BMJ Open

The efficacy of activated protein C for the treatment of sepsis: incorporating observational evidence with Bayesian approach

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-006524
Article Type:	Research
Date Submitted by the Author:	03-Sep-2014
Complete List of Authors:	Zhang, Zhongheng; Jinhua municipal central hospital, Department of critical care medicine
Primary Subject Heading:	Intensive care
Secondary Subject Heading:	Anaesthesia, Emergency medicine, Intensive care
Keywords:	Adult intensive & critical care < ANAESTHETICS, BACTERIOLOGY, INFECTIOUS DISEASES

The efficacy of activated protein C for the treatment of sepsis: incorporating observational evidence with Bayesian approach

Zhongheng ZHANG (MMed)

Affiliation: Department of critical care medicine, Jinhua municipal central hospital, Jinhua hospital of Zhejiang university, Zhejiang, P.R.China

Corresponding author: Zhongheng Zhang

Address: 351#, Mingyue Road, Jinhua, Zhejiang province, China, 321000

Phone number: 86-579-82552667

Email: zh_zhang1984@hotmail.com

Key words: Bayesian analysis; observational evidence, activated protein C, sepsis, septic shock

There are no conflicts of interest.

Abstract

Background: activated protein C (aPC) has been extensively studies for its efficacy on sepsis but results from randomized controlled trials (RCT) were disappointing. However, many observational studies suggest that aPC is effective in reducing mortality.

Objective: The present study aimed to combine observational evidence with RCTs by using Bayesian approach.

Data sources: Electronic databases including Pubmed, Cochrane Central Register of Contrlled Trials (CENTRAL), ISI Web of Science, EMBASE and EBSCO were searched from inception to January 2014.

Study eligibility: Randomized controlled trail (RCTs) and observational studies (OS) investigating the effectiveness of aPC on mortality reduction were included for analysis.

Participants: patients with sepsis.

Intervention: aPC

Synthesis methods: Observational evidence was incorporated into analysis by using power transform priors in Bayesian framework. Trial sequential analysis (TSA) was performed to quantify the reliability of data in meta-analysis of RCTs.

Main results: a total of 7 RCTs and 12 observational studies were included for analysis. There was no significant heterogeneity among included RCTs (I2=48.6%, p=0.07). The pooled OR for mortality from RCTs was 1.00 (95% CI: 0.84-1.19). In observational studies, the pooled OR for mortality with the use of aPC was 0.66 (95% CI: 0.57-0.75). The pooled treatment effect of aPC from RCTs could be changed by using different power transform priors derived from observational evidence. When observational evidence was used at its “face value”, the treatment effect of aPC was statistically significant in reducing mortality.

Conclusion: while RCT evidence showed no beneficial effect of aPC on sepsis, observational evidence showed significant treatment effect of aPC. By using power transform priors in Bayesian model, we explicitly demonstrated how RCT evidence could be changed by observational evidence.

Registration: The protocol for the current study was registered in PROSPERO (registration number: CRD42014009562).

Article summary

1. There is considerable disparity between observational and RCT evidence.

2. While observational evidence shows beneficial effect of aPC on mortality reduction, RCTs failed to identify any such treatment effect.

3. By using power transform priors in Bayesian model, we explicitly demonstrated how RCT evidence could be changed by observational evidence.

4. Strengths: the study employed Bayesian approach to explicitly demonstrate how the result of RCTs can be influenced by observational evidence.

5. Limitations: it is still unknown how to discount observational evidence, namely, how to assign a value to the power of prior.

Introduction

Treatment of sepsis or septic shock is a major challenge for clinicians in intensive care unit (ICU).(1, 2) Many strategies and drugs have been developed for their potential beneficial effect on clinical outcomes. Most famous interventions include the early goal directed therapy (EGDT) for early resuscitation of septic shock, protective ventilation strategy for sepsis-induced acute lung injury,(3) intensive dose renal replacement therapy for sepsis-induced acute kidney injury, and activated protein C for immunomodulation.(4) However, these interventions experienced a wax and wean of enthusiasm for their clinical utility. For instance, the EGDT has been a standard of care for septic shock resuscitation in the first 6 hours, which however is challenged by a recent large randomized controlled trial published in the New England Journal of medicine (NEJM).(5) The same situation occurred in the field of CRRT dose. In 2000, a landmark study by Ronco C and coworkers(6) demonstrated mortality reduction in patients treated with high dose CRRT. However, the study cannot be replicated in subsequent mega-trials and systematic reivew.(7)

Activated protein C is a drug with pleiotrophic biological effects and is thought to play an important role in the modulation of inflammatory response.(8) Early observational studies, as well as a large randomized controlled trial (RCT) demonstrated remarkable mortality reduction by using this drug.(9-11) The famous PROWESS trial has urged approval of this drug by the Food and Drug administration (FDA) for septic shock patients.(9) However, the beneficial effect of aPC cannot be replicated in subsequent RCTs.(12, 13) Several meta-analyses including one published in Cochrane library have consistently refute the effectiveness of aPC for septic patients, and now it has been withdrawn from the market by the company.(14, 15) Although RCT is considered to be the gold standard of the test of biological efficacy of certain intervention, it has been criticized for multiple limitations. RCT is not conducted in “real world” setting as reflected by its strict inclusion/exclusion criteria, performance in specialized centers, and complicated intervention protocol. In contrast, observational studies are considered to be performed in “real world” setting

that patients being studied are just as they are treated in practice. Thus, some authors have suggested that observational studies should be considered in evidence synthesis, particularly when the intervention or clinical condition is complicated. Our previous analysis also showed that there is significant difference in treatment effect size between RCTs and OS.

In the present study we performed evidence synthesis by incorporating evidence from observational studies, and the observational evidence was down-weighted across of a wide range. Bayesian analysis allowed such calculation by using observational evidence as the informative prior. The main purpose of the study is to examine how results derived from RCTs can be changed by assigning different degrees of skepticism to observational evidence.

Methods

Searching strategy and study selection

Electronic databases including Pubmed, Cochrane Central Register of Contrlled Trials (CENTRAL), ISI Web of Science, EMBASE and EBSCO were searched from inception to January 2014. Our core search consists of terms related to activated protein C and sepsis. References of systematic reviews were reviewed for identifying additional eligible articles.

Randomized controlled trail (RCTs) and observational studies (OS) investigating the effectiveness of aPC on mortality reduction were included for analysis. OS included: 1) cohort studies using multivariable analysis with aPC treatment as one of the covariates; 2) cohort studies using propensity analysis; 3) case-control studies; 4) both prospective and retrospective designs were considered eligible.

The following data were extracted from original articles: name of the first author, year of publication, sample size, number of death in each arm, total number of participants in each arm, major bleeding events in each arm, odds ratio of treatment versus non-treatment for mortality, the method used for covariate adjustment (propensity

score analysis, logistic regression model), and design of observational study (prospective vs. retrospective).

Included RCTs were assessed for their quality by using Delphi list. Publication bias were assessed using the Egger regression test and Begg rank correlation test. Contour enhanced funnel plot were depicted to visually assess the presence of publication bias.

Statistical analysis

Observational evidence was used as the informative prior in Bayesian analysis. The model involved power transformation of observational data likelihood as proposed by Chen and Ibrahim.(16) Full details of calculations and the WinBugs codes were described elsewhere.(17) Trial sequential analysis (TSA) was also performed to quantify the reliability of data in meta-analysis adjusting significance levels for sparse data and multiple testing on accumulating trials.(18) Statistical analysis was performed by using WinBUGS (Imperial College & MRC, UK) and Stata 12.0 (College Station, Texas 77845 USA). Trial sequential analyses were performed by using the software TSA version 0.9 Beta (Copenhagen Trial Unit, 2011).

Results

Our initial search identified a total of 531 distinct citations, and 456 of them were excluded immediately after inspection of the title and abstract (figure 1). The remaining 75 clinical studies were potentially eligible and were examined for full text. Fifty-six studies were excluded because: 1) eight studies used duplicated report; 2) 18 studies used inappropriate control arm; 3) 19 did not report mortality as the endpoint; and 4) 11 had inappropriate intervention. As a result, a total of 7 RCTs(9, 12, 13, 19-22) and 12 observational studies(10, 11, 23-32) were included for analysis.

Figure 2 shows the efficacy of aPC on mortality reduction, reported separately for RCTs and observational studies. Results were pooled by using conventional meta-analytic approach. Of the 7 RCTs, only the PROWESS study showed significant mortality reduction with aPC (OR: 0.74, 95% CI: 0.59-0.91),(9) and the other six studies failed to conclude a beneficial effect. There was no significant heterogeneity

among included RCTs (I^2=48.6%, p=0.07). The pooled OR for mortality was 1.00 (95% CI: 0.84-1.19). In contrast, 7 out of the 12 observational studies showed significant mortality reduction with the use of aPC; and the remaining five studies showed a trend towards better lower mortality rate in aPC group. The heterogeneity was statistically significant with an I^2 of 68.4% (p<0.001). The pooled OR for mortality with the use of aPC was 0.66 (95% CI: 0.57-0.75). Publication bias was identified for observational studies as reflected by the asymmetrically distributed component studies (figure 3).

The result of sequential trial analysis is shown in figure 4. Studies were displayed sequentially by their publication year from left to the right of the horizontal line. After publication of the first and second studies (PROWESS 2001 and rhAPC sepsis 2001), the Z score crossed the conventional significance boundary (Z=1.96) but did not cross the O'Brien-Fleming boundaries. With the publication of the study ADDRESS 2005, the Z-score reached and crossed the futility line, indicating no effect of the aPC for mortality reduction in septic patients.

Figure 5 shows the caterpillar plot of individual and pooled ORs for observational studies. The posterior distribution of individual OR was shrunken, as reflected by the narrower credible interval of study level estimates as compared to the observed estimates. For instance, the credible interval of OR in the study de Pont AC 2005 was 0.40-1.06, which was significantly wider than the observed confidence interval of 0.04-6.70 (figure 2). This was because each component study borrowed evidence from the overall effect by using Bayesian approach. The overall OR was 0.67 (credible interval: 0.56, 0.78).

Figure 6 shows the mean OR and 95% credible interval (CrI) for different power transformation priors to down-weight observational evidence on the risk of death with aPC. To the left of the figure when alpha took negligible values, the observational evidence was totally discounted and the mean OR was 1, which was consistent with the pooled result from RCTs. Increasing weight was assigned to observational evidence with increasing alpha values. We could see from the figure that the upper limit of CrI crossed the reference line. When observational evidence was combined at

its face value (alpha=1), the aPC group showed significant mortality reduction as compared with the control group.

Discussion

Key findings of the present analysis are 1) aPC appears to be able to reduce mortality rate when evidence is pooled from observational studies, and the results are consistent by using conventional Bayesian approaches; 2) RCTs failed to identify any beneficial effect of aPC; 3) observational evidence, when discounted by different power transformation priors, can alter the conclusion derived from RCTs. 4) With trial sequential analysis, the positive result (significant beneficial effect of aPC) as shown in the PROWESS study should be interpreted with caution.

One explanation for the positive findings in observational studies is the publication bias as shown in figure 3. The funnel contour plot showed that most observational studies located in the region with p<1%, indicating that the asymmetrical distribution was more likely due to publication bias. It is not surprising that observational studies are more subject to publication bias in that they are less likely to be registered a priori.(33) In contrast, RCTs are usually registered and there are many online registration sites.(34) The value of observational studies is usually discounted in evidence synthesis, and the conventional view is that observational evidence can only serve as hypothesis-generating. In such context, if the finding of an observational study is neutral, it will be less interesting to readers and journals, making it less likely to be published. In contrast, because RCTs are always registered and requires large amount of cost and other resources, their results even when negative can be published and is equally important to those with positive findings.

Activated protein C (aPC) for the treatment of sepsis is a good example illustrating the importance of using sequential trial analysis in evidence synthesis. aPC was approved by the food and drug administration after publication of PROWESS trial, which seemed too hasty when viewed retrospectively. Although the initial trial was

positive at conventional significance level of p=0.05 (Z=1.96), it was subject to repeated measurement error. This problem can be addressed by using adjusted alpha level. In sequential trial analysis, this is achieved by using alpha-spending function and constructing the O'Brien-Fleming boundaries. If sequential trial analysis was performed at the conclusion of PROWESS trial, the approval of aPC for sepsis would not be so hasty. Someone argued that the disparity between PROWESS trial and subsequent trials such as PROWESS-SHOCK could be explained by the heterogeneity of enrolled subjects.(35) However, we propose that other than heterogeneity, the exaggerated type Ⅰ error with repeated measurement may partly explain the spurious positive result.

RCT is considered as the gold standard for clinical practice and the evidence derived from such design is at the top of the evidence pyramid. However, RCT is not without shortcomings. The biggest problem is that RCT is usually conducted in non-real world setting, that is, it is always performed in specialized academic centers with strict inclusion/exclusion criteria. For instance, in the Dhainaut 2009 study there was a long list of exclusion criteria, including expected surgical procedure in the next 3 days, platelet count<30,000/nm3, receiving therapeutic heparin, moribund, withdrawn from aggressive management by patients' family, and pregnant or breast feeding. Such strict exclusion criteria would exclude most of patients with septic shock. Therefore, it appears unfair to treat our septic shock patients based on evidence derived from a minority of the population. In this situation, observational studies have its advantage in testing the clinical effectiveness of aPC on mortality reduction.(36, 37) Observational study included wider range of patients with septic shock and the setting is just like what we will encounter in routine clinical practice. Therefore, the observational evidence cannot be simply ignored in evidence synthesis for decision making. Since there was no consensus on how to combine observational evidence with RCTs, we discounted observational evidence with power transform priors taking advantage of the flexibility in Bayesian modeling.(16) In this model, we found that the treatment effect of aPC increased with more weight assigning to observational

evidence (figure 6). A value of 0 for alpha implies that the observational evidence is ignored, and a value of 1 for alpha means that observational evidence is accepted at its "face value". This approach gives a full picture of how pooled evidence can be altered by observational studies, by explicitly showing the power transform priors.

In aggregate, our study demonstrates that there is considerable disparity between observational and RCT evidence. While observational evidence shows beneficial effect of aPC on mortality reduction, RCTs failed to identify any such treatment effect.

Figure legends

Figure 1. Flow chart of study selection.

Figure 2. Forest plots showing the efficacy of aPC on mortality reduction, reported separately for RCTs and observational studies. Results were pooled by using conventional meta-analytic approach.

Figure 3. Contour funnel plots showing the publication bias in RCTs and observational studies. Publication bias was identified for observational studies as reflected by the asymmetrically distributed component studies.

Figure 4. Sequential trial analysis showing that the Z-score crossed the futility line after the study ADRESS 2005. Parameters used for the creation of boundaries were: type: Two-sided; type 1 Error: 5.0%; alpha spending: O'Brien-Fleming; information axis: sample size; power: 80.0%; effect type intervention: RRR User Defined (21.25%); heterogeneity correction: user defined (0.5%).

Figure 5. Caterpillar plot of individual and pooled ORs for observational studies. The study level estimates were shrunken as compared to those obtained by estimating each study in isolation (typically drawn in forest plot), because the Bayesian estimates borrow information/strength from each other.

Figure 6. Mean OR and 95% credible interval (CrI) for different power transformation priors to down-weight observational evidence on the risk of death with aPC.

Data sharing: No additional data available.

Competing Interests: None

Reference

1. Schorr CA, Zanotti S, Dellinger RP. Severe sepsis and septic shock: management and performance improvement. Virulence. 2014;5(1):190-9.

2. Dellinger RP, Levy MM, Rhodes A, Annane D, Gerlach H, Opal SM, et al. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012. Critical care medicine. 2013;41(2):580-637.

3. MacIntyre N. Ventilatory Management of ALI/ARDS. Seminars in respiratory and critical care medicine. 2006;27(4):396-403.

4. Rivers E, Nguyen B, Havstad S, Ressler J, Muzzin A, Knoblich B, et al. Early goal-directed therapy in the treatment of severe sepsis and septic shock. The New England journal of medicine. 2001;345(19):1368-77.

5. Pro CI, Yealy DM, Kellum JA, Huang DT, Barnato AE, Weissfeld LA, et al. A randomized trial of protocol-based care for early septic shock. The New England journal of medicine. 2014;370(18):1683-93.

6. Ronco C, Bellomo R, Homel P, Brendolan A, Dan M, Piccinni P, et al. Effects of different doses in continuous veno-venous haemofiltration on outcomes of acute renal failure: a prospective randomised trial. Lancet. 2000;356(9223):26-30.

7. Zhang Z, Xu X, Zhu H. Intensive- vs less-intensive-dose continuous renal replacement therapy for the intensive care unit-related acute kidney injury: a meta-analysis and systematic review. Journal of critical care. 2010;25(4):595-600.

8. Christiaans SC, Wagener BM, Esmon CT, Pittet JF. Protein C and acute inflammation: a clinical and biological perspective. American journal of physiology Lung cellular and molecular physiology. 2013;305(7):L455-66.

9. Bernard GR, Vincent JL, Laterre PF, LaRosa SP, Dhainaut JF, Lopez-Rodriguez A, et al. Efficacy and safety of recombinant human activated protein C for severe sepsis. The New England journal of medicine. 2001;344(10):699-709.

10. Dhainaut JF, Payet S, Vallet B, Franca LR, Annane D, Bollaert PE, et al. Cost-effectiveness of activated protein C in real-life clinical practice. Critical care (London, England). 2007;11(5):R99.

11. Kubler A, Mayzner-Zawadzka E, Durek G, Gaszynski W, Karpel E, Mikaszewska-Sokolewicz M, et al. Results of severe sepsis treatment program using recombinant human activated protein C in Poland. Medical science monitor : international medical journal of experimental and clinical research. 2006;12(3):CR107-12.

12. Annane D, Timsit JF, Megarbane B, Martin C, Misset B, Mourvillier B, et al. Recombinant human activated protein C for adults with septic shock: a randomized controlled trial. American journal of respiratory and critical care medicine. 2013;187(10):1091-7.

13. Ranieri VM, Thompson BT, Barie PS, Dhainaut JF, Douglas IS, Finfer S, et al. Drotrecogin alfa (activated) in adults with septic shock. The New England journal of medicine. 2012;366(22):2055-64.

14. Lai PS, Matteau A, Iddriss A, Hawes JC, Ranieri V, Thompson BT. An updated meta-analysis to understand the variable efficacy of drotrecogin alfa (activated) in severe sepsis and septic shock. Minerva anestesiologica. 2013;79(1):33-43.

15. Marti-Carvajal AJ, Sola I, Gluud C, Lathyris D, Cardona AF. Human recombinant protein C for severe sepsis and septic shock in adult and paediatric patients. The Cochrane database of systematic reviews. 2012;12:CD004388.

16. Chen MH, Ibrahim JG. Power prior distributions for regression models. statistical science.

2000;15(1):46-60.

17. Zhang Z. Recombinant human activated protein C for the treatment of severe sepsis and septic shock: a study protocol for incorporating observational evidence using a Bayesian approach. BMJ Open. 2014;4(7):e005622.

18. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. Journal of clinical epidemiology. 2008;61(1):64-75.

19. Bernard GR, Ely EW, Wright TJ, Fraiz J, Stasek JE, Jr., Russell JA, et al. Safety and dose relationship of recombinant human activated protein C for coagulopathy in severe sepsis. Critical care medicine. 2001;29(11):2051-9.

20. Abraham E, Laterre PF, Garg R, Levy H, Talwar D, Trzaskoma BL, et al. Drotrecogin alfa (activated) for adults with severe sepsis and a low risk of death. The New England journal of medicine. 2005;353(13):1332-41.

21. Nadel S, Goldstein B, Williams MD, Dalton H, Peters M, Macias WL, et al. Drotrecogin alfa (activated) in children with severe sepsis: a multicentre phase III randomised controlled trial. Lancet. 2007;369(9564):836-43.

22. Dhainaut JF, Antonelli M, Wright P, Desachy A, Reignier J, Lavoue S, et al. Extended drotrecogin alfa (activated) treatment in patients with prolonged septic shock. Intensive care medicine. 2009;35(7):1187-95.

23. de Pont AC, Bakhtiari K, Hutten BA, de Jonge E, Vroom MB, Meijers JC, et al. Recombinant human activated protein C resets thrombin generation in patients with severe sepsis - a case control study. Critical care (London, England). 2005;9(5):R490-7.

24. Bertolini G, Rossi C, Anghileri A, Livigni S, Addis A, Poole D. Use of Drotrecogin alfa (activated) in Italian intensive care units: the results of a nationwide survey. Intensive care medicine. 2007;33(3):426-34.

25. Rowan KM, Welch CA, North E, Harrison DA. Drotrecogin alfa (activated): real-life use and outcomes for the UK. Critical care (London, England). 2008;12(2):R58.

26. Vincent JL, Laterre PF, Decruyenaere J, Spapen H, Raemaekers J, Damas F, et al. A registry of patients treated with drotrecogin alfa (activated) in Belgian intensive care units--an observational study. Acta clinica Belgica. 2008;63(1):25-30.

27. Ferrer R, Artigas A, Suarez D, Palencia E, Levy MM, Arenzana A, et al. Effectiveness of treatments for severe sepsis: a prospective, multicenter, observational study. American journal of respiratory and critical care medicine. 2009;180(9):861-6.

28. Martin G, Brunkhorst FM, Janes JM, Reinhart K, Sundin DP, Garnett K, et al. The international PROGRESS registry of patients with severe sepsis: drotrecogin alfa (activated) use and patient outcomes. Critical care (London, England). 2009;13(3):R103.

29. Lindenauer PK, Rothberg MB, Nathanson BH, Pekow PS, Steingrub JS. Activated protein C and hospital mortality in septic shock: a propensity-matched analysis. Critical care medicine. 2010;38(4):1101-7.

30. Sadaka F, O'Brien J, Migneron M, Stortz J, Vanston A, Taylor RW. Activated protein C in septic shock: a propensity-matched analysis. Critical care (London, England). 2011;15(2):R89.

31. Casserly B, Gerlach H, Phillips GS, Marshall JC, Lemeshow S, Levy MM. Evaluating the use of recombinant human activated protein C in adult severe sepsis: results of the Surviving Sepsis Campaign. Critical care medicine. 2012;40(5):1417-26.

32. Rimmer E, Kumar A, Doucette S, Marshall J, Dial S, Gurka D, et al. Activated protein C and septic

shock: a propensity-matched cohort study*. Critical care medicine. 2012;40(11):2974-81.

33. Onukwugha E. Improving confidence in observational studies : should statistical analysis plans be made publicly available? PharmacoEconomics. 2013;31(3):177-9.

34. Anand V, Scales DC, Parshuram CS, Kavanagh BP. Registration and design alterations of clinical trials in critical care: a cross-sectional observational study. Intensive care medicine. 2014;40(5):700-22.

35. Kalil AC, Florescu DF. Severe sepsis: are PROWESS and PROWESS-SHOCK trials comparable? A clinical and statistical heterogeneity analysis. Critical care (London, England). 2013;17(4):167.

36. Albert RK. "Lies, damned lies ..." and observational studies in comparative effectiveness research. Am J Respir Crit Care Med. 2013;187(11):1173-7.

37. Nallamothu BK, Hayward RA, Bates ER. Beyond the randomized clinical trial: the role of effectiveness studies in evaluating cardiovascular therapies. Circulation. 2008;118(12):1294-303.

