

Analysis of the use of the Delphi technique for decision-making in critically ill patients: a systematic review

A análise do uso da técnica Delphi na tomada de decisão em pacientes críticos: uma revisão sistemática

Ana Paula Freitas Bahia dos Santos¹, Jessika de Freitas Andrade¹, Geisa Cristina da Silva Alves², Samuel Dutra da Silva³, Cristina Sanches⁴, Farah Maria Drumond Chequer⁴

Santos APFB, Andrade JF, Alves GCS, Silva SD, Sanches C, Chequer FMD. Analysis of the use of the Delphi technique for decision-making in critically ill patients: a systematic review / A análise do uso da técnica Delphi na tomada de decisão em pacientes críticos: uma revisão sistemática. Rev Med (São Paulo). 2020 May-June;99(3):291-304.

ABSTRACT: *Introduction:* The Delphi technique is a method used to generate consensus among experts on a given subject. *Objectives:* To develop a systematic review on the contribution and use of the Delphi method exclusively by physicians for decision-making in critically ill patients. *Methods:* The study was conducted in the databases: PubMed, SciELO, *Biblioteca Virtual em Saúde*, Scopus, CAPES Journals and ClinicalTrials.gov, using the descriptor: “*Delphi technique*”, “*Decision making*”, “*Critical care*” and “*Physicians*”. Clinical trials or observational studies were selected to answer the guiding question: “What is the contribution of the Delphi technique for the decision-making process of physicians caring for critically ill patients?” The scale of the Oxford Centre for Evidence-Based Medicine was used to evaluate the level of evidence and grade of recommendation of the studies included. *Results:* Eighteen articles were included in this study. The contributions of the Delphi method were considered important for decision-making in fifteen of the studies analyzed. The method contributed to define management strategies, treatment and recommendations; manage and diagnose patients; generate consensus for examinations; and identify topics that cause differences in opinion and their underlying reasons. Among the limitations of the method, the most frequent was the selection of specialists with different levels of knowledge. *Conclusion:* This study presents the applicability of the Delphi technique for the management of critically ill patients, drug treatment, recommendations for specific symptoms, guidelines for the management and diagnosis of patients, consensus for tests and examinations and identification of topics that cause differences in opinion and their underlying reasons in situations where there is no consensus in the medical literature.

Keywords: Delphi technique; Decision making; Critical care; Physicians; Intensive Care Units.

RESUMO: *Introdução:* A técnica Delphi é um método utilizado para alcançar consenso entre especialistas sobre determinado assunto. *Objetivos:* realizar revisão sistemática sobre a contribuição e utilização do método Delphi exclusivamente por médicos, para tomada de decisões em pacientes críticos. *Métodos:* A pesquisa foi realizada nas bases de dados: PubMed, SciELO, Biblioteca Virtual em Saúde, Scopus, Periódicos CAPES e ClinicalTrials.gov, utilizando os descritores: “*Delphi technique*”, “*Decision making*”, “*Critical care*” e “*Physicians*”. Selecionouse ensaios clínicos ou estudos observacionais para responder à pergunta norteadora: “Qual a contribuição da técnica Delphi na tomada de decisão pelos médicos em pacientes críticos?” Para avaliação do nível de evidência e grau de recomendação dos estudos incluídos, foi utilizado a escala do Centro de Medicina Baseada em Evidências de Oxford. *Resultados:* Foram incluídas dezoito referências para compor este estudo. As contribuições do método Delphi foram consideradas importantes para a tomada de decisões em quinze estudos analisados, para: definir estratégias de manejo, tratamento e recomendações; gerenciar e diagnosticar pacientes; definir consenso para exames; identificar ideias controversas e seus motivos subjacentes. Das limitações do método, a mais recorrente foi a seleção de especialistas de diferentes níveis de conhecimento. *Conclusão:* Este estudo mostra a aplicabilidade da técnica Delphi para manejo de pacientes críticos, tratamento medicamentoso, recomendações para sintomas específicos, diretrizes para gerenciar e diagnosticar pacientes, consenso para testes e exames e identificação de tópicos de ideias controversas e seus motivos subjacentes em situações onde não há consenso na literatura médica.

Descritores: Técnica Delfos; Tomada de decisões; Cuidados críticos; Unidades de Terapia Intensiva.

Presentation: IX Academic Day of Medicine at the University of Itaúna, Itaúna, MG, November 9, 2017.

1. Universidade de Itaúna, Itaúna, MG. Acadêmica de medicina. ORCID: Santos APFB - <https://orcid.org/0000-0001-6720-5367>; Andrade JF - <https://orcid.org/0000-0002-2868-0883>. Email: anapaulafreitasb@gmail.com, jessikafreitas08@gmail.com
2. Universidade Federal de São João del-Rei, Campus Centro-Oeste Dona Lindu (UFSJ-CCO), Divinópolis, MG. Enfermeira - Doutoranda em Ciências da Saúde. ORCID: <https://orcid.org/0000-0002-2023-8011>. Email: geisa.cristina@gmail.com
3. Universidade de Itaúna, Itaúna, MG. Médico, preceptor de medicina e mestre pela Universidade Federal de São João del-Rei, Campus Centro-Oeste Dona Lindu (UFSJ-CCO), Divinópolis, MG. ORCID: <https://orcid.org/0000-0002-4287-3542>. Email: samueldutradasilva@gmail.com
4. Universidade Federal de São João del-Rei, Campus Centro-Oeste Dona Lindu (UFSJ-CCO), Divinópolis, MG. – Farmacêutica - Professora Adjunto. ORCID: Sanches C - <https://orcid.org/0000-0002-8562-1337>; Chequer FMD - <https://orcid.org/0000-0003-3514-2132>. Email: csanches@ufsj.edu.br, farahchequer@ufsj.edu.br.

Correspondence: Farah Maria Drumond Chequer. Universidade Federal de São João del-Rei, Campus Centro-Oeste Dona Lindu (UFSJ-CCO). Rua Sebastião Gonçalves Coelho, 400. Chanadour Divinópolis, MG, Brasil. CEP: 35.501-296. E-mail: farahchequer@ufsj.edu.br.

INTRODUCTION

The Delphi technique was developed by Dalkey and Helmer at the Rand Corporation in the 1950s and was initially used to obtain experts' opinions on military defense projects¹. It is a widely used and accepted method to acquire consensus of a group of experts on a given subject². It can be used for the development of diagnostic criteria in the absence of a gold standard and it is a recognized, accepted and used method for gathering data from specialists within their domain of expertise^{3,4}.

The Delphi technique is conducted with participants with knowledge and experience in a given area. Essentially, there is no specific number of participants for the technique. The method starts with the distribution of a questionnaire on the subject for each specialist. After gathering their opinions, the results are summarized. The technique uses a series of questionnaires and the provision of feedback on the participant's opinions, with the objective of gaining consensus on a given subject. The percentage of approval that determines that an answer to a given question is an agreement must be previously established. If an item is not accepted by a minimum number of professionals, it must be altered according to the participants' suggestions until consensus is reached. After all the rounds required to reach consensus on the suggested items, they can be classified as approved or not¹.

