PERSPECTIVE

The Zeitgeist of Challenging the Evidence. A Perspective on the International Consensus Meeting on Periprosthetic Joint Infection

Hangama C. Fayaz, MD, PhD; Jesse B. Jupiter, MD

Research performed at Department of Orthopaedic Surgery, Harvard Medical School, Massachusetts General Hospital, Boston, MA, USA

Received: 10 January 2016 Accepted: 05 June 2016

Abstract

Background: The economic burden of the treatment of periprosthetic joint infection (PJI) is high and the treatment of PJI has a high degree of international controversy. Several papers have declared the International Consensus Meeting on Periprosthetic Joint Infection (ICMPJI) to be the “flawless pledge of international academics” to overcome the challenges of musculoskeletal infections. The purpose of this paper is to highlight for the first time some essential insights into the key dilemmas that are associated with this international consensus process.

Methods: The proceedings of the ICMPJI was reviewed, and the critical consensus agreements that were reached were communicated via e-mail to 48 leading orthopaedic surgeons, microbiologists and statisticians around the world. Of these, 30 responded, 8 did not, and 10 of respondents were not aware of the ICMPJI.

Results: A thorough review of the ICMPJI proceedings identified a clear need to resolve some of the dilemmas that we highlight in this paper. The Delphi procedure has been described as a survey technique that enables a group dynamic-based practice. Although there have been several published reports on this procedure, its scientific merit is still being debated. Several challenges and questions have been raised regarding the application of the Delphi technique, but there is no doubt that it is a vital approach for achieving consensus on subjects where none currently exists.

Conclusion: Performing prospective clinical studies in this area is currently the best and only option to overcome this challenge. In the long term, this approach will not only incorporate the standard of clinical evidence but also adopt regional mores for treating infection, which include patient values, cultural differences and local financial resources.

Keywords: Delphi Procedure, International Consensus Meeting, Level of Evidence, Periprosthetic Joint Infection, Quality of patient care, Trans-continental controversy

Why we do what we do

During grand rounds in 2010, Dr. Thomas Thornhill (Brigham and Women’s Hospital, Boston) was emphasizing the developing controversy surrounding PJIs and explaining the thresholds for the erythrocyte sedimentation rate and C-reactive protein level in infected prosthetic joints. At the same time, Dr. Arlen Hanssen (Mayo Clinic, Rochester) and Dr. Javad Parvizi (Rothman Institute, Philadelphia) were diligently publishing papers on PJI. Their work has attracted significant attention because of the breadth of their publications and their tireless effort to shed light on a topic that has lacked a high level of evidence to guide clinical decision-making. Their continuing efforts have undoubtedly required a tremendous amount of dedication and hard work to achieve a certain amount of clarification on an international level.

To me (first author) personally, this topic has become the most indicative parameter of the orthopaedic knowledge of any orthopaedic chief surgeon as well as a
symbol of the quality of patient care at a hospital. It is our belief that the quality of patient care varies greatly by geographic location. To evaluate this hypothesis, I joined orthopaedic surgery programs throughout Europe, the United States and Asia.

During my visits to various hospitals, I observed that even in high patient-volume centers with joint replacement programs, the understanding of PJI was significantly lacking.

In the majority of these centers, PJI was handled as an anecdote. Only a few attending physicians were willing to offer insights into their decision-making process.

Consequently, in 2013, more than 400 delegates from 51 countries contributed to a staged, Delphi-based international multidisciplinary consensus meeting in Philadelphia, U.S.A., to discuss both what we know and do not know and to define what we need to know. The participants were involved in active discussions and voted on the questions and consensus statements. The delegates first met on July 31, in small workgroups, which was followed by a general assembly for further discussion of the questions and consensus statements. Once the consensus statements were revised, the finalized consensus statement was loaded into the Audience Response System for voting to begin the next day. On August 1, the delegates came into the general assembly and voted on the 207 questions/consensus statements that were being presented (1).

