PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Effect of Corticosteroid Administration on Neurologically Deceased Organ Donors and Transplant Recipients : A Systematic Review and Meta-Analysis
AUTHORS	D'Aragon, Frédérick; Belley-Côté, Émilie; Agarwal, Arnav; Frenette, Anne-Julie; Lamontagne, Francois; Guyatt, Gordon; Dhanani, Sonny; Meade, Maureen

VERSION 1 - REVIEW

REVIEWER	Rolando Rebolledo A. Department of Surgery Pontificia Universidad de Chile Chile
REVIEW RETURNED	20-Nov-2016

GENERAL COMMENTS	The authors investigate the impact of corticosteroids on donation rates and transplant outcomes based on available RCTs. In a very clear manuscript they expose the lack of good-quality evidence about this topic. Clinical recommendations are adjusted to available evidence. I have no important criticism. Congratulations!

REVIEWER	Sheila M Bird MRC Biostatistics Unit (as visiting scientist), Cambridge UK Long-standing interest in how to increase cadaveric organ donation in UK - not by NHS Organ Donor register but changing from opting-in to opt-out from organ donation, as now in Wales.
REVIEW RETURNED	13-Dec-2016

GENERAL COMMENTS	Abstract concludes that, on the available, evidence administration or withholding steroids are equally reasonable. I do not so conclude from the evidence presented. I conclude that RCTs to date have been too small to discern modest benefits or harms YET differentiating modest benefit from modest harm is precisely what the transplant community needs to do to make a decision in the face of dire shortage of cadaveric organ donations. I also conclude, as do authors [p17, last sentence before conclusion] that design of RCTs to date has been hampered because of concerns about apparent inability to pre-consent recipients for the possibility of receiving organ from donor managed by method A versus B. RCT seem not to have been properly analysed to take into account that DONOR is unit of randomization and so we have, in effect, a cluster randomized trial with donor as cluster and, for

	example, the number of donated kidneys being - FROM that DONOR - zero, 1 or 2. Likewise, analysis of outcome for recipients needs to respect DONOR-cluster. This is non-trivial and unlikely to have been done in papers published pre-1995. In my view , authors need to comment on correct method of analysis - were the data available to them - and to indicate that the likely effect of so doing is CIs that are wider still. RCTs should state a priori likely effect-size that they have been designed to discern: also unlikely for RCTs that reported pre-1985. However, meta-analysts should use their own subject-matter knowledge (of costs, likely benefits, possible harms) to propose the effect-sizes that future RCTs need to tilt for. Basically, 20% benefit=increase in the number of functioning transplants at 6 months post-Tp would be better than anything yet achieved recently in cadaveric organ donors . . . it might not be plausible however as a steroid-related goal. Even 10% benefit would probably be cost-effective [authors to CHECK/CONSIDER]. If we assume that 60% of potential brain stem dead donors become actual donors, then RCTs to differentiate 60% versus 66% (ie 10% increase which translates to increase in number of donated organs) requires randomization of over 2100 (nearest 100) brain-stem dead potential donors. Compared to the target, the meta-analysis demonstrates merely that every study has been too small be a factor of around 10 at least and the totality does not make for a half-decently-designed study with around 50% power to discern 10% benefit/harm. Continuing not to know is not a safe option, I'd suggest - unless expert opinion has moved on and ICU teams have decided on first principles. A further issue is that there have been major changes (surgical + cyclosporin in mid 1980 & immunological such as beneficial matching and its successors in the 1990s) that have radically changed graft outcome. Frankly, I would not admit into meta-analysis RCTs for with ROB cannot be assessed and which are so outdated in terms of modern transplantation. Of course, I do not hold with the Cochrane doctrine of inclusion of all RCTs in meta-analyses: doctrines lead to unthinking science. Comments on paper are therefore: 1. Abstract's conclusion verses last sentence in text before CONCLUSION on p17 sends different messages, the latter better. 2. Abstract does not sufficiently put paucity of RCT evidence into context because appropriate effect-sizes not cited/proposed. 3. Abstract does not explain that correct analysis of individuals trials [DONOR => many organs] needs to respect clustering by DONOR; and that meta-analysis ails to do so . . . 4. Abstract/Discussion should revert to question of "when to randomize" - some RCTs only after family consent, which may be too late fro effectiveness . . . (even if steroids were effective). 5. I find confusing (n) in the column in Table 1 headed Author,Year as I had expected n to accord with text-referencing of the same study, which it appears not to do. 6. Table 2 is simplistic in terms if its summary of "No of Patients" which appears to be [sum of numerators/sum of denominators] for

