Elaboration of Consensus Clinical Endpoints to Evaluate Antimicrobial Treatment Efficacy in Future Hospital-acquired/Ventilator-associated Bacterial Pneumonia Clinical Trials

Emmanuel Weiss,[1,2] Jean-Ralph Zahar,[3,4] Jeff Alder,[5] Karim Asehnoune,[6] Matteo Bassetti,[7] Marc J. M. Bonten,[8] Jean Chastre,[9] Jan De Waele,[10] George Dimopoulos,[11] Philippe Eggimann,[12] Marc Engelhardt,[13] Santiago Ewig,[14] Marin Kollef,[15] Jeffrey Lipman,[16,17] Carlos Luna,[18] Ignacio Martin-Loeches,[19] Leonardo Pagani,[20] Lucy B. Palmer,[21] Laurent Papazian,[22] Garyphallia Poulakou,[23,24] Philippe Prokocimer,[25] Jordi Rello,[26] John H. Rex,[27] Andrew F. Shorr,[28] George H. Talbot,[29] Visanu Thamlikitkul,[30] Antoni Torres,[31] Richard G. Wunderink,[32] and Jean-François Timsit[33,34]

[1]Department of Anesthesiology and Critical Care, Assistance Publique–Hôpitaux de Paris (AP-HP), Beaujon Hospital, Clichy, [2]Unité Mixte de Recherche (UMR) 1149, Centre for Research on Inflammation, Institut national de la santé et de la recherche médicale (INSERM)/Université Paris Diderot, Paris, [3]Department of Clinical Microbiology and Infection Control Unit, Avicennes Hospital, AP-HP, Bobigny, and [4]Infection, Antibiotics, Modelisation, Epidemiology (IAME), UMR 1137, Université Paris 13, Sorbonne Paris Cité, France; [5]Bayer US LLC, Parsippany, New Jersey; [6]University Hospital of Nantes, Intensive Care Unit, Anesthesia and Critical Care Department, Hôtel Dieu, Nantes, France; [7]Infectious Diseases Division, Department of Medicine, University of Udine and Santa Maria Misericordia University Hospital, Italy; [8]Department of Medical Microbiology and Julius Center for Health Science and Primary Care, University Medical Center Utrecht, The Netherlands; [9]Service de Réanimation Médicale, Institut de Cardiologie, Groupe Hospitalier Pitié-Salpêtrière, AP-HP, Paris, France; [10]Department of Critical Care Medicine, Ghent University Hospital, Belgium; [11]Department of Critical Care, University Hospital Attikon, National and Kapodistrian University of Athens, Greece; [12]Department of Critical Care, Centre Hospitalier Universitaire Vaudois, Lausanne, and [13]Basilea Pharmaceutica International Ltd, Basel, Switzerland; [14]Department of Respiratory Medicine and Infectious Diseases, Evangelic Hospital in Herne and Augusta Hospital, Bochum, Germany; [15]Division of Pulmonary and Critical Care Medicine, Washington University School of Medicine, St Louis, Missouri; [16]Royal Brisbane and Womens Hospital, Australia; [17]University of the Witwatersrand, Johannesburg, South Africa; [18]Department of Medicine, Pulmonary Diseases Division, Hospital de Clínicas, Universidad de Buenos Aires, Argentina; [19]Department of Clinical Medicine, Multidisciplinary Intensive Care Research Organization, St James's Hospital, Trinity Centre for Health Sciences, Dublin, Ireland; [20]Infectious Diseases Unit, Bolzano Central Hospital, Italy; [21]Pulmonary, Critical Care and Sleep Division, State University of New York at Stony Brook; [22]Médecine Intensive-Réanimation, Assistance Publique–Hôpitaux de Marseille, Hôpital Nord, France; [23]Third Department of Medicine, Sotiria General Hospital, and [24]Medical School, National and Kapodistrian University of Athens, Greece; [25]Merck & Co, Inc, Kenilworth, New Jersey; [26]Centro Investigacion Biomedica En Red de Enfermedades Respiratorias (CIBERES), Vall d'Hebron Barcelona Hospital Campus, Spain; [27]F2G, Ltd, Eccles, United Kingdom; [28]Medstar Washington Hospital Center, Washington, District of Columbia; [29]Talbot Advisors LLC, Anna Maria, Florida; [30]Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand; [31]Servei de Pneumologia, Hospital Clinic, Universitat de Barcelona, Institut De Investigacio Biomedica Agusti Pi i Sunyer, CIBERES, Spain; [32]Feinberg School of Medicine, Northwestern University, Chicago, Illinois; and [33]AP-HP, Medical and Infectious Diseases Intensive Care Unit, Bichat Hospital, Paris, and [34]UMR 1137 IAME, INSERM, Université Paris Diderot, France

Background. Randomized clinical trials (RCTs) in hospital-acquired and ventilator-associated bacterial pneumonia (HABP and VABP, respectively) are important for the evaluation of new antimicrobials. However, the heterogeneity in endpoints used in RCTs evaluating treatment of HABP/VABP may puzzle clinicians. The aim of this work was to reach a consensus on clinical endpoints to consider in future clinical trials evaluating antimicrobial treatment efficacy for HABP/VABP.