Among the advantages of using the Delphi technique, it is possible to mention: it allows the inclusion of several specialists from different areas and geographic regions, without the need to meet in person; participants are anonymous, which eliminates the influence of prominent people in the group^{4,5-6}; it is effective for promoting reflection and developing guidelines⁶; saves time and money by not requiring personal meetings⁷.

The Delphi technique is a fundamental tool when it comes to topics that still generate controversies in the literature^{3,4} and it can be used to determine the appropriate clinical care⁸. This justifies studies that, like this one, clearly and directly address the effectiveness of this technique in the medical field. Despite being a field where the method has great potential, it is important to evaluate its effectiveness and feasibility for reaching consensus in the medical context.

Therefore, this article aimed to develop a systematic review on the contribution of the use of the Delphi method for the decision-making process of physicians caring for critically ill patients admitted to the Intensive Care Unit (ICU)/Critical Care Unit (CCU). This setting (care of critical patients) was chosen because it requires urgency in the decision-making of physicians and because there may be doubts about which interventions and/or diagnosis to perform.

METHODS

A systematic review of published articles with information on the Delphi technique was developed with the objective of analyzing the applicability of the technique.

Study design: The PRISMA Statement was used to elaborate the systematic review⁹.

Guiding question of the research: "What is the contribution of the Delphi technique for the decision-making process of physicians caring for critically ill patients?"

Definition of Population, Intervention, Comparison, Outcomes and Studies (PICOS): the "PICOS" was established according to the guiding question¹⁰: "P" (population): Use of the Delphi technique by physicians; "I" (intervention): use of the Delphi technique; "C" (comparison): there is no comparison; "O" (outcomes): effectiveness of the use of the Delphi technique in critical care; "S" (types of studies): observational and clinical studies.

Sources of information: The search for studies to be included in this review was carried out in the databases: PubMed, SciELO, *Biblioteca Virtual em Saúde* (BVS), CAPES Journals and Scopus. Clinical trial records of studies in progress or recently completed were also searched in the ClinicalTrials.gov. The Medical Subject Headings (MeSH) was used to define the following descriptors: "Delphi technique", "Decision making", "Critical care" and "Physicians", which were combined with the Boolean operator "AND" in the databases Pubmed, SciELO, *Biblioteca Virtual em Saúde* (BVS), CAPES Journals, Scopus and ClinicalTrials.gov. The search in Pubmed included the MeSH descriptors and their synonyms: "*Delphi Technique*"; *Delphi Techniques*; *Technique, Delphi*; *Techniques, Delphi*; *Delphi Technic*; *Delphi Technics*; *Technic, Delphi*; *Technics, Delphi*; *Delphi Studies*; *Delphi Study*; *Studies, Delphi*; *Study, Delphi*; "*Critical care*"; *Care, Critical*; *Intensive Care*; *Care, Intensive*; *Surgical Intensive Care*; *Care, Surgical Intensive*; *Intensive Care, Surgical*, "*Decision making*"; *Decision Making, Shared*; *Decision Makings, Shared*; *Making, Shared Decision*; *Makings, Shared Decision*; *Shared Decision Making*; *Shared Decision Makings*, "*physicians*".

Studies obtained from other sources: The references of the articles selected were also analyzed after applying the inclusion and exclusion criteria.

Selection of studies: After selecting the articles by applying the search strategy in each database, the articles were compared to remove duplicates.

Eligibility criteria: The studies considered eligible for this review were clinical trials or observational studies conducted in humans, published until June 2017, in English, Spanish and Portuguese, describing the applicability of the Delphi technique and its use to assist the decision-

making process of physicians caring for critically ill patients. Review articles, editorials, letters to the editor, news, comments and results of dissertations, theses and abstracts published in conference proceedings or scientific journals were excluded. The use of the Delphi technique by other health professionals was considered as an exclusion criterion. The criteria for the inclusion and exclusion of the studies found were applied independently and blindly by two authors. In case of disagreement between the reviewers, other team members were convened to develop consensus.

Assessment of the level of evidence and additional analyses: The scale of the Oxford Centre for Evidence-Based Medicine was used to assess the level of evidence and grade of recommendation of the studies included¹¹. This

scale evaluates each study according to its methodology, as shown in Table 1.

The Kappa Coefficient¹² was used to assess the degree of agreement between the two evaluators (APFBS and JFA). The confidence interval was set at 95% and the test was performed in the program Stata, version 10.0.

Data collection and analysis: The articles that met the inclusion criteria were read in full and, during this phase, the following variables were collected: author/year/location; objective of the study; sample size (number of physicians participating); duration of the Delphi method; description of the contribution of the use of the Delphi for decision-making in critically ill patients; effectiveness; limitations; applicability of the technique.

Table 1. Level of Scientific Evidence by Type of Study - “Oxford Centre for Evidence-based Medicine”

| Grade of recommendation | Level of evidence | Treatment – Prevention – Etiology |
|--|-------------------|--|
| B Studies that recommend an action; significant evidence is found in the outcome, and the conclusion is that it is beneficial to choose the action even considering the damage risks. There is reasonable evidence to support the recommendation | 2B | Decision analysis |
| C It consists of level 4 studies or extrapolations from level 2 or 3 studies. There is no satisfactory evidence in the analysis of outcomes, and the conclusion is that the benefits and risks of the procedure do not justify the generalization of the recommendation. There is insufficient evidence, either for or against it | 2C | Observation of therapeutic results (<i>outcomes research</i>). |

Adapted from: <http://www.cebm.net/index.aspx?o=1025>. Fonte: Phillips et al.¹¹

RESULTS

A total of 33 studies were selected after reading the abstracts. Of these, 15 were excluded for not meeting the

eligibility criteria, as described in Figure 1.

Only 18 studies met the inclusion criteria and were eligible for this review, as shown in Figure 2.