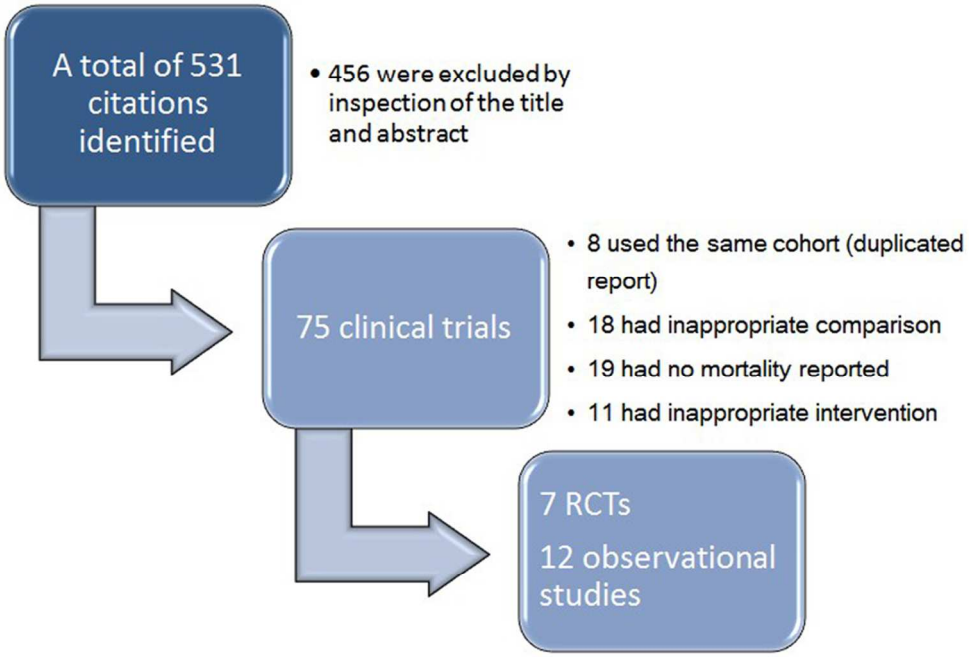

Figure 1. Flow chart of study selection.
84x58mm (300 x 300 DPI)

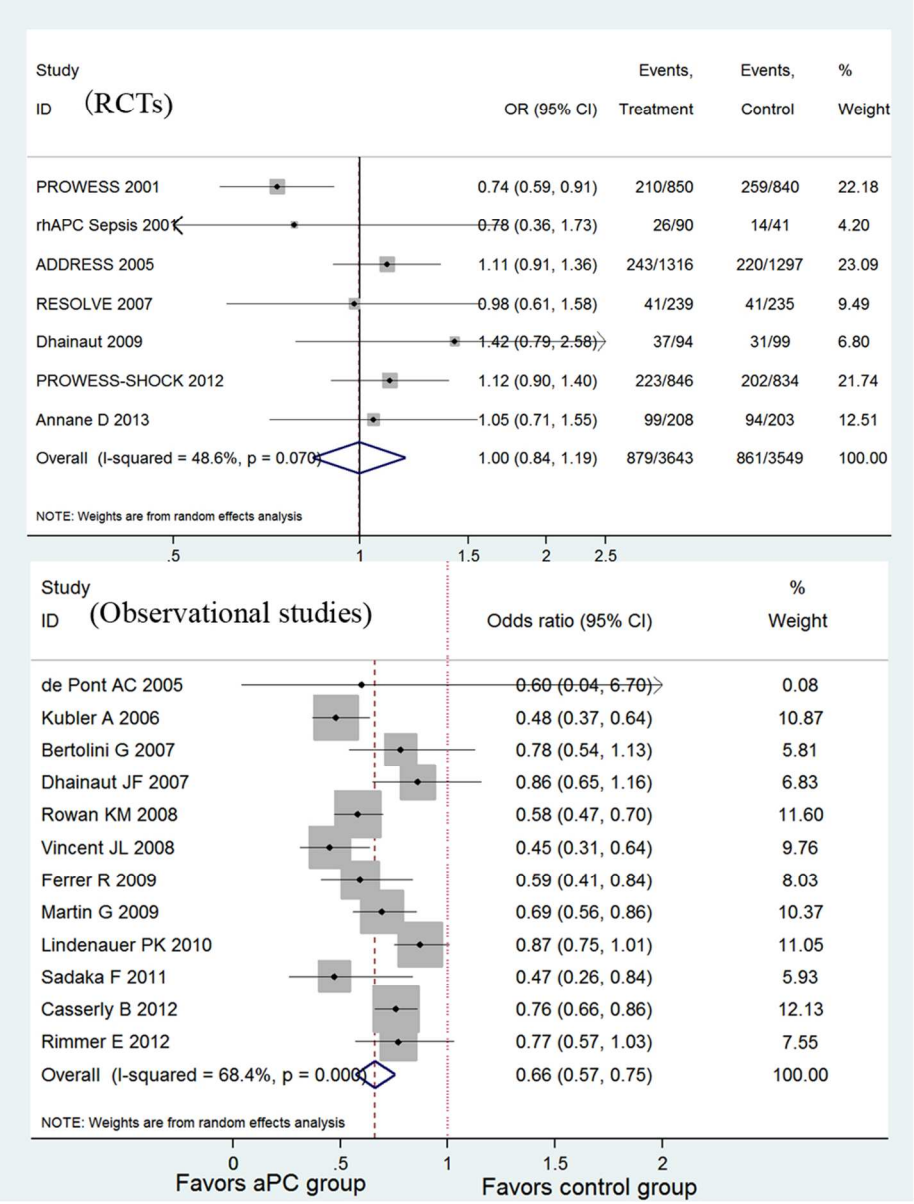

Figure 2. Forest plots showing the efficacy of aPC on mortality reduction, reported separately for RCTs and observational studies. Results were pooled by using conventional meta-analytic approach.
84x111mm (300 x 300 DPI)

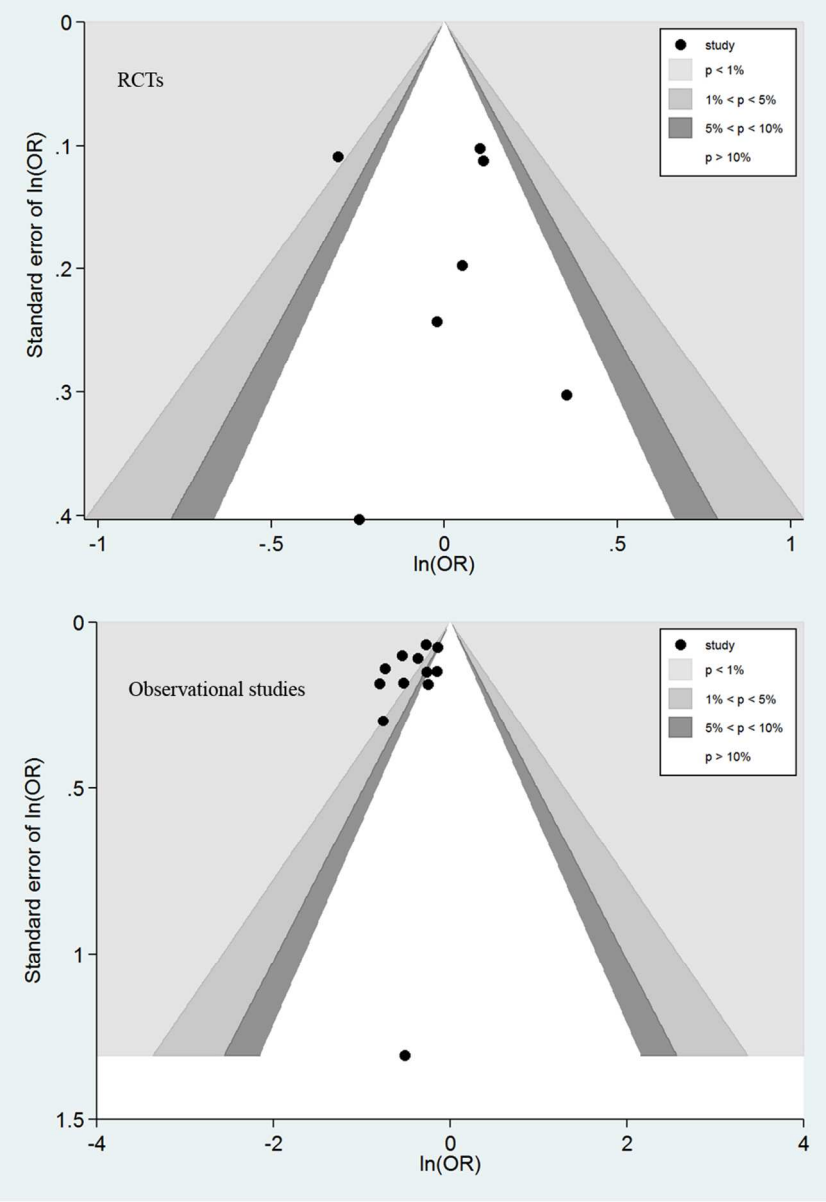

Figure 3. Contour funnel plots showing the publication bias in RCTs and observational studies. Publication bias was identified for observational studies as reflected by the asymmetrically distributed component studies.
84x123mm (300 x 300 DPI)

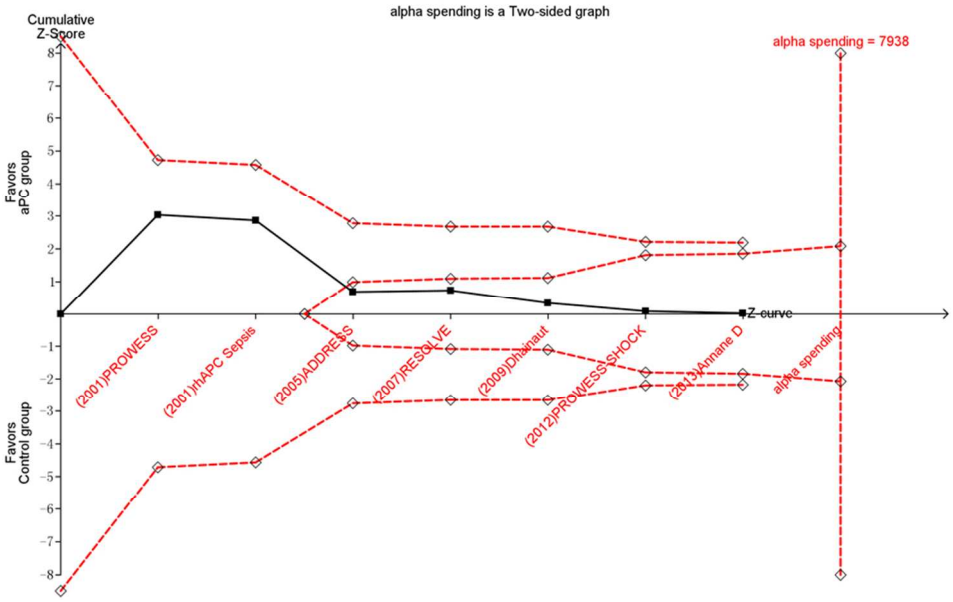

Figure 4. Sequential trial analysis showing that the Z-score crossed the futility line after the study ADRESS 2005. Parameters used for the creation of boundaries were: type: Two-sided; type 1 Error: 5.0%; alpha spending: O'Brien-Fleming; information axis: sample size; power: 80.0%; effect type intervention: RRR User Defined (21.25%); heterogeneity correction: user defined (0.5%).
84x51mm (300 x 300 DPI)

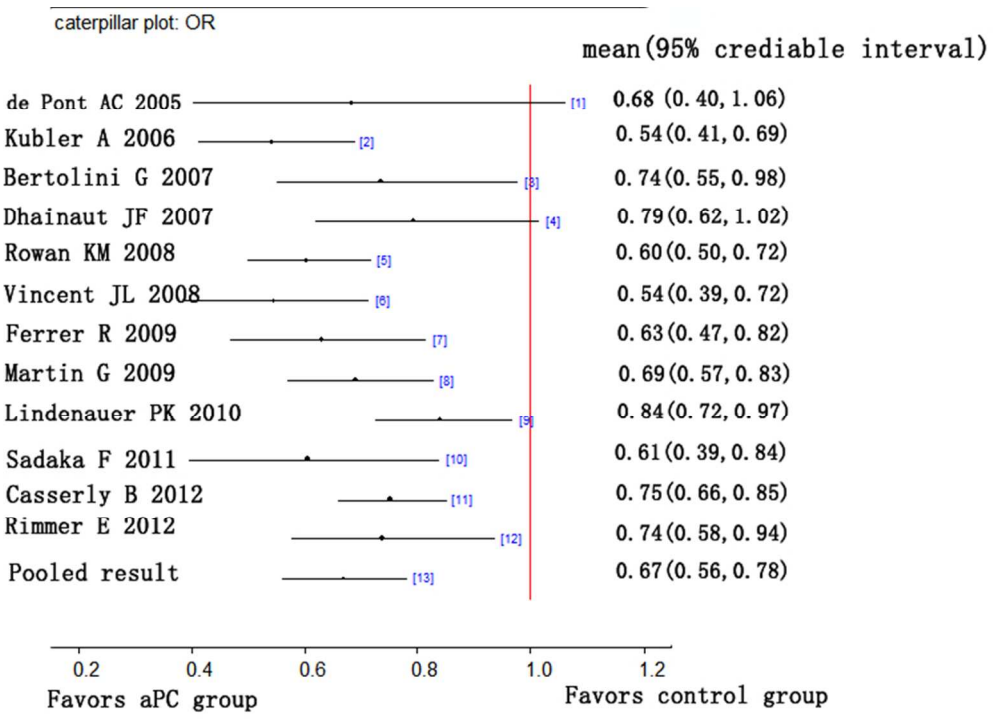

Figure 5. Caterpillar plot of individual and pooled ORs for observational studies. The study level estimates were shrunken as compared to those obtained by estimating each study in isolation (typically drawn in forest plot), because the Bayesian estimates borrow information/strength from each other.
127x90mm (150 x 150 DPI)

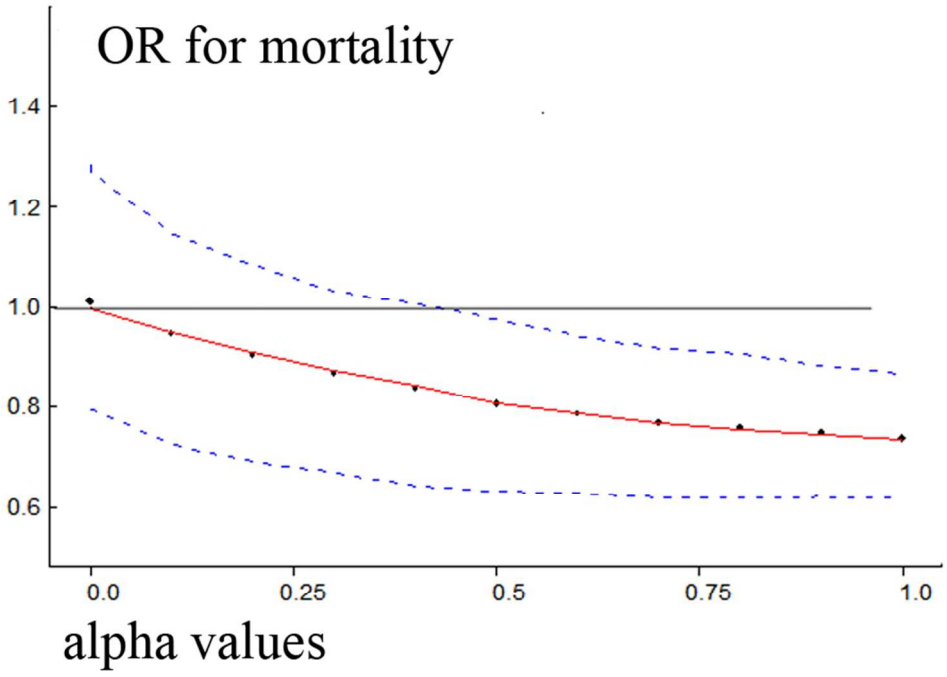

Figure 6. Mean OR and 95% credible interval (CrI) for different power transformation priors to down-weight observational evidence on the risk of death with aPC.
159x109mm (150 x 150 DPI)

BMJ Open

The efficacy of activated protein C for the treatment of sepsis: incorporating observational evidence with Bayesian approach

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-006524.R1
Article Type:	Research
Date Submitted by the Author:	05-Nov-2014
Complete List of Authors:	Zhang, Zhongheng; Jinhua municipal central hospital, Department of critical care medicine
Primary Subject Heading:	Intensive care
Secondary Subject Heading:	Anaesthesia, Emergency medicine, Intensive care
Keywords:	Adult intensive & critical care < ANAESTHETICS, BACTERIOLOGY, INFECTIOUS DISEASES

The efficacy of activated protein C for the treatment of sepsis: incorporating observational evidence with Bayesian approach

Zhongheng ZHANG (MMed)

Affiliation: Department of critical care medicine, Jinhua municipal central hospital, Jinhua hospital of Zhejiang university, Zhejiang, P.R.China

Corresponding author: Zhongheng Zhang

Address: 351#, Mingyue Road, Jinhua, Zhejiang province, China, 321000

Phone number: 86-579-82552667

Email: zh_zhang1984@hotmail.com

Key words: Bayesian analysis; observational evidence, activated protein C, sepsis, septic shock

There are no conflicts of interest.

Abstract

Objective: The present study aimed to combine observational evidence with RCTs by using Bayesian approach.

Data sources: Electronic databases including Pubmed, Cochrane Central Register of Contrlled Trials (CENTRAL), ISI Web of Science, EMBASE and EBSCO were searched from inception to January 2014.

Study eligibility: Randomized controlled trail (RCTs) and observational studies (OS) investigating the effectiveness of aPC on mortality reduction were included for analysis.

Participants: patients with sepsis.

Intervention: aPC

Synthesis methods: Observational evidence was incorporated into analysis by using power transform priors in Bayesian framework. Trial sequential analysis (TSA) was performed to quantify the reliability of data in meta-analysis of RCTs.

Main results: a total of 7 RCTs and 12 observational studies were included for analysis. There was moderate heterogeneity among included RCTs (I^2=48.6%, p=0.07). The pooled OR for mortality from RCTs was 1.00 (95% CI: 0.84-1.19). In observational studies, there was potential publication bias as indicated by funnel plot and the pooled OR for mortality with the use of aPC was 0.67 (95% CI: 0.62-0.72). Pooled effects sizes of RCTs were changed by using different power transform priors derived from observational evidence. When observational evidence was used at its "face value", the treatment effect of aPC was statistically significant in reducing mortality.

Conclusion: while RCT evidence showed no beneficial effect of aPC on sepsis, observational evidence showed significant treatment effect of aPC. By using power transform priors in Bayesian model, we explicitly demonstrated how RCT evidence could be changed by observational evidence.

Registration: The protocol for the current study was registered in PROSPERO (registration number: CRD42014009562).

Article summary

1. There is considerable disparity between observational and RCT evidence.

2. While observational evidence shows beneficial effect of aPC on mortality reduction, RCTs failed to identify any such treatment effect.

3. By using power transform priors in Bayesian model, we explicitly demonstrated how RCT evidence could be changed by observational evidence.

4. Strengths: the study employed Bayesian approach to explicitly demonstrate how the result of RCTs can be influenced by observational evidence.

5. Limitations: it is still unknown how to discount observational evidence, namely, how to assign a value to the power of prior. The most appropriate prior will vary from study to study.

Introduction

Treatment of sepsis or septic shock is a major challenge for clinicians in intensive care unit (ICU).(1, 2) Many strategies and drugs have been developed for their potential beneficial effects on clinical outcomes. Well-known interventions include the early goal directed therapy (EGDT) for early resuscitation of septic shock, protective ventilation strategy for sepsis-induced acute lung injury,(3) intensive dose renal replacement therapy for sepsis-induced acute kidney injury, and activated protein C for immunomodulation.(4) However, these interventions experienced a wax and wean of enthusiasm for their clinical utility. For instance, the EGDT has been a standard of care for septic shock resuscitation in the first 6 hours, which however is challenged by a recent large randomized controlled trial published in the New England Journal of medicine (NEJM).(5) This RCT was done 10 years after the original landmark EGDT on sepsis trial so it is a totally different time frame and different current practice. The same situation occurred in the field of CRRT dose. In 2000, a landmark study by Ronco C and coworkers(6) demonstrated mortality reduction in patients treated with high dose CRRT. However, the study could not be replicated in subsequent mega-trials and systematic reivew.(7)

Activated protein C is a drug with pleiotrophic biological effects and is thought to play an important role in the modulation of inflammatory response.(8) Early observational studies, as well as a large randomized controlled trial (RCT) demonstrated remarkable mortality reduction by using this drug.(9-11) The well-known PROWESS trial has urged approval of this drug by the Food and Drug administration (FDA) for septic shock patients.(9) However, the beneficial effect of aPC cannot be replicated in subsequent RCTs (12, 13). Several meta-analyses including one published in Cochrane library have consistently refuted the effectiveness of aPC for septic patients. As a result, it was withdrawn from the market (14, 15). Although RCTs are considered to be the gold standard for testing treatment efficacy, they have limitations. RCTs are often not conducted in “real world” settings as reflected by strict inclusion/exclusion criteria, performance in specialized centers,

and complicated intervention protocol. In contrast, observational studies are often performed in "real world" setting that patients enrolled in studies are just as they are treated in practice. Thus, some authors have suggested that observational studies should be considered in evidence synthesis, particularly when the intervention or clinical condition is complicated. Our previous analysis also showed that there is significant difference in treatment effect sizes between RCTs and OS (16).

In the present study we performed evidence synthesis by incorporating evidence from observational studies, and the observational evidence was down-weighted by using alphas ranging from 0 to 1. Bayesian analysis allowed such calculation by using observational evidence as the informative prior. The main purpose of the study is to examine how results derived from RCTs can be changed by assigning different degrees of skepticism to observational evidence. Another purpose of the study was to perform trial sequential analysis (TSA) to quantify the reliability of data in meta-analysis adjusting significance levels for sparse data and multiple testing on accumulating trials.

Methods

Amendment to the protocol

The study protocol has been published previously and amendment to the protocol was made during data analysis (17). The protocol for the current study was registered in PROSPERO (registration number: CRD42014009562). Herein, we explicitly listed the amendment to the protocol.

1) Quality assessment was not performed in the present analysis because the quality has been well described in a previous Cochrane systematic review (15).
2) Sensitivity analysis by excluding poor quality studies was not performed because the present study was aimed to explicitly display how the evidence derived from RCT could be modified by observational evidence. Sensitivity analysis of this kind belonged to the realm of systematic review involving only RCTs.

Searching strategy and study selection

Electronic databases including Pubmed, Cochrane Central Register of Contrlled Trials (CENTRAL), ISI Web of Science, EMBASE and EBSCO were searched from inception to January 2014. Our core search consists of terms related to activated protein C and sepsis. References of systematic reviews were reviewed for identifying additional eligible articles.

Randomized controlled trail (RCTs) and observational studies (OS) investigating the effectiveness of aPC on mortality reduction were included for analysis. OS included: 1) cohort studies using multivariable analysis with aPC treatment as one of the covariates; 2) cohort studies using propensity analysis; 3) case-control studies; 4) both prospective and retrospective designs were considered eligible.

The following data were extracted from original articles: name of the first author, year of publication, sample size, number of death in each arm, total number of participants in each arm, major bleeding events in each arm, odds ratio of treatment versus non-treatment for mortality, the method used for covariate adjustment (propensity score analysis, logistic regression model), and design of observational study (prospective vs. retrospective).

Publication bias was assessed using the Egger regression test and Begg rank correlation test. Contour enhanced funnel plot were depicted to visually assess the presence of publication bias.

Statistical analysis

Observational evidence was used as the informative prior in Bayesian analysis. The model involved power transformation of observational data likelihood as proposed by Chen and Ibrahim.(18) Full details of calculations and the WinBugs codes were described elsewhere.(17) Trial sequential analysis (TSA) was also performed to quantify the reliability of data in meta-analysis adjusting significance levels for sparse data and multiple testing on accumulating trials.(19) Statistical analysis was performed by using WinBUGS (Imperial College & MRC, UK) and Stata 12.0 (College Station, Texas 77845 USA).

Trial sequential monitoring boundaries were employed to control the risks for type I

and II errors and to indicate whether additional trials are needed. The information size calculation requires the mortality rate in the control group and the minimal effect size for the intervention. We predefined that the mortality in the control group is 30%, and the intervention is able to reduce the relative risk by 10%. The conventional α and β are 0.05 and 0.2, respectively. Meta-analysis will be updated by adding component studies sequentially in the order of publication. β -spending function was constructed to indicate futility of intervention. Trial sequential analysis was performed by using the software TSA version 0.9 Beta (Copenhagen Trial Unit, 2011).

Results

Our initial search identified a total of 531 distinct citations, and 456 of them were excluded immediately after inspection of the title and abstract (figure 1). The remaining 75 clinical studies were potentially eligible and were examined for full text. Fifty-six studies were excluded because: 1) eight studies used duplicated report; 2) 18 studies used inappropriate control arm (e.g. single arm, all patients received aPC); 3) 19 did not report mortality as the endpoint; and 4) 11 did not include aPC as an intervention. As a result, a total of 7 RCTs(9, 12, 13, 20-23) and 11 observational studies(10, 11, 24-32) were included for analysis. Characteristics of RCTs are shown in table 1.

Figure 2 shows the efficacy of aPC on mortality reduction, reported separately for RCTs and observational studies. Of the 7 RCTs, only the PROWESS study showed significant mortality reduction with aPC (OR: 0.74, 95% CI: 0.59-0.91),(9) and the other six studies failed to conclude a beneficial effect. There was moderate heterogeneity among included RCTs (I^2=48.6%, p=0.07). The pooled OR for mortality was 1.00 (95% CI: 0.84-1.19). In contrast, 6 out of the 11 observational studies showed significant mortality reduction with the use of aPC; and the remaining five studies showed a trend towards better lower mortality rate in aPC group. The heterogeneity was statistically significant with an I^2 of 70.8% (p<0.001). The pooled OR for mortality with the use of aPC was 0.67 (95% CI: 0.62-0.72). Egger's test did not show evidence of publication bias in RCTs (p=0.808) and observational studies

(p=0.145). Similarly, Begg's test did not show evidence of publication bias in RCTs (p=0.293) and observational studies (p=0.337). However, publication bias was suspected for observational studies as suggested by the funnel plot in which each dot represents a study and they gathered at the upper left corner (figure 3).