Methods. Twenty-six international experts from intensive care, infectious diseases, and the pharmaceutical industry were polled using the Delphi method.

Results. The panel recommended a hierarchical composite endpoint including, by priority order, (1) survival at day 28, (2) mechanical ventilation–free days through day 28, and (3) clinical cure between study days 7 and 10 for VABP; and (1) survival (day 28) and (2) clinical cure (days 7–10) for HABP. Clinical cure was defined as the combination of resolution of signs and symptoms present at enrollment and improvement or lack of progression of radiological signs. More than 70% of the experts agreed to assess survival and mechanical ventilation–free days though day 28, and clinical cure between day 7 and day 10 after treatment initiation. Finally, the hierarchical order of endpoint components was reached after 3 Delphi rounds (72% agreement).

Conclusions. We provide a multinational expert consensus on separate hierarchical composite endpoints for VABP and HABP, and on a definition of clinical cure that could be considered for use in future HABP/VABP clinical trials.

Keywords. hospital-acquired bacterial pneumonia; multinational consensus; Delphi method; hierarchical composite endpoint; clinical cure.

Hospital-acquired bacterial pneumonia (HABP) and ventilator-associated bacterial pneumonia (VABP) are the second most common nosocomial infections in adults and are leading causes of death among critical care infections [1, 2]. Consequently, indications for treatment of HABP and VABP are of considerable interest in the development of new drugs. Nine randomized clinical trials (RCTs) of antibiotic therapy for HABP and VABP are currently registered as recruiting patients (ClinicalTrials.

gov, accessed on 3 January 2019). However, a universal, well-accepted, and sensitive-to-change criterion for objective and reproducible assessment of drug benefit in HABP/VABP other than all-cause mortality (ACM) at a fixed time (usually 28 days) is still lacking [3, 4]. Appropriate endpoints for HABP/VABP trials have been a subject of debate for several years and remain controversial [3–6].

In the most recent US Food and Drug Administration (FDA) guidance for industry on studies evaluating antibiotics for HABP/VABP treatment, ACM at a fixed timepoint between 14 and 28 days after therapy start is the accepted primary endpoint [4]. In contrast, the most recent European Medicines Agency (EMA) guidelines (EMA/CHMP/351889/2013) recommend as the primary endpoint clinical outcome timed from therapy start, so that it occurs 7–14 days after the last possible day of treatment [5]. As shown recently, such a controversy has led to a significant heterogeneity among endpoints reported in RCTs, thereby confounding their results [7]. Based on this uncertainty, our aim was to achieve a consensus on the choice and definition of relevant and plausible primary nonmortality clinical endpoints to address antimicrobial efficacy in HABP/VABP that could be used as a component of future trials. Our approach used the Delphi method [8].

METHODS

A panel of 26 international experts from intensive care, infectious diseases, and the pharmaceutical industry were consulted using the Delphi method [8] from January 2016 to January 2017 (5 questionnaires). More than 70% of similar answers to a question were necessary to reach a consensus [8]. The iterative process followed is shown in Figure 1. More details about initial literature review, expert selection, and Delphi process are provided in the Supplementary Appendix.

RESULTS

Which Endpoint?
We first asked the panelists to rank 10 potential primary endpoints commonly reported in the literature [7] by assigning a score from 1 (worst) to 10 (best) according to their preference (Figure 1 and Table 1). According to 60% of the experts, clinical cure was the most desirable primary outcome; its mean ranking was 8.5 (standard deviation [SD], 2.7). Two other endpoints were highly rated: ACM and mechanical ventilation (MV)–free days, with mean rankings of 7 (SD, 2.8) and 6.5 (SD, 2.4), respectively (Table 1). Ranking of primary endpoints varied according to the specialty and the nationality of the expert: ACM was more frequently chosen as most desirable among non-European experts (55% vs 7%) and/or experts from industry (75% vs 18%).