The dilemma regarding the delphi procedure

The Greek word Delphoi means "womb" or "hollow" and historically refers to Gaia, "the great mother of all creatures on earth" or "the primordial Earth goddess" (2). A half-century ago, the first Delphi experiment took place at the RAND Corporation headquarters in Santa Monica, California, and for security reasons, the first publication followed a decade later (3). The procedure was applied to formulate predictions and to support decision-making. Since its first secret application in the military, the Delphi procedure has been modified to become more sophisticated over the years based on criticism and re-evaluations until it attained its most consistent consensus view of a group of skilled experts. The participants should preferably represent a "broad sampling" of experts from diverse populations (9).

Therefore, the accuracy of the Delphi technique requires that more attention be paid to the appropriate selection of expert participants, the development of statements and methods of analysis, and the determination of consensus to circumvent bias and misrepresentation.

In conclusion, the Delphi method possesses some methodological limitations, as does any method. Specifically, these include the definition of an expert, the biases of experts (10), the restriction of interactions in written and controlled responses (10), the limitation of the opportunity for social rewards for individual input to the group (11), the simplicity associated with the methodology used for detailed analysis by the individual conducting the study (10), the complexity of examining the method's accuracy (3) and reliability (12), and the time needed to perform the procedure (13).

In addition, there are application-related deficiencies (4, 14), including the inaccurate selection of experts; questions and the wording of these questions. Moreover, there can be inaccuracies with the order in which the questions are presented within the survey. There may also be questions that are not well formulated and inadequately analyzed outcomes. The benefits of consensus include the anonymity of the participants, the expression of consensus by summary measures, and controlled responses, which enable individuals to change their view in light of the group's response. In accordance with the principles of a Delphi survey, each expert should remain blind to the identity of the other experts.

Considering the limitations of this method, several papers have described the Delphi technique not as a standalone approach but rather as a technique that may be improved by other approaches or used in conjunction with other methods (15). All of these publications highlighted the limitations of the Delphi method, including the relatively restricted interaction.


The following four essential features characterize the Delphi consensus method:

1. The development of a group consensus view; 2. consultation with experts at least twice on the same question; 3. "anonymity" of involvement; and 4. statistical forecasts of the possible time course for the settings (9).

This amalgamation of procedures generates forecasts that are more precise than a prediction from an individual expert would be (5).

A thorough review of the literature identified only one published study of Delphic polling accuracy, which was conducted by Parente and Parente over a 30-year period. The authors evaluated the overall accuracy of long-term Delphic polling and found that it correctly predicted the occurrence of 14 out of 18 scenarios. The authors indicated that the selection of participants must include a "representative sampling" of skilled experts. The participants should preferably represent a "broad sampling" of experts from diverse populations (9).

The following four essential features characterize the Delphi consensus method:

1. The development of a group consensus view; 2. consultation with experts at least twice on the same question; 3. "anonymity" of involvement; and 4. statistical forecasts of the possible time course for the settings (9).

This amalgamation of procedures generates forecasts that are more precise than a prediction from an individual expert would be (5).

A thorough review of the literature identified only one published study of Delphic polling accuracy, which was conducted by Parente and Parente over a 30-year period. The authors evaluated the overall accuracy of long-term Delphic polling and found that it correctly predicted the occurrence of 14 out of 18 scenarios. The authors indicated that the selection of participants must include a "representative sampling" of skilled experts. The participants should preferably represent a "broad sampling" of experts from diverse populations (9).

In conclusion, the Delphi method possesses some methodological limitations, as does any method. Specifically, these include the definition of an expert, the biases of experts (10), the restriction of interactions in written and controlled responses (10), the limitation of the opportunity for social rewards for individual input to the group (11), the simplicity associated with the methodology used for detailed analysis by the individual conducting the study (10), the complexity of examining the method's accuracy (3) and reliability (12), and the time needed to perform the procedure (13).