	however many RCTs were relevant. However, such summary is NOT what the meta-analysis does in weighting the per-RCT estimates. Consider a different summary. NOR can I make sense of the numbers in Table 1 & 2, which do not accord. Please check also against Figures, eg on Organ Recovery. 7. When all studies are too small by far (see para ahead of Fig 5), how can analysts assess "large variation in effect between studies" when each effect is highly variable in & of itself? 8. In DISCUSSION, is it worth remarking that central estimates point in different directions for number of organs recovered [fewer] & graft rejection [lower] . . . which is the ultimate goal for future RCTs?

REVIEWER	Prof Geoffrey J Dobb Royal Perth Hospital and University of Western Australia, Australia
REVIEW RETURNED	02-Jan-2017

GENERAL COMMENTS	You have undertaken a well conducted and well written systematic review of corticosteroids given to brain dead organ donors and their effect on transplant related outcomes. This is a challenging area for research because of the ethical uncertainties around interventions in donors that may benefit or harm recipients who are unknown at the time of the intervention and the need to explain research interventions to families who may already be overwhelmed by the amount of information that have been given in relation to organ donation at a time of grief in the usual context of unexpected death. Your findings confirm and largely duplicate those of Dupuis and colleagues (ref 16). Both systematic reviews highlight that the evidence supporting corticosteroid use is of overall poor quality and that the randomised controlled trials fail to demonstrate benefit. Your abstract conclusion and the emphasis in the Conclusion (line 337) is that the studies also do not identify significant harm from corticosteroid administration, though the relatively small patient base and poor study quality for one of the two studies that evaluated steroid related adverse events I would suggest means that less emphasis and confidence should be placed on this conclusion. - Was the plan for this systematic review and meta-analysis registered or published, and if so where? - the reporting format appears to comply with the PRISMA checklist. Can you please confirm that this is the case and include a statement in the methods? - The reasoning behind the use of a fixed effects model for pooled analyses is noted and supported. - Was a sensitivity analysis comparing the studies in which only corticosteroid was given with all studies considered for the main outcomes? If so, what was the result? This would also help to address the issue identified in the discussion at line 289. - Line 210: multiple factors influence the number of organs retrieved from each donor. Was there sufficient detail in any or some of the studies to enable any controlling for such factors? - Line 214: the use of the extravascular lung water index by Venkateswaran et al was very much a surrogate (physiological) end point. Is there actual evidence that this measurement affects consideration of the suitability of lungs for transplantation, as opposed to the more routine measures such as PaO2/FO2 ratio? - You emphasis the importance of use of the GRADE system in your review (line 272) but does this add very much to the risk of bias assessment and quality assessment using the Cochrane Collaboration's Tool us used by Dupuis et al, particularly since you

	point out that this is a main point of differentiation between your review and theirs - A minor point, line 163 should read participants (not participant). Despite addressing an important issue, being well conducted and well written, the main reason for my recommendation is that it substantially duplicates the material and conclusions previously published by Dupuis and colleagues.

REVIEWER	DR JUNE GOH SINGAPORE GENERAL HOSPITAL SINGAPORE
REVIEW RETURNED	06-Jan-2017

GENERAL COMMENTS	The review is clear and concise. 3500 articles were screened out of which just the 11 were selected. Is it possible to give more detail as to why the rest were rejected? A similar review in the BJA in 2014 has been done before : Corticosteroids in the management of brain-dead potential organ donors, a systematic review. The difference would be the use of the GRADE methodology. If the studies chosen for the review was more recent or greater in number and dissimilar, then there would be greater value and less redundancy. The reviewer also provided a file in addition to these comments. Please contact the publisher for full details.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

2.1 The authors investigate the impact of corticosteroids on donation rates and transplant outcomes based on available RCTs. In a very clear manuscript they expose the lack of good-quality evidence about this topic. Clinical recommendations are adjusted to available evidence. I have no important criticism. Congratulations!