Recognizing that several endpoints might be relevant simultaneously, 6 experts suggested using a composite endpoint and even a hierarchical composite endpoint to combine these items during round 1. This idea was approved by 88% of the panelists who confirmed the 3 previously mentioned items (ie, clinical
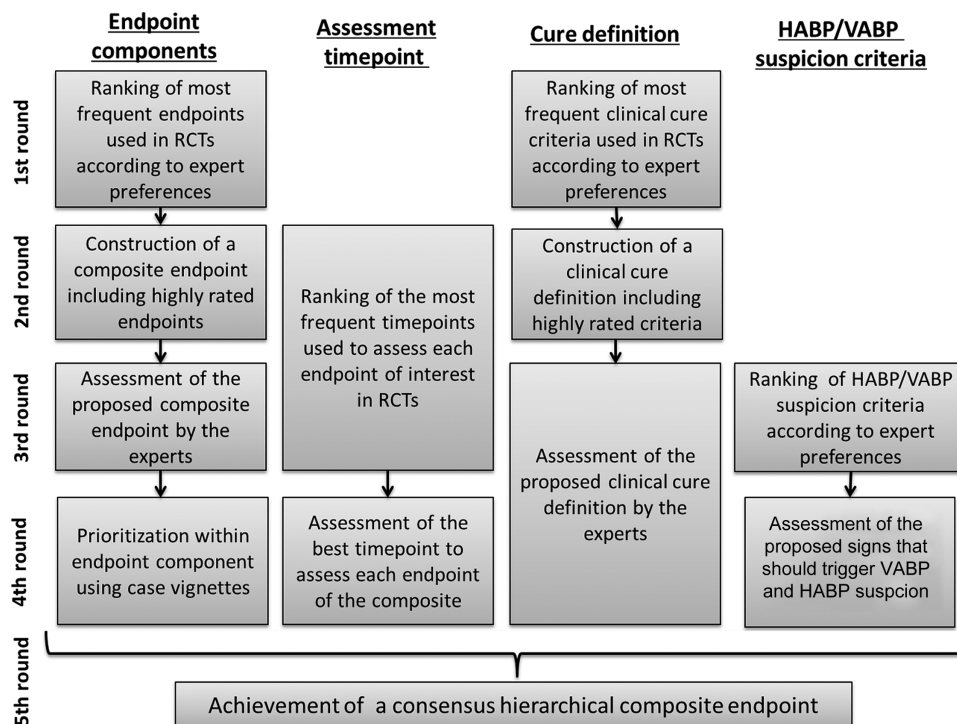

Figure 1. Iterative process followed to achieve the proposed hierarchical composite endpoint. Abbreviations: HABP, hospital-acquired bacterial pneumonia; RCT, randomized controlled trial; VABP, ventilator-associated bacterial pneumonia.

Table 1. Ranking of Endpoints Found in the Literature by the Expert Panel

Rank	Item	Mean (SD) Score
1	Clinical cure	8.5 (2.7)
2	ACM	7.0 (2.8)
3	MV-free days	6.5 (2.4)
4	Improvement in oxygenation parameters	5.8 (2.4)
5	No. of days before resolution	5.6 (2.4)
6	CPIS [9] decrease	5.2 (2.4)
7	Microbiological cure	5.1 (2.4)
8	Safety	3.8 (2.4)
9	PCT decrease	3.7 (3.1)
10	Acquisition of antimicrobial resistance	3.0 (2.3)

A score from 1 (worst) to 10 (best) was attributed to every criterion by each expert according to his or her preference. Results are given as mean (SD) score obtained by each endpoint.

Abbreviations: ACM, all-cause mortality; CPIS, Clinical Pulmonary Infection Score; MV, mechanical ventilation; PCT, procalcitonin; SD, standard deviation.

cure, ACM, and MV-free days) as the best components of this hierarchical composite endpoint.

Which Definition for Clinical Cure?

Once again, a ranking of the various clinical cure criteria found in the literature was first performed by the experts (Supplementary Table 1) [7]. Only 3 items obtained a mean score >3 (across the range 1–4) according to the experts: resolution of signs and symptoms of infection, improvement of oxygenation parameters, and no appearance of new signs of sepsis (Supplementary Table 1). Given these results, 96% of the experts agreed to combine resolution of signs and symptoms of infection, improvement of oxygenation parameters, and no appearance of new signs of sepsis together under the heading "resolution at the end of therapy of signs and symptoms present at enrollment" and to include this item in clinical cure definition (88% agreement). Strikingly, a low score was initially attributed to the improvement of chest radiography (Supplementary Table 1) because the experts thought that a lack of progression of radiographic images could be sufficient to qualify for cure. Finally, 72% of the experts agreed that resolution or lack of progression of radiological signs of pneumonia could be part of a clinical cure definition. Altogether, a consensus was reached on a clinical and radiological definition of cure: (1) clinical: resolution of signs and symptoms present at enrollment; (2) radiography: resolution or lack of progression of radiological signs of pneumonia.