In addition, there are application-related deficiencies (4, 14), including the inaccurate selection of experts; questions and the wording of these questions. Moreover, there can be inaccuracies with the order in which the questions are presented within the survey. There may also be questions that are not well formulated and inadequately analyzed outcomes. The benefits of consensus include the anonymity of the participants, the expression of consensus by summary measures, and controlled responses, which enable individuals to change their view in light of the group's response. In accordance with the principles of a Delphi survey, each expert should remain blind to the identity of the other experts.

Considering the limitations of this method, several papers have described the Delphi technique not as a standalone approach but rather as a technique that may be improved by other approaches or used in conjunction with other methods (15). All of these publications highlighted the limitations of the Delphi method, including the relatively restricted interaction.
among participants. These publications indicated that these limitations in both “input and output” might be addressed through the addition of alternative techniques (16-19).

Additional research is needed to evaluate the effectiveness of the Delphi method (20).

Unfortunately, until now, there have been no standard guidelines for determining a cut-off point in a Delphi process. In 2006, Keeney et al. concluded that there were no standard guidelines on the appropriate level of consensus. According to their findings, 75% appears to be the minimum level, although there is no scientific validation for this. They proposed that researchers should determine the consensus level before initiating a study and apply confidence intervals to values with a high level of importance (21).

In conclusion, the Delphi technique is mainly supported by the concept that several people are less likely to arrive at a wrong conclusion than a single person would be (22).

However, we need to emphasize that clinicians and researchers must be aware that the extent to which experts agree with one another does not indicate that the ‘correct’ answer has been reached (21). Indeed, there is even a high possibility of spreading false information.

To illustrate this assertion, here we reflect on some of the questions originated from the proceedings that highlight this quandary in the setting of a Delphi-based contemplated consensus procedure:

1. **Question 9, workgroup 7: Is there a role for molecular techniques, such as polymerase chain reaction (PCR), for the diagnosis of PJI? If so, in which group of patients should this be done?**

   **Consensus:** Nucleic acid-based testing is not currently a recommended routine diagnostic test for PJI. In cases with high clinical suspicion of infection but negative cultures or other diagnostic tests, molecular techniques with or without sonication may help to identify the unknown pathogens or antibiotic sensitivity for targeting antimicrobial therapies.

   **Delegate Vote:** Agree: 96%, Disagree: 3%, Abstain: 1% (Strong Consensus)

2. **Question 1A, workgroup 7: What is the definition of PJI?**

   **Consensus:** PJI is defined as:
   - Two positive periprosthetic cultures with phenotypically identical organisms; or
   - A sinus tract communicating with the joint; or
   - Having three of the following minor criteria:
     - elevated serum C-reactive protein (CRP) AND erythrocyte sedimentation rate (ESR),
     - elevated synovial fluid white blood cell (WBC) count OR ++ change on leukocyte esterase test strip,
     - elevated synovial fluid polymorphonuclear neutrophil percentage (PMN%),
     - positive histological analysis of periprosthetic tissue, and
     - a single positive culture.

   **Delegate Vote:** Agree: 85%, Disagree: 13%, Abstain: 2% (Strong Consensus)

2. **The question that arises is how clinicians should rectify the fact that finding two positive cultures is considered a main criterion for PJI while at the same time, a single positive culture can fulfill the criteria if it is accompanied by an elevated CRP or ESR, elevated WBC/change in LET, elevated PMN, or positive histological analysis of periprosthetic tissue.**

Many European centers have abandoned ESR and now rely solely on CRP.

In our experience, ESR should not be included into the algorithm. We assume it is included because ESR and CRP are always used together in the U.S., despite our increasing knowledge that the ESR only increases the chance of false positives. Therefore, ESR should no longer be used.

Two positive periprosthetic cultures should not be the only criteria. Indeed, the absence of a standard definition for PJI makes it difficult to compare studies.

The diagnosis of infection must be confirmed prior to surgery to guide the surgeon: CRP and preoperative aspiration (culture and WBC) are therefore crucial diagnostic tools. In the future, more specific biomarkers will be mandatory to diagnose PJI.

Considering these factors, it is concerning that 85% of delegates with a strong consensus agreed with this approach.
3. Question 14, workgroup 9: Which antibiotic should be used, and how much should be added to the cement spacers?