Thank you for your positive comment.

Reviewer 2:

2.2 Abstract concludes that, on the available, evidence administration or withholding steroids are equally reasonable. I do not so conclude from the evidence presented. I conclude that RCTs to date have been too small to discern modest benefits or harms YET differentiating modest benefit from modest harm is precisely what the transplant community needs to do to make a decision in the face of dire shortage of cadaveric organ donations.

We agree with the reviewer that, from a clinical perspective, differentiating modest benefit from modest harm is of utmost importance. However, we respectfully maintain that, based on the current evidence, strong recommendations in favor of administering or withholding steroids should be avoided as both options seem equally reasonable.

2.3 I also conclude, as do authors [p17, last sentence before conclusion] that design of RCTs to date has been hampered because of concerns about apparent inability to pre-consent recipients for the possibility of receiving organ from donor managed by method A versus B. RCT seem not to have been properly analysed to take into account that DONOR is unit of randomization and so we have, in effect, a cluster randomized trial with donor as cluster and, for example, the number of donated kidneys being - FROM that DONOR - zero, 1 or 2. Likewise, analysis of outcome for recipients needs to respect DONOR-cluster. This is non-trivial and unlikely to have been done in papers published pre-1995.

In my view, authors need to comment on correct method of analysis - were the data available to them - and to indicate that the likely effect of so doing is CIs that are wider still.

We would like to thank the reviewer for bringing up this very important statistical issue. We agree that the effect of taking into account clustering would be a widening of the confidence intervals if there is a correlation in outcome within donors (i.e. multiple organs from one donor are likely to do better or worse then organs from another donor). We have now added the following to the discussion: "Another limitation is that studies did not take the clustering of organs within donors (a single donor can contribute up to 7 organs) into account in the analysis. To the extent that organs from some donors do systematically better than organs from other donors, the confidence intervals presented in the studies are narrower than would be the case in an analysis that took clustering into account."

2.4 RCTs should state a priori likely effect-size that they have been designed to discern: also unlikely for RCTs that reported pre-1985. However, meta-analysts should use their own subject-matter knowledge (of costs, likely benefits, possible harms) to propose the effect-sizes that future RCTs need to tilt for.

Basically, 20% benefit=increase in the number of functioning transplants at 6 months post-Tp would be better than anything yet achieved recently in cadaveric organ donors . . . it might not be plausible however as a steroid-related goal. Even 10% benefit would probably be cost-effective [authors to CHECK/CONSIDER].

If we assume that 60% of potential brain stem dead donors become actual donors, then RCTs to differentiate 60% versus 66% (ie 10% increase which translates to increase in number of donated organs) requires randomization of over 2100 (nearest 100) brain-stem dead potential donors. Compared to the target, the meta-analysis demonstrates merely that every study has been too small be a factor of around 10 at least and the totality does not make for a half-decently-designed study with around 50% power to discern 10% benefit/harm.

Continuing not to know is not a safe option, I'd suggest - unless expert opinion has moved on and ICU teams have decided on first principles.

A further issue is that there have been major changes (surgical + cyclosporin in mid 1980 & immunological such as beneficial matching and its successors in the 1990s) that have radically changed graft outcome. Frankly, I would not admit into meta-analysis RCTs for with ROB cannot be assessed and which are so outdated in terms of modern transplantation. Of course, I do not hold with the Cochrane doctrine of inclusion of all RCTs in meta-analyses: doctrines lead to unthinking science.

The reviewer makes interesting points about the clinical relevance of a 10% relative increase in the number of organs donated. We agree that all clinical trials conducted to date on steroids in potential organ donors have been highly underpowered to assess for anything close to what would be minimally clinical relevant and that, even when meta-analyzed, the current body of literature does not

allow for conclusions regarding the benefit or harm of steroid administration in potential organ donors.

Comments on paper are therefore:
2.5 Abstract's conclusion verses last sentence in text before CONCLUSION on p17 sends different messages, the latter better.