To decrease the risk of incorrect pneumonia diagnosis that would impact clinical cure, we next aimed to specify which signs and symptoms should trigger the suspicion of VABP and HABP (Figure 1 and Supplementary Table 2). Interestingly, signs leading to VABP suspicion differed from those leading to HABP suspicion (Supplementary Table 2). For VABP, 4 signs were selected by the experts: worsening of gas exchange, hypotension and/or vasopressor requirements, fever or hypothermia, and purulent tracheal secretions (Supplementary Table 2). For HABP, the panelist added dyspnea and leukocytosis or leukopenia to the VABP suspicion criteria (Supplementary Table 2). When we asked the experts to confirm these choices, 88% and 73% of them agreed for VABP and HABP suspicion criteria, respectively.

Which Timepoint for Endpoint Assessment?

In parallel, the optimal time frame for endpoint assessment was selected: Expert choices were day 28 for ACM, the period from the diagnosis to day 28 for MV-free days, and the period between days 7 and 10 after treatment initiation for clinical cure. More details on expert choices regarding timepoints are provided in the Supplementary Results.

Consensus Hierarchy Between Components of a Possible Composite Endpoint

When considered together, the 3 endpoints of ACM at day 28, MV-free days from diagnosis to day 28, and clinical cure between study days 7 and 10 seemed sufficiently independent as to suggest a possible composite endpoint. The panel was thus asked to rank these 3 items according to their importance. ACM was the most important component for 72% of the experts, but the relative ranking of the 2 remaining components was more difficult to establish: 60% of the experts placed clinical cure in the second position, whereas 40% of them thought that MV-free days was more important for this rank. To address this question, we used 2 small case vignettes reflecting the possible situations and asked the experts which is most desirable: (1) a survivor at day 28, not cured 7–10 days after treatment initiation, and who required a short duration of MV before day 28; or (2) a survivor at day 28, clinically cured 7–10 days after treatment initiation and who required a long duration of MV before day 28. Using this strategy, 75% of the experts judged that the first situation with short duration of MV was the more desirable scenario, demonstrating that the panel valued MV-free days as more important than time to clinical cure.

Of note, some experts suggested differentiating HABP from VABP, and 85% of the panel agreed to a dual composite endpoint for nonventilated HABP consisting of ACM at day 28 and clinical cure 7–10 days after treatment initiation (Figure 2). Conversely, in cases of HABP requiring MV (ie, ventilated HABP), experts recommended adopting the same rules as for VABP.

Finally, some experts suggested a discrepancy in the direction of the composite endpoint that included ACM as a negative endpoint and MV-free days and clinical cure as positive ones. To cope with that situation, 96% of the experts agreed to revise this endpoint to "survival at day 28."

In conclusion, we reached a consensus providing separate hierarchical composite endpoints for VABP and HABP that assess by priority order: (1) day 28 survival; (2) any number of MV-free days from the diagnosis of pneumonia to day 28 (only for VABP); and (3) clinical cure measured between study days
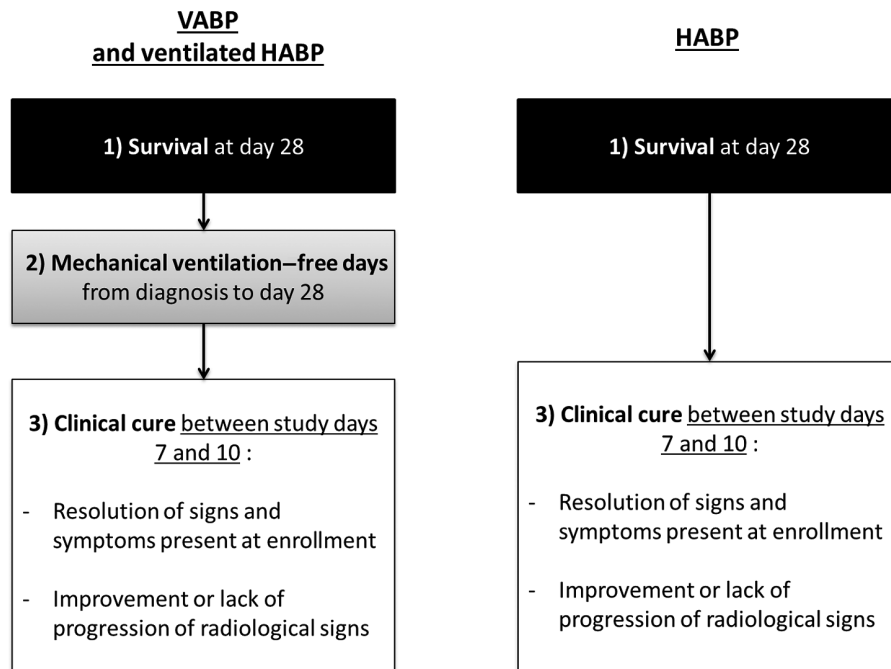
VABP and ventilated HABP