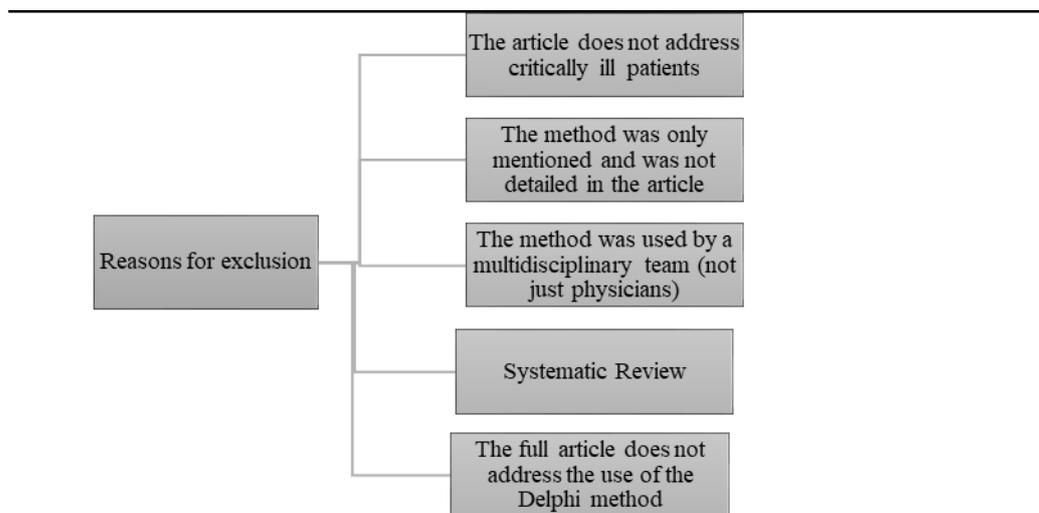
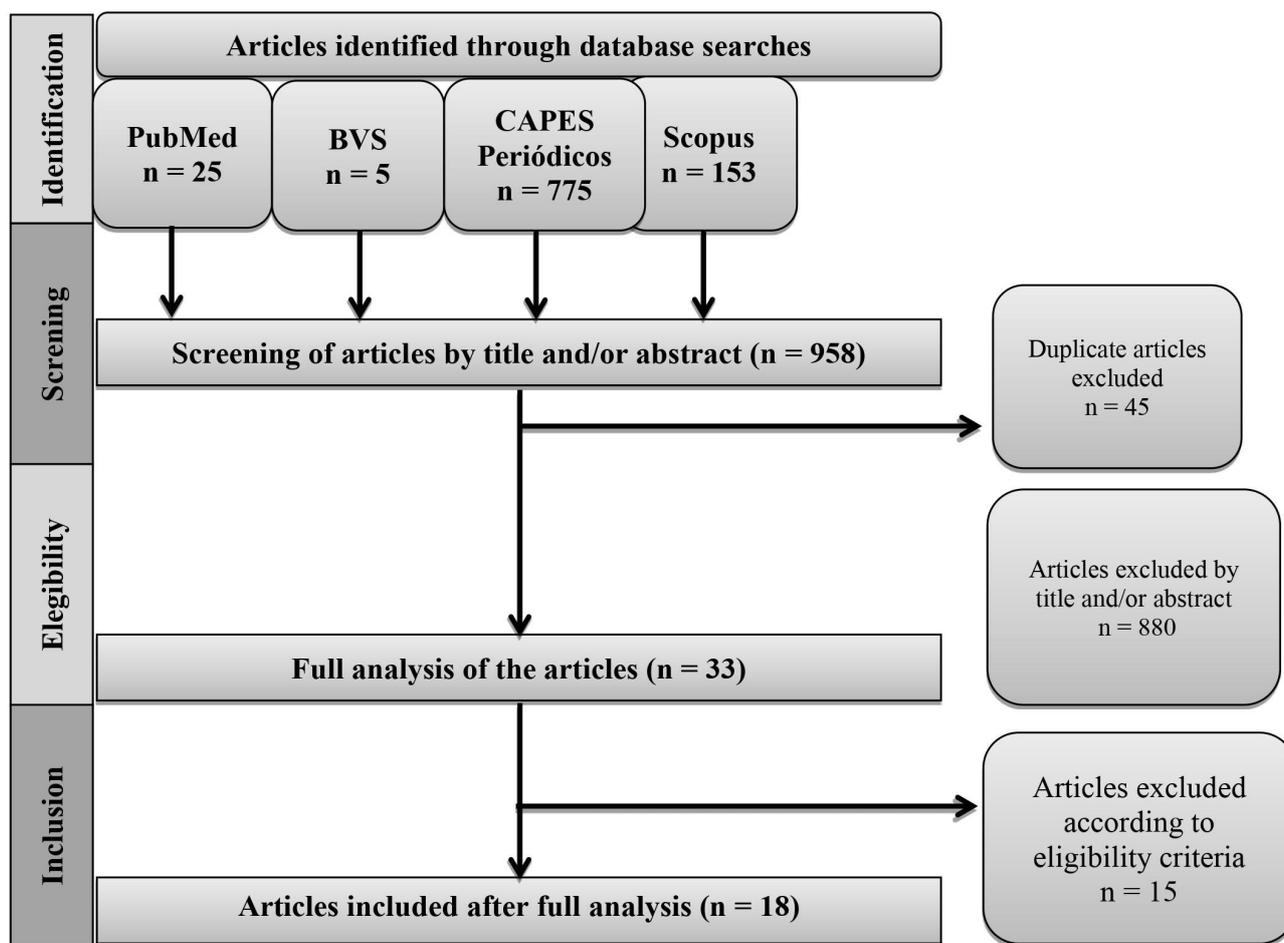


Figure 1: Reasons for the exclusion of studies that did not meet the inclusion criteria (n = 15)



Source: Flowchart based on the PRISMA guideline

Figure 2: Flowchart of the phases of the systematic review.

The degree of agreement between the two authors was considered substantial according to the kappa coefficient = 0.7150¹². All the 18 articles selected according to the eligibility criteria were observational studies, carried out from 2001 to 2016 in different countries such as the USA¹³⁻¹⁹, the United Kingdom,^{20,21} Canada^{3,22,23}, South Korea², Portugal⁵, Australia²⁴, France⁷ and Spain^{25,26}.

Regarding the level of evidence and grade of recommendation of the studies selected, fourteen (78%) were classified as level of evidence 2c and grade of recommendation IV, meaning that they provided little evidence to analyze the outcomes and concluded that the benefits and risks of the procedure do not justify the generalization of the recommendation. Another 4 studies were classified as level of evidence 2b and grade of recommendation II, with good reference standards for

decisions.

Regarding the objectives of the Delphi technique, in two studies, the method was used to develop checklists of verification²⁴ or bypass separation¹⁸. In two of the articles, the method was used to propose guidelines for the management of patients^{13,15}. Another article aimed to establish guidelines for the use of bedside ultrasound¹⁹. In three studies, the method was used to develop criteria for admission in the ICU⁷ and for cases of chronic obstructive pulmonary disease (COPD) exacerbations^{25,26}. Three articles applied the method to develop consensus, whether on appropriate decisions regarding treatment and diagnosis^{14,16,20}, on the need for consultation with specialists and evaluation of important tests and exams in the emergency room², or on the prescription of analgesia and sedation²⁰. In one study, the method was used to

evaluate the best technique for mechanical ventilation²¹. Another article used the method to develop criteria to identify the thresholds in monitoring equipment²³. Three articles presented more specific applications, such as the analysis of topics that cause differences in opinion between anesthesiologists and obstetricians³, the use of palliative care at the end of life⁵ and important practices for patients with respiratory failure¹⁷.