Thank you for bringing this oversight to our attention. we have made the correction.

2.6 Abstract does not sufficiently put paucity of RCT evidence into context because appropriate effect-sizes not cited/proposed.

We would like to add more context on the paucity of high quality evidence in the abstract as suggested by the reviewer. Unfortunately, the word count does not allow for sufficient space to properly explain the assumptions underlying the choice of an appropriate effect size.

2.7 Abstract does not explain that correct analysis of individuals trials [DONOR => many organs] needs to respect clustering by DONOR; and that meta-analysis ails to do so . . .

Please note our response to issue 2.3. Although we agree that this issue is important, the most important limitations are the number and size of studies, and these we have highlighted in the abstract. Highlighting the issue of clustering in the discussion is, in our view, sufficient.

2.8 Abstract/Discussion should revert to question of "when to randomize" - some RCTs only after family consent, which may be too late for effectiveness . . . (even if steroids were effective).

We agree, and have added the following to the discussion: “Variation in timing of randomization and subsequent administration of study intervention also have affected treatment effect presuming that later administration (i.e. 5-8 hours before organ recovery) may be less effective.” Once again, we do not think it necessary to add to the abstract (though, if the editors do, we will be happy to do so).

2.9 I find confusing (n) in the column in Table 1 headed Author,Year as I had expected n to accord with text-referencing of the same study, which it appears not to do.

We modified Table 1 based on these suggestions.

2.10 Table 2 is simplistic in terms if its summary of "No of Patients" which appears to be [sum of numerators/sum of denominators] for however many RCTs were relevant. However, such summary is NOT what the meta-analysis does in weighting the per-RCT estimates. Consider a different summary. NOR can I make sense of the numbers in Table 1 & 2, which do not accord. Please check also against Figures, eg on Organ Recovery.

We thank the reviewer for this comment. We changed our table of summary. We also corrected numbers where applicable.

2.11 When all studies are too small by far (see para ahead of Fig 5), how can analysts assess "large variation in effect between studies" when each effect is highly variable in & of itself?

The reviewer raises an interesting point. If we understand it correctly, the issue is that the small studies yield imprecise estimates and chance may explain apparent differences between studies. The p-value of the test for heterogeneity reflects the probability of chance explaining the differences. It has long been argued, however, that this test will, if studies are small, be underpowered, and one should be concerned about large differences between studies even if the test for heterogeneity is not

significant. That is the reason Julian Higgins came up with the I2 as an alternative statistical test for looking at heterogeneity. There are therefore two reasonable positions:
i)Rate down for inconsistency only if chance cannot explain differences between study results (the position, as we understand it, implied in the reviewer's comment).
ii)Consider not only the p-value for the test for heterogeneity, but also the magnitude of the difference in results, and the I2
We have taken the latter approach. We have, however, rated down for inconsistency only once, for acute graft rejection. Here one study suggests appreciable benefit, another appreciable harm (see forest plot 5), the confidence intervals in these two studies overlap very little, the p-value for the test for heterogeneity approaches statistical significance (0.08), and the I2 is high (61%). Although it is a judgment call, all things considered, we continue to think that rating down for inconsistency is the right decision here.

2.12 In DISCUSSION, is it worth remarking that central estimates point in different directions for number of organs recovered [fewer] & graft rejection [lower] . . . which is the ultimate goal for future RCTs?

The reviewer makes a good point that not only is variability across studies and issue, but also variability across outcomes

Reviewer 3:

3.1 You have undertaken a well conducted and well written systematic review of corticosteroids given to brain dead organ donors and their effect on transplant related outcomes. This is a challenging area for research because of the ethical uncertainties around interventions in donors that may benefit or harm recipients who are unknown at the time of the intervention and the need to explain research interventions to families who may already be overwhelmed by the amount of information that have been given in relation to organ donation at a time of grief in the usual context of unexpected death.

We thank the reviewer for these positive comments.