1) Survival at day 28

2) Mechanical ventilation–free days from diagnosis to day 28

3) Clinical cure between study days 7 and 10 :

- Resolution of signs and symptoms present at enrollment

- Improvement or lack of progression of radiological signs

HABP

1) Survival at day 28

3) Clinical cure between study days 7 and 10 :

- Resolution of signs and symptoms present at enrollment

- Improvement or lack of progression of radiological signs

Figure 2. Final hierarchical composite endpoint proposed. Abbreviations: HABP, hospital-acquired bacterial pneumonia; VABP, ventilator-associated bacterial pneumonia.

7 and 10 (Figure 2). As an example, we ranked the outcome of 4 representative patients according to 5 different primary endpoints including our hierarchical composite endpoint (Figure 3) and found good discriminatory ability of the latter.

DISCUSSION

Several studies are currently investigating the efficacy of new antimicrobials against multidrug-resistant pathogens or will do so in the near future. Given their frequency, HABP and VABP are key testing grounds for new agents against gram-negative pathogens, despite the lack of validated nonmortality, patient-centered, and physician-centered clinical endpoints, which would specifically assess the effectiveness of a new drug. Using the Delphi method to elicit opinions from a panel of 26 international experts, we have developed consensus clinical endpoints with the potential of the evaluation of antimicrobial treatment for HABP/VABP. This carefully implemented iterative process required a notable effort of the panel to obtain a hierarchical composite endpoint that includes survival, MV-free days, and clinical cure and should provide a better discrimination of patient outcomes.

Taken separately, each component of this composite endpoint has strengths and weaknesses. A high agreement was obtained regarding the inclusion of ACM (or survival) as an outcome measure. As suggested by the FDA, ACM offers numerous advantages: It is verifiable and reproducible, and it is the only unambiguously and entirely objective "patient-centered" outcome [3, 4, 10]. However, the EMA raised arguments against it, for the reason that a substantial part of ACM (or survival) is

not related to HABP/VABP in the intensive care unit (ICU) but rather to underlying disease [3, 5].

The second criterion selected by the experts was ventilator-free days. This patient-oriented endpoint is easy to measure but remains susceptible to unrelated events. While its impact on patient's benefit is unambiguous, the process around discontinuing MV includes a subjective judgment by the care provider.

Clinical cure, the last criterion included in our composite, is disease oriented [10]. While very useful (especially if mortality is low), it is prone to subjective clinical considerations and geographic or practice-setting variability. Acknowledging the lack of a validated clinical cure definition as an important drawback, we developed a consensus definition of this endpoint and then validated a list of signs and symptoms of HABP and/or VABP that should be resolved to qualify for cure.

Besides ACM, MV-free days and clinical cure, several other outcomes may have deserved to be included in the composite endpoint but were not suggested or chosen by the panelists. Among them, complications of the disease are of particular interest because they reflect how patients feel and function and because better tolerability may confer an advantage to a new treatment [11]. Clinical Pulmonary Infection Score acceptance was very variable, with 19% of the panel considering it a highly desirable outcome measure, and 30% judging it not useful at all [1, 2]. The experts also very readily rejected microbiological cure, probably because of the lack of correlation with patient outcome. Finally, some medium- to long-term patient-centered outcomes, such as quality of life or hospital discharge [12], were not suggested as they were deemed difficult to measure.

