. • Ensure a short time frame between rounds to maintain participant interest and reduce attrition. • Ensure feedback is part of the agreement to take part.

Key: IQR, inter-quartile range; ICC, intraclass correlation coefficient.