The result of sequential trial analysis is shown in figure 4. Studies were displayed sequentially by their publication year from left to the right of the horizontal line. After publication of the first and second studies (PROWESS 2001 and rhAPC sepsis 2001), the Z score crossed the conventional significance boundary (Z=1.96) but did not cross the O'Brien-Fleming boundaries. With the publication of the study ADDRESS 2005, the Z-score reached and crossed the futility line, indicating no effect of the aPC for mortality reduction in septic patients.

Meta-analysis of observational studies was performed by using Bayesian approach. The posterior distribution of individual OR was shrunken, as reflected by the narrower credible interval of study level estimates as compared to the observed estimates. For instance, the credible interval of OR in the study de Pont AC 2005 was 0.40-1.06, which was significantly narrower than the observed confidence interval of 0.04-6.70 (figure 2). This was because each component study borrowed evidence from the overall effect by using Bayesian approach. The overall OR was 0.67 (credible interval: 0.56-0.78).

Figure 5 shows the mean OR and 95% credible interval (CrI) for different power transformation priors to down-weight observational evidence on the risk of death with aPC. To the left of the figure when alpha took negligible values, the observational evidence was totally discounted and the mean OR was 1, which was consistent with the pooled result from RCTs. Increasing weight was assigned to observational evidence with increasing alpha values. We could see from the figure that the upper limit of CrI crossed the reference line. When observational evidence was combined at its face value (alpha=1), the aPC group showed significant mortality reduction as compared with the control group. The alpha value influenced the precision of prior evidence. As shown in figure 6, the precision of prior increased with increasing value of alpha from 0.000001 to 1.

Discussion

Key findings of the present analysis are 1) aPC appears to be able to reduce mortality rate when evidence is pooled from observational studies, and the results are consistent by using conventional Bayesian approaches; 2) RCTs failed to identify any beneficial effect of aPC; 3) observational evidence, when discounted by different power transformation priors, can alter the conclusion derived from RCTs. 4) With trial sequential analysis, the positive result (significant beneficial effect of aPC) as shown in the PROWESS study should be interpreted with caution.

One potential explanation for the positive findings in observational studies is the publication bias as shown in figure 3. The funnel contour plot showed that most observational studies located in the region with p<1%, indicating that the asymmetrical distribution was more likely due to publication bias. It is not surprising that observational studies are more subject to publication bias in that they are less likely to be registered a priori.(33) In contrast, RCTs are usually registered and there are many online registration sites.(34) The value of observational studies is usually discounted in evidence synthesis, and the conventional view is that observational evidence can only serve as hypothesis-generating. In such context, if the finding of an observational study is neutral, it will be less interesting to readers and journals, making it less likely to be published. In contrast, because RCTs are always registered and requires large amount of cost and other resources, studies with negative findings can be published and is equally important to those with positive findings.

Activated protein C (aPC) for the treatment of sepsis is a good example illustrating the importance of using sequential trial analysis in evidence synthesis. aPC was approved by the food and drug administration after publication of PROWESS trial, which seemed too hasty when viewed retrospectively. Although the initial trial was positive at conventional significance level of p=0.05 (Z=1.96), it was subject to repeated measurement error. This problem can be addressed by using adjusted alpha

level. In sequential trial analysis, this is achieved by using alpha-spending function and constructing the O'Brien-Fleming boundaries. If sequential trial analysis had been performed at the conclusion of PROWESS trial, the approval of aPC for sepsis would not be so hasty. Someone argued that the disparity between PROWESS trial and subsequent trials such as PROWESS-SHOCK could be explained by the heterogeneity of enrolled subjects.(35)

In the translation of research into clinical practice, there are a lot of influence factors that to consider. RCT is generally accepted as a gold standard. However, there are some limitations in real clinical practice that RCT cannot simulate all the clinical situations. The biggest problem is that RCT is usually conducted in non-real world setting, that is, it is always performed in specialized academic centers with strict inclusion/exclusion criteria. For example, in the Dhainaut 2009 study there was a long list of exclusion criteria, including expected surgical procedure in the next 3 days, platelet count<30,000/nm^3, receiving therapeutic heparin, moribund, withdrawn from aggressive management by patients' family, and pregnant or breast feeding. Such strict exclusion criteria would exclude most of patients with septic shock. Therefore, it appears unfair to treat our septic shock patients based on evidence derived from a minority of the population. In this situation, observational studies have its advantage in testing the clinical effectiveness of aPC on mortality reduction (36-38). Observational study included wider range of patients with septic shock and the setting is just like what we will encounter in routine clinical practice. Therefore, the observational evidence cannot be simply ignored in evidence synthesis for decision-making. On the other hand, observational trial could be misleading by more clinical bias as reflect by asymmetrically distribute component studies so it should be interpret very cautiously and in conjunction with other evidence. Since there was no consensus on how to combine observational evidence with RCTs, we discounted observational evidence with power transform priors taking advantage of the flexibility in Bayesian modeling (18). In this model, we found that the treatment effect of aPC increased with more weight assigning to observational evidence (figure 6). A value of

0 for alpha implies that the observational evidence is ignored, and a value of 1 for alpha means that observational evidence is accepted at its "face value". This approach gives a full picture of how pooled evidence can be altered by observational studies, by explicitly showing the power transform priors.

Several limitations of the study need to be acknowledged. First, there are substantial heterogeneity among included RCTs, which may be explained by the differences in study population, timing of intervention and definition of study endpoint. As expected, observational studies showed substantial heterogeneity. Considering the very different study criteria and the various geographic sources of each observational study (representing different standards of care), heterogeneity should be expected due to the more generalizable (real-world) evidence. Second, it is still largely unknown on how to discount the observational evidence. The most appropriate prior will vary from study to study. The present study only displays a wide range of possible alpha values and explicitly demonstrates how RCT evidence can be modified by observational evidence.

In aggregate, our study demonstrates that there is considerable disparity between observational and RCT evidence. While observational evidence shows beneficial effect of aPC on mortality reduction, RCTs failed to identify any such treatment effect.

Figure legends

Figure 1. Flow chart of study selection.

Figure 2. Forest plots showing the efficacy of aPC on mortality reduction, reported separately for RCTs and observational studies. Results were pooled by using conventional meta-analytic approach.

Figure 3. Contour funnel plots showing the publication bias in RCTs and observational studies. Publication bias was identified for observational studies as reflected by the asymmetrically distributed component studies.

Figure 4. Sequential trial analysis involving randomized controlled trials showing that the Z-score crossed the futility line after the study ADRESS 2005. Parameters used for the creation of boundaries were: type: Two-sided; type 1 Error: 5.0%; alpha spending: O'Brien-Fleming; information axis: sample size; power: 80.0%; effect type intervention: RRR User Defined (21.25%); heterogeneity correction: user defined (0.5%).

Figure 5. Mean OR and 95% credible interval (CrI) for different power transformation priors to down-weight observational evidence on the risk of death with aPC.

Figure 6. Prior distribution derived by discounting observational evidence with alpha from 0.000001 to 1. The plots shows that the precision of prior increases with increasing alpha values.

Reference

1. Schorr CA, Zanotti S, Dellinger RP. Severe sepsis and septic shock: management and performance improvement. Virulence. 2014;5(1):190-9.

2. Dellinger RP, Levy MM, Rhodes A, Annane D, Gerlach H, Opal SM, et al. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012. Critical care medicine. 2013;41(2):580-637.

3. MacIntyre N. Ventilatory Management of ALI/ARDS. Seminars in respiratory and critical care medicine. 2006;27(4):396-403.

4. Rivers E, Nguyen B, Havstad S, Ressler J, Muzzin A, Knoblich B, et al. Early goal-directed therapy in the treatment of severe sepsis and septic shock. The New England journal of medicine. 2001;345(19):1368-77.

5. Pro CI, Yealy DM, Kellum JA, Huang DT, Barnato AE, Weissfeld LA, et al. A randomized trial of protocol-based care for early septic shock. The New England journal of medicine. 2014;370(18):1683-93.

6. Ronco C, Bellomo R, Homel P, Brendolan A, Dan M, Piccinni P, et al. Effects of different doses in continuous veno-venous haemofiltration on outcomes of acute renal failure: a prospective randomised trial. Lancet. 2000;356(9223):26-30.

7. Zhang Z, Xu X, Zhu H. Intensive- vs less-intensive-dose continuous renal replacement therapy for the intensive care unit-related acute kidney injury: a meta-analysis and systematic review. Journal of critical care. 2010;25(4):595-600.

8. Christiaans SC, Wagener BM, Esmon CT, Pittet JF. Protein C and acute inflammation: a clinical and biological perspective. American journal of physiology Lung cellular and molecular physiology. 2013;305(7):L455-66.

9. Bernard GR, Vincent JL, Laterre PF, LaRosa SP, Dhainaut JF, Lopez-Rodriguez A, et al. Efficacy and safety of recombinant human activated protein C for severe sepsis. The New England journal of medicine. 2001;344(10):699-709.

10. Dhainaut JF, Payet S, Vallet B, Franca LR, Annane D, Bollaert PE, et al. Cost-effectiveness of activated protein C in real-life clinical practice. Critical care (London, England). 2007;11(5):R99.

11. Kubler A, Mayzner-Zawadzka E, Durek G, Gaszynski W, Karpel E, Mikaszewska-Sokolewicz M, et al. Results of severe sepsis treatment program using recombinant human activated protein C in Poland. Medical science monitor : international medical journal of experimental and clinical research. 2006;12(3):CR107-12.

12. Annane D, Timsit JF, Megarbane B, Martin C, Misset B, Mourvillier B, et al. Recombinant human activated protein C for adults with septic shock: a randomized controlled trial. American journal of respiratory and critical care medicine. 2013;187(10):1091-7.

13. Ranieri VM, Thompson BT, Barie PS, Dhainaut JF, Douglas IS, Finfer S, et al. Drotrecogin alfa (activated) in adults with septic shock. The New England journal of medicine. 2012;366(22):2055-64.

14. Lai PS, Matteau A, Iddriss A, Hawes JC, Ranieri V, Thompson BT. An updated meta-analysis to understand the variable efficacy of drotrecogin alfa (activated) in severe sepsis and septic shock. Minerva anestesiologica. 2013;79(1):33-43.

15. Marti-Carvajal AJ, Sola I, Gluud C, Lathyris D, Cardona AF. Human recombinant protein C for severe sepsis and septic shock in adult and paediatric patients. The Cochrane database of systematic reviews. 2012;12:CD004388.

16. Zhang Z, Ni H, Xu X. Do the observational studies using propensity score analysis agree with

randomized controlled trials in the area of sepsis? Journal of critical care. 2014;29(5):886 e9-15.

17. Zhang Z. Recombinant human activated protein C for the treatment of severe sepsis and septic shock: a study protocol for incorporating observational evidence using a Bayesian approach. BMJ Open. 2014;4(7):e005622.

18. Chen MH, Ibrahim JG. Power prior distributions for regression models. statistical science. 2000;15(1):46-60.

19. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. Journal of clinical epidemiology. 2008;61(1):64-75.

20. Bernard GR, Ely EW, Wright TJ, Fraiz J, Stasek JE, Jr., Russell JA, et al. Safety and dose relationship of recombinant human activated protein C for coagulopathy in severe sepsis. Critical care medicine. 2001;29(11):2051-9.

21. Abraham E, Laterre PF, Garg R, Levy H, Talwar D, Trzaskoma BL, et al. Drotrecogin alfa (activated) for adults with severe sepsis and a low risk of death. The New England journal of medicine. 2005;353(13):1332-41.

22. Nadel S, Goldstein B, Williams MD, Dalton H, Peters M, Macias WL, et al. Drotrecogin alfa (activated) in children with severe sepsis: a multicentre phase III randomised controlled trial. Lancet. 2007;369(9564):836-43.

23. Dhainaut JF, Antonelli M, Wright P, Desachy A, Reignier J, Lavoue S, et al. Extended drotrecogin alfa (activated) treatment in patients with prolonged septic shock. Intensive care medicine. 2009;35(7):1187-95.

24. de Pont AC, Bakhtiari K, Hutten BA, de Jonge E, Vroom MB, Meijers JC, et al. Recombinant human activated protein C resets thrombin generation in patients with severe sepsis - a case control study. Critical care (London, England). 2005;9(5):R490-7.

25. Bertolini G, Rossi C, Anghileri A, Livigni S, Addis A, Poole D. Use of Drotrecogin alfa (activated) in Italian intensive care units: the results of a nationwide survey. Intensive care medicine. 2007;33(3):426-34.

26. Rowan KM, Welch CA, North E, Harrison DA. Drotrecogin alfa (activated): real-life use and outcomes for the UK. Critical care (London, England). 2008;12(2):R58.

27. Vincent JL, Laterre PF, Decruyenaere J, Spapen H, Raemaekers J, Damas F, et al. A registry of patients treated with drotrecogin alfa (activated) in Belgian intensive care units--an observational study. Acta clinica Belgica. 2008;63(1):25-30.

28. Martin G, Brunkhorst FM, Janes JM, Reinhart K, Sundin DP, Garnett K, et al. The international PROGRESS registry of patients with severe sepsis: drotrecogin alfa (activated) use and patient outcomes. Critical care (London, England). 2009;13(3):R103.

29. Lindenauer PK, Rothberg MB, Nathanson BH, Pekow PS, Steingrub JS. Activated protein C and hospital mortality in septic shock: a propensity-matched analysis. Critical care medicine. 2010;38(4):1101-7.

30. Sadaka F, O'Brien J, Migneron M, Stortz J, Vanston A, Taylor RW. Activated protein C in septic shock: a propensity-matched analysis. Critical care (London, England). 2011;15(2):R89.

31. Casserly B, Gerlach H, Phillips GS, Marshall JC, Lemeshow S, Levy MM. Evaluating the use of recombinant human activated protein C in adult severe sepsis: results of the Surviving Sepsis Campaign. Critical care medicine. 2012;40(5):1417-26.

32. Rimmer E, Kumar A, Doucette S, Marshall J, Dial S, Gurka D, et al. Activated protein C and septic shock: a propensity-matched cohort study*. Critical care medicine. 2012;40(11):2974-81.

33. Onukwugha E. Improving confidence in observational studies : should statistical analysis plans be made publicly available? PharmacoEconomics. 2013;31(3):177-9.

34. Anand V, Scales DC, Parshuram CS, Kavanagh BP. Registration and design alterations of clinical trials in critical care: a cross-sectional observational study. Intensive care medicine. 2014;40(5):700-22.

35. Kalil AC, Florescu DF. Severe sepsis: are PROWESS and PROWESS-SHOCK trials comparable? A clinical and statistical heterogeneity analysis. Critical care (London, England). 2013;17(4):167.

36. Albert RK. "Lies, damned lies ..." and observational studies in comparative effectiveness research. Am J Respir Crit Care Med. 2013;187(11):1173-7.

37. Nallamothu BK, Hayward RA, Bates ER. Beyond the randomized clinical trial: the role of effectiveness studies in evaluating cardiovascular therapies. Circulation. 2008;118(12):1294-303.

38. Zhang Z. Big data and clinical research: focusing on the area of critical care medicine in mainland China. Quantitative Imaging in Medicine and Surgery. 2014;4(5):426-9.

Funding

There was no funding for the present study.

Table 1 Characteristics of included randomized controlled trials

Studies	Patients (n)	Mean age (years)	Population	Mean APACHE II score	Control	Primary outcome	Baseline mortality (%)
Bernard GR 2001 (rhAPC)	131	59.3	Severe shock	17.3	Placebo	Coagulopathy	34.2
Bernard GR 2001 (PROWESS)	1690	60.5	Systemic inflammation and organ failure	24.8	Placebo (saline or albumin)	28-day all cause mortality	30.8
Ranieri VM 2012	1697	63.1	Sepsis and shock receiving fluids and vasopressor	25.3	Placebo (saline)	28-day all cause mortality	24.2
Abraham E 2005	2613	58.7	Severe sepsis and single organ failure or Mean APACHE II<25	18.2	Placebo (saline)	28-day all cause mortality	17
Nadel S 2007	477	2.5	Children with sepsis induced	-	Placebo (saline)	CTCOFR	17.5

			cardiac or respiratory failure				
Annane D 2013	411	63	Sepsis with >2 organ failure	-	Placebo (saline)	90-day mortality	46.3
Dhainaut JF 2009	193	62.4	Severe sepsis with vasopressor dependent hypotension	28.1	Placebo	Resolution of vasopressor dependent hypotension	32.3

APACHE: Acute Physiology and Chronic Health Evaluation; CTCOFR: Composite Time to Complete Organ Failure Resolution.

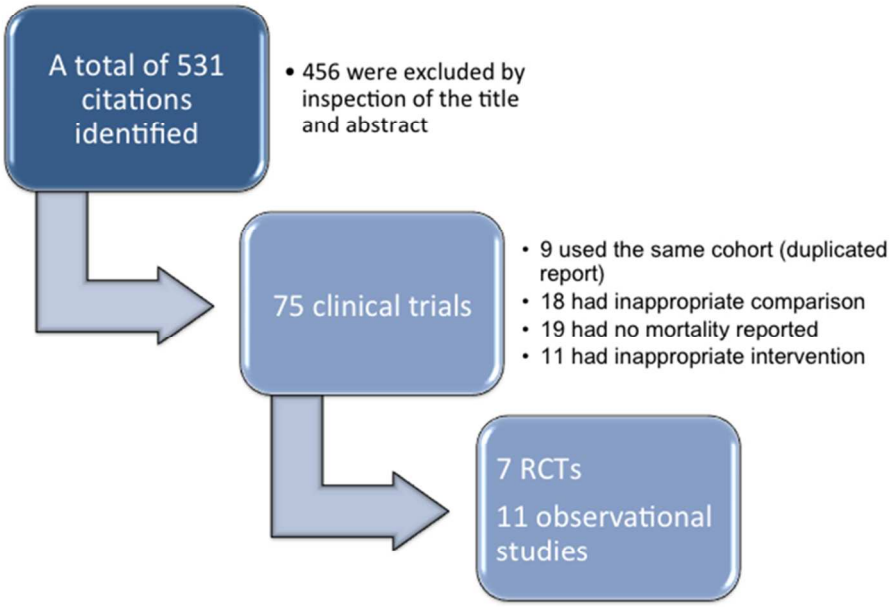

Figure 1. Flow chart of study selection.
84x63mm (300 x 300 DPI)

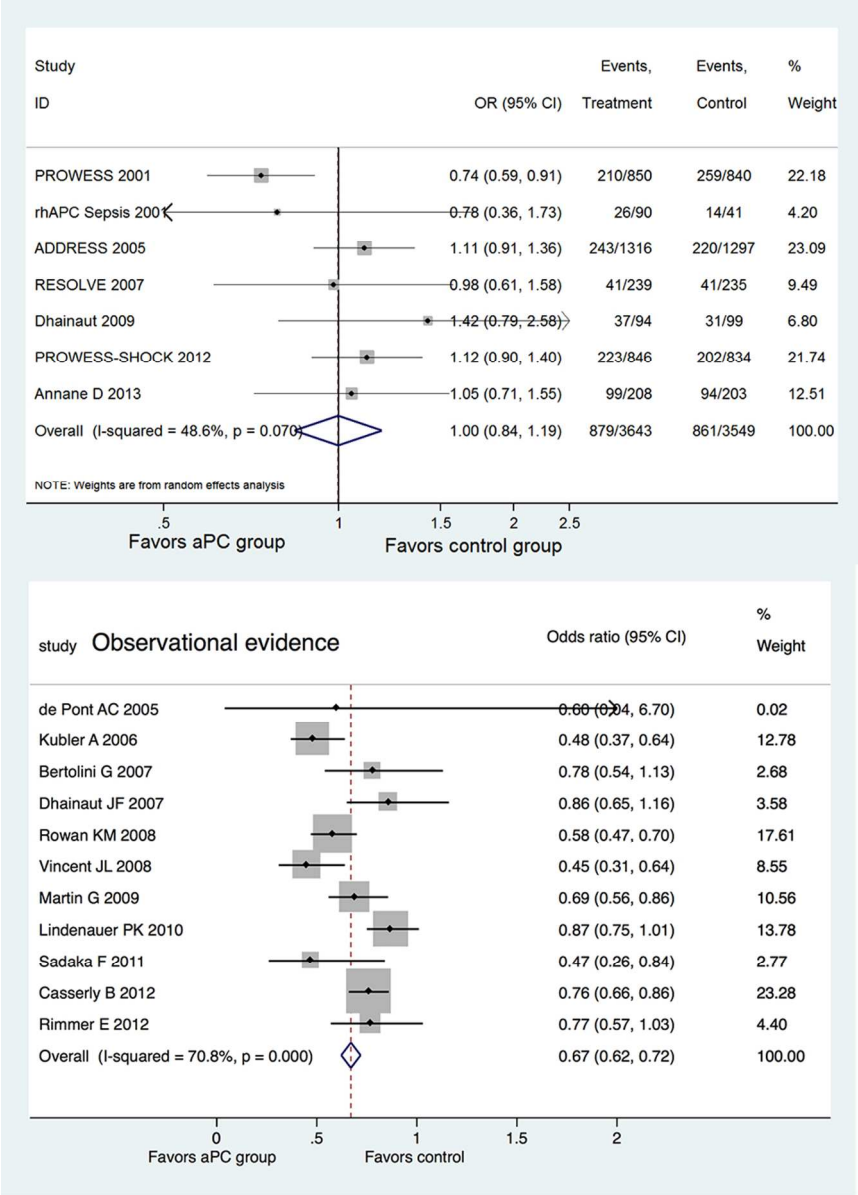

Figure 2. Forest plots showing the efficacy of aPC on mortality reduction, reported separately for RCTs and observational studies. Results were pooled by using conventional meta-analytic approach.
84x118mm (300 x 300 DPI)

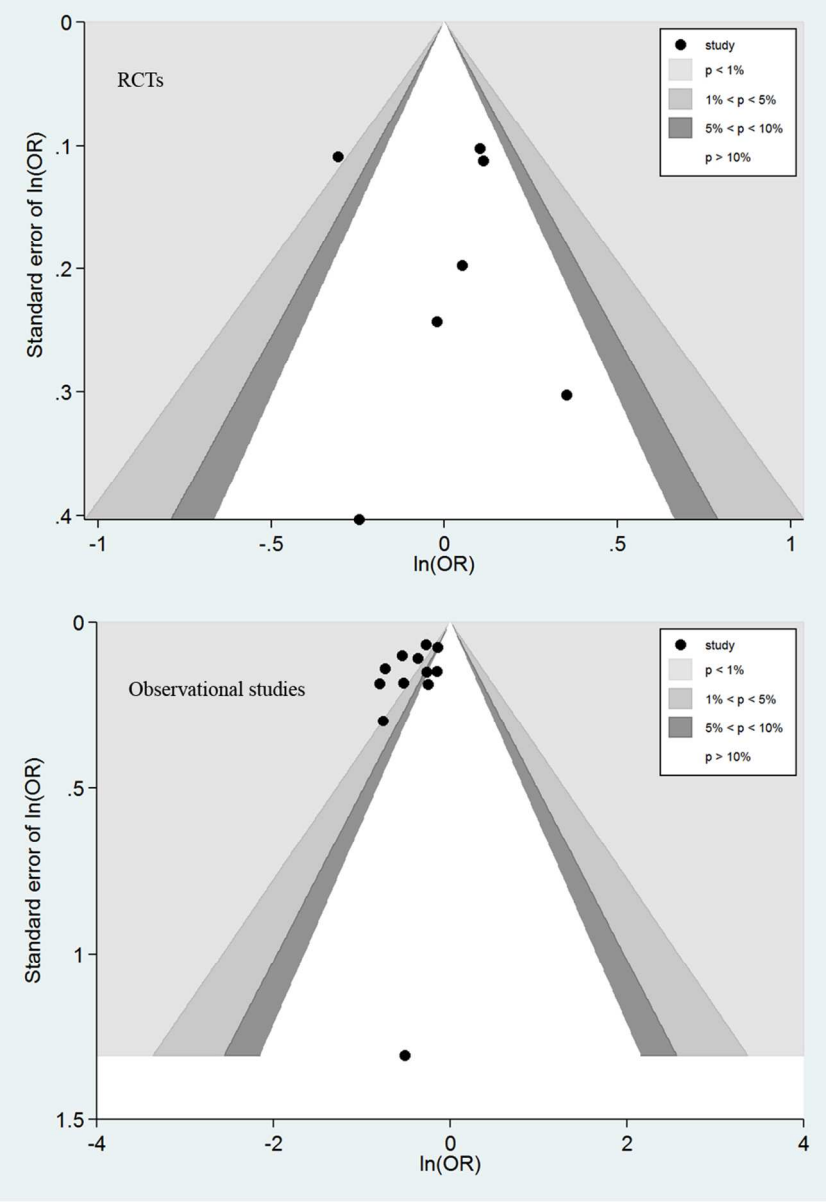

Figure 3. Contour funnel plots showing the publication bias in RCTs and observational studies. Publication bias was identified for observational studies as reflected by the asymmetrically distributed component studies.
84x123mm (300 x 300 DPI)