The consensus composite endpoint that we provide includes the 3 most significant outcomes according to the experts. As shown in the case vignettes in Figure 3, reporting of a composite may improve the assessment of patient's outcome compared to a single endpoint. However, an inherent limitation of composite endpoints is that it emphasizes each patient's first event, which is often the outcome of lesser importance. For example, a patient who was cured at day 7 but died during his ICU stay could be considered as having a better outcome than a second patient who was discharged alive from the ICU but was cured from his VABP later (Figure 3). Here, as previously suggested by Pocock et al [13] with the win ratio, and Evans et al [14] with the desirability of outcome ranking concept, we introduced a hierarchy to take the different priorities of the components into account and allow a better patient differentiation according to the desirability of the outcome. Those hierarchical composite endpoints were recently used in several studies comparing different antibiotic strategies in community-acquired pneumonia [15], intra-abdominal infections [16], or mixed infections caused by carbapenem-resistant Enterobacteriaceae [12]. However, statistical design/analysis must be anticipated [17] to operationalize the use of the rank-based composite endpoint [18]. While conventional composite could increase power by combining some or all variables included in the hierarchy as a primary endpoint, a rank-based composite (as suggested here) would test the components of the primary endpoint in a prespecified, hierarchical order with termination rules (depending on null hypothesis rejection) of the procedure at each step. In this case, multistate models or hierarchical nested-trial designs should be considered [17].

Several areas of uncertainty remain. First, the debate about the relative merits of utilizing clinical cure or ACM remains unresolved. The RCTs assessing antibiotic therapy for HABP and VABP currently reported as recruiting on ClinicalTrials.gov illustrate this uncertainty with ACM and clinical cure chosen as primary endpoint in two-thirds and one-third of cases, respectively. In our study, although the sample size may limit the meaning of the results, clinical cure seemed to be especially recommended by European experts, consistent with EMA recommendations [5], while others, including experts from the pharmaceutical industry, showed a preference for ACM, consistent with FDA guidance [4].

The timepoint to assess each component of the proposed composite endpoint was difficult to define. While a consensus was rapidly reached for ACM, probably because choosing a late timepoint (day 28) increases event number [11], it required 2 rounds for MV-free days. The highest difficulties were encountered in clinical cure, for which the experts were divided between early and late timepoints, requiring 3 Delphi rounds to reach a consensus about day 7 to day 10 after treatment
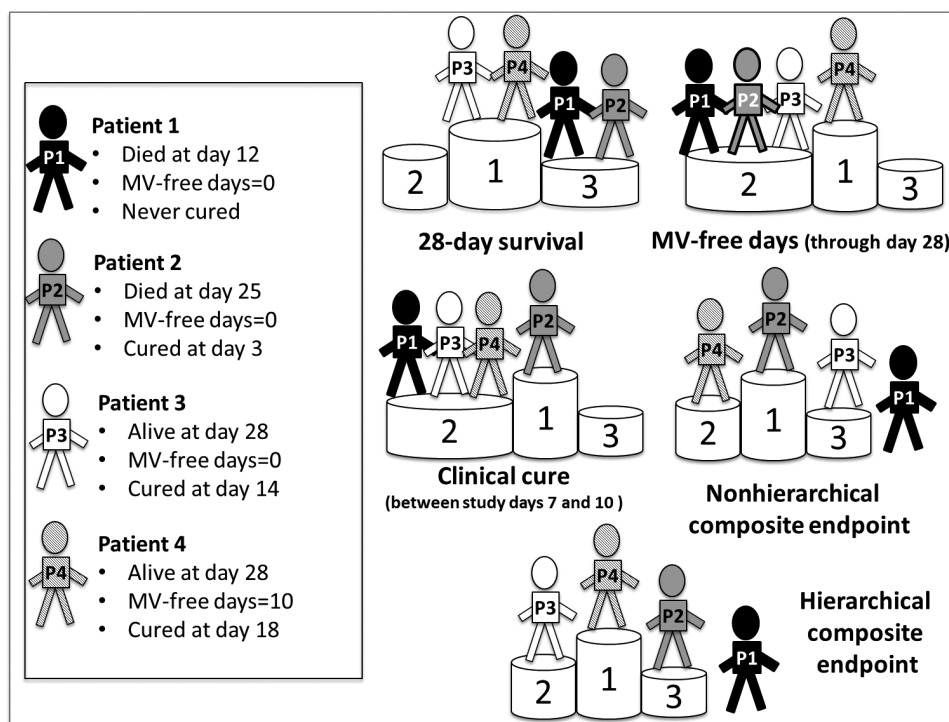

Figure 3. Ranking of desirability of the outcome of 4 representative patients (P1–P4) according to the chosen primary endpoint. Patient outcomes were ranked from most desirable (first position of the podium) to least desirable (third position of the podium or even outside the podium). Classical composite endpoint: no hierarchy between components (all-cause mortality [ACM] at day 28, MV-free day from diagnosis to day 28, and clinical cure within 7–10 days of study initiation); the earliest event goes first. Hierarchical composite endpoint: ACM at day 28 is first assessed. Then, MV-free day from diagnosis to day 28 is assessed, and finally, clinical cure within 7–10 days of study initiation. Abbreviation: MV, mechanical ventilation.

initiation, with only 74% of similar answers. Such uncertainty reflects the opaqueness of our understanding of "clinical cure." Of note, an early timepoint reflecting initial response to treatment was added beside the later timepoints used for ACM and MV-free days that may be more representative of global outcome. In addition, this timepoint was earlier than the one chosen for clinical cure evaluation by regulatory agencies (7–14 days after antibacterial treatment completion).

