Unexplained mortality differences between septic shock trials: A systematic analysis of population characteristics and control-group mortality rates.

Harm-Jan de Grooth, Jonne Postema, Stephan A. Loer, Jean-Jacques Parienti, Heleen M. Oudemans-van Straaten, Armand R. Girbes

ELECTRONIC SUPPLEMENTARY MATERIAL - ADDITIONAL METHODS AND RESULTS

Table of contents

1. PRISMA checklist

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(7): e1000097. doi:10.1371/journal.pmed1000097

2. Search and selection strategy

PubMed, Embase and the Cochrane Central Register of Controlled Trials were queried using the search term [*"septic shock" AND (random* or rct)*]. Embase was additionally queried using the search term "septic shock" with the randomized controlled trial filter. The queries were limited to publications from 01-01-2006 and the queries were last performed on January 20, 2018. The list of excluded trials is appended at the end of this supplementary materials document.

eFigure 1. Results of the search and selection strategy. * Irrelevant search results were reports that were clearly and unambiguously irrelevant to the research question for a variety of reasons (too many to count individually). These reasons included, but were not limited to: Animal studies, other preclinical studies, healthy volunteer studies and letters. Authors of trials that did not report mortality or baseline