Regarding the number of physicians participating in the Delphi method, in nine of the studies analyzed all members participated in all stages^{2,3,7,15-17,19,25,26}. In eight other articles, there was a loss of participants during the rounds^{5,13,14,18,20,21,23,25}, which accounted for approximately 26.16% of specialists. One article had three rounds, each round with two questionnaires, and there were variations in the number of participants within the same round; the mean number of participants in this study was 10.66. The number of professionals in the beginning of each study ranged from 10 to 79, with a mean of 30.17 physicians invited. One of the articles did not mention the number of physicians participating¹⁹.

Regarding the duration of the Delphi method, 15 of the 18 articles selected did not report the duration of the procedure.^{2,7,13,14,16-26} In one article, the method was applied from January 2006 to 2007¹⁵. In the other two studies, it was applied from February to April 2012⁵, and from October 2012 to January 2013³.

Regarding the contributions of the Delphi method to the decision-making process of medical professionals, in 14 studies, the use of the technique was important to define management strategies, drug treatment, recommendations for specific symptoms and use of ultrasound, guidelines for the management and diagnosis of patients; to develop consensus on tests and examinations; to identify topics that cause differences in opinion and their underlying reasons; and to establish the determinants of outcomes of COPD^{2,3,5,13-20,22-26}. In one of the studies, the author emphasized the importance of the Delphi technique to foster a broad movement of reflection and to challenge professionals⁵. Furthermore, in three articles, the method did not offer a significant contribution^{7,21,24}. In one of these studies, the checklist elaborated was not associated with a significant difference when used for diagnosis²⁴. In another study, the author argued that the method is not effective because the thought process and the ideas behind an opinion are not described²¹. One of these articles reported that the criteria developed were not relevant for application⁷.

The effectiveness of the Delphi technique was evaluated by the authors Elliot²⁴, Baumann¹³, Lee², Mendes⁵,

Ogden¹⁸ and Hawryluck²² through a LIKERT scale, and by the authors Dellinger¹⁵ and Frankel 2015¹⁹ through the GRADE system. The authors Garroust⁷, Quintana²⁶ and Garcia Gutierrez 2011²⁵ used decision trees with scores from 1 to 9, where different intervals demonstrated different levels of agreement. Among the 18 articles, five established that consensus was reached when the idea was accepted by 80% or more of the participants^{5,14,19,22,25}. In three articles, consensus was determined by at least 70% agreement^{2,16}. In two articles, consensus was reached when values equal to or greater than 75% were obtained^{17,23}. In three other articles, values equal to or greater than 60% and 66% were considered a consensus^{3,4,13,18,21}. In three articles^{15,17,20,24}, these values were not mentioned; only a "strong agreement" was mentioned in one of them⁷. Six out of the 18 studies highlighted that the strategies approved by consensus were selected in more than one round.^{5,16,17,18,20,23} The other 12 studies^{2,3,7,13,14,15,19,21,22,24,25,26} did not report that.

Two out of the 18 articles did not report the limitations related to the use of the Delphi technique.^{5,15} Seven studies^{2,3,7,13,14,18,20}, mentioned the selection of specialists as an important limitation. The reasons included possible differences in knowledge³, the volunteer character¹⁴, the presence of only emergency room physicians⁷, or the selection of professionals from a single location^{2,3,7,20}. Four studies mentioned the lack of scientific rigor of the method as a limitation^{2,3,13,20}. One article stated that the recommendations, despite being from experts, were not sufficient evidence^{13,23} or could not be applied to all patients¹⁶. One article mentioned the inability to exchange ideas with other participants as one of limitations of the technique³.

The other limitations were mentioned only once in each article; they were: the cases identified differed from the medical diagnosis²⁴; the Delphi technique did not represent sufficiently strong evidence to form the basis for health-care policy¹³; the method did not allow comparing referral rates and data that could influence decisions were not collected;⁷ there may have been errors in transcription and perception of responses²¹; the arbitrary definition of consensus¹⁶; specialists¹⁸ and patients²⁶ were lost to follow-up; lack of data for calculation²⁶; failure to address issues of politics and costs¹⁹; effectiveness depends on appropriate implementation of the protocol¹⁷; feedback provided to participants may have influenced decisions²³; guidelines based only on the perspective of physicians²² and the time restriction of each round of the Delphi method³.

The results of the analysis of the articles selected for the systematic review are described in Table 2.

Table 2. Characteristics of the studies included in the systematic review

| Objectives | | Methods | | Results | | |
|--|--|--|--|---|--|--|
| Author/year/ country | Objective of the study | Sample size (physicians) | Use of the Delphi method: (time) | Description of the contribution of using the Delphi method for decision making in critically ill patients | Agreement and consensus of the method | Limitations |
| Tan et al. ²¹ 2009 United Kingdom | To achieve consensus on the use of neonatal mechanical ventilation using the Delphi technique | 15 physicians 1° round: 15 physicians 2° round: 10 physicians 3° round: 07 physicians | Unreported | In complex areas of medicine, such as management of neonatal ventilation, techniques such as focus groups, in which there is a possibility of describing the thought process behind a decision, may be better for reaching consensus than totally anonymous techniques such as the Delphi method | Consensus was defined as 66% of participants using a particular strategy. There were some differences between responses in the second round, but consensus was reached in 20 out of 40 scenarios | Wide variation in the strategies used by neonatologists. The results can be affected by the artificiality of the exercise. There may have been errors in transcription and perception of responses. The group dynamics can make participants of focus group more prone to agreeing |
| DeLinger et al. ¹⁵ 2008 USA | To provide an update to the clinical guidelines for management of sepsis and septic shock | 55 physicians | January 2006 to 2007 | Evidence-based recommendations regarding the acute management of sepsis and septic shock are the first step to improve outcomes for this important group of critically ill patients | GRADE methodology was used to guide the assessment of quality of evidence from high (A) to very low (D) and to determine the strength of the recommendations. There was a strong agreement among a cohort of international experts | Unreported |
| Li et al. ¹⁶ 2013 USA | To achieve national consensus on warning indicators and recommended action plans for patients after colorectal surgery | 11 physicians | Unreported | There were experts' recommendations for 12 symptoms that were considered warning signs for patients with deteriorating conditions after discharge for colorectal surgery | In the second round consensus of 8/11 specialists (>70%) was reached. At least 70% of specialists | It was necessary to create an arbitrary definition of consensus, and the indicators may not apply to all patients |
| Lee et al. ² 2016 South Korea | To suggest experts' opinions to clinicians in the emergency room, which clinical, laboratory or radiological tests should be considered whether arthrocentesis or specialized consultation is necessary or not | 18 physicians | Unreported | This guideline can help clinicians determine the best diagnostic strategy for patients who need a differential diagnosis between septic arthritis and transient synovitis. The guideline proposes a structure to ensure that the necessary steps to obtain a relevant medical history and perform an effective physical examination are taken. It can be used in conjunction with other diagnostic methods that are considered appropriate in any given institution | Consensus is reached when there is 70% approval. Ten items with more than 70% approval were chosen. In one round there was consensus on 1/4 items related to general assessment and 9/14 items on the need for laboratory tests. Finally, consensus was reached on three of 11 items on general assessment and need for laboratory tests | Selection of participants for the Delphi method. The experts were all from school or university hospitals in South Korea. This could have affected the selection process. The Delphi method has limited validity in terms of scientific evidence. It is not a substitute for basic and clinical research, which must be carried out to confirm any conclusions |