Whether treatment efficacy of HABP, ventilated HABP, and VABP could be studied together remains debated [11]. Although combining these entities is convenient for analyzing data (eg, by ensuring an adequate power), differences in outcome between these subcategories may preclude such pooling. In this work, while the experts considered HABP requiring MV as equivalent to VABP, the majority of them suggested a specific endpoint for nonventilated HABP. This separation is consistent with a recent statement from the Foundation of the National Institutes of Health [11]. Based on the analysis of 10 contemporary HABP/VABP clinical trials, the authors found that the ACM rates are similar for ventilated HABP and VABP, and higher than for nonventilated HABP, suggesting inclusion of ventilated HABP and VABP in the same clinical trial. Similarly, suspicion criteria for nonventilated HABP differed from those of the 2 other entities. Despite variability in expert opinion, a large majority of them agreed with 4 VABP suspicion criteria (gas exchange worsening, purulent tracheal secretions, hypotension and/or vasopressor requirements, and fever or hypothermia) and 6 HABP suspicion criteria. Notably, hypotension and/or vasopressor requirements were added to VABP suspicion criteria, whereas they are not included in "usual" criteria recommended by the Infectious Diseases Society of America and the American Thoracic Society [1]. These findings suggest that in case of severe ICU-acquired sepsis, VABP should be systematically suspected given its frequency. Conversely, leukocyte count was removed from VABP suspicion criteria, probably because of the numerous factors that could affect this inflammatory parameter in the ICU.

Some limitations of our novel endpoint have to be underlined. Consensus on the hierarchical ranking of the elements of the composite was difficult to obtain. While the experts gave the maximum weight to ACM, 3 Delphi rounds and the use of case vignettes were required to choose the more clinically meaningful outcome between MV-free days and clinical cure, with a final panel agreement limited to 75%. Such difficulties in ranking of the components demonstrate potential subjectivity in expert opinion, given the paucity of evidence, which was not explained by expert background. Furthermore, consensus-obtained timepoints for endpoint assessment were difficult to obtain, especially for clinical cure, which was a source of intense debate between clinicians and experts from the pharmaceutical industry. Another possibility would have

been to study the kinetics of disease progression instead of a binary cure/noncure at a fixed timepoint. In particular, multistate endpoints could be used to demonstrate a treatment effect over the complete treatment process instead of merely at the end of follow-up [17]. The choice of the statistical design (ie, superiority or noninferiority), which was not predetermined, may also change expert opinion about endpoints. Use of the composite endpoint in noninferiority trials will require assessment of its effect size relative to effective therapy. Concerning our panel, which included both industry and academic experts in the field, a third group composed of patient representatives would have been interesting. Including representatives from regulatory agencies was also considered but was not done to avoid already formed ideas or potential conflicts of interest. Finally, as for any new method, the use of such an endpoint may be initially limited by lack of familiarity among trialists.

In conclusion, judgment-based criteria for assessing antimicrobial efficacy in HABP/VABP are heterogenous. Therefore, a convergence on approaches for the evaluation of antibacterial drugs is urgently required [19]. We reached a consensus on a set of hierarchical composite endpoints for VABP and HABP using the Delphi method by drawing on the expertise of 26 international experts. A pilot validation of our proposed endpoint by use of data from an already completed clinical trial, which was conducted with the classic endpoints, should be performed first. Then, if validated, it could be considered for inclusion in future trials and, if further validated, may allow a pragmatic and homogenous assessment of antimicrobial efficacy in future HABP/VABP RCTs.

Supplementary Data

Notes

References

1. Kalil AC, Metersky ML, Klompas M, et al. Management of adults with hospital-acquired and ventilator-associated pneumonia: 2016 clinical practice guidelines by the Infectious Diseases Society of America and the American Thoracic Society. Clin Infect Dis **2016**; 63:e61–111.