Consensus: The type of antibiotic and the dose need to be individualized for each patient based on the organism profile and antibiogram (if available) as well as on the patient’s renal function and allergy profile. However, most infections can be treated with a spacer with vancomycin (1 to 4 g per 40 g package of cement) and gentamicin or tobramycin (2.4 to 4.8 g per 40 g package of cement).

Delegate Vote: Agree: 89%, Disagree: 7%, Abstain: 4% (Strong Consensus)

3. This is a topic with international variations. It has not yet been proven that individualized cements are better than standard, industry-mixed cements.

In Europe, depending on the local epidemiology, most infections do not require vancomycin. Rather, the standard revision cement containing gentamicin (gram-negative bacteria) and clindamycin (gram-positive bacteria) is sufficient. In addition, in vitro data suggest an antibiotic film effect with clindamycin, unlike vancomycin. In 15-20% of cases, Staphylococcus and Streptococcus strains are resistant to clindamycin. In those cases, vancomycin should be used, and we recommend limiting vancomycin to only those cases. When selecting antibiotics for use in the cement, clinicians should consider the known or suspected organisms, local antimicrobial susceptibility patterns, and thermal stability of the organism. We are not aware of any human studies comparing different antibiotics (e.g., clindamycin vs. vancomycin) in cement, although there are industry-sponsored studies on the use of antibiotics containing Polymethyl methacrylate. As far as we know, no studies have prospectively evaluated the need for a second drug in addition to the aminoglycoside. In the U.S., however, not all Staphylococci are sensitive to gentamicin. In Boston, for example, there is a 20% gentamicin resistance rate among coagulase-negative Staphylococci. In the U.S., rates of clindamycin-resistant Staphylococcus and Streptococcus are also higher than those observed in Europe. Therefore, when gram-positive spacer coverage is required, vancomycin is generally used.

4. Question 11, workgroup 11: Can oral antibiotic therapy be used instead of intravenous (IV) therapy for the initial treatment of PJI following resection?

Consensus: There is evidence to support pathogen-specific, highly bioavailable oral antibiotic therapy for the treatment of PJI.

Delegate Vote: Agree: 79%, Disagree: 11%, Abstain: 1% (Strong Consensus)

4. This is a topic with international variations. In the U.S., almost all cases of PJI are treated with IV antibiotic therapy for 4 to 6 weeks, followed by oral antibiotic suppression therapy depending on the surgical strategy. In cases of infection due to susceptible Staphylococci, oral Rifampicin is usually added to the treatment regimen of IV antibiotics. Oral antibiotics are not used as monotherapy, after debridement, or after another antibiotic has reached therapeutic levels. When oral antibiotics are employed (usually in the suppressive phase in the U.S.), options with good bioavailability include Quinolones, Cotrimoxazole, Clindamycin, Linezolid, and Doxycycline/Minocycline.

In Europe, due to a preoperatively limited enteral resorption capacity and the emergence of resistance, particularly with Rifampicin and Quinolones, clinicians do not recommend oral antibiotics as an initial therapy. Rather, they recommend the following protocols.

In an acute infection requiring emergent intervention when an antibiogram is not available, they recommend a high dose of Co-amoxicillin. Antibiotics should not be initiated before cultures have been taken. As soon as an antibiogram is available, a 14-day course of IV antibiotics should be initiated. These antibiotics should be pathogen specific, have high bioavailability and be able to inhibit biofilm formation.

As contrast to the U.S., Rifampicin should never be given as monotherapy, and it should be initiated only after wounds have dried. Monotherapy with Chinolones against staphylococci should not be used. An oral antibiotic regimen based on an antibiogram for a total treatment duration of 6 to 12 weeks should be followed.

In conclusion, there appears to be no evidence that a treatment duration of longer than 2 weeks is more effective than a shorter course of treatment. However, 6 weeks is still the standard at most institutions.