Continúa

Table 2. Characteristics of the studies included in the systematic review

| Objectives | | Methods | | Results | | |
|--|---|--|--|---|--|--|
| Author/year/ country | Objective of the study | Sample size (physicians) | Use of the Delphi method: (time) | Description of the contribution of using the Delphi method for decision making in critically ill patients | Agreement and consensus of the method | Limitations |
| Tan et al. ²¹ 2009 United Kingdom | To achieve consensus on the use of neonatal mechanical ventilation using the Delphi technique | 15 physicians 1° round: 15 physicians 2° round: 10 physicians 3° round: 07 physicians | Unreported | In complex areas of medicine, such as management of neonatal ventilation, techniques such as focus groups, in which there is a possibility of describing the thought process behind a decision, may be better for reaching consensus than totally anonymous techniques such as the Delphi method | Consensus was defined as 66% of participants using a particular strategy. There were some differences between responses in the second round, but consensus was reached in 20 out of 40 scenarios | Wide variation in the strategies used by neonatologists. The results can be affected by the artificiality of the exercise. There may have been errors in transcription and perception of responses. The group dynamics can make participants of focus group more prone to agreeing |
| DeLinger et al. ¹⁵ 2008 USA | To provide an update to the clinical guidelines for management of sepsis and septic shock | 55 physicians | January 2006 to 2007 | Evidence-based recommendations regarding the acute management of sepsis and septic shock are the first step to improve outcomes for this important group of critically ill patients | GRADE methodology was used to guide the assessment of quality of evidence from high (A) to very low (D) and to determine the strength of the recommendations. There was a strong agreement among a cohort of international experts | Unreported |
| Li et al. ¹⁶ 2013 USA | To achieve national consensus on warning indicators and recommended action plans for patients after colorectal surgery | 11 physicians | Unreported | There were experts' recommendations for 12 symptoms that were considered warning signs for patients with deteriorating conditions after discharge for colorectal surgery | In the second round consensus of 8/11 specialists (>70%) was reached. At least 70% of specialists | It was necessary to create an arbitrary definition of consensus, and the indicators may not apply to all patients |
| Lee et al. ² 2016 South Korea | To suggest experts' opinions to clinicians in the emergency room, which clinical, laboratory or radiological tests should be considered, whether arthrocentesis or specialized consultation is necessary or not | 18 physicians | Unreported | This guideline can help clinicians determine the best diagnostic strategy for patients who need a differential diagnosis between septic arthritis and transient synovitis. The guideline proposes a structure to ensure that the necessary steps to obtain a relevant medical history and perform an effective physical examination are taken. It can be used in conjunction with other diagnostic methods that are considered appropriate in any given institution | Consensus is reached when there is 70% approval. Ten items with more than 70% approval were chosen. In one round there was consensus on 1/4 items related to general assessment and 9/14 items on the need for laboratory tests. Finally, consensus was reached on three of 11 items on general assessment and need for laboratory tests | Selection of participants for the Delphi method. The experts were all from school or university hospitals in South Korea. This could have affected the selection process. The Delphi method has limited validity in terms of scientific evidence. It is not a substitute for basic and clinical research, which must be carried out to confirm any conclusions |

Continua

Table 2. Characteristics of the studies included in the systematic review

| Objectives | | Methods | | Results | | |
|--|---|--|--|---|---|---|
| Author/year/ country | Objective of the study | Sample size (physicians) | Use of the Delphi method: (time) | Description of the contribution of using the Delphi method for decision making in critically ill patients | Agreement and consensus of the method | Limitations |
| Mendes et al. ⁵ 2013 Portugal | To identify the main areas for the protocol, to build consensus and to develop a protocol for neonatal palliative care that could be used in Portuguese neonatal ICUs | 57 physicians 1° round: 45 physicians 2° round: 45 physicians 3° round: 44 physicians | February to April 2012 | The Delphi method fostered a broad movement, challenging neonatologists to address neonatal palliative care and end-of-life problems | For each statement, the participants had to position on a Likert scale (1 strongly disagree and 10 strongly agree). It was decided that the proposed protocol would be accepted if 80% of the answers were equal to or greater than 8 on the Likert scale. In the first round, 34 experts added 466 comments for content analysis, and 27 new ideas were added in the second round. After the three rounds, 92% statements reached the defined consensus criteria | Unreported |
| Garroutte-Orgeas et al. ⁷ 2009 France | To describe ICU admission criteria for patients aged 80 and over. To assess ICU referral and admission rates of patients with at least one of these criteria. To describe the factors associated with emergency physicians' decisions to refer patients | 30 emergency physicians All 30 physicians participated in the 3 rounds | Unreported | The guidelines of the Society of Critical Care Medicine for patients over 80 years old, adapted using the Delphi technique, were not relevant and were not followed by emergency physicians in France | The specialists classified each criterion on a scale of 1 (inappropriate admission to the ICU) to 9 (extremely appropriate admission to the ICU), where 5 indicated that the adequacy of admission to the ICU was not clear. When all answers fell within a single interval (7-9, 4-6 or 1-3), agreement was strong, when answers spanned two intervals, agreement was fair, and when answers spanned all three intervals, agreement was poor | The list was created by emergency room physicians, and it could be different if it were created by physicians working in the ICU. It was not possible to compare the influence on the referral rates of patients over 80 years versus younger patients. The list may have changed the emergency physician's standard practice. Data on ethnicity, socioeconomic level and educational level were not collected and it was not possible to assess whether these factors influenced referral decisions. All study centers were in the metropolitan area of Paris, and the findings may not reflect practices in other countries |