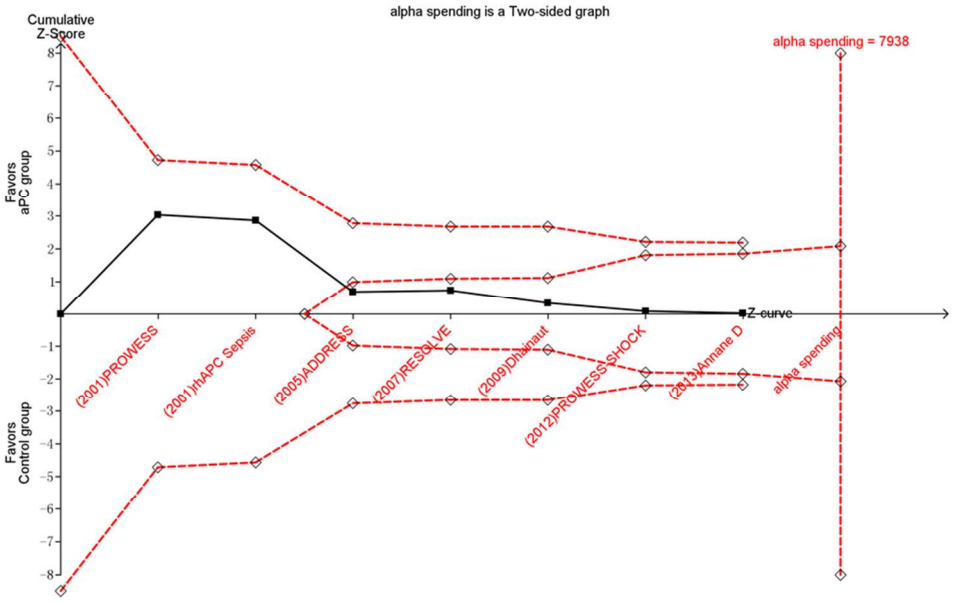

Figure 4. Sequential trial analysis showing that the Z-score crossed the futility line after the study ADRESS 2005. Parameters used for the creation of boundaries were: type: Two-sided; type 1 Error: 5.0%; alpha spending: O'Brien-Fleming; information axis: sample size; power: 80.0%; effect type intervention: RRR User Defined (21.25%); heterogeneity correction: user defined (0.5%).
84x51mm (300 x 300 DPI)

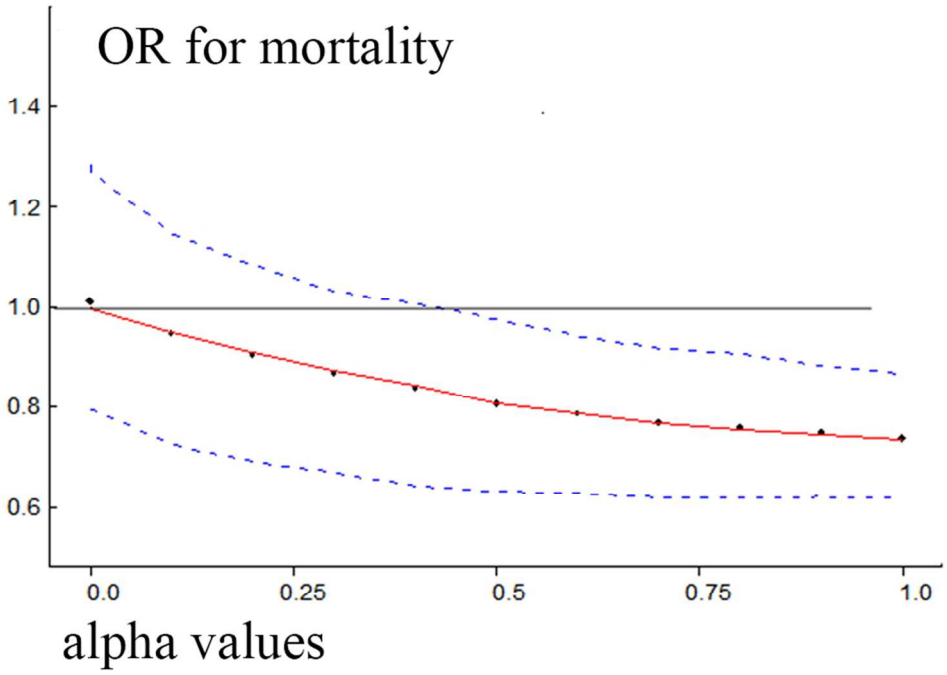

Figure 5. Mean OR and 95% credible interval (CrI) for different power transformation priors to down-weight observational evidence on the risk of death with aPC.
84x58mm (300 x 300 DPI)

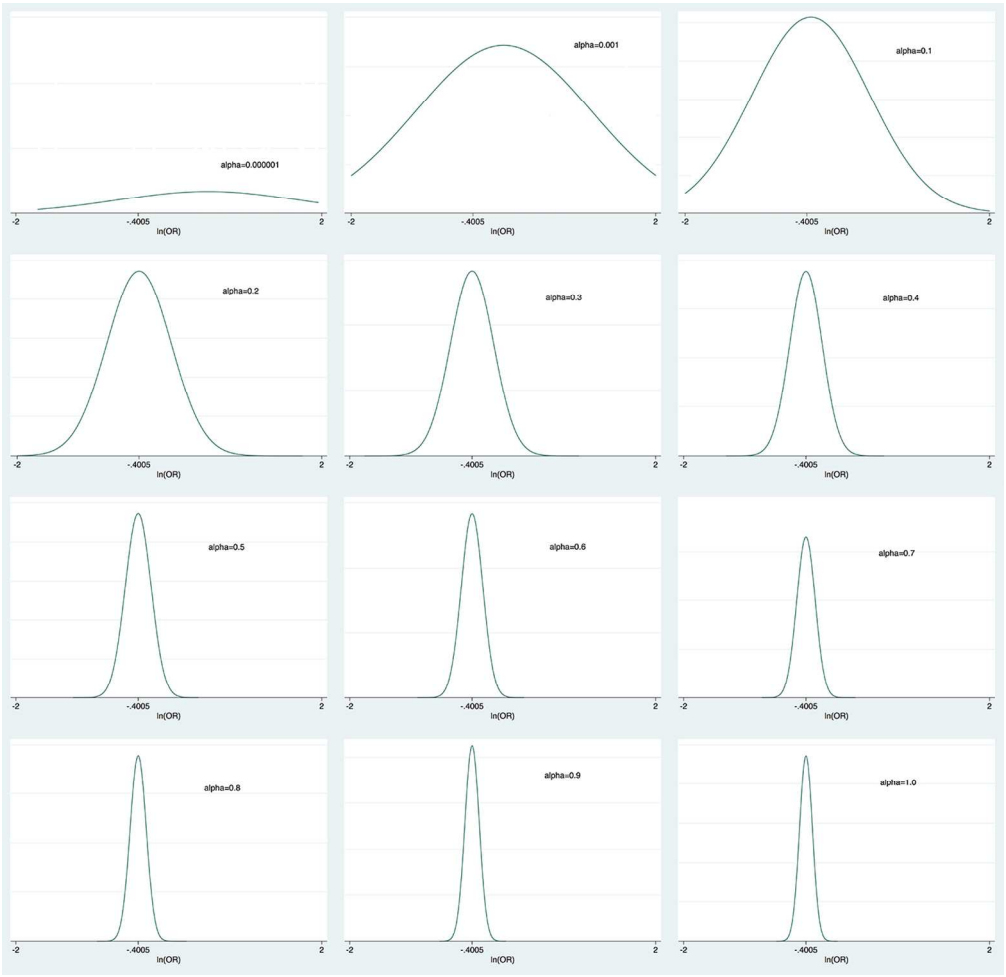

Figure 6. Prior distribution derived by discounting observational evidence with alpha from 0.000001 to 1. The plots shows that the precision of prior increases with increasing alpha values.
127x123mm (300 x 300 DPI)

BMJ Open

Recombinant human activated protein C for the treatment of severe sepsis and septic shock: a study protocol for incorporating observational evidence using a Bayesian approach

Zhongheng Zhang

To cite: Zhang Z. Recombinant human activated protein C for the treatment of severe sepsis and septic shock: a study protocol for incorporating observational evidence using a Bayesian approach. *BMJ Open* 2014;**4**:e005622. doi:10.1136/bmjopen-2014-005622

► Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2014-005622).

Received 5 May 2014
Revised 8 July 2014
Accepted 9 July 2014

Department of Critical Care Medicine, Jinhua Municipal Central Hospital, Jinhua Hospital of Zhejiang University, Zhejiang, PR China

Correspondence to
Dr Zhongheng Zhang;
zh_zhang1984@hotmail.com

ABSTRACT

Introduction: Activated protein C (aPC) plays a pivotal role in modulating a severe inflammatory response and is thought to be beneficial for patients with sepsis. However, several meta-analyses of randomised controlled trials (RCTs) show that aPC is not significantly associated with improved survival in critically ill patients with sepsis. One suggestion is that these analyses simply ignored observational evidence. The present study aims to quantitatively demonstrate how observational data can alter the findings derived from synthesised evidence from RCTs by using a Bayesian approach.

Methods and analysis: RCTs and observational studies investigating the effect of aPC on mortality outcome in critically ill patients with sepsis will be included. The quality of included RCTs will be assessed by using the Delphi list. Publication bias will be quantitatively analysed by using the traditional Egger regression test and the Begg rank correlation test. Observational data will be used as the informative prior for the distribution of OR. A power transformation of the observational data likelihood will be considered. Observational evidence will be down-weighted by a power of α which takes values from 0 to 1. Trial sequential analysis will be performed to quantify the reliability of data in meta-analysis adjusting significance levels for sparse data and multiple testing on accumulating trials.

Trial registration number: PROSPERO (CRD42014009562).

INTRODUCTION

Sepsis is defined as systematic inflammatory response syndrome (SIRS) caused by infection.[1] Levels of severity vary widely depending on the presence of shock and organ failure. Sepsis is a leading cause of morbidity and mortality in intensive care units. In the USA alone, there were over 750 000 estimated cases in 1995,[2] and sepsis accounts for over 25% of admissions to ICUs in Europe.[3] Due to its significant impact on global health, every effort has been made to improve the survival of patients with sepsis. One such initiative is the Surviving Sepsis Campaign (SSC) with the objective of reducing mortality from sepsis by 25%.[4] Various strategies have been implemented to achieve this aim, such as early goal directed therapy, early use of broad spectrum antibiotics, source control and low tidal volume ventilation. Although the sepsis mortality rate has subsequently declined, the SSC goal is far from being achieved.[5]

Activated protein C (aPC) has pleiotropic biological effects and plays a pivotal role in modulating the severe inflammatory response which occurs in sepsis. Its biological effects include, but are not limited to, reduction of thrombin production by inactivating factors Va and VIII, and inhibition of IL-1, IL-6 and TNF-α production by monocytes.[6] Many observational studies (OS) have shown significantly improved survival outcomes in patients with sepsis treated with aPC compared with controls. Furthermore, these encouraging results have been confirmed in the milestone clinical trial PROWESS. However, the findings have not been replicated in subsequent randomised clinical trials, and thus enthusiasm for aPC has declined.

Randomised controlled trials (RCT) are designed to test the biological efficacy of a particular treatment, while observational studies test the effectiveness of that treatment in the real world setting.[7] Differences in efficacy and effectiveness may result from issues related to trial design, patient selection and therapeutic implementation. Some systematic reviews exploring the effect of aPC on sepsis

exclusively focused on RCTs while ignoring evidence from OS, and consistently showed that aPC had a neutral effect on survival outcomes.[8 9] We propose that although RCTs are the 'gold standard' for the definite determination of the clinical efficacy of an intervention, OS cannot simply be ignored in evidence synthesis. Kalil and LaRosa provided a frequentist analysis of both observational and randomised studies, but no Bayesian analyses were performed.[10] From the Bayesian perspective, OS can be incorporated into the analysis and an informative prior distribution on the treatment effect derived from the observational data.[11] In contrast to previous meta-analysis, we will incorporate observational data into analysis using the Bayesian approach. Furthermore, additional RCTs will be incorporated in order to update the systematic review.

METHODS

Search strategy

We will search electronic databases including the Cochrane Central Register of Controlled Trials (CENTRAL), PubMed, EBSCO, EMBASE and ISI Web of Science from inception to January 2014. Our core search consists of terms related to aPC and sepsis (see table 1 for the detailed search strategy to be used in PubMed). Strategies will be adapted to other databases. There will be no language restriction. The references of systematic reviews will be reviewed to identify additional eligible articles.

Studies to be included

We will include RCTs and OS for analysis. OS will include: (1) cohort studies using multivariable analysis with aPC treatment as one of the covariates; (2) cohort studies using propensity analysis; (3) case–control studies; (4) studies with both prospective and retrospective designs; and (5) all OS irrespective of their methodological design quality.

Table 1 Search strategy performed in PubMed

Items	Search terms	Number of citations
1#	((activated protein C[Title/Abstract]) OR xigris[Title/Abstract]) OR drotrecogin alfa [Title/Abstract]	4460
2#	(sepsis[Title/Abstract]) OR septic shock[Title/Abstract]	72 635
3#	(((mortality[Title/Abstract]) OR safety[Title/Abstract]) OR adverse events[Title/Abstract]) OR bleeding[Title/Abstract]	875 580
1# AND 2# AND 3#		531

Studies to be excluded

We will exclude studies that: (1) do not report mortality as an endpoint; (2) are a secondary analysis of a primary study whose data have been published elsewhere; and (3) only include a single arm so that no comparison can be made between different treatment strategies (eg, such as analysis of risk factors).

Data extraction

A custom-made form will be used to extract the following data from eligible studies: name of the first author, year of publication, sample size, illness severity scores (APACHE II, SOFA and SAPS), number of deaths in each arm, total number of participants in each arm, bleeding or haemorrhage events in each arm, OR of treatment versus non-treatment for mortality, the method used for covariate adjustment (propensity score analysis, logistic regression model) and the design of the OS (prospective vs retrospective). The adverse event of bleeding will be divided into two categories: major bleeding (terms consist of combinations of 'massive', 'major' and 'bleeding', 'haemorrhage') and any bleeding (terms consist of combinations of 'minor' and 'bleeding', 'haemorrhage'). If only the risk ratio (RR) is reported, we will transform it into the OR by using standard formula (described elsewhere[12]):

$$\mathrm{OR} = \frac{\mathrm{RR} \times (1 - \mathrm{CER})}{1 - \mathrm{CER} \times \mathrm{RR}}$$

where CER indicates control event rate (same as control group risk). Mortality is defined variably across studies (eg, 28-day, in-hospital, 60-day or 90-day) and we will include all types of definitions for analysis.

Quality assessment of RCTs and OS

Quality assessment of included RCTs will be performed by using the Delphi list, which consists of nine items: sequence generation, allocation concealment, baseline characteristics, eligibility criteria, blindness to outcome assessor, blindness to care provider, blindness to patient, use of point estimate and variability for outcome measures, and use of intention to treat analysis.[13] The explanation and rating for each item are given in table 2. Quality assessment of OS will be performed by using the modified Newcastle–Ottawa scale which has been described elsewhere (table 3).[14]

Publication bias

Publication bias will be quantitatively analysed by using the traditional Egger regression test and Begg rank correlation test.[15 16] The Begg rank correlation test investigates the relationship between the standardised OR and sample size or variance by using the Spearman rank correlation.[17] In the Egger regression test, the standard normal deviate (the OR divided by its SE) is regressed against the estimates precision. The intercept of the regression line is an estimate of asymmetry: the larger its

Table 2 Quality assessment of randomised controlled trials using tools adapted from the Delphi list

Items	Explanation	Rating
Sequence generation	Is the method of sequence generation clearly reported?	Yes/no/unclear
Allocation concealment	Is treatment allocation concealment (using an opaque envelope, central allocation) performed?	Yes/no/unclear
Baseline characteristics	Are the groups similar at baseline regarding the most important prognostic factors?	Yes/no/unclear
Eligibility criteria	Are eligibility criteria clearly specified?	Yes/no/unclear
Blindness to outcome assessor	Is the outcome (mortality) assessor blinded?	Yes/no/unclear
Blindness to care provider	Is the allocation unknown to the treating physician?	Yes/no/unclear
Blindness to patient	Is the patient blinded?	Yes/no/unclear
Point estimate and variability	Are the point estimate and variability reported for the outcome measure?	Yes/no/unclear
Intention-to-treat	Does the analysis include intention to treat analysis?	Yes/no/unclear

deviation from origin, the more significant the asymmetry.[18] A contour enhanced funnel plot will be used to visually assess the presence of publication bias. OR is plotted on the horizontal axis, and precision is plotted on the vertical axis, with asymmetric distribution of component studies representing potential publication bias. Contour lines are added to the plot at conventional statistical significance levels of <0.01, <0.05 and <0.1. A funnel contour enhanced plot can aid interpretation of the funnel plot. If studies are missing in the non-significance area, it is likely that the asymmetry is caused by publication bias. Conversely, if studies are in the significance area, the asymmetry is more likely caused by factors other than publication bias, such as study quality.[19]

Sensitivity or subgroup analysis

Sensitivity analysis will be performed by excluding studies with poor methodological design. Subgroup analysis will be performed to explore confounding factors such as shock versus non-shock, and the effect of aPC modified by disease severity. If there are enough studies with the same definition of mortality (n>5), subgroup analysis will be performed by different mortality definitions.

Statistical analysis

Three key components of Bayesian analysis are prior, likelihood and posterior. The quantity of interest in our study is the OR for mortality. Observational data are used as the informative prior for the distribution of OR. For studies using a logistic regression model for risk adjustment, we will extract adjusted OR and relevant 95% CI for analysis. For studies using propensity matched analysis, the OR from matched samples are calculated. Random effects meta-analysis will be performed to combine the results obtained from OS, by using a Bayesian approach.[20] The WinBUGS code for performing the calculation is shown in table 3. The pooled OR will be transformed by natural log to ln(OR) to improve normality. The SE in the natural log

Table 3 Quality assessment of included observational studies using the modified Newcastle–Ottawa scale

Selection	Representativeness of the exposed cohort	This item will be assigned a '★' when all eligible patients with severe sepsis or septic shock are included in the analysis during the study period
	Selection of the non-exposed cohort	This item will be assigned a '★' when all eligible patients without aPC treatment are included in the analysis during the study period
	Ascertainment of exposure	This item will be assigned a '★' when aPC administration is directly obtained from a medical chart, not from reporting by the patient
	Outcome of interest is not present at the start of the study	This item will be assigned a '★' when the subject is alive at the time of enrolment
Comparability	Comparability of cohorts on the basis of design or analysis	Baseline characteristics of aPC and control groups are comparable. Usually this can be found in table 1 of the original article.
Outcome	Assessment of outcome	This item will be assigned a '★'when mortality is assessed by the investigator, not by the report of the patient's family or next-of-kin
	Is follow-up long enough for outcome to occur?	Adequate follow-up is carried out during hospital stay, ICU stay or redefined study time
	Adequacy of follow-up of the cohort	This item will be assigned a '★' when the follow-up rate is >80%

aPC, activated protein C.

scale can be transformed from the 95% credible interval by using the equation:

$$\text{Standard error}\,(\sigma) = \frac{L_{up} - L_{lo}}{2 \times 1.96}$$

where L_{up} and L_{lo} represent the upper and lower limits of the 95% credible interval. The precision is the reciprocal of SE.

The framework to incorporate observational data as informative prior is presented by Chen and Ibrahim.[11] Model development has been described elsewhere but we repeat it here for the reader's benefit. Let the data from RCTs be denoted by D, and the likelihood of RCTs be denoted by $L(\theta|D)$. Suppose we have data from OS which are denoted by D_0. Furthermore, let $P(\theta)$ denote the prior distribution for θ before OS are incorporated. $P(\theta)$ is the initial prior distribution for θ. Given α, the power prior distribution of θ is defined as:

$$P(\theta|D_0, \alpha) \propto L(\theta|D_0)^{\alpha} \times P(\theta|c_0)$$

where c_0 is the hyperparameter for initial prior, and α is used to weight observational evidence relative to the likelihood of RCT evidence. The value of α controls the impact of observational evidence on $P(\theta|D_0, \alpha)$. When evidence from RCTs is added to the model, a power transformation of the observational data likelihood is considered:

$$P(\theta|\text{Data}) = L(\theta|\text{RCTs}) \times [L(\theta|\text{Obs})]^{\alpha} \times P(\theta)$$

where $P(\theta|\text{Data})$ is the posterior distribution for model quantities, $[L(\theta|\text{Obs})]$ is the likelihood function derived from observational data, and $L(\theta|\text{RCTs})$ is the likelihood function from RCT data. The weight of observational data is counted by the power α. The power takes values from 0 to 1. If α=0, the observational data are essentially removed from analysis and only RCTs are used for evidence synthesis; if α=1, observational data are taken at their 'face value' and not discounted at all. Traditional meta-analyses such as those done in The Cochrane Collaboration included only RCTs that actually render α=0. In our analysis, α will take 12 values ranging between 0 and 1 (0.000001, 0.001, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0), resulting in a series of posterior distributions for OR. As shown in table 4, the WinBUGS code is composed of three parts. Part (1) is to repeat meta-analysis of RCTs 12 times, once for each value of α to discount the observational evidence. Part (2) is the meta-analysis model. In this section, i represents the component studies and k indices each of the 12 meta-analyses. These meta-analyses differ from each other only in the prior distribution for the overall pooled effect d, which is represented by:

$$d[k] \sim \text{dnorm}(0.33, \text{prec.d}[k]).$$

The mean of prior distribution (the figure 0.33 in the expression is used for illustration purposes, and is not obtained from real analysis) is the natural log of the pooled OR (LOR) estimated from observational data. The pooled OR is estimated with a Bayesian approach with a random effects model. The code for the random effects meta-analysis is shown in table 4. The precision of the prior distribution, prec.d[k], is determined in part (3). Part (3) is to calculate precision of the prior discounted by using α.[21]

Convergence diagnostics will be explored by running two chains. Simulated values will be compared to identify when they become similar. History plots with different chains superimposed (in different colours) will help to determine convergence. Furthermore, we will use the Brooks–Gelman–Rubin diagnostic to test convergence. The procedure will produce three coloured lines (red, blue and green). Convergence is deemed to occur when the red line settles close to 1 and the blue and green lines converge together.

Trial sequential analysis (TSA) is performed to quantify the reliability of data in meta-analysis adjusting significance levels for sparse data and multiple testing on accumulating trials.[22] Trial sequential monitoring boundaries are used to control the risks for type I and II errors and to indicate whether additional trials are needed. A zero-event trial will be handled by the constant continuity correction method with a correction factor of 0.5, that is, 0.5 is added to each cell of the 2×2 table.[23] The information size calculation requires the mortality rate in the control group and the minimal effect size for the intervention. We predefined that the mortality in the control group is 30%, and the intervention is able to reduce the relative risk by 10%. The conventional α and β are 0.05 and 0.2, respectively. Meta-analysis will be updated by adding component studies sequentially in the order of publication.

Statistical analysis will be performed by using WinBUGS (Imperial College and MRC, UK) and Stata V.12.0 (College Station, Texas, USA). TSA will be performed by using the software TSA V.0.9 Beta (Copenhagen Trial Unit, 2011).

Results to be reported

Search results will be displayed in a flowchart. Pooled results from conventional meta-analysis techniques will be displayed in forest plots separately for RCTs and OS. Publication bias as shown in funnel plots will also be displayed, again separately for RCTs and OS. The results of TSA will be reported graphically. Random effects meta-analysis using a Bayesian approach will be used to pool summary effects for observational evidence and the results will be reported by using a caterpillar plot. Summary OR will also be plotted against different values of α to examine how observational evidence influences the summary effect. The Brooks–Gelman–Rubin plot will be used to display convergence diagnostics.