5. Question 6A, workgroup 7: Is there a role for routine acid-fast bacillus (AFB) and fungal testing in suspected PJI?

Consensus: In cases of proven or suspected PJI, AFB and fungal cultures should be limited to those patients at risk for such infections or to those cases in which other traditional pathogens have not been identified and clinical suspicion persists.

Delegate Vote: Agree: 92%, Disagree: 6%, Abstain: 1% (Strong Consensus)

5. Although factors such as prior bacterial PJI, earlier antimicrobial coverage, immunosuppressive treatment, and diabetes have been implicated as risk factors for fungal PJI (24, 25), studies performed by Azzam et al. (2009) and Phelan et al. (2002) have indicated that the majority of fungal PJIs occur after revision arthroplasty (24, 26). Several of the published cases of Aspergillus PJI occurred in immunocompetent individuals, unlike fungal pulmonary infections, which often arise in immunocompromised patients (27, 28).

In conclusion, Aspergillus PJI may also occur in immunocompetent individuals, and more importantly, it occurs after revision arthroplasty. Considering these factors, it is concerning that 92% of delegates with a strong consensus agreed with this approach.

6. Question 7C, workgroup 7: Should antibiotics be withheld prior to obtaining samples for culture in all cases?

Consensus: No. Perioperative prophylactic antibiotics
should be withheld only in cases with a high suspicion for PJI in which an infecting organism has not been isolated. **Delegate Vote:** Agree: 87%, Disagree: 12% Abstain: 1% (Strong Consensus)

6. This question is not well formulated and is not adequately answered. While the question refers to whether antibiotics should be withheld prior to obtaining samples, i.e., prior antimicrobial treatment, the consensus statement refers solely to perioperative prophylactic antibiotics. A precise time frame should have been indicated.

The incidence of culture-negative PJI varies from 5 to 35% (29, 30). The most important risk factor for culture-negative PJI appears to be prior antimicrobial administration. In 2010, Malekzadeh et al. conducted a case-controlled study to identify risk factors associated with the occurrence of culture-negative PJI, and they paid special attention to the administration of prior antimicrobial therapy. They concluded that prior antimicrobial treatment and postoperative wound drainage were related to an increased risk of negative cultures among PJI patients (31).

Because the identification of one or more pathogens is crucial for selecting the antimicrobial regimen, any prior antibiotic treatment should be discontinued at least 2 weeks prior to surgery, and perioperative antimicrobial exposure should be withheld until culture specimens have been collected (32).

In standard revision cases, in which an aseptic situation is likely, the answer is no: Perioperative prophylactic antibiotics should be administered to decrease the risk of a new infection. However, in revision cases with a suspected infection and without an isolated microorganism, the answer is yes: Withholding antimicrobials prior to surgery will definitely decrease the number of culture-negative infections, and it will contribute to a decreased rate of PJI in revision cases.

Considering these factors, it is questionable how 87% of the delegates agreed on this approach. According to Dr. Parvizi, numerous level I studies are already being conducted in research areas identified by the consensus project. In the next consensus, a level of evidence should be assigned to each “recommendation” based on the weight of published data related to each specific question and the Oxford Center for Evidence-Based Medicine criteria may be applied.

7. **Question 8, workgroup 7:** Is there a role for routine sonication of the prosthesis? If so, in which groups of patients should this be done?

**Consensus:** No. We do not recommend routine sonication of explants. Its use should be limited to cases of suspected or proven PJI (based on presentation and other testing) in which preoperative aspiration does not yield a positive culture, and antibiotics have been administered within the previous 2 weeks.

**Delegate Vote:** Agree: 84%, Disagree: 9%, Abstain: 7% (Strong Consensus)

7. We prefer to perform routine sonication of explants. Sonication should be performed in all cases, and particularly in those with otherwise unexplained early implant loosening. Early implant loosening may be due to a low-grade infection, which is difficult to detect with standard intraoperative cultures. The sonication of explanted prosthesis components enables a sampling of the implant-associated biofilm, which is not possible when performing an aspiration or taking an intraoperative culture. By culturing the sonication fluid, the number of detected pathogens and the sensitivity can be improved (33). This high level of sensitivity is particularly advantageous in cases that employ perioperative antibiotics.