Continua

Table 2. Characteristics of the studies included in the systematic review

| Objectives | | Methods | | Results | | |
|---|---|---|--|--|---|--|
| Author/year/ country | Objective of the study | Sample size (physicians) | Use of the Delphi method: (time) | Description of the contribution of using the Delphi method for decision making in critically ill patients | Agreement and consensus of the method | Limitations |
| García-Gutiérrez et al. ²⁵ 2011 Spain | To develop explicit criteria for admission in COPD exacerbation and a preliminary validation | First group: 12 physicians Second group: 10 physicians | Unreported | This algorithm could be used in the ED to help physicians in their medical decision-making process, in field studies and for the development of practice guidelines | Physicians rated each scenario using a 9-point scale (1 point for highly inappropriate admission and 9 for highly appropriate admission). A scenario was considered appropriate if the score was between 7 and 9 without disagreement, and inappropriate if the score was between 1 and 3 without disagreement. Agreement was defined as less than one third of the participants rating an indication outside a 3- point region (1 to 3, 4 to 6, or 7 to 9). 54.7% of the scenarios were classified as appropriate | The guidelines were based on opinions |
| Quintana et al. ²⁶ 2011 Spain | To create criteria for hospital admission for patients with COPD exacerbations, to validate these criteria in clinical situations; to develop a severity scale for patients with COPD exacerbations | 12 physicians | Unreported | The study delineated the requirements for admission or discharge of patients with COPD exacerbations and provided a better understanding of the determinants of outcomes of COPD exacerbations | The study used a 9-point scale, ranging from 1 point for inappropriate admissions to 9 points for appropriate admissions. A scenario was defined as appropriate if the median score was between 7 and 9 without disagreement, and inappropriate if the median score was between 1 and 3 without disagreement. Agreement was defined as less than one third of the participants rating an indication outside a 3- point region (1 to 3, 4 to 6, or 7 to 9) | The final list was adjusted to 10 items due to redundancy of topics, 19 specialists were lost to follow-up, the best order of items in the final list was not tested |
| Hawryluck et al. ²⁷ 2002 Canada | To develop consensus to guide physicians on analgesia and sedation and help distinguish palliative care from euthanasia | In the first round: 19 intensivists. In the second: 5 coroners. In the third: 12 intensive care residents | Unreported | The guidelines formulated can guide the physician when prescribing analgesia and sedation for dying patients, helping to avoid high or low doses in ICU patients | A Likert scale was used, and consensus was determined as 80% agreement. Consensus was achieved on 16 statements. All guidelines achieved high degrees of consensus | A small group of physicians participated; the guidelines are only for sedation and analgesia, so other types of care should be included for integrated attention to the dying patient in the ICU; the guidelines were based on the perspective of physicians and not patients |

Continúa

Table 2. Characteristics of the studies included in the systematic review

| Objectives | | Methods | | Results | | |
|---|---|---|----------------------------------|---|---|---|
| Author/ year/ country | Objective of the study | Sample size (physicians) | Use of the Delphi method: (time) | Description of the contribution of using the Delphi method for decision making in critically ill patients | Agreement and consensus of the method | Limitations |
| Pearl et al. ¹⁷ 2016 USA | To develop consensus on what care practices are important to implement for patients that have developed or are progressing toward respiratory failure | 13 physicians | Unreported | These expert recommendations can be used anywhere in the hospital for treatment of patients with respiratory failure | Consensus was determined by 75% agreement between physicians. In the first stage of the process, 25 items were proposed as initial recommendations. Upon further discussion, eight new items were added. After the rounds, nine items were excluded by not receiving over 75% consensus | Its usefulness and effectiveness will depend largely on appropriate implementation of the protocol and correct identification of patients who are at high risk for respiratory failure |
| Ogden et al. ¹⁸ 2016 USA | To develop a list of key items to prepare for the separation of cardiopulmonary bypass | 90 anesthesiologists completed the first round, 77 completed the second, 73 completed the third and 71 completed the fourth round | Unreported | The checklist developed was considered by the authors as essential for a safe bypass separation during cardiac surgery. Due to the large number of participants from many different institutions and locations, this checklist was considered appropriate for most practice sites | The modified Delphi method was used. A Likert scale was used. Consensus was determined by a mean ≥ 4 on the 5-point Likert scale. Of the 49 necessary items proposed for bypass separation, 21 were selected at the end of the rounds. Since there were similarities between the items, they were grouped into a final list of 10 topics | The final list was adjusted to 10 items due to redundancy of topics; 19 specialists were lost to follow-up; the best order of items in the final list was not tested |
| Nguyen-Lu et al. ³ 2015 Canada | To determine the main controversial medical problems in the ward that cause differences in opinion between anesthesiologists and obstetricians | 20 physicians 1st round: 20 2nd round: 20 3rd round: 20 | October 2012 to January 2013 | The controversial topics and underlying reasons identified in this study may serve as a basis to develop education and training in areas where concordance is lacking between anesthesiologists and obstetricians and could have a significant impact on patient care and outcome | Agreement by 60% of the participants in each discipline was interpreted as the group's agreement with a certain topic or reason. Eight controversial topics were identified | Single centre. The one-month time restriction on each round could have meant that complete data saturation was not achieved. Some participants may have more in-depth knowledge of certain topics than others within their specialty. This may lead to identifying a series of general statements rather than accomplishing an in-depth exploration of each topic |

Continúa

Table 2. Characteristics of the studies included in the systematic review

| Author/year/ country | Objectives | | Methods | | Description of the contribution of using the Delphi method for decision making in critically ill patients | Results | |
|---|---|--|--|--|--|--|--|
| | Objective of the study | Sample size (physicians) | Use of the Delphi method: (time) | Agreement and consensus of the method | | Limitations | |
| Ansermino et al. ²³ 2008 Canada | To develop criteria to identify the thresholds of monitoring systems for three ventilatory events: disconnection or significant leak in the anesthesia circuit (1), decreased lung compliance or increased resistance (2), and anesthetic overdose from inhaled anesthetics (3) | Questionnaire (1): 2 anesthetists completed the 1st round, 16 completed both. Questionnaire (2): 4 completed the 1st round, 18 completed both. Questionnaire (3): 2 completed the 1st round, 22 completed both | Unreported | This study provided consensus-based criteria for the identification of three critical ventilatory events and has presented insight into the decision insights used by clinicians. In addition, the authors stated that the rules identified will have wide applicability | The modified Delphi method was used. Consensus was determined by 75% agreement. In the questionnaire on disconnection or significant leak in the anesthesia circuit, consensus was reached on the threshold values of the 12 items after two rounds. In the questionnaire on decreased lung compliance or increased resistance, consensus was reached on the threshold values of the 8 items after 2 rounds. In the questionnaire on anesthetic overdose, consensus was reached on the threshold values of the 30 items after 3 rounds | The nature of the feedback provided to participants could influence the likelihood of reaching consensus; the provision of additional feedback in the second round influenced the participants' subsequent responses; only physicians from Canada participated; the sample was small | |
| Frankel et al. ¹⁹ 2015 USA | To establish evidence-based guidelines for the use of bedside ultrasound by intensivists in the ICU and equivalent care sites for diagnostic and therapeutic purposes for organs of the chest, abdomen, pelvis, neck, and extremities | Clinical physicians, surgeons and intensive care anesthetists and one emergency physician | Unreported | Recommendations and the level of evidence for the use of bedside ultrasound for the diagnosis of diseases and procedures were established. It is expected that these guidelines will arm intensivists with the evidence to confront the political and economic challenges of implementing a critical care ultrasound program | The modified Delphi technique was used. Strong recommendations required 80% majority. Twenty-four statements regarding the use of ultrasound were considered; three did not achieve agreement and nine were approved as conditional recommendations. The remaining 12 statements were approved as strong recommendations | This work did not address issues of politics and the cost of acquiring sufficient equipment and training to implement these guidelines that some providers may encounter. | |
| Goodrich et al. ¹⁴ 2013 USA | To propose a set of clinical guidelines for use by optometrists and other eye care providers when screening or examining patients with mild traumatic brain injury (mTBI) | 16 physicians 1 ^o round: 12 2 ^o round: 13 | Unreported | Four examination procedure items had 77 percent of the participants rate the test as important. The clinical guidelines produced by this effort can be modified to suit individual patient needs | Any question with 80 percent or more of the responses marked strongly agree or agree was accepted. Questions with less than 50 percent of the responses rated strongly agree or agree were rejected | The first potential source of bias was in recruiting the panel of experts. Since participation was voluntary, inclusion may have been biased. It is also unknown whether other experts who chose not to participate would have provided similar answers | |