DISCUSSION

aPC was once the only approved drug for the treatment of sepsis. However, it was withdrawn from the market

Table 4 WinBUGS codes for performing random effects meta-analysis and meta-analysis incorporating observational data

	Random effects meta-analysis	**Informative prior with observational data**

Model†	model {
	for (i in 1:N)
	{
	P[i]<-1/V[i] Y[i]~dnorm(delta[i], P[i]) delta[i]~dnorm(d, prec)
	OR[i]<-exp(delta[i])
	} d~dnorm(0, 1.0E-5) OR[13]<-exp(d) tau~dunif(0,10) tau.sq<-tau*tau prec<-1/tau.sq
	}
		model {
		# (1) create multiple datasets
		for (i in 1:5) {
		for (k in 1:12) {
		rc[i, k]<- rc.dat[i] rt[i, k]<-rt.dat[i] nc[i, k]<-nc.dat[i] nt[i, k]<-nt.dat[i] }
		}
		# (2) estimate RCT meta-analysis model for each value of data
		for (k in 1:12) {
		for (i in 1:5) {
		rc[i,k]~dbin(pc[i,k], nc[i,k])
		rt[i,k]~dbin(pt[i,k], nt[i,k]) logit(pc[i,k])<-mu[i,k]
		logit(pt[i,k])<-mu[i,k]+delta[i,k]
		mu[i,k]~dnorm(0.0, 1.0E-6)
		delta[i,k]~dnorm(d[k], prec[k])
		or[i,k]<-exp(delta[i,k])
		}
		d[k]~dnorm(0.33, prec.d[k])
		OR[k]<-exp(d[k])
		prec[k]<-1/tau.sq[k]
		tau.sq[k]<-tau[k]*tau[k]
		tau[k]~dunif(0,5)
		}
		# (3) calculate precision of prior (from meta-analysis of obs studies) downweighted using alpha
		for (k in 1:12) {
		prec.d[k]<-alpha[k]*271.3
		}
		}

Data‡	list(Y=c(-0.51083, -0.73397, -0.24846, -0.15082, -0.54473, -0.52763, -0.36817, -0.13926, -0.75502, -0.27444, -0.26136), V=c(1.706611, 0.01954, 0.035483, 0.021832, 0.010326, 0.033478, 0.011817, 0.005765, 0.089499, 0.004559, 0.022782), N=11)	list(rt.dat=c(0,2,3,2,3), nt.dat=c(67,45,34,56,34), rc.dat=c(2,3,4,2,0), nc.dat=c(44,56,78,123,35), alpha=c(0.0001, 0.2, 0.8))
Initials§	list(d = 472.0235128342391, delta = c(470.6994400270435, 472.3980455275865, 472.201137881263, 472.0198057372273, 471.8605396435204, 470.2850099832592, 469.5829735618464, 473.0258057826344, 470.3932238143316, 469.5792223324207, 469.6419041364815), tau = 0.8303798133648838) list(d = 0, delta = c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), tau = 1)	list(d = c(0,0,0), delta = structure(.Data = c(**place 5*12=60 initial values here**), .Dim = c(5,12)), mu = structure(.Data = c(**place 5*12=60 initial values here**), .Dim = c(5,12)), tau = c(1,1,1,1,1,1,1,1,1,1,1,1))

†Contents following # are not syntax used for analysis, but are used to annotate corresponding codes.
‡Data are used for illustration purpose and are not obtained from the real analysis.
§Initial values are randomly generated and do not represent the actual values used in analysis.

after the large clinical trial PROWESS-SHOCK failed to identify any beneficial effect in patients with sepsis. However, in the first place, aPC was approved for use in patients with sepsis because the PROWESS study demonstrated a significant beneficial effect, with the study being stopped early because of its efficacy.[24] Furthermore, a large number of OS also showed a large beneficial effect with the use of aPC. Clinicians may be confused by these seemingly differing results. It is still largely unknown whether aPC is beneficial for specific subgroups of patients with sepsis. In this situation, the synthesis of evidence for decision making may help to address these conflicting findings. As a result, a few study groups have conducted systematic reviews and meta-analyses to provide comprehensive and up-to-date evidence for clinical use. The Cochrane Collaboration has also published the results of an updated meta-analysis on the effectiveness of aPC for sepsis, which however showed a neutral effect.[8] However, this meta-analysis only included RCTs. There is no doubt that the RCT is the gold standard for supplying evidence for medical decision making and can provide high level evidence on the comparative effectiveness of interventions. However, there are some circumstances where non-randomised evidence should be incorporated in order to estimate effectiveness. These include situations where there are concerns about internal and external validity (only effective in specialised centres or highly selected subjects) and size (estimates are imprecision). Many RCTs in critically ill patients showed a neutral effect of the intervention under investigation. In other situations, initial trials showed a beneficial effect of the intervention which, however, was refuted by a subsequent meta-trial. Reasons for these negative results include timing of enrolment, endpoint selection and heterogeneous subjects.[25 26]

When both RCTs and OS are available, common practice is to combine data by equally weighting these two types of studies. When evaluating protective ventilation for non-acute respiratory distress syndrome (ARDS) patients, Serpa Neto *et al*[27] combined both RCTs and observational data with equal weights. The use of such a practice is partly due to difficulties in model building under the conventional statistical framework. However, there will be more flexibility for model building under the framework of a Bayesian perspective. The advantages of Bayesian analysis include but are not limited to: (1) it allows for evidence derived from a variety of sources including RCTs and observational data; (2) it enables a direct probability statement regarding the quantity of interest; and (3) all parameter uncertainties can be automatically accounted for.[28] We believe that the present study will provide new evidence for the effectiveness of aPC on mortality in patients with sepsis.

Competing interests None.

Ethics approval The study was approved by the ethics committee of Jinhua Municipal Central Hospital.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Jawad I, Luksic I, Rafnsson SB. Assessing available information on the burden of sepsis: global estimates of incidence, prevalence and mortality. *J Glob Health* 2012;2:010404.
2. Angus DC, Linde-Zwirble WT, Lidicker J, *et al*. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. *Crit Care Med* 2001;29:1303–10.
3. Vincent JL, Sakr Y, Sprung CL, *et al*. Sepsis in European intensive care units: results of the SOAP study. *Crit Care Med* 2006;34:344–53.
4. Slade E, Tamber PS, Vincent JL. The Surviving Sepsis Campaign: raising awareness to reduce mortality. *Crit Care (London, England)* 2003;7:1–2.
5. Levy MM, Dellinger RP, Townsend SR, *et al*. The Surviving Sepsis Campaign: results of an international guideline-based performance improvement program targeting severe sepsis. *Crit Care Med* 2010;38:367–74.
6. Christiaans SC, Wagener BM, Esmon CT, *et al*. Protein C and acute inflammation: a clinical and biological perspective. *Am J Physiol Lung Cell Mol Physiol* 2013;305:L455–66.
7. Nallamothu BK, Hayward RA, Bates ER. Beyond the randomized clinical trial: the role of effectiveness studies in evaluating cardiovascular therapies. *Circulation* 2008;118:1294–303.
8. Marti-Carvajal AJ, Sola I, Gluud C, *et al*. Human recombinant protein C for severe sepsis and septic shock in adult and paediatric patients. *Cochrane Database Syst Rev* 2012;12:CD004388.
9. Lai PS, Matteau A, Iddriss A, *et al*. An updated meta-analysis to understand the variable efficacy of drotrecogin alfa (activated) in severe sepsis and septic shock. *Minerva Anestesiol* 2013;79:33–43.
10. Kalil AC, LaRosa SP. Effectiveness and safety of drotrecogin alfa (activated) for severe sepsis: a meta-analysis and metaregression. *Lancet Infect Dis* 2012;12:678–86.
11. Chen MH, Ibrahim JG. Power prior distributions for regression models. *Stat Sci* 2000;15:46–60.
12. Prasad K, Jaeschke R, Wyer P, *et al*. Tips for teachers of evidence-based medicine: understanding odds ratios and their relationship to risk ratios. *J Gen Intern Med* 2008;23:635–40.
13. Verhagen AP, de Vet HC, de Bie RA, *et al*. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *J Clin Epidemiol* 1998;51:1235–41.
14. Zhang Z, Xu X, Chen K. Lactate clearance as a useful biomarker for the prediction of all-cause mortality in critically ill patients: a systematic review study protocol. *BMJ Open* 2014;4:e004752-e.
15. Macaskill P, Walter SD, Irwig L. A comparison of methods to detect publication bias in meta-analysis. *Stat Med* 2001;20:641–54.
16. Peters JL, Sutton AJ, Jones DR, *et al*. Comparison of two methods to detect publication bias in meta-analysis. *JAMA* 2006;295:676–80.
17. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994;50:1088–101.
18. Egger M, Davey Smith G, *et al*. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629–34.
19. Peters JL, Sutton AJ, Jones DR, *et al*. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J Clin Epidemiol* 2008;61:991–6.
20. Warn DE, Thompson SG, Spiegelhalter DJ. Bayesian random effects meta-analysis of trials with binary outcomes: methods for the absolute risk difference and relative risk scales. *Stat Med* 2002;21:1601–23.
21. Welton NJ, Sutton AJ, Cooper NJ, *et al*. *Generalized evidence synthesis. Evidence synthesis for decision making in healthcare*. 1st edn. Wiley, 2012:227–33.
22. Wetterslev J, Thorlund K, Brok J, *et al*. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *J Clin Epidemiol* 2008;61:64–75.
23. Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Stat Med* 2004;23:1351–75.

24. Angus DC. Drotrecogin alfa (activated)…a sad final fizzle to a roller-coaster party. *Crit Care* 2012;16:107.
25. Vincent J-L. We should abandon randomized controlled trials in the intensive care unit. *Crit Care Med* 2010;38:S534–8.
26. Han S-H, Sevransky J. Another "negative" acute lung injury trial. *Crit Care Med* 2012;40:315–16.
27. Serpa Neto A, Cardoso SO, Manetta JA, *et al*. Association between use of lung-protective ventilation with lower tidal volumes and clinical outcomes among patients without acute respiratory distress syndrome: a meta-analysis. *JAMA* 2012;308:1651–9.
28. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. *Stat Methods Med Res* 2001;10:277–303.

BMJ Open

Recombinant human activated protein C for the treatment of severe sepsis and septic shock: a study protocol for incorporating observational evidence using a Bayesian approach

Zhongheng Zhang

BMJ Open 2014 4:
doi: 10.1136/bmjopen-2014-005622

Notes

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Potential diagnostic value of serum p53 antibody for detecting esophageal cancer: a meta-analysis	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	6
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	7
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	7
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7

PRISMA 2009 Checklist

Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency ($e.g., I^2$) for each meta-analysis.	7-8

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	8
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	9
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	9
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	8-9
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	9
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	9
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	10
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	12
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	12
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	13

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

PRISMA 2009 Checklist

For more information, visit: **www.prisma-statement.org**.

BMJ Open

The efficacy of activated protein C for the treatment of sepsis: incorporating observational evidence with Bayesian approach

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-006524.R2
Article Type:	Research
Date Submitted by the Author:	08-Dec-2014
Complete List of Authors:	Zhang, Zhongheng; Jinhua municipal central hospital, Department of critical care medicine
Primary Subject Heading:	Intensive care
Secondary Subject Heading:	Anaesthesia, Emergency medicine, Intensive care
Keywords:	Adult intensive & critical care < ANAESTHETICS, BACTERIOLOGY, INFECTIOUS DISEASES

The efficacy of activated protein C for the treatment of sepsis: incorporating observational evidence with Bayesian approach

Zhongheng ZHANG (MMed)

Affiliation: Department of critical care medicine, Jinhua municipal central hospital, Jinhua hospital of Zhejiang university, Zhejiang, P.R.China

Corresponding author: Zhongheng Zhang

Address: 351#, Mingyue Road, Jinhua, Zhejiang province, China, 321000

Phone number: 86-579-82552667

Email: zh_zhang1984@hotmail.com

Key words: Bayesian analysis; observational evidence, activated protein C, sepsis, septic shock

There are no conflicts of interest.

Abstract

Objective: The present study aimed to combine observational evidence with RCTs by using Bayesian approach.

Data sources: Electronic databases including Pubmed, Cochrane Central Register of Contrlled Trials (CENTRAL), ISI Web of Science, EMBASE and EBSCO were searched from inception to January 2014.

Study eligibility: Randomized controlled trail (RCTs) and observational studies (OS) investigating the effectiveness of aPC on mortality reduction were included for analysis.

Participants: patients with sepsis.

Intervention: aPC

Synthesis methods: Observational evidence was incorporated into analysis by using power transformed priors in a Bayesian. Trial sequential analysis (TSA) was performed to examine changes over time and whether further studies need to be conducted.

Main results: a total of 7 RCTs and 12 observational studies were included for analysis. There was moderate heterogeneity among included RCTs (I^2=48.6%, p=0.07). The pooled OR for mortality from RCTs was 1.00 (95% CI: 0.84-1.19). In observational studies, there was potential publication bias as indicated by funnel plot and the pooled OR for mortality with the use of aPC was 0.67 (95% CI: 0.62-0.72). The pooled effects sizes of RCTs were changed by using different power transform priors derived from observational evidence. When observational evidence was used at its "face value", the treatment effect of aPC was statistically significant in reducing mortality.

Conclusion: while RCT evidence showed no beneficial effect of aPC on sepsis, observational evidence showed a significant treatment effect of aPC. By using power transform priors in Bayesian model, we explicitly demonstrated how RCT evidence could be changed by observational evidence.

Registration: The protocol for the current study was registered in PROSPERO (registration number: CRD42014009562).

Article summary

1. While observational evidence shows beneficial effect of aPC on mortality reduction, RCTs failed to identify any such treatment effect.

2. By using power transform priors in Bayesian model, we explicitly demonstrated how RCT evidence could be changed by observational evidence.

3. Strengths: the study employed Bayesian approach to explicitly demonstrate how the result of RCTs can be influenced by observational evidence.

4. Limitations: it is still unknown how to discount observational evidence, namely, how to assign a value to the power of prior. The most appropriate prior will vary from study to study.

Introduction

Treatment of sepsis or septic shock is a major challenge for clinicians in the intensive care unit (ICU).(1, 2) Many strategies and drugs have been developed for their potential beneficial effects on clinical outcomes. Well-known interventions include the early goal directed therapy (EGDT) for early resuscitation of septic shock, protective ventilation strategy for sepsis-induced acute lung injury,(3) intensive dose renal replacement therapy for sepsis-induced acute kidney injury, and activated protein C for immunomodulation.(4) However, these interventions experienced a wax and wane of enthusiasm for their clinical utility. For instance, the EGDT has been a standard of care for septic shock resuscitation in the first 6 hours, which however is challenged by a recent large randomized controlled trial published in the New England Journal of medicine (NEJM).(5) This RCT was done 10 years after the original landmark EGDT on sepsis trial so it is a totally different time frame and different current practice. The same situation occurred in the field of CRRT dose. In 2000, a landmark study by Ronco C and coworkers(6) demonstrated mortality reduction in patients treated with high dose CRRT. However, the study could not be replicated in subsequent mega-trials and systematic reivew.(7)

Activated protein C is a drug with pleiotrophic biological effects and is thought to play an important role in the modulation of inflammatory response.(8) Early observational studies, as well as a large randomized controlled trial (RCT) demonstrated remarkable mortality reduction by using this drug.(9-11) The well-known PROWESS trial has urged approval of this drug by the Food and Drug administration (FDA) for septic shock patients.(9) However, the beneficial effect of aPC cannot be replicated in subsequent RCTs (12, 13). Several meta-analyses including one published in Cochrane library have consistently refuted the effectiveness of aPC for septic patients. As a result, it was withdrawn from the market (14, 15). Although RCTs are considered to be the gold standard for testing treatment efficacy, they have limitations. RCTs are often not conducted in “real world” settings as reflected by strict inclusion/exclusion criteria, performance in specialized centers,

and complicated intervention protocol. In contrast, observational studies are often performed in "real world" setting that patients enrolled in studies are just as they are treated in practice. Thus, some authors have suggested that observational studies should be considered in evidence synthesis, particularly when the intervention or clinical condition is complicated. Our previous analysis also showed that there is significant difference in treatment effect sizes between RCTs and OS (16).

In the present study we performed evidence synthesis by incorporating evidence from observational studies, and the observational evidence was down-weighted by using alphas ranging from 0 to 1. While there is no prior weighting for observational evidence with alpha value equals 0, observational evidence is incorporated at its face value (equal prior weighting) with alpha value equals 1. Bayesian analysis allowed such calculation by using observational evidence as the informative prior. The main purpose of the study is to examine how results derived from RCTs can be changed by assigning different degrees of skepticism to observational evidence. Another purpose of the study was to perform trial sequential analysis (TSA) to quantify the reliability of data in meta-analysis adjusting significance levels for sparse data and multiple testing on accumulating trials.

Methods

Amendment to the protocol

The study protocol has been published previously and amendment to the protocol was made during data analysis (17). The protocol for the current study was registered in PROSPERO (registration number: CRD42014009562). Herein, we explicitly listed the amendment to the protocol.

1) Quality assessment was not performed in the present analysis because the quality has been well described in a previous Cochrane systematic review (15).
2) Sensitivity analysis by excluding poor quality studies was not performed because the present study was aimed to explicitly display how the evidence derived from RCT could be modified by observational evidence. Sensitivity

analysis of this kind belonged to the realm of systematic review involving only RCTs.

Searching strategy and study selection

Electronic databases including Pubmed, Cochrane Central Register of Contrlled Trials (CENTRAL), ISI Web of Science, EMBASE and EBSCO were searched from inception to January 2014. Our core search consists of terms related to activated protein C and sepsis. References of systematic reviews were reviewed for identifying additional eligible articles.

Randomized controlled trial (RCTs) and observational studies (OS) investigating the effectiveness of aPC on mortality reduction were included for analysis. OS included: 1) cohort studies using multivariable analysis with aPC treatment as one of the covariates; 2) cohort studies using propensity analysis; 3) case-control studies; 4) both prospective and retrospective designs were considered eligible.

The following data were extracted from original articles: name of the first author, year of publication, sample size, number of deaths in each arm, total number of participants in each arm, odds ratio of treatment versus non-treatment for mortality, the method used for covariate adjustment (propensity score analysis, logistic regression model), and design of observational study (prospective vs. retrospective).

Publication bias was assessed using the Egger regression test and Begg rank correlation test. Contour enhanced funnel plot were depicted to visually assess the presence of publication bias.

Statistical analysis

Observational evidence was used as the informative prior in Bayesian analysis. The model involved power transformation of observational data likelihood as proposed by Chen and Ibrahim.(18) Full details of calculations and the WinBugs codes were described elsewhere.(17) Trial sequential analysis (TSA) was also performed to quantify the reliability of data in meta-analysis adjusting significance levels for sparse data and multiple testing on accumulating trials.(19) Statistical analysis was performed by using WinBUGS (Imperial College & MRC, UK) and Stata 12.0

(College Station, Texas 77845 USA).

Trial sequential monitoring boundaries were employed to control the risks for type I and II errors and to indicate whether additional trials are needed. The information size calculation requires the mortality rate in the control group and the minimal effect size for the intervention. We predefined that the mortality in the control group is 30%, and the intervention is able to reduce the relative risk by 15%. The conventional α and β are 0.05 and 0.2, respectively. Meta-analysis will be updated by adding component studies sequentially in the order of publication. The β-spending function was constructed to indicate futility of intervention. Trial sequential analysis was performed by using the software TSA version 0.9 Beta (Copenhagen Trial Unit, 2011).

Results

Our initial search identified a total of 531 distinct citations, and 456 of them were excluded immediately after inspection of the title and abstract (figure 1). The remaining 75 clinical studies were potentially eligible and were examined for full text. Fifty-six studies were excluded because: 1) eight studies used duplicated report; 2) 18 studies used inappropriate control arm (e.g. single arm, all patients received aPC); 3) 19 did not report mortality as the endpoint; and 4) 11 did not include aPC as an intervention. As a result, a total of 7 RCTs(9, 12, 13, 20-23) and 11 observational studies(10, 11, 24-32) were included for analysis. Characteristics of RCTs are shown in table 1.

Figure 2 shows the efficacy of aPC on mortality reduction, reported separately for RCTs and observational studies. Of the 7 RCTs, only the PROWESS study showed significant mortality reduction with aPC (OR: 0.74, 95% CI: 0.59-0.91),(9) and the other six studies failed to conclude a beneficial effect. There was moderate heterogeneity among included RCTs (I^2=48.6%, p=0.07). The pooled OR for mortality was 1.00 (95% CI: 0.84-1.19).

In contrast, 6 out of the 11 observational studies showed significant mortality reduction with the use of aPC; and the remaining five studies showed a trend towards better lower mortality rate in aPC group. The heterogeneity was statistically

significant with an I^2 of 70.8% (p<0.001). The pooled OR for mortality with the use of aPC was 0.67 (95% CI: 0.62-0.72). Egger's test did not show evidence of publication bias in RCTs (p=0.808) and observational studies (p=0.145). Similarly, Begg's test did not show evidence of publication bias in RCTs (p=0.293) and observational studies (p=0.337). However, publication bias was suspected for observational studies as suggested by the funnel plot in which each dot represents a study and they gathered at the upper left corner (figure 3).

The result of sequential trial analysis is shown in figure 4. Studies were displayed sequentially by their publication year from left to the right of the horizontal line. After publication of the first and second studies (PROWESS 2001 and rhAPC sepsis 2001), the Z score crossed the conventional significance boundary (Z=1.96) but did not cross the O'Brien-Fleming boundaries. With the publication of the study Dhainaut 2009, the Z-score reached and crossed the futility line, indicating no effect of the aPC for mortality reduction in septic patients.

Meta-analysis of observational studies was performed by using Bayesian approach. The posterior distribution of individual OR was shrunken, as reflected by the narrower credible interval of study level estimates as compared to the observed estimates. For instance, the credible interval of OR in the study de Pont AC 2005 was 0.40-1.06, which was significantly narrower than the observed confidence interval of 0.04-6.70 (figure 2). This was because each component study borrowed evidence from the overall effect by using Bayesian approach. The overall OR was 0.67 (credible interval: 0.56-0.78).

Figure 5 shows the mean OR and 95% credible interval (CrI) for different power transformation priors to down-weight observational evidence on the risk of death with aPC. To the left of the figure when alpha took negligible values, the observational evidence was totally discounted and the mean OR was 1, which was consistent with the pooled result from RCTs. Increasing weight was assigned to observational evidence with increasing alpha values. We could see from the figure that the upper limit of CrI crossed the reference line with alpha values<0.4. When observational evidence was combined at its face value (alpha=1), the aPC group showed significant

mortality reduction as compared with the control group. The alpha value influenced the precision of prior evidence. As shown in figure 6, the precision of prior increased with increasing value of alpha from 0.000001 to 1.

Discussion

Key findings of the present analysis are 1) aPC appears to be able to reduce mortality rate when evidence is pooled from observational studies, and the results are consistent by using conventional Bayesian approaches; 2) RCTs failed to identify any beneficial effect of aPC; 3) observational evidence, when discounted by different power transformation priors, can alter the conclusion derived from RCTs. 4) With trial sequential analysis, the positive result (significant beneficial effect of aPC) as shown in the PROWESS study should be interpreted with caution.

One potential explanation for the positive findings in observational studies is the publication bias as shown in figure 3. The funnel contour plot showed that most observational studies located in the region with p<1%, indicating that the asymmetrical distribution was more likely due to publication bias. It is not surprising that observational studies are more subject to publication bias in that they are less likely to be registered a priori.(33) In contrast, RCTs are usually registered and there are many online registration sites.(34) The value of observational studies is usually discounted in evidence synthesis, and the conventional view is that observational evidence can only serve as hypothesis-generating. In such context, if the finding of an observational study is neutral, it will be less interesting to readers and journals, making it less likely to be published. In contrast, because RCTs are always registered and requires large amount of cost and other resources, studies with negative findings can be published and is equally important to those with positive findings.