3.2 Your findings confirm and largely duplicate those of Dupuis and colleagues (ref 16). Both systematic reviews highlight that the evidence supporting corticosteroid use is of overall poor quality and that the randomised controlled trials fail to demonstrate benefit.
Your abstract conclusion and the emphasis in the Conclusion (line 337) is that the studies also do not identify significant harm from corticosteroid administration, though the relatively small patient base and poor study quality for one of the two studies that evaluated steroid related adverse events I would suggest means that less emphasis and confidence should be placed on this conclusion.

As suggested by the reviewer, we have diminished the emphasis placed on this conclusion. It is now reads: "Current clinical trials do not identify benefits of corticosteroid therapy for deceased organ donors or their transplant recipients. The quality of this evidence is insufficient, however, to rule out the possibility of benefits or harms with respect to donation rates or transplant outcomes for any organ. In light of these results, there is no imperative to modify current recommendations for clinical care, based on observational studies, to consider corticosteroid therapy in the management of organ donors."

3.3 Was the plan for this systematic review and meta-analysis registered or published, and if so where?

Unfortunately, this systematic review was not registered. However, a detailed protocol was developed a priori.

3.4 the reporting format appears to comply with the PRISMA checklist. Can you please confirm that this is the case and include a statement in the methods?

We confirm that our reporting complies with the PRISMA checklist. We added a statement on PRISMA in the methods section.

3.5 The reasoning behind the use of a fixed effects model for pooled analyses is noted and supported.

We thank the reviewer for this comment.

3.6 Was a sensitivity analysis comparing the studies in which only corticosteroid was given with all studies considered for the main outcomes? If so, what was the result? This would also help to address the issue identified in the discussion at line 289.

We did not plan and have not conducted an a posteriori sensitivity analysis comparing studies with steroids alone vs. steroids administrated with other medications. Although very interesting in theory, the few small trials available for pooling make any subgroup analysis difficult to interpret.

3.7 Line 210: multiple factors influence the number of organs retrieved from each donor. Was there sufficient detail in any or some of the studies to enable any controlling for such factors?

A minimal set of informations was available on factors that influence the number of organs recovered in the articles reporting organ retrieval. For example, comorbities, results of test to assess organ suitability assessment were not available in both studies. For this reason, we were not able to conduct subgroup analyses related to these factors.

3.8 Line 214: the use of the extravascular lung water index by Venkateswaran et al. was very much a surrogate (physiological) end point. Is there actual evidence that this measurement affects consideration of the suitability of lungs for transplantation, as opposed to the more routine measures such as PaO2/FO2 ratio?

We agree with the reviewer that EVWLI is a surrogate end point. Only one study by Venkateswaran et al. assessed the association between high (>10 mL/kg) and low (<10 mL/kg) extravascular lung water index (EVLWI) and lung recipient survival. In a pre-post study (N=60 deceased donors), a total of 27 lung transplants were performed. The 30-day survival was lower in recipients of lungs with high EVWLI vs. those who received lungs with low EVLWI (60% vs. 100%, p=0.03). This study has important limitations and results must therefore be interpreted with caution.

3.9 You emphasis the importance of use of the GRADE system in your review (line 272) but does this add very much to the risk of bias assessment and quality assessment using the Cochrane Collaboration's Tool us used by Dupuis et al, particularly since you point out that this is a main point of differentiation between your review and theirs

The use of the GRADE framework is a strength of our systematic review. It provides a transparent assessment of our confidence in the estimates of the effect of steroids on key clinical outcomes in potential organ donors. The GRADE assessment is definitely an added value as it will provide knowledge users with evaluations of the quality of evidence underlying the use of steroids in potential organ donors. Risk of bias is only one of the domains evaluated by GRADE.

3.10 A minor point, line 163 should read participants (not participant).

Thank you for highlighting this typo. We have made the correction.

3.11 Despite addressing an important issue, being well conducted and well written, the main reason for my recommendation is that it substantially duplicates the material and conclusions previously published by Dupuis and colleagues.

Although our main results are substantially the same as Dupuis et al., we included quantitative pooled estimates and used the GRADE framework to provide knowledge users with evaluations of the quality of evidence underlying the use of steroids in potential organ donors. Moreover, although no new evidence was published, our review covers 3 more years.