The validity of the diagnosis can be improved by a combination of sonication results and a histologically classified periprosthetic membrane, particularly in cases of low-grade infection with prosthetic loosening (34). Considering these factors, it is concerning that 84% of delegates with a strong consensus agreed with this approach.

**Summary**

The economic burden of PJI is considerable: the approximate cost to the U.S. health care system to treat PJI was $566 million in 2009 alone, a number that is expected to reach $1.62 billion in 2020. Yet, this estimation is a gross underestimate, as it considered only the estimated hospital cost, excluding many other direct and indirect costs (35).

Although the concept of consensus in other medical disciplines is well respected, we do not yet believe that the optimum outcome related to patient care, i.e., the ICMPJI consensus, has been accomplished for PJI for the following reasons. First, the education level of orthopaedic surgeons worldwide is heterogeneous. Second, cultural beliefs that we should use the best available evidence and patient consultation to guide treatment of essential orthopaedic conditions vary. Third, the Delphi technique is highly vulnerable to selection bias.

A second ICMPJI will be held by 2018. We hope that the initiators will take into account the concerns that we have expressed and will address some of the notable gaps in the literature.

According to Dr. Parvizi, numerous level I studies are already being conducted in research areas identified by the consensus project. In the next consensus, a level of evidence should be assigned to each “recommendation” based on the weight of published data related to each specific question and the Oxford Center for Evidence-Based Medicine criteria may be applied. Ideally, when planning the next meeting, the initiators will consider the diverse nature of their project. They will be more inclusive of countries and experts and focus more on panel diversity. Unlike the recent ICMPJI, most of the liaisons should not be recruited from the same institution. In accordance with the principles of a Delphi survey, panel diversity will definitely contribute next time to a reduced amount of bias in terms of selection, self-interest, prospective, and the desirability of the outcomes (19, 20).

Considering the overarching goal of achieving optimal patient care, a consensus reached by a panel of experts should preferably reflect the application
of current evidence to surgical practice. In addition, this international consensus meeting should be considered as a starting point for developing a framework that initiates and encourages clinicians to perform a higher level of evidence-based research on topics that appear to be unclear or controversial in this area. This consensus has successfully identified unanswered questions in the field of PJI, and these questions require additional investigation because the answer to a clinical question must be based on a combination of all types of evidence. “No single study provides a definitive answer” (36).

Acknowledgments

The authors acknowledge the contribution of following colleagues for their efforts in manuscript preparation: Thomas S. Thornhill, MD, John B. and Buckminster Brown Professor of Orthopaedic Surgery, Harvard Medical School, Boston, USA. R. Malcolm Smith, MD, FRCS, Associate Professor of Orthopaedic Surgery, Harvard Medical School, Boston, USA. Frank M. Klenke, MD, PhD, Associate Professor of Orthopaedic Surgery, Department of Orthopedic Surgery, Inselspital, Bern University Hospital, Switzerland. Arne C. Rodloff, MD, PhD, Professor of Microbiology and Infection, Department of Microbiology, University Hospital, Leipzig, Germany. Klaus A. Siebenrock, MD, PhD, Professor of Orthopaedic Surgery, Department of Orthopedic Surgery, Inselspital, Bern University Hospital, Switzerland. Fares S. Haddad, MD, FRCS, Professor of Orthopaedic Surgery, HYPERLINK “http://www.uclh.org/Pages/home.aspx” University College London Hospitals, England. Daniel Kendoff, MD, Professor of Orthopaedic Surgery, Department of Orthopedic Surgery, Helios Klinik, Berlin, Germany.

Hangama C. Fayaz MD PhD
Jesse B. Jupiter MD
Department of Orthopaedic Surgery, Harvard Medical School, Massachusetts General Hospital, Boston, MA, USA

References

19. Landeta J, Barrutia J, Lertxundi A. Hybrid Delphi: a methodology to facilitate contribution from experts