Activated protein C (aPC) for the treatment of sepsis is a good example illustrating the importance of using sequential trial analysis in evidence synthesis. aPC was approved by the food and drug administration after publication of PROWESS trial,

which seemed too hasty when viewed retrospectively. Although the initial trial was positive at conventional significance level of p=0.05 (Z=1.96), its statistical significance should be tested by using adjusted alpha level. In sequential trial analysis, this is achieved by using alpha-spending function and constructing the O'Brien-Fleming boundaries. If sequential trial analysis had been performed at the conclusion of PROWESS trial, the approval of aPC for sepsis would not be so hasty. It has been argued that the disparity between PROWESS trial and subsequent trials such as PROWESS-SHOCK could be explained by the heterogeneity of enrolled subjects.(35)

In the translation of research into clinical practice, there are a lot of influence factors that to consider. RCT is generally accepted as a gold standard. However, there are some limitations in real clinical practice that RCT cannot simulate all the clinical situations. The biggest problem is that RCT is usually conducted in non-real world setting, that is, it is often performed in specialized academic centers with strict inclusion/exclusion criteria. For example, in the Dhainaut 2009 study there was a long list of exclusion criteria, including expected surgical procedure in the next 3 days, platelet count<30,000/nm^3, receiving therapeutic heparin, moribund, withdrawn from aggressive management by patients' family, and pregnant or breast feeding. Such strict exclusion criteria would exclude most of patients with septic shock. Therefore, it appears unfair to treat our septic shock patients based on evidence derived from a minority of the population. In this situation, observational studies have its advantage in testing the clinical effectiveness of aPC on mortality reduction (36-38). Observational study included wider range of patients with septic shock and the setting is just like what we will encounter in routine clinical practice. For example, in prospective RCTs patients with comorbidities were strictly screened and excluded, but in retrospective studies it is often unreliable to exclude certain comorbidities based on medical records. Therefore, the observational evidence cannot be simply ignored in evidence synthesis for decision-making. However, the result of observational studies could be misleading due to inherent bias. In our study the funnel plot showed

asymmetrically distributed component studies, indicating potential publication bias. In this regard, the observational evidence should be interpreted with caution and in conjunction with other evidence. Since there was no consensus on how to combine observational evidence with RCTs, we discounted observational evidence with power transform priors taking advantage of the flexibility in Bayesian modeling (18). In this model, we found that the treatment effect of aPC increased with more weight assigning to observational evidence (figure 6). A value of 0 for alpha implies that the observational evidence is ignored, and a value of 1 for alpha means that observational evidence is accepted at its "face value". This approach gives a full picture of how pooled evidence can be altered by observational studies, by explicitly showing the power transform priors.

Several limitations of the study need to be acknowledged. First, there are substantial heterogeneity among included RCTs, which may be explained by the differences in study population, timing of intervention and definition of study endpoint. As expected, observational studies showed substantial heterogeneity. Considering the very different study criteria and the various geographic sources of each observational study (representing different standards of care), heterogeneity should be expected due to the more generalizable (real-world) evidence. Second, it is still largely unknown on how to discount the observational evidence. The most appropriate prior will vary from study to study. The present study only displays a wide range of possible alpha values and explicitly demonstrates how RCT evidence can be modified by observational evidence.

In summary, our study demonstrates that there is considerable disparity between observational and RCT evidence. While observational evidence shows beneficial effect of aPC on mortality reduction, RCTs failed to identify any such treatment effect.

Figure legends

Figure 1. Flow chart of study selection.

Figure 2. Forest plots showing the efficacy of aPC on mortality reduction, reported separately for RCTs and observational studies.

Figure 3. Contour funnel plots showing the publication bias in RCTs and observational studies. Publication bias was identified for observational studies as reflected by the asymmetrically distributed component studies.

Figure 4. Sequential trial analysis involving randomized controlled trials showing that the Z-score crossed the futility line after the study Dhainaut 2009. Parameters used for the creation of boundaries were: type: Two-sided; type 1 Error: 5.0%; alpha spending: O'Brien-Fleming; information axis: sample size; power: 80.0%; effect type intervention: RRR User Defined (15%). The shaded area indicates futility area.

Figure 5. Mean OR and 95% credible interval (CrI) for different power transformation priors to down-weight observational evidence on the risk of death with aPC. Lower values of alpha down-weight the observational evidence

Figure 6. Prior distribution derived by discounting observational evidence with alpha from 0.000001 to 1. The plots shows that the precision of prior increases with increasing alpha values.

Reference

1. Schorr CA, Zanotti S, Dellinger RP. Severe sepsis and septic shock: management and performance improvement. Virulence. 2014;5(1):190-9.
2. Dellinger RP, Levy MM, Rhodes A, Annane D, Gerlach H, Opal SM, et al. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012. Critical care medicine. 2013;41(2):580-637.
3. MacIntyre N. Ventilatory Management of ALI/ARDS. Seminars in respiratory and critical care medicine. 2006;27(4):396-403.
4. Rivers E, Nguyen B, Havstad S, Ressler J, Muzzin A, Knoblich B, et al. Early goal-directed therapy in the treatment of severe sepsis and septic shock. The New England journal of medicine. 2001;345(19):1368-77.
5. Pro CI, Yealy DM, Kellum JA, Huang DT, Barnato AE, Weissfeld LA, et al. A randomized trial of protocol-based care for early septic shock. The New England journal of medicine. 2014;370(18):1683-93.
6. Ronco C, Bellomo R, Homel P, Brendolan A, Dan M, Piccinni P, et al. Effects of different doses in continuous veno-venous haemofiltration on outcomes of acute renal failure: a prospective randomised trial. Lancet. 2000;356(9223):26-30.
7. Zhang Z, Xu X, Zhu H. Intensive- vs less-intensive-dose continuous renal replacement therapy for the intensive care unit-related acute kidney injury: a meta-analysis and systematic review. Journal of critical care. 2010;25(4):595-600.
8. Christiaans SC, Wagener BM, Esmon CT, Pittet JF. Protein C and acute inflammation: a clinical and biological perspective. American journal of physiology Lung cellular and molecular physiology. 2013;305(7):L455-66.
9. Bernard GR, Vincent JL, Laterre PF, LaRosa SP, Dhainaut JF, Lopez-Rodriguez A, et al. Efficacy and safety of recombinant human activated protein C for severe sepsis. The New England journal of medicine. 2001;344(10):699-709.
10. Dhainaut JF, Payet S, Vallet B, Franca LR, Annane D, Bollaert PE, et al. Cost-effectiveness of activated protein C in real-life clinical practice. Critical care (London, England). 2007;11(5):R99.
11. Kubler A, Mayzner-Zawadzka E, Durek G, Gaszynski W, Karpel E, Mikaszewska-Sokolewicz M, et al. Results of severe sepsis treatment program using recombinant human activated protein C in Poland. Medical science monitor : international medical journal of experimental and clinical research. 2006;12(3):CR107-12.
12. Annane D, Timsit JF, Megarbane B, Martin C, Misset B, Mourvillier B, et al. Recombinant human activated protein C for adults with septic shock: a randomized controlled trial. American journal of respiratory and critical care medicine. 2013;187(10):1091-7.
13. Ranieri VM, Thompson BT, Barie PS, Dhainaut JF, Douglas IS, Finfer S, et al. Drotrecogin alfa (activated) in adults with septic shock. The New England journal of medicine. 2012;366(22):2055-64.
14. Lai PS, Matteau A, Iddriss A, Hawes JC, Ranieri V, Thompson BT. An updated meta-analysis to understand the variable efficacy of drotrecogin alfa (activated) in severe sepsis and septic shock. Minerva anestesiologica. 2013;79(1):33-43.
15. Marti-Carvajal AJ, Sola I, Gluud C, Lathyris D, Cardona AF. Human recombinant protein C for severe sepsis and septic shock in adult and paediatric patients. The Cochrane database of systematic reviews. 2012;12:CD004388.
16. Zhang Z, Ni H, Xu X. Do the observational studies using propensity score analysis agree with

randomized controlled trials in the area of sepsis? Journal of critical care. 2014;29(5):886 e9-15.
17. Zhang Z. Recombinant human activated protein C for the treatment of severe sepsis and septic shock: a study protocol for incorporating observational evidence using a Bayesian approach. BMJ Open. 2014;4(7):e005622.
18. Chen MH, Ibrahim JG. Power prior distributions for regression models. statistical science. 2000;15(1):46-60.
19. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. Journal of clinical epidemiology. 2008;61(1):64-75.
20. Bernard GR, Ely EW, Wright TJ, Fraiz J, Stasek JE, Jr., Russell JA, et al. Safety and dose relationship of recombinant human activated protein C for coagulopathy in severe sepsis. Critical care medicine. 2001;29(11):2051-9.
21. Abraham E, Laterre PF, Garg R, Levy H, Talwar D, Trzaskoma BL, et al. Drotrecogin alfa (activated) for adults with severe sepsis and a low risk of death. The New England journal of medicine. 2005;353(13):1332-41.
22. Nadel S, Goldstein B, Williams MD, Dalton H, Peters M, Macias WL, et al. Drotrecogin alfa (activated) in children with severe sepsis: a multicentre phase III randomised controlled trial. Lancet. 2007;369(9564):836-43.
23. Dhainaut JF, Antonelli M, Wright P, Desachy A, Reignier J, Lavoue S, et al. Extended drotrecogin alfa (activated) treatment in patients with prolonged septic shock. Intensive care medicine. 2009;35(7):1187-95.
24. de Pont AC, Bakhtiari K, Hutten BA, de Jonge E, Vroom MB, Meijers JC, et al. Recombinant human activated protein C resets thrombin generation in patients with severe sepsis - a case control study. Critical care (London, England). 2005;9(5):R490-7.
25. Bertolini G, Rossi C, Anghileri A, Livigni S, Addis A, Poole D. Use of Drotrecogin alfa (activated) in Italian intensive care units: the results of a nationwide survey. Intensive care medicine. 2007;33(3):426-34.
26. Rowan KM, Welch CA, North E, Harrison DA. Drotrecogin alfa (activated): real-life use and outcomes for the UK. Critical care (London, England). 2008;12(2):R58.
27. Vincent JL, Laterre PF, Decruyenaere J, Spapen H, Raemaekers J, Damas F, et al. A registry of patients treated with drotrecogin alfa (activated) in Belgian intensive care units--an observational study. Acta clinica Belgica. 2008;63(1):25-30.
28. Martin G, Brunkhorst FM, Janes JM, Reinhart K, Sundin DP, Garnett K, et al. The international PROGRESS registry of patients with severe sepsis: drotrecogin alfa (activated) use and patient outcomes. Critical care (London, England). 2009;13(3):R103.
29. Lindenauer PK, Rothberg MB, Nathanson BH, Pekow PS, Steingrub JS. Activated protein C and hospital mortality in septic shock: a propensity-matched analysis. Critical care medicine. 2010;38(4):1101-7.
30. Sadaka F, O'Brien J, Migneron M, Stortz J, Vanston A, Taylor RW. Activated protein C in septic shock: a propensity-matched analysis. Critical care (London, England). 2011;15(2):R89.
31. Casserly B, Gerlach H, Phillips GS, Marshall JC, Lemeshow S, Levy MM. Evaluating the use of recombinant human activated protein C in adult severe sepsis: results of the Surviving Sepsis Campaign. Critical care medicine. 2012;40(5):1417-26.
32. Rimmer E, Kumar A, Doucette S, Marshall J, Dial S, Gurka D, et al. Activated protein C and septic shock: a propensity-matched cohort study*. Critical care medicine. 2012;40(11):2974-81.

33. Onukwugha E. Improving confidence in observational studies : should statistical analysis plans be made publicly available? PharmacoEconomics. 2013;31(3):177-9.

34. Anand V, Scales DC, Parshuram CS, Kavanagh BP. Registration and design alterations of clinical trials in critical care: a cross-sectional observational study. Intensive care medicine. 2014;40(5):700-22.

35. Kalil AC, Florescu DF. Severe sepsis: are PROWESS and PROWESS-SHOCK trials comparable? A clinical and statistical heterogeneity analysis. Critical care (London, England). 2013;17(4):167.

36. Albert RK. "Lies, damned lies ..." and observational studies in comparative effectiveness research. Am J Respir Crit Care Med. 2013;187(11):1173-7.

37. Nallamothu BK, Hayward RA, Bates ER. Beyond the randomized clinical trial: the role of effectiveness studies in evaluating cardiovascular therapies. Circulation. 2008;118(12):1294-303.

38. Zhang Z. Big data and clinical research: focusing on the area of critical care medicine in mainland China. Quantitative Imaging in Medicine and Surgery. 2014;4(5):426-9.

Funding

There was no funding for the present study.

Table 1 Characteristics of included randomized controlled trials

Studies	Patients (n)	Mean age (years)	Population	Mean APACHE II score	Control	Primary outcome	Baseline mortality (%)
Bernard GR 2001 (rhAPC)	131	59.3	Severe shock	17.3	Placebo	Coagulopathy	34.2
Bernard GR 2001 (PROWESS)	1690	60.5	Systemic inflammation and organ failure	24.8	Placebo (saline or albumin)	28-day all cause mortality	30.8
Ranieri VM 2012	1697	63.1	Sepsis and shock receiving fluids and vasopressor	25.3	Placebo (saline)	28-day all cause mortality	24.2
Abraham E 2005	2613	58.7	Severe sepsis and single organ failure or Mean APACHE II<25	18.2	Placebo (saline)	28-day all cause mortality	17
Nadel S 2007	477	2.5	Children with sepsis induced	-	Placebo (saline)	CTCOFR	17.5

			cardiac or respiratory failure				
Annane D 2013	411	63	Sepsis with >2 organ failure	-	Placebo (saline)	90-day mortality	46.3
Dhainaut JF 2009	193	62.4	Severe sepsis with vasopressor dependent hypotension	28.1	Placebo	Resolution of vasopressor dependent hypotension	32.3

APACHE: Acute Physiology and Chronic Health Evaluation; CTCOFR: Composite Time to Complete Organ Failure Resolution.

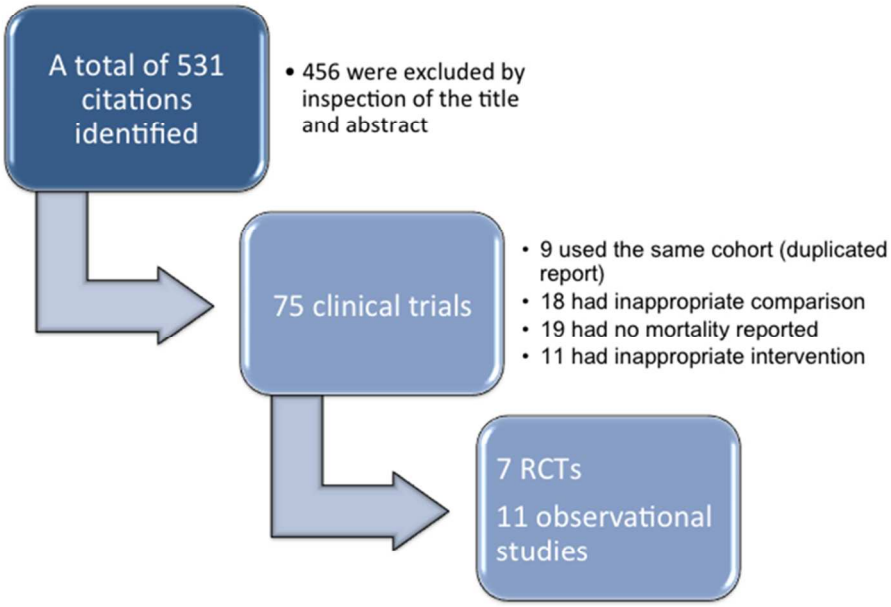

Figure 1. Flow chart of study selection.
84x63mm (300 x 300 DPI)

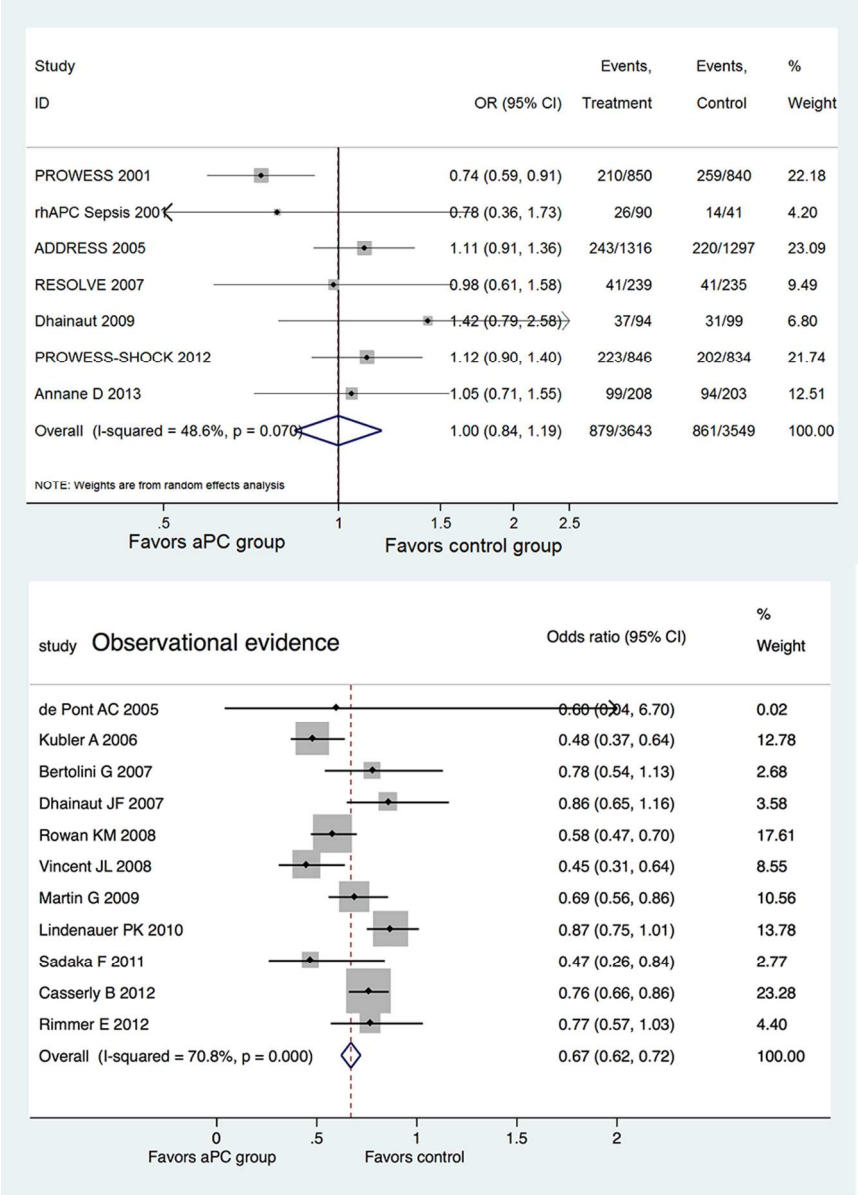

Figure 2. Forest plots showing the efficacy of aPC on mortality reduction, reported separately for RCTs and observational studies. Results were pooled by using conventional meta-analytic approach.
84x118mm (300 x 300 DPI)

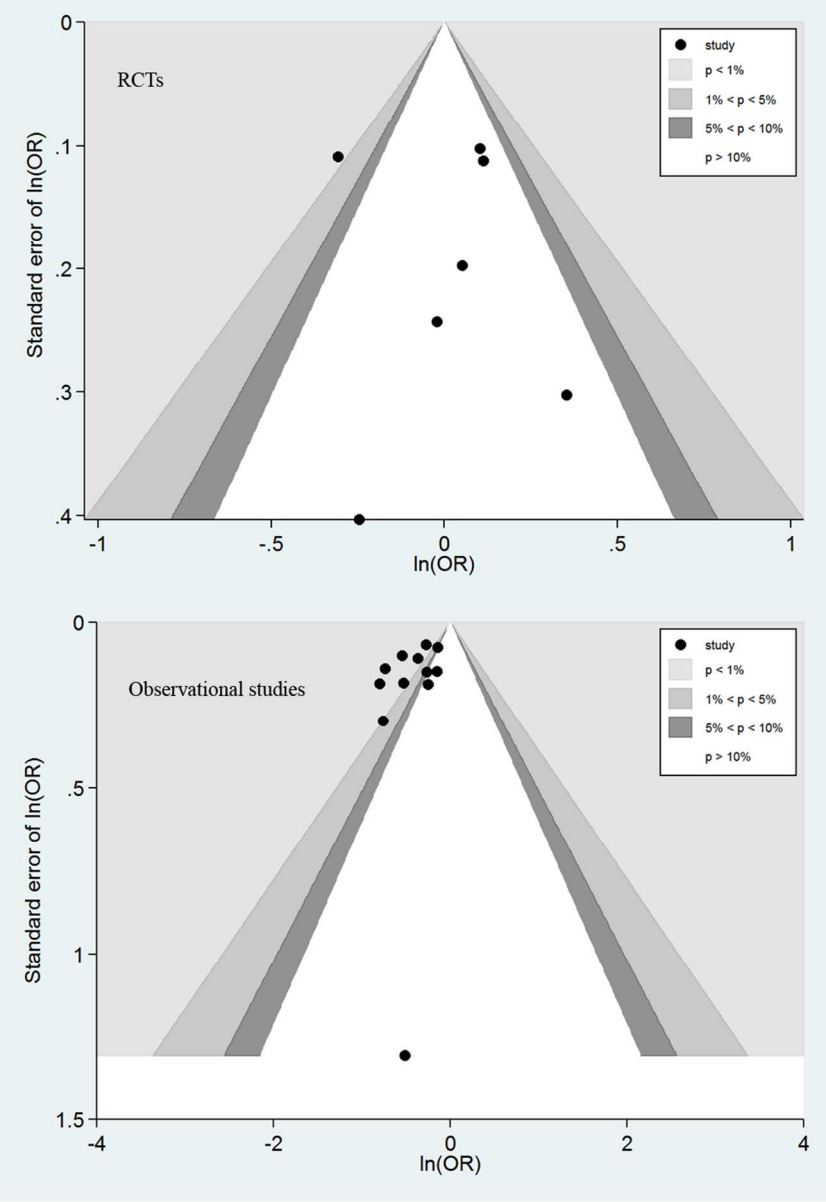

Figure 3. Contour funnel plots showing the publication bias in RCTs and observational studies. Publication bias was identified for observational studies as reflected by the asymmetrically distributed component studies.
84x123mm (300 x 300 DPI)

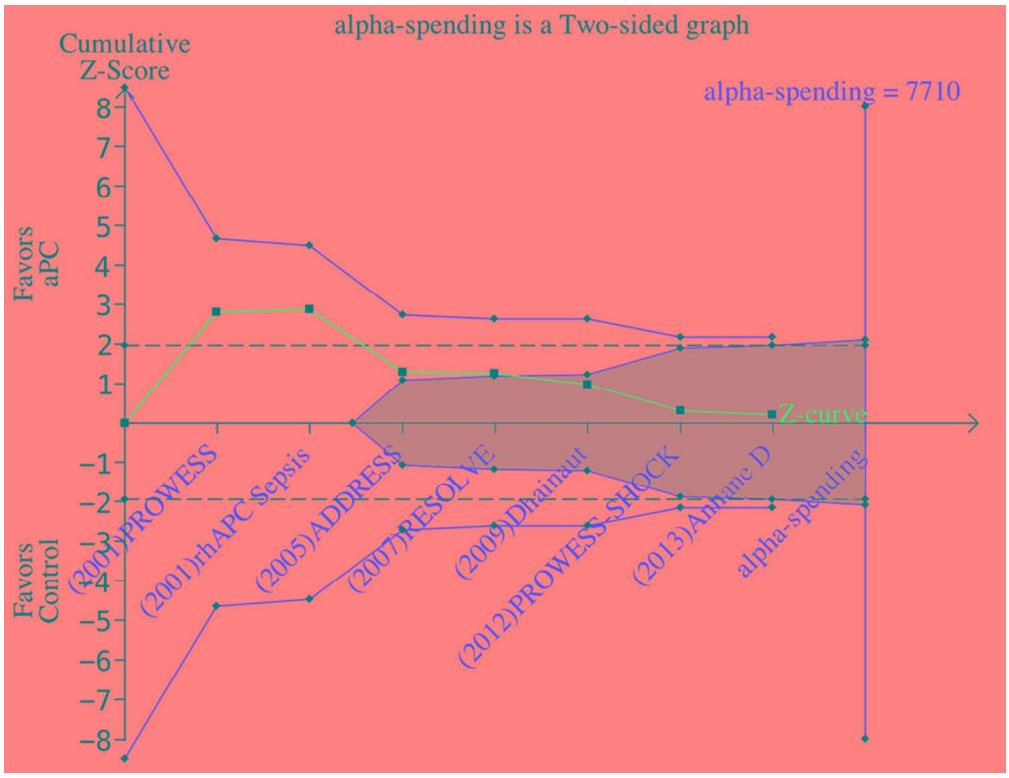

Figure 4. Sequential trial analysis involving randomized controlled trials showing that the Z-score crossed the futility line after the study Dhainaut 2009. Parameters used for the creation of boundaries were: type: Two-sided; type 1 Error: 5.0%; alpha spending: O'Brien-Fleming; information axis: sample size; power: 80.0%; effect type intervention: RRR User Defined (15%). The shaded area indicates futility area.
127x97mm (200 x 200 DPI)