2. Torres A, Niederman MS, Chastre J, et al. International ERS/ESICM/ESCMID/ ALAT guidelines for the management of hospital-acquired pneumonia and ventilator-associated pneumonia: guidelines for the management of hospital-acquired pneumonia (HAP)/ventilator-associated pneumonia (VAP) of the European Respiratory Society (ERS), European Society of Intensive Care Medicine (ESICM), European Society of Clinical Microbiology and Infectious Diseases (ESCMID) and Asociacion Latinoamericana del Torax (ALAT). Eur Respir J **2017**; 50. doi:10.1183/13993003.00582-2017.

3. Timsit JF, de Kraker MEA, Sommer H, et al; COMBACTE-NET Consortium. Appropriate endpoints for evaluation of new antibiotic therapies for severe infections: a perspective from COMBACTE's STAT-Net. Intensive Care Med **2017**; 43:1002–12.

4. US Food and Drug Administration Center for Drug Evaluation and Research. Guidance for industry hospital-acquired bacterial pneumonia and ventilator-associated bacterial pneumonia: developing drugs for treatment. XXX: US Department of Health and Human Services, **2014**.

5. European Medicines Agency. Addendum to the guideline on the evaluation of medicinal products indicated for treatment of bacterial infections. London, UK: EMA, **2013**.

6. Rex JH, Talbot GH, Goldberger MJ, et al. Progress in the fight against multidrug-resistant bacteria 2005-2016: modern noninferiority trial designs enable antibiotic development in advance of epidemic bacterial resistance. Clin Infect Dis **2017**; 65:141–6.

7. Weiss E, Essaied W, Adrie C, Zahar JR, Timsit JF. Treatment of severe hospital-acquired and ventilator-associated pneumonia: a systematic review of inclusion and judgment criteria used in randomized controlled trials. Crit Care **2017**; 21:162.

8. Okoli C, Pawlowski SD. The Delphi method as a research tool: an example, design considerations and applications. Inform Manag **2004**; 42:15–29.

9. Luna CM, Blanzaco D, Niederman MS, et al. Resolution of ventilator-associated pneumonia: prospective evaluation of the Clinical Pulmonary Infection Score as an early clinical predictor of outcome. Crit Care Med **2003**; 31:676–82.

10. de Grooth HJ, Parienti JJ, Oudemans-van Straaten HM. Should we rely on trials with disease- rather than patient-oriented endpoints? Intensive Care Med **2018**; 44:464–6.

11. Foundation for the National Institutes of Health Biomarkers Consortium HABP/ VABP Project Team. Considerations for clinical trial design for the study of hospital-acquired bacterial pneumonia and ventilator-associated bacterial pneumonia. For submission to docket FDA-2013-N-0556. Available at: https://www.regulations.gov/document?D=FDA-2010-D-0589-0027. Accessed 26 May 2017.

12. van Duin D, Lok JJ, Earley M, et al. Colistin vs. ceftazidime-avibactam in the treatment of infections due to carbapenem-resistant Enterobacteriaceae. Clin Infect Dis **2017**; 66:163–71.

13. Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. Eur Heart J **2012**; 33:176–82.

14. Evans SR, Rubin D, Follmann D, et al. Desirability of outcome ranking (DOOR) and response adjusted for duration of antibiotic risk (RADAR). Clin Infect Dis **2015**; 61:800–6.

15. Schweitzer VA, van Smeden M, Postma DF, Oosterheert JJ, Bonten MJM, van Werkhoven CH. Response adjusted for days of antibiotic risk (RADAR): evaluation of a novel method to compare strategies to optimize antibiotic use. Clin Microbiol Infect **2017**; 23:980–5.

16. Celestin AR, Odom SR, Angelidou K, et al. Novel method suggests global superiority of short-duration antibiotics for intra-abdominal infections. Clin Infect Dis **2017**; 65:1577–9.

17. de Kraker MEA, Sommer H, de Velde F, et al; COMBACTE-NET Consortium. Optimizing the design and analysis of clinical trials for antibacterials against multidrug-resistant organisms: a white paper from COMBACTE's STAT-Net. Clin Infect Dis **2018**; 67:1922–31.

18. Schaller SJ, Anstey M, Blobner M, et al; International Early SOMS-guided Mobilization Research Initiative. Early, goal-directed mobilisation in the surgical intensive care unit: a randomised controlled trial. Lancet **2016**; 388:1377–88.

19. European Medicines Agency. Tripartite meeting held between the PMDA, EMA, and FDA in Kyoto, Japan to discuss convergence on approaches for the evaluation of antibacterial drugs. **2017**. Available at: https://www.fda.gov/downloads/drugs/guidances/ucm234907.pdf. Accessed 15 February 2019.