Reviewer 4:
Reviewer Name: DR JUNE GOH
Institution and Country: SINGAPORE GENERAL HOSPITAL, SINGAPORE
Competing Interests: NONE DECLARED

4.1 The review is clear and concise.

Thank you for this positive comment.

4.2 3500 articles were screened out of which just the 11 were selected. Is it possible to give more detail as to why the rest were rejected?

A substantial portion (n=2500) of the 3500 citations were excluded before the step of the full text review. At that stage, we did not formally collect reasons for exclusion, but most studies assessed steroids in different populations or had non-RCT designs. As for the exclusion of articles after full-text review, the reasons are reported in Figure 1.

4.3 A similar review in the BJA in 2014 has been done before : Corticosteroids in the management of brain-dead potential organ donors, a systematic review. The difference would be the use of the GRADE methodology. If the studies chosen for the review was more recent or greater in number and dissimilar, then there would be greater value and less redundancy.

Yes, the main differences are the utilization of the GRADE approach for rating the quality of evidence and the meta-analysis. See previous responses.

VERSION 2 – REVIEW

REVIEWER	Sheila M Bird, visiting professor Department of Mathematics and Statistics, Strathclyde University, UK
REVIEW RETURNED	24-Feb-2017

GENERAL COMMENTS	I am pleased that authors have included comments about need for analysis to recognise that organs are clustered within donor and widening of CIs which correct analysis might entail; and remarks about timing of consent across RCTs. However, authors put much greater store on GRADE classification than I do because all meta-analysed RCTs, individually (& even summed) are far too small for even 50% power to discern modest 10% increase in organ retrieval. Hence, the current evidence-base gives no useful guidance because the intervention has been inadequately trialled. Authors do not

	suggest design criteria for an appropriately powered RCT.

REVIEWER	Geoffrey J Dobb Intensive Care Unit, Royal Perth Hospital, Western Australia and School of Medicine and Pharmacology, University of Western Australia
REVIEW RETURNED	10-Mar-2017

GENERAL COMMENTS	You have addressed the points raised by the reviewers and the editorial team in detail and made consequent changes to the manuscript. Within the constraints of the material available for the systematic review and meta-analysis I do not believe it would be possible to significantly improve the content. It remains an incremental advance on the previously published review that is referenced.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Reviewer Name: Sheila M Bird, visiting professor
Institution and Country: Department of Mathematics and Statistics, Strathclyde University, UK
Competing Interests: none declared

I am pleased that authors have included comments about need for analysis to recognise that organs are clustered within donor and widening of CIs which correct analysis might entail; and remarks about timing of consent across RCTs. However, authors put much greater store on GRADE classification than I do because all meta-analysed RCTs, individually (& even summed) are far too small for even 50% power to discern modest 10% increase in organ retrieval.

Hence, the current evidence-base gives no useful guidance because the intervention has been inadequately trialled. Authors do not suggest design criteria for an appropriately powered RCT.

Response: In the discussion section, we modified the last paragraph: "This systematic review highlights three types of challenges to research addressing the medical management of deceased organ donors: the scarcity of donors; practical challenges of studying therapeutic interventions and subsequent outcomes among very separate study populations, (i.e., organ donors and transplant recipients); and the complexity of definitions of graft function. To better guide clinical management of deceased donors will require strong research collaborations among donation and transplantation communities at a national or even international level. Scientifically sound, large clinical trials ideally will enrol consecutive eligible deceased donors, administer a single experimental steroid therapy in a blinded fashion, and measure outcomes not only among donors but also transplant recipients in a manner that allows the integration of transplant outcomes across organ groups. To achieve these goals may even require modification of current health services in donation and transplantation."

Reviewer: 3
Reviewer Name: Geoffrey J Dobb
Institution and Country: Intensive Care Unit, Royal Perth Hospital, Western Australia and School of Medicine and Pharmacology, University of Western Australia, Australia
Competing Interests: None declared

You have addressed the points raised by the reviewers and the editorial team in detail and made consequent changes to the manuscript. Within the constraints of the material available for the

systematic review and meta-analysis I do not believe it would be possible to significantly improve the content. It remains an incremental advance on the previously published review that is referenced.

Response: We thank the reviewer for this comment.