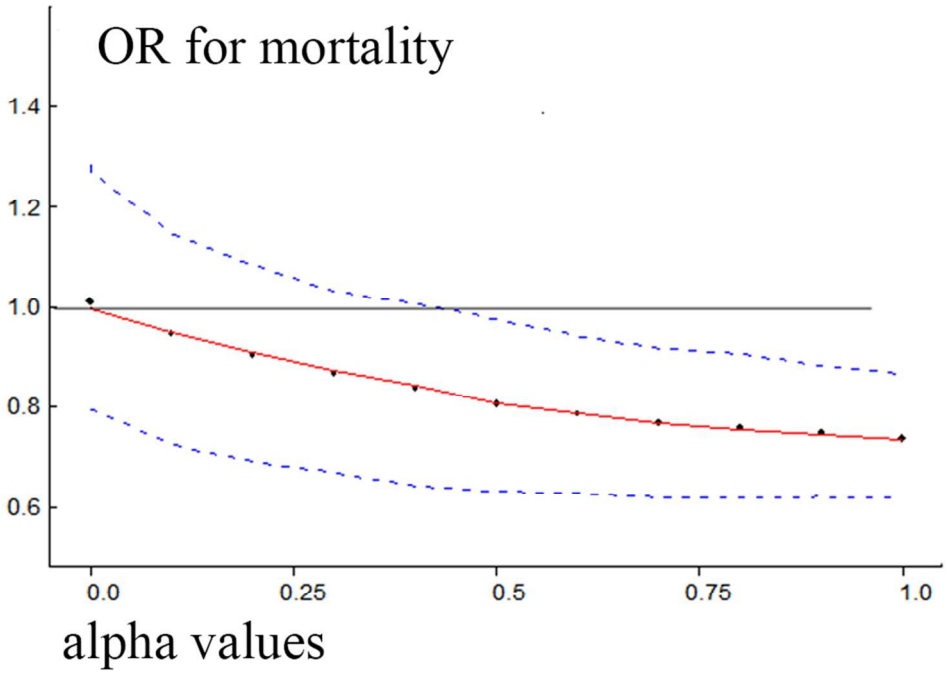

Figure 5. Mean OR and 95% credible interval (CrI) for different power transformation priors to down-weight observational evidence on the risk of death with aPC.
84x58mm (300 x 300 DPI)

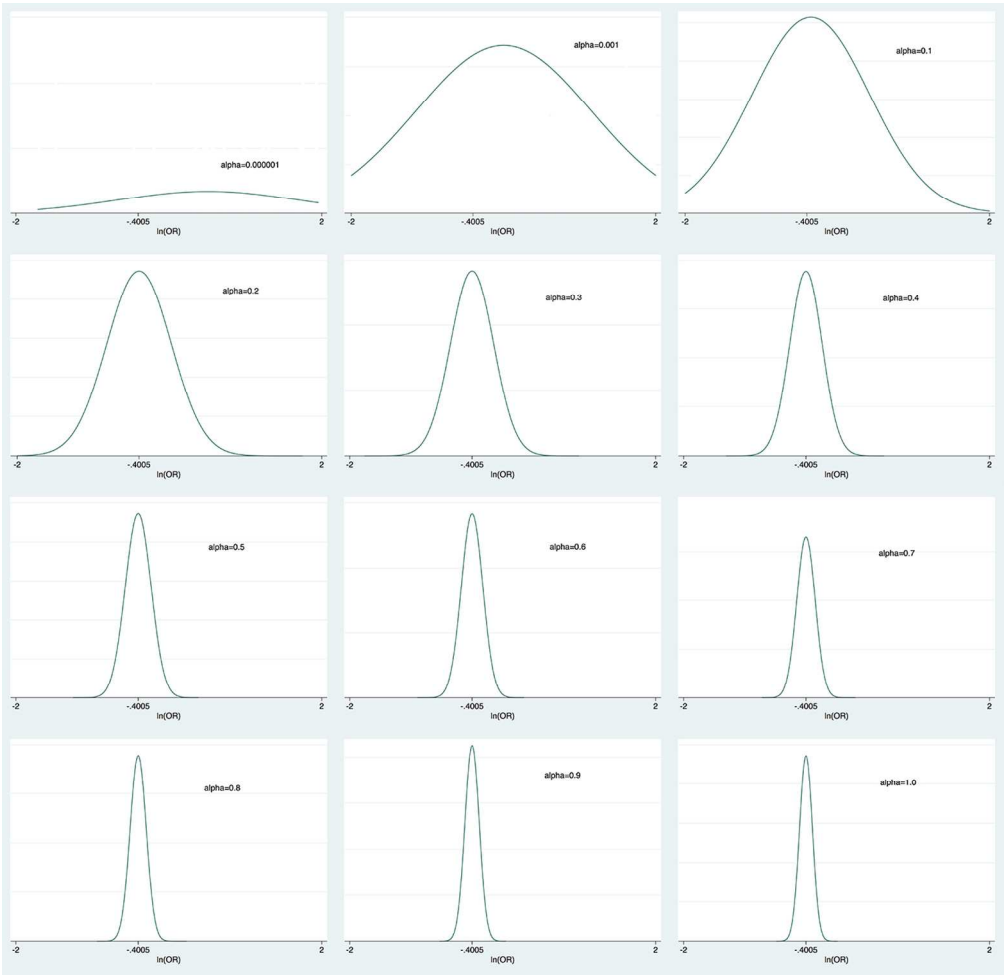

Figure 6. Prior distribution derived by discounting observational evidence with alpha from 0.000001 to 1. The plots shows that the precision of prior increases with increasing alpha values.
127x123mm (300 x 300 DPI)

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	The efficacy of activated protein C for the treatment of sepsis: incorporating observational evidence with Bayesian approach	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	6
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	7
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	7
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency	7-8

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	8
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	9
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	9
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	8-9
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	9
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	9
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	10
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	12
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	12
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	16

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

BMJ Open

The efficacy of activated protein C for the treatment of sepsis: incorporating observational evidence with a Bayesian approach

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-006524.R3
Article Type:	Research
Date Submitted by the Author:	21-Dec-2014
Complete List of Authors:	Zhang, Zhongheng; Jinhua municipal central hospital, Department of critical care medicine
Primary Subject Heading:	Intensive care
Secondary Subject Heading:	Anaesthesia, Emergency medicine, Intensive care
Keywords:	Adult intensive & critical care < ANAESTHETICS, BACTERIOLOGY, INFECTIOUS DISEASES

The efficacy of activated protein C for the treatment of sepsis: incorporating observational evidence with a Bayesian approach

Zhongheng ZHANG (MMed)

Affiliation: Department of critical care medicine, Jinhua municipal central hospital, Jinhua hospital of Zhejiang university, Zhejiang, P.R.China

Corresponding author: Zhongheng Zhang

Address: 351#, Mingyue Road, Jinhua, Zhejiang province, China, 321000

Phone number: 86-579-82552667

Email: zh_zhang1984@hotmail.com

Key words: Bayesian analysis; observational evidence, activated protein C, sepsis, septic shock

There are no conflicts of interest.

Abstract

Objective: The present study aimed to combine observational evidence with RCTs by using Bayesian approach.

Data sources: Electronic databases including Pubmed, Cochrane Central Register of Contrlled Trials (CENTRAL), ISI Web of Science, EMBASE and EBSCO were searched from inception to January 2014.

Study eligibility: Randomized controlled trail (RCTs) and observational studies (OS) investigating the effectiveness of aPC on mortality reduction were included for analysis.

Participants: patients with sepsis.

Intervention: aPC

Synthesis methods: Observational evidence was incorporated into analysis by using power transformed priors in a Bayesian. Trial sequential analysis (TSA) was performed to examine changes over time and whether further studies need to be conducted.

Main results: a total of 7 RCTs and 12 observational studies were included for analysis. There was moderate heterogeneity among included RCTs (I^2=48.6%, p=0.07). The pooled OR for mortality from RCTs was 1.00 (95% CI: 0.84-1.19). In observational studies, there was potential publication bias as indicated by funnel plot and the pooled OR for mortality with the use of aPC was 0.67 (95% CI: 0.62-0.72). The pooled effects sizes of RCTs were changed by using different power transform priors derived from observational evidence. When observational evidence was used at its "face value", the treatment effect of aPC was statistically significant in reducing mortality.

Conclusion: while RCT evidence showed no beneficial effect of aPC on sepsis, observational evidence showed a significant treatment effect of aPC. By using power transform priors in Bayesian model, we explicitly demonstrated how RCT evidence could be changed by observational evidence.

Registration: The protocol for the current study was registered in PROSPERO (registration number: CRD42014009562).

Article summary

1. While observational evidence shows a beneficial effect of aPC on mortality reduction, RCTs failed to identify any such treatment effect.

2. By using power transform priors in a Bayesian model, we explicitly demonstrated how RCT evidence could be changed by observational evidence.

3. Strengths: the study employed Bayesian approach to explicitly demonstrate how the result of RCTs can be influenced by observational evidence.

4. Limitations: it is still unknown how to discount observational evidence, namely, how to assign a value to the power of prior. The most appropriate prior will vary from study to study.

Introduction

Treatment of sepsis or septic shock is a major challenge for clinicians in the intensive care unit (ICU).(1, 2) Many strategies and drugs have been developed for their potential beneficial effects on clinical outcomes. Well-known interventions include the early goal directed therapy (EGDT) for early resuscitation of septic shock, protective ventilation strategy for sepsis-induced acute lung injury,(3) intensive dose renal replacement therapy for sepsis-induced acute kidney injury, and activated protein C for immunomodulation.(4) However, these interventions experienced a wax and wane of enthusiasm for their clinical utility. For instance, the EGDT has been a standard of care for septic shock resuscitation in the first 6 hours, which however is challenged by a recent large randomized controlled trial published in the New England Journal of medicine (NEJM).(5) This RCT was done 10 years after the original landmark EGDT on sepsis trial so it is a totally different time frame and different current practice. The same situation occurred in the field of CRRT dose. In 2000, a landmark study by Ronco C and coworkers(6) demonstrated mortality reduction in patients treated with high dose CRRT. However, the study could not be replicated in subsequent mega-trials and systematic reivew.(7)

Activated protein C is a drug with pleiotrophic biological effects and is thought to play an important role in the modulation of inflammatory response.(8) Early observational studies, as well as a large randomized controlled trial (RCT) demonstrated remarkable mortality reduction by using this drug.(9-11) The well-known PROWESS trial has urged approval of this drug by the Food and Drug administration (FDA) for septic shock patients.(9) However, the beneficial effect of aPC cannot be replicated in subsequent RCTs (12, 13). Several meta-analyses including one published in Cochrane library have consistently refuted the effectiveness of aPC for septic patients. As a result, it was withdrawn from the market (14, 15). Although RCTs are considered to be the gold standard for testing treatment efficacy, they have limitations. RCTs are often not conducted in “real world” settings as reflected by strict inclusion/exclusion criteria, performance in specialized centers,

and complicated intervention protocol. In contrast, observational studies are often performed in "real world" setting that patients enrolled in studies are just as they are treated in practice. Thus, some authors have suggested that observational studies should be considered in evidence synthesis, particularly when the intervention or clinical condition is complicated. Our previous analysis also showed that there is significant difference in treatment effect sizes between RCTs and OS (16).

In the present study we performed evidence synthesis by incorporating evidence from observational studies, and the observational evidence was down-weighted by using alphas ranging from 0 to 1. No prior weighting for observational evidence uses an alpha of 0, and observational evidence is incorporated at its face value (equal prior weighting) with an alpha value of 1. Bayesian analysis allowed such calculation by using observational evidence as the informative prior. The main purpose of the study is to examine how results derived from RCTs can be changed by assigning different degrees of skepticism to observational evidence. Another purpose of the study was to perform trial sequential analysis (TSA) to examine the changes over time and whether further studies need to be conducted, by adjusting significance levels for sparse data and multiple testing on accumulating trials.

Methods

Amendment to the protocol

The study protocol has been published previously and amendment to the protocol was made during data analysis (17). The protocol for the current study was registered in PROSPERO (registration number: CRD42014009562). Herein, we explicitly listed the amendment to the protocol.

1) Quality assessment was not performed in the present analysis because the quality has been well described in a previous Cochrane systematic review (15).
2) Sensitivity analysis by excluding poor quality studies was not performed because the present study was aimed to explicitly display how the evidence derived from RCT could be modified by observational evidence. Sensitivity analysis of this kind belonged to the realm of systematic review involving only

RCTs.

Searching strategy and study selection

Electronic databases including Pubmed, Cochrane Central Register of Controlled Trials (CENTRAL), ISI Web of Science, EMBASE and EBSCO were searched from inception to January 2014. Our core search consists of terms related to activated protein C and sepsis. References of systematic reviews were reviewed for identifying additional eligible articles.

Randomized controlled trial (RCTs) and observational studies (OS) investigating the effectiveness of aPC on mortality reduction were included for analysis. OS included: 1) cohort studies using multivariable analysis with aPC treatment as one of the covariates; 2) cohort studies using propensity analysis; 3) case-control studies; 4) both prospective and retrospective designs were considered eligible.

The following data were extracted from original articles: name of the first author, year of publication, sample size, number of deaths in each arm, total number of participants in each arm, odds ratio of treatment versus non-treatment for mortality, the method used for covariate adjustment (propensity score analysis, logistic regression model), and design of observational study (prospective vs. retrospective).

Publication bias was assessed using the Egger regression test and Begg rank correlation test. Contour enhanced funnel plot were depicted to visually assess the presence of publication bias.

Statistical analysis

Observational evidence was used as the informative prior in Bayesian analysis. The model involved power transformation of observational data likelihood as proposed by Chen and Ibrahim.(18) Full details of calculations and the WinBugs codes were described elsewhere.(17) Trial sequential analysis (TSA) was also performed to quantify the reliability of data in meta-analysis adjusting significance levels for sparse data and multiple testing on accumulating trials.(19) Statistical analysis was performed by using WinBUGS (Imperial College & MRC, UK) and Stata 12.0 (College Station, Texas 77845 USA).

Trial sequential monitoring boundaries were employed to control the risks for type I and II errors and to indicate whether additional trials are needed. The information size calculation requires the mortality rate in the control group and the minimal effect size for the intervention. We predefined that the mortality in the control group is 30%, and the intervention is able to reduce the relative risk by 15%. The conventional α and β are 0.05 and 0.2, respectively. Meta-analysis will be updated by adding component studies sequentially in the order of publication. The β-spending function was constructed to indicate futility of intervention. Trial sequential analysis was performed by using the software TSA version 0.9 Beta (Copenhagen Trial Unit, 2011).

Results

Our initial search identified a total of 531 distinct citations, and 456 of them were excluded immediately after inspection of the title and abstract (figure 1). The remaining 75 clinical studies were potentially eligible and were examined for full text. Fifty-six studies were excluded because: 1) eight studies used duplicated report; 2) 18 studies used inappropriate control arm (e.g. single arm, all patients received aPC); 3) 19 did not report mortality as the endpoint; and 4) 11 did not include aPC as an intervention. As a result, a total of 7 RCTs(9, 12, 13, 20-23) and 11 observational studies(10, 11, 24-32) were included for analysis. Characteristics of RCTs are shown in table 1.

Figure 2 shows the efficacy of aPC on mortality reduction, reported separately for RCTs and observational studies. Of the 7 RCTs, only the PROWESS study showed significant mortality reduction with aPC (OR: 0.74, 95% CI: 0.59-0.91),(9) and the other six studies failed to conclude a beneficial effect. There was moderate heterogeneity among included RCTs (I^2=48.6%, p=0.07). The pooled OR for mortality was 1.00 (95% CI: 0.84-1.19).

In contrast, 6 out of the 11 observational studies showed significant mortality reduction with the use of aPC; and the remaining five studies showed a trend towards better lower mortality rate in aPC group. The heterogeneity was statistically significant with an I^2 of 70.8% (p<0.001). The pooled OR for mortality with the use

of aPC was 0.67 (95% CI: 0.62-0.72). Egger's test did not show evidence of publication bias in RCTs (p=0.808) and observational studies (p=0.145). Similarly, Begg's test did not show evidence of publication bias in RCTs (p=0.293) and observational studies (p=0.337). However, publication bias was suspected for observational studies as suggested by the funnel plot in which each dot represents a study and they gathered at the upper left corner (figure 3).

The result of sequential trial analysis is shown in figure 4. Studies were displayed sequentially by their publication year from left to the right of the horizontal line. After publication of the first and second studies (PROWESS 2001 and rhAPC sepsis 2001), the Z score crossed the conventional significance boundary (Z=1.96) but did not cross the O'Brien-Fleming boundaries. With the publication of the study Dhainaut 2009, the Z-score reached and crossed the futility line, indicating no effect of the aPC for mortality reduction in septic patients.

Meta-analysis of observational studies was performed by using Bayesian approach. The posterior distribution of individual OR was shrunken, as reflected by the narrower credible interval of study level estimates as compared to the observed estimates. For instance, the credible interval of OR in the study de Pont AC 2005 was 0.40-1.06, which was significantly narrower than the observed confidence interval of 0.04-6.70 (figure 2). This was because each component study borrowed evidence from the overall effect by using Bayesian approach. The overall OR was 0.67 (credible interval: 0.56-0.78).

Figure 5 shows the mean OR and 95% credible interval (CrI) for different power transformation priors to down-weight observational evidence on the risk of death with aPC. To the left of the figure when alpha took negligible values, the observational evidence was totally discounted and the mean OR was 1, which was consistent with the pooled result from RCTs. Increasing weight was assigned to observational evidence with increasing alpha values. We could see from the figure that the upper limit of CrI crossed the reference line with alpha values<0.4. When observational evidence was combined at its face value (alpha=1), the aPC group showed significant mortality reduction as compared with the control group. The alpha value influenced

the precision of prior evidence. As shown in figure 6, the precision of prior increased with increasing value of alpha from 0.000001 to 1.

Discussion

Key findings of the present analysis are 1) aPC appears to be able to reduce mortality rate when evidence is pooled from observational studies, and the results are consistent by using conventional Bayesian approaches; 2) RCTs failed to identify any beneficial effect of aPC; 3) observational evidence, when discounted by different power transformation priors, can alter the conclusion derived from RCTs. 4) With trial sequential analysis, the positive result (significant beneficial effect of aPC) as shown in the PROWESS study should be interpreted with caution.

One potential explanation for the positive findings in observational studies is the publication bias as shown in figure 3. The funnel contour plot showed that most observational studies located in the region with p<1%, indicating that the asymmetrical distribution was more likely due to publication bias. It is not surprising that observational studies are more subject to publication bias in that they are less likely to be registered a priori.(33) In contrast, RCTs are usually registered and there are many online registration sites.(34) The value of observational studies is usually discounted in evidence synthesis, and the conventional view is that observational evidence can only serve as hypothesis-generating. In such context, if the finding of an observational study is neutral, it will be less interesting to readers and journals, making it less likely to be published. In contrast, because RCTs are always registered and requires large amount of cost and other resources, studies with negative findings can be published and is equally important to those with positive findings.

Activated protein C (aPC) for the treatment of sepsis is a good example illustrating the importance of using sequential trial analysis in evidence synthesis. aPC was approved by the food and drug administration after publication of PROWESS trial, which seemed too hasty when viewed retrospectively. Although the initial trial was

positive at conventional significance level of p=0.05 (Z=1.96), its statistical significance should be tested by using adjusted alpha level. In sequential trial analysis, this is achieved by using alpha-spending function and constructing the O'Brien-Fleming boundaries. If sequential trial analysis had been performed at the conclusion of PROWESS trial, the approval of aPC for sepsis would not be so hasty. It has been argued that the disparity between PROWESS trial and subsequent trials such as PROWESS-SHOCK could be explained by the heterogeneity of enrolled subjects.(35)

In the translation of research into clinical practice, there are a lot of important factors to consider. RCT is generally accepted as a gold standard. However, there are some limitations in real clinical practice that RCT cannot simulate all the clinical situations. The biggest problem is that RCTs are usually conducted in non-real world setting, that is, often performed in specialized academic centers with strict inclusion/exclusion criteria. For example, in the Dhainaut 2009 study there was a long list of exclusion criteria, including expected surgical procedure in the next 3 days, platelet count<30,000/nm^3, receiving therapeutic heparin, moribund, withdrawn from aggressive management by patients' family, and pregnant or breast feeding. Such strict exclusion criteria would exclude most of patients with septic shock. Therefore, it appears unfair to treat our septic shock patients based on evidence derived from a minority of the population. In this situation, observational studies generally have an advantage in testing the clinical effectiveness of aPC on mortality reduction (36-38). Observational studies usually include a wider range of patients with septic shock and the setting is just like what we will encounter in routine clinical practice. For example, in prospective RCTs patients with comorbidities were strictly screened and excluded, but in retrospective studies it is often unreliable to exclude certain comorbidities based on medical records. Therefore, the observational evidence cannot be simply ignored in evidence synthesis for decision-making. However, the

result of observational studies could be misleading due to inherent bias. In our study the funnel plot showed asymmetrically distributed component studies, indicating potential publication bias. In this regard, the observational evidence should be interpreted with caution and in conjunction with other evidence. Since there was no consensus on how to combine observational evidence with RCTs, we discounted observational evidence with power transform priors taking advantage of the flexibility in Bayesian modeling (18). In this model, we found that the treatment effect of aPC increased with more weight assigning to observational evidence (figure 6). A value of 0 for alpha implies that the observational evidence is ignored, and a value of 1 for alpha means that observational evidence is accepted at its "face value". This approach gives a full picture of how pooled evidence can be altered by observational studies, by explicitly showing the power transform priors.

Several limitations of the study need to be acknowledged. First, there are substantial heterogeneity among included RCTs, which may be explained by the differences in study population, timing of intervention and definition of study endpoint. As expected, observational studies showed substantial heterogeneity. Considering the very different study criteria and the various geographic sources of each observational study (representing different standards of care), heterogeneity should be expected due to the more generalizable (real-world) evidence. Second, it is still largely unknown on how to discount the observational evidence. The most appropriate prior will vary from study to study. The present study only displays a wide range of possible alpha values and explicitly demonstrates how RCT evidence can be modified by observational evidence.

In summary, our study demonstrates that there is considerable disparity between observational and RCT evidence. While observational evidence shows beneficial effect of aPC on mortality reduction, RCTs failed to identify any such treatment effect.

Figure legends

Figure 1. Flow chart of study selection.

Figure 2. Forest plots showing the efficacy of aPC on mortality reduction, reported separately for RCTs and observational studies.

Figure 3. Contour funnel plots showing the publication bias in RCTs and observational studies. Publication bias was identified for observational studies as reflected by the asymmetrically distributed component studies.

Figure 4. Sequential trial analysis involving randomized controlled trials showing that the Z-score crossed the futility line after the study Dhainaut 2009. Parameters used for the creation of boundaries were: type: Two-sided; type 1 Error: 5.0%; alpha spending: O'Brien-Fleming; information axis: sample size; power: 80.0%; effect type intervention: RRR User Defined (15%). The shaded area indicates futility area.

Figure 5. Mean OR and 95% credible interval (CrI) for different power transformation priors to down-weight observational evidence on the risk of death with aPC. Lower values of alpha down-weight the observational evidence

Figure 6. Prior distribution derived by discounting observational evidence with alpha from 0.000001 to 1. The plots shows that the precision of prior increases with increasing alpha values.