Legend: Level of evidence*: assessed according to the recommendation score of the Oxford Centre for Evidence-based Medicine. Abbreviations: VAP-ventilator-associated pneumonia. ICU- Intensive care unit. SA- Septic Arthritis. TS- Transient Synovitis. ED- Emergency Department. COPD- Chronic Obstructive Pulmonary Disease

DISCUSSION

The use of the Delphi technique for decision-making in critically ill patients effectively contributed to solve medical issues in 15 of the 18 articles selected in this review. The contributions of the Delphi technique in the studies selected were in the areas of management strategies, drug treatment, recommendations for specific symptoms, use of equipment, guidelines for the management and diagnosis of patients, consensus for tests and examinations and identification of topics that cause differences in opinion and their underlying reasons^{2,3,5,13-20,22-26}.

Thus, this study analyzed the contribution of the Delphi technique to the management of critical patients by physicians. However, only three studies mentioned the final results obtained with the implementation of the guidelines developed: Garrouste-Orgeas⁷, Elliot et al.²⁴ and Quintana²⁶. This fact limits the evaluation of the survival rates of these patients.

In ten of the 18 articles published from 2001 to 2016, the analysis addressed the use of a modified Delphi technique. The changes in the application of the method that were mentioned were face-to-face meetings^{2,15,25,26} and meeting on a first-round schedule determined by the moderator²¹ and not by the answers of the first round, as in the classical technique. The studies of Elliot²⁴, Goodrich¹⁴, Pearl¹⁷, Ogden¹⁸, Frankel¹⁹ and Arsemio²³ reported the use of a method with modifications but did not describe these modifications. In these cases, an incompletely described Delphi method can affect the overall quality of the final consensus, since the decisions made will probably not have the credibility required for use in medical practice. In the other studies, the Delphi technique was performed according to the original protocol.

Nine out of the 18 studies evaluated did not report any physicians as lost to follow up during the application of the technique.^{2,3,7,15,16,17,19,25,26} In the other studies, the high percentage of permanence of the participants (71.99%) during the process contributed to the reliability of the consensus decisions and helped to consolidate the structure of the method. In addition, among all the studies included in this review, ten established consensus as greater than 70%, which is a value already used in the literature,

as in Harmsen²⁷. This guarantees that the method has a valid structure.

Sixteen out of the 18 articles evaluated in this systematic review mentioned the main limitations related to the use of the Delphi method^{2,3,7,13,14,16-26}. Among these limitations, the most recurrent were the selection of specialists and the lack of scientific rigor of the method. Thus, the present study was important to recognize and group the main limitations that may occur during the process. However, further studies are necessary to validate mechanisms that can solve these issues.

This is the first Brazilian systematic review that elucidates the use of the Delphi technique for decision-making in critically ill patients. The strengths of the study include the recovery of studies published over a 15-year period (2001-2016) and the analysis of the contribution of the method to medical decisions in critical situations. However, this study has its limitations. There is no consensus on how to assess the applicability of a Delphi method. Consequently, in three of the studies analyzed, the method was not effective^{7,21,24}. In addition, it is difficult to accurately assess the effectiveness of the method, as there is no minimum and ideal percentage to reach consensus defined in the literature. Finally, it was not possible to find clinical trials using the Delphi technique, since the time required to reach consensus can be considered relatively long for a patient in a critical situation who needs early and rapid intervention. In three studies^{3,5,15}, the time required to complete the rounds ranged from three months to one year.

CONCLUSION

This study presents the applicability of the Delphi technique for the management of critically ill patients, drug treatment, recommendations for specific symptoms, guidelines for the management and diagnosis of patients, use of equipment, consensus for tests and examinations and identification of topics that cause differences in opinion and their underlying reasons in situations where there is no consensus in the medical literature. In 15 of the articles selected, the technique contributed to the construction of medical guidelines.

Acknowledgements: We thank the Itaúna University and the Federal University of São João del-Rei for supporting this study. This study was carried out with the support of the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES) - Financing Code 001 and Minas Gerais State Foundation for Research Support (FAPEMIG).

Author's contributions: Santos APFB - contributed substantially to the conception, data collection, analysis and interpretation of data; Andrade JF - contributed substantially to the conception, data collection, analysis and interpretation of data; Alves GCS - contributed substantially to the conception, data collection, analysis and interpretation of data, discussion of results and critical review of the article regarding technical content; Silva SD - contributed substantially to the conception, data collection, analysis and interpretation of data; Cristina Sanches - contributed to the preparation of the study, supervision of data collection, guidance and analysis of data, discussion of results and critical review of the article regarding technical content; Chequer FMD - contributed to the preparation of the study, supervision of data collection, guidance and analysis of data, discussion of results and critical review of the article regarding technical content.