Reference

1. Schorr CA, Zanotti S, Dellinger RP. Severe sepsis and septic shock: management and performance improvement. Virulence. 2014;5(1):190-9.
2. Dellinger RP, Levy MM, Rhodes A, Annane D, Gerlach H, Opal SM, et al. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012. Critical care medicine. 2013;41(2):580-637.
3. MacIntyre N. Ventilatory Management of ALI/ARDS. Seminars in respiratory and critical care medicine. 2006;27(4):396-403.
4. Rivers E, Nguyen B, Havstad S, Ressler J, Muzzin A, Knoblich B, et al. Early goal-directed therapy in the treatment of severe sepsis and septic shock. The New England journal of medicine. 2001;345(19):1368-77.
5. Pro CI, Yealy DM, Kellum JA, Huang DT, Barnato AE, Weissfeld LA, et al. A randomized trial of protocol-based care for early septic shock. The New England journal of medicine. 2014;370(18):1683-93.
6. Ronco C, Bellomo R, Homel P, Brendolan A, Dan M, Piccinni P, et al. Effects of different doses in continuous veno-venous haemofiltration on outcomes of acute renal failure: a prospective randomised trial. Lancet. 2000;356(9223):26-30.
7. Zhang Z, Xu X, Zhu H. Intensive- vs less-intensive-dose continuous renal replacement therapy for the intensive care unit-related acute kidney injury: a meta-analysis and systematic review. Journal of critical care. 2010;25(4):595-600.
8. Christiaans SC, Wagener BM, Esmon CT, Pittet JF. Protein C and acute inflammation: a clinical and biological perspective. American journal of physiology Lung cellular and molecular physiology. 2013;305(7):L455-66.
9. Bernard GR, Vincent JL, Laterre PF, LaRosa SP, Dhainaut JF, Lopez-Rodriguez A, et al. Efficacy and safety of recombinant human activated protein C for severe sepsis. The New England journal of medicine. 2001;344(10):699-709.
10. Dhainaut JF, Payet S, Vallet B, Franca LR, Annane D, Bollaert PE, et al. Cost-effectiveness of activated protein C in real-life clinical practice. Critical care (London, England). 2007;11(5):R99.
11. Kubler A, Mayzner-Zawadzka E, Durek G, Gaszynski W, Karpel E, Mikaszewska-Sokolewicz M, et al. Results of severe sepsis treatment program using recombinant human activated protein C in Poland. Medical science monitor : international medical journal of experimental and clinical research. 2006;12(3):CR107-12.
12. Annane D, Timsit JF, Megarbane B, Martin C, Misset B, Mourvillier B, et al. Recombinant human activated protein C for adults with septic shock: a randomized controlled trial. American journal of respiratory and critical care medicine. 2013;187(10):1091-7.
13. Ranieri VM, Thompson BT, Barie PS, Dhainaut JF, Douglas IS, Finfer S, et al. Drotrecogin alfa (activated) in adults with septic shock. The New England journal of medicine. 2012;366(22):2055-64.
14. Lai PS, Matteau A, Iddriss A, Hawes JC, Ranieri V, Thompson BT. An updated meta-analysis to understand the variable efficacy of drotrecogin alfa (activated) in severe sepsis and septic shock. Minerva anestesiologica. 2013;79(1):33-43.
15. Marti-Carvajal AJ, Sola I, Gluud C, Lathyris D, Cardona AF. Human recombinant protein C for severe sepsis and septic shock in adult and paediatric patients. The Cochrane database of systematic reviews. 2012;12:CD004388.
16. Zhang Z, Ni H, Xu X. Do the observational studies using propensity score analysis agree with

randomized controlled trials in the area of sepsis? Journal of critical care. 2014;29(5):886 e9-15.
17. Zhang Z. Recombinant human activated protein C for the treatment of severe sepsis and septic shock: a study protocol for incorporating observational evidence using a Bayesian approach. BMJ Open. 2014;4(7):e005622.
18. Chen MH, Ibrahim JG. Power prior distributions for regression models. statistical science. 2000;15(1):46-60.
19. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. Journal of clinical epidemiology. 2008;61(1):64-75.
20. Bernard GR, Ely EW, Wright TJ, Fraiz J, Stasek JE, Jr., Russell JA, et al. Safety and dose relationship of recombinant human activated protein C for coagulopathy in severe sepsis. Critical care medicine. 2001;29(11):2051-9.
21. Abraham E, Laterre PF, Garg R, Levy H, Talwar D, Trzaskoma BL, et al. Drotrecogin alfa (activated) for adults with severe sepsis and a low risk of death. The New England journal of medicine. 2005;353(13):1332-41.
22. Nadel S, Goldstein B, Williams MD, Dalton H, Peters M, Macias WL, et al. Drotrecogin alfa (activated) in children with severe sepsis: a multicentre phase III randomised controlled trial. Lancet. 2007;369(9564):836-43.
23. Dhainaut JF, Antonelli M, Wright P, Desachy A, Reignier J, Lavoue S, et al. Extended drotrecogin alfa (activated) treatment in patients with prolonged septic shock. Intensive care medicine. 2009;35(7):1187-95.
24. de Pont AC, Bakhtiari K, Hutten BA, de Jonge E, Vroom MB, Meijers JC, et al. Recombinant human activated protein C resets thrombin generation in patients with severe sepsis - a case control study. Critical care (London, England). 2005;9(5):R490-7.
25. Bertolini G, Rossi C, Anghileri A, Livigni S, Addis A, Poole D. Use of Drotrecogin alfa (activated) in Italian intensive care units: the results of a nationwide survey. Intensive care medicine. 2007;33(3):426-34.
26. Rowan KM, Welch CA, North E, Harrison DA. Drotrecogin alfa (activated): real-life use and outcomes for the UK. Critical care (London, England). 2008;12(2):R58.
27. Vincent JL, Laterre PF, Decruyenaere J, Spapen H, Raemaekers J, Damas F, et al. A registry of patients treated with drotrecogin alfa (activated) in Belgian intensive care units--an observational study. Acta clinica Belgica. 2008;63(1):25-30.
28. Martin G, Brunkhorst FM, Janes JM, Reinhart K, Sundin DP, Garnett K, et al. The international PROGRESS registry of patients with severe sepsis: drotrecogin alfa (activated) use and patient outcomes. Critical care (London, England). 2009;13(3):R103.
29. Lindenauer PK, Rothberg MB, Nathanson BH, Pekow PS, Steingrub JS. Activated protein C and hospital mortality in septic shock: a propensity-matched analysis. Critical care medicine. 2010;38(4):1101-7.
30. Sadaka F, O'Brien J, Migneron M, Stortz J, Vanston A, Taylor RW. Activated protein C in septic shock: a propensity-matched analysis. Critical care (London, England). 2011;15(2):R89.
31. Casserly B, Gerlach H, Phillips GS, Marshall JC, Lemeshow S, Levy MM. Evaluating the use of recombinant human activated protein C in adult severe sepsis: results of the Surviving Sepsis Campaign. Critical care medicine. 2012;40(5):1417-26.
32. Rimmer E, Kumar A, Doucette S, Marshall J, Dial S, Gurka D, et al. Activated protein C and septic shock: a propensity-matched cohort study*. Critical care medicine. 2012;40(11):2974-81.

33. Onukwugha E. Improving confidence in observational studies : should statistical analysis plans be made publicly available? PharmacoEconomics. 2013;31(3):177-9.
34. Anand V, Scales DC, Parshuram CS, Kavanagh BP. Registration and design alterations of clinical trials in critical care: a cross-sectional observational study. Intensive care medicine. 2014;40(5):700-22.
35. Kalil AC, Florescu DF. Severe sepsis: are PROWESS and PROWESS-SHOCK trials comparable? A clinical and statistical heterogeneity analysis. Critical care (London, England). 2013;17(4):167.
36. Albert RK. "Lies, damned lies ..." and observational studies in comparative effectiveness research. Am J Respir Crit Care Med. 2013;187(11):1173-7.
37. Nallamothu BK, Hayward RA, Bates ER. Beyond the randomized clinical trial: the role of effectiveness studies in evaluating cardiovascular therapies. Circulation. 2008;118(12):1294-303.
38. Zhang Z. Big data and clinical research: focusing on the area of critical care medicine in mainland China. Quantitative Imaging in Medicine and Surgery. 2014;4(5):426-9.

Funding

There was no funding for the present study.

Competing Interest

None

Table 1 Characteristics of included randomized controlled trials

Studies	Patients (n)	Mean age (years)	Population	Mean APACHE II score	Control	Primary outcome	Baseline mortality (%)
Bernard GR 2001 (rhAPC)	131	59.3	Severe shock	17.3	Placebo	Coagulopathy	34.2
Bernard GR 2001 (PROWES	1690	60.5	Systemic inflammation and	24.8	Placebo (saline or	28-day all cause mortality	30.8

S)			organ failure		albumin)		
Ranieri VM 2012	1697	63.1	Sepsis and shock receiving fluids and vasopressor	25.3	Placebo (saline)	28-day all cause mortality	24.2
Abraham E 2005	2613	58.7	Severe sepsis and single organ failure or Mean APACHE II<25	18.2	Placebo (saline)	28-day all cause mortality	17
Nadel S 2007	477	2.5	Children with sepsis induced cardiac or respiratory failure	-	Placebo (saline)	CTCOFR	17.5
Annane D 2013	411	63	Sepsis with >2 organ failure	-	Placebo (saline)	90-day mortality	46.3
Dhainaut JF 2009	193	62.4	Severe sepsis with vasopressor dependent	28.1	Placebo	Resolution of vasopressor dependent	32.3

			hypotensio n			hypotension	

APACHE: Acute Physiology and Chronic Health Evaluation; CTCOFR: Composite Time to Complete Organ Failure Resolution.

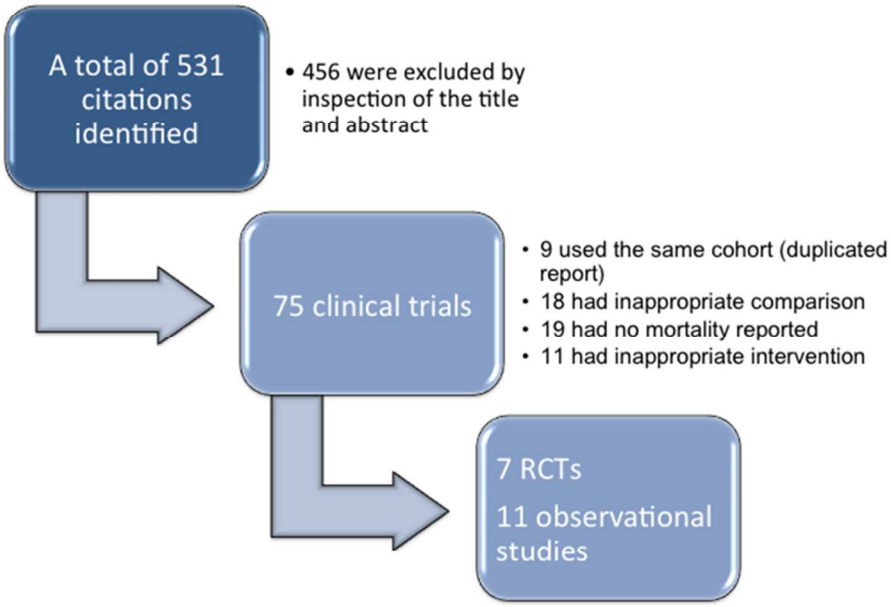

Figure 1. Flow chart of study selection.
84x63mm (300 x 300 DPI)

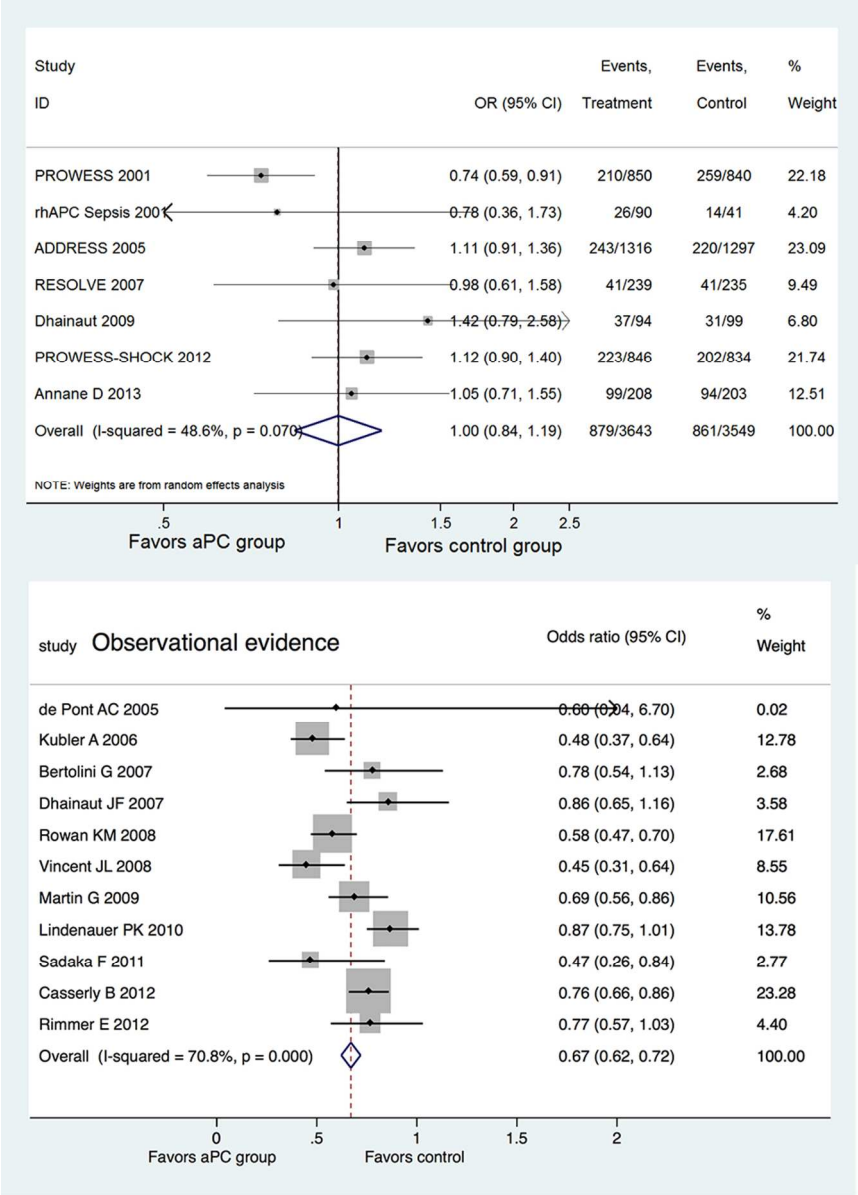

Figure 2. Forest plots showing the efficacy of aPC on mortality reduction, reported separately for RCTs and observational studies. Results were pooled by using conventional meta-analytic approach.
84x118mm (300 x 300 DPI)

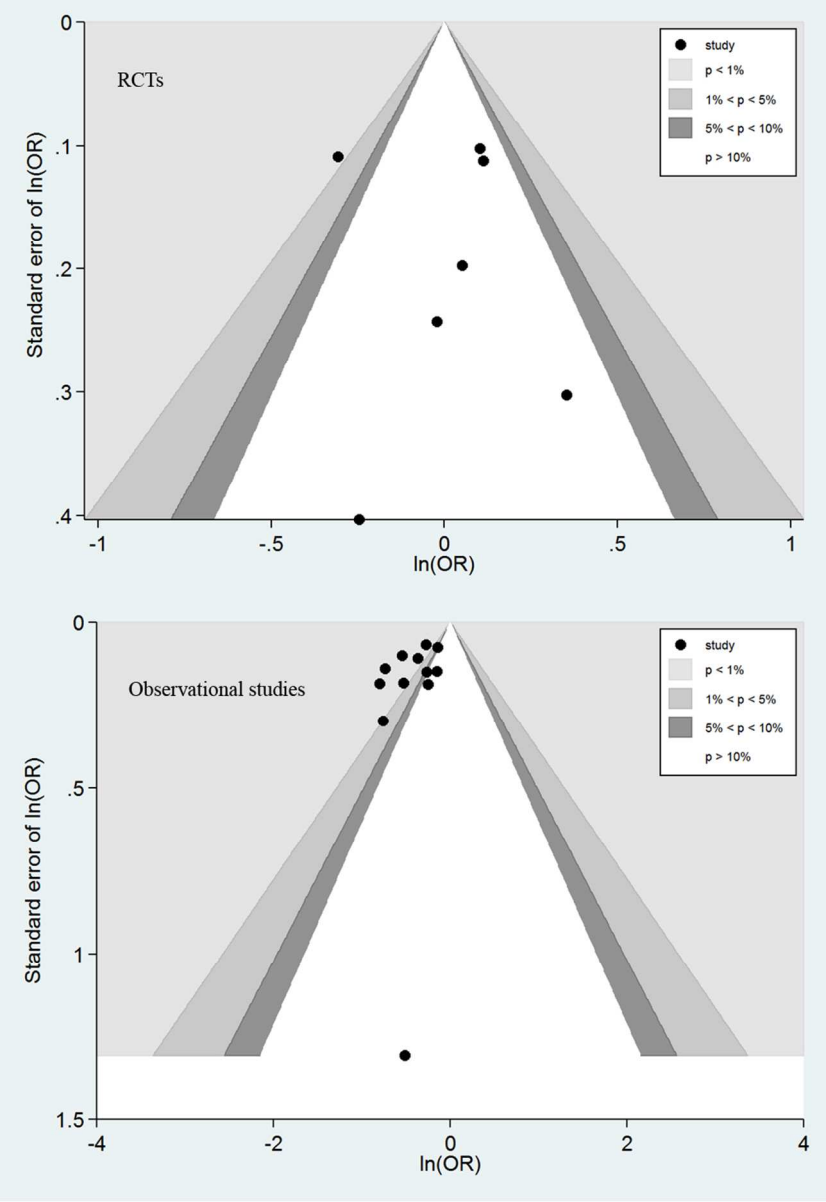

Figure 3. Contour funnel plots showing the publication bias in RCTs and observational studies. Publication bias was identified for observational studies as reflected by the asymmetrically distributed component studies.
84x123mm (300 x 300 DPI)

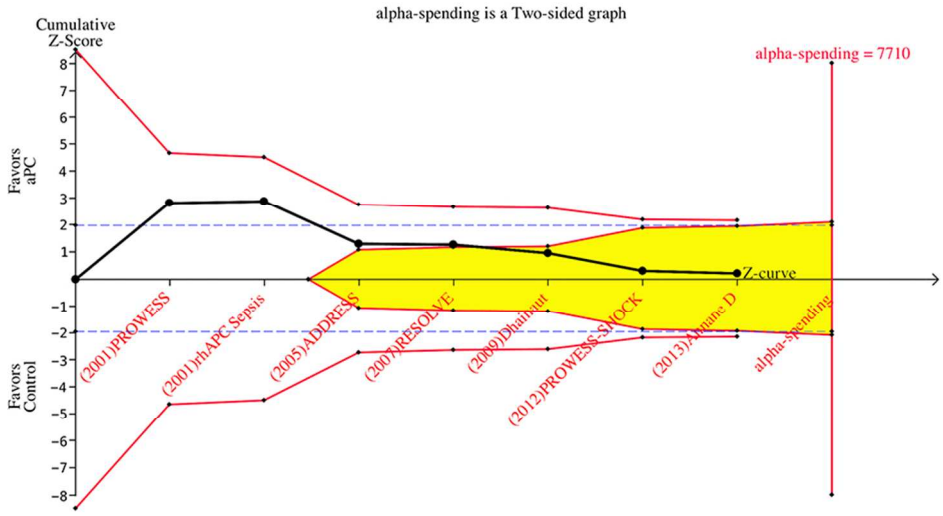

Figure 4. Sequential trial analysis involving randomized controlled trials showing that the Z-score crossed the futility line after the study Dhainaut 2009. Parameters used for the creation of boundaries were: type: Two-sided; type 1 Error: 5.0%; alpha spending: O'Brien-Fleming; information axis: sample size; power: 80.0%; effect type intervention: RRR User Defined (15%). The shaded area indicates futility area.
84x45mm (300 x 300 DPI)

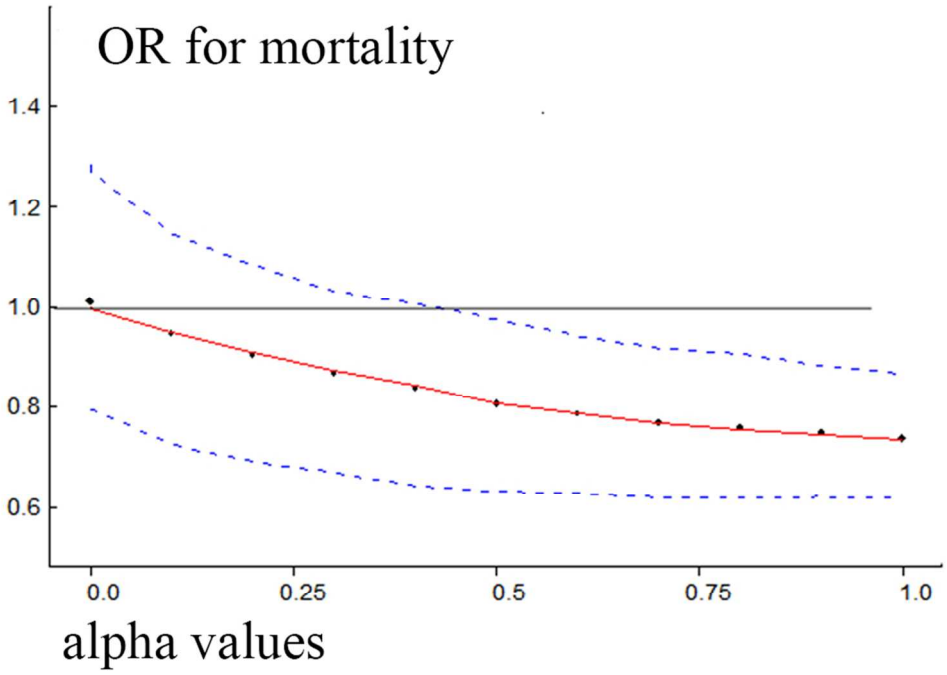

Figure 5. Mean OR and 95% credible interval (CrI) for different power transformation priors to down-weight observational evidence on the risk of death with aPC.
84x58mm (300 x 300 DPI)

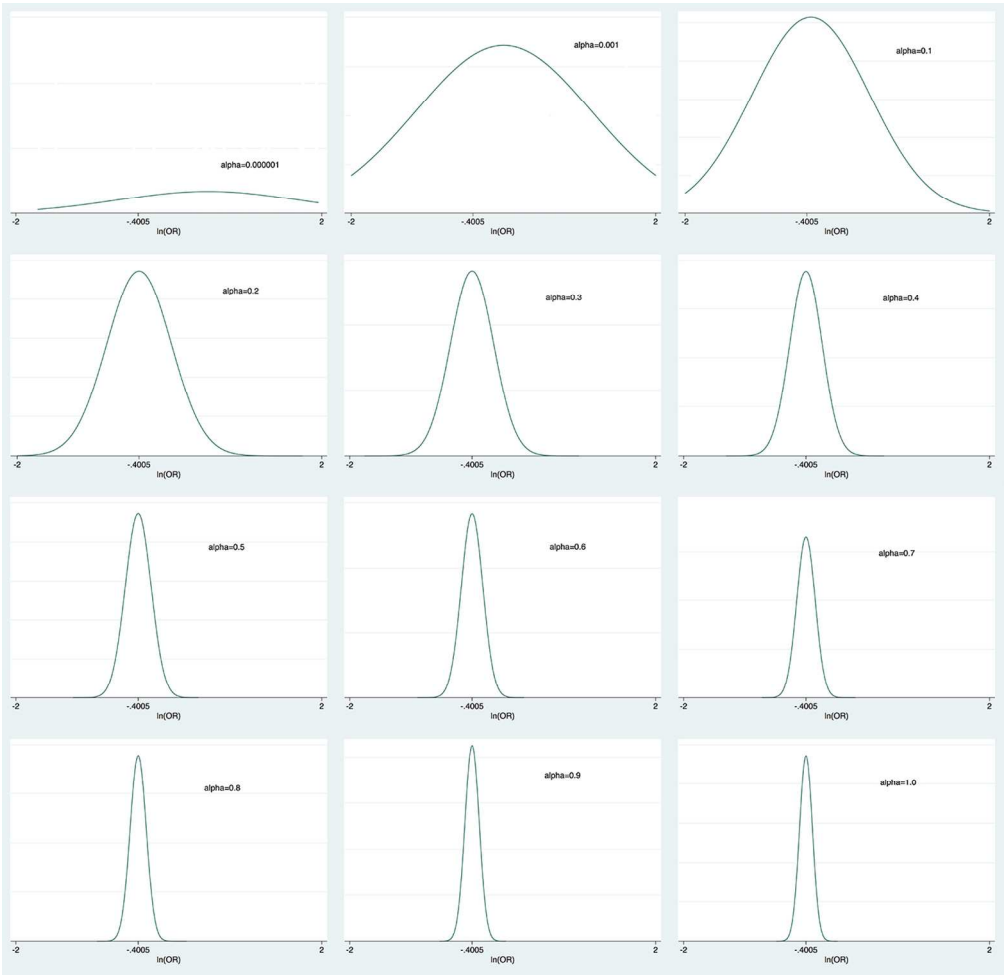

Figure 6. Prior distribution derived by discounting observational evidence with alpha from 0.000001 to 1. The plots shows that the precision of prior increases with increasing alpha values.
127x123mm (300 x 300 DPI)

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	The efficacy of activated protein C for the treatment of sepsis: incorporating observational evidence with Bayesian approach	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	6
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	7
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	7
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency	7-8

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	8
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	9
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	9
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	8-9
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	9
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	9
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	10
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	12
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	12
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	16

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.