REFERENCES

1. Habibi A, Sarafrazi A, Izadyar S. Delphi Technique Theoretical Framework in Qualitative Research. *Int J Eng Sci*. 2014;3(4):8-13. Available from: <http://theijes.com/papers/v3-i4/Version-4/B03404008013.pdf>.
2. Lee JH, Park MS, Kwon H, Chung CY, Lee KM, Kim YJ, et al. A guideline for differential diagnosis between septic arthritis and transient synovitis in the emergency department: a Delphi survey. *Am J Emerg Med*. 2016;34(8):1631-6. doi: 10.1016/j.ajem.2016.06.006.
3. Nguyen-Lu N, Downey K, Carvalho JC. Controversy between anesthesiologists and obstetricians on the labour ward: the Delphi method is used as a consensus-building technique. *Can J Anesth*. 2015;62(3):271-7. doi: 10.1007/s12630-014-0294-7.
4. Wassenaar A, Van den Boogaard M, Underpin-Icu Study Group, Schoonhoven L, Pickkers P. Determination of the feasibility of a multicomponent intervention program to prevent delirium in the Intensive Care Unit: a modified RAND Delphi study. *Aust Crit Care*. 2017;30(6):321-7. doi: 10.1016/j.aucc.2016.12.004.
5. Mendes JC, Silva LJ. Neonatal palliative care: developing consensus among neonatologists using the Delphi Technique in Portugal. *Adv Neonatal Care*. 2013;13(6):408-14. doi: 10.1097/ANC.0000000000000037.
6. Strang JF, Meagher H, Kenworthy L, de Vries AC, Menvielle E, Leibowitz S, et al. Initial Clinical Guidelines for Co-Occurring Autism Spectrum Disorder and Gender Dysphoria or Incongruence in Adolescents. *J Clin Child Adolesc Psychol*. 2018;47(1):105-15. doi: 10.1080/15374416.2016.1228462.
7. Garrouste-Orgeas M, Boumendil A, Pateron D, Aegerter P, Somme D, Simon T, et al. Selection of intensive care unit admission criteria for patients aged 80 years and over and compliance of emergency and intensive care unit physicians with the selected criteria: An observational, multicenter, prospective study. *Crit Care Med*. 2009;37(11):2919-28. doi: 10.1097/ccm.0b013e3181b019f0.
8. Sousa AS, Cunha CP, Magalhães LC, Kaiser SE, Saraiva JK. Posicionamentos, diretrizes e normatizações. Veículo de auxílio à prática médica. *Arq Bras Cardiol*. 2017;109(4):368-9. doi:10.5935/abc.20170133.
9. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097. doi: 10.1371/journal.pmed.1000097.
10. Galvão TF, Pereira MG. Revisões sistemáticas da literatura: passos para sua elaboração. *Epidemiol Serv Saúde*. 2014;23(1):183-4. doi: 10.5123/S1679-49742014000100018.
11. Phillips B, Ball C, Sackett D, Badenoch D, Straus S, Haynes B, et al. Oxford Centre for Evidence-based Medicine - Levels of evidence (March 2009). Oxford: Centre for Evidence-Based Medicine (CEBM); 2009 [cited 2018 Feb 18]. Available from: <http://www.cebm.net/index.aspx?o=1025>.
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-74. doi: 10.2307/2529310
13. Baumann MH, Strange C, Heffner JE, Light R, Kirby TJ, Klein J, et al. Management of Spontaneous Pneumothorax: an American College of Chest Physicians Delphi Consensus Statement. *Chest*. 2001;119(2):590-602. doi: 10.1378/chest.119.2.590.
14. Goodrich GL, Martinsen GL, Flyg HM, Kirby J, Asch SM, Brahm KD, et al. Development of a mild traumatic brain injury-specific vision screening protocol: a Delphi study. *J Rehabil Res Dev*. 2013;50(6):757-68. doi: 10.1682/JRRD.2012.10.0184.
15. Dellinger RP, Levy MM, Carlet JM, Bion J, Parker MM, Jaeschke R, et al. Surviving Sepsis Campaign: International guidelines for management of severe sepsis and septic shock: 2008. *Intens Care Med*. 2008;34(1):17-60. doi: 10.1007/s00134-007-0934-2.
16. Li LT, Mills WL, Gutierrez AM, Herman LI, Berger DH, Naik AD. A patient-centered early warning system to prevent readmission after colorectal surgery: a national consensus using the Delphi method. *J Am Coll Surg*. 2013;216(2):210-6. doi: 10.1016/j.jamcollsurg.2012.10.011.
17. Pearl JS, Gajic O, Dong Y, Herasevich V, Gong MN. Creation of the Prevention of Organ Failure Checklist. A Multidisciplinary Approach Using the Modified Delphi Technique. *Ann Am Thorac Soc*. 2016;13(6):910-6. doi: 10.1513/AnnalsATS.201509-626BC.
18. Ogden SR, Culp WC, Villamaria FJ, Ball TR. Developing a checklist: Consensus Via a Modified Delphi Technique. *J Cardiothorac Vasc Anesth*. 2016;30(4):855-8. doi: 10.1053/j.jvca.2016.02.022.
19. Frankel HL, Kirkpatrick AW, Elbarbary M, Blaivas M, Desai H, Evans D, et al. Guidelines for the Appropriate Use of Bedside General and Cardiac Ultrasonography in the Evaluation of Critically Ill Patients-Part I: General Ultrasonography. *Crit Care Med*. 2015;43(11):2479-502. doi: 10.1097/CCM.0000000000001216
20. Maher TM, Whyte MK, Hoyles RK, Parfrey H, Ochiai Y, Mathieson N, et al. Development of a Consensus Statement for the Definition, Diagnosis, and Treatment of Acute Exacerbations of Idiopathic Pulmonary Fibrosis Using the Delphi Technique. *Adv Ther*. 2015;32(10):929-43. doi: 10.1007/s12325-015-0249-6.
21. Tan K, Baxter G, Newell S, Smye S, Dear P, Brownlee K, et al. Knowledge elicitation for validation of a neonatal ventilation expert system utilising modified Delphi and focus group techniques. *Int J Hum Comput Stud*. 2009;68(6):344-54. doi: 10.1016/j.ijhcs.2009.08.003.
22. Hawryluck LA, Harvey WR, Lemieux-Charles L, Singer PA. Consensus guidelines on analgesia and sedation in dying intensive care unit patients. *BMC Med Ethics*. 2002;12;3:E3. doi: 10.1186/1472-6939-3-3.
23. Ansermino JM, Dosani M, Amari E, Choi PT, Schwarz SW. Defining rules for the identification of critical ventilatory events. *Can J Anesth*. 2008;55(10):702-714. doi: 10.1007/bf03017747
24. Elliott D, Elliott R, Burrell A, Harrigan P, Murgo M, Rolls K, et al. Incidence of ventilator-associated pneumonia in Australasian intensive care units: use of a consensus-

- developed clinical surveillance checklist in a multisite prospective audit. *BMJ Open*. 2015;5(10):e008924. doi: 10.1136/bmjopen-2015-008924.
25. Garcia-Gutierrez S, Quintana JM, Aguirre U, Esteban C, Bilbao A, Escobar A, et al. Explicit criteria for hospital admission in exacerbations of chronic obstructive pulmonary disease. *Int J Tuberc Lung Dis*. 2011;15(5):680-6. doi: 10.5588/ijtld.10.0408.
26. Quintana JM, Esteban C, Barrio I, Garcia-Gutierrez S, Gonzalez N, Arostegui I, et al. The IRYSS-COPD appropriateness study: objectives, methodology, and description of the prospective cohort. *BMC Health Serv Res*. 2011;11:322. doi: 10.1186/1472-6963-11-322.
27. Harmsen AM, Geeraedts LM, Giannakopoulos GF, Terra M, Christiaans HM, Mookink LB, et al. Protocol of the DENIM study: a Delphi-procedure on the identification of trauma patients in need of care by physician-staffed Mobile Medical Teams in the Netherlands. *Scand J Trauma Resusc Emerg Med*. 2015;23:15. doi: 10.1186/s13049-015-0089-z.

Received: August 04, 2019

Accepted: March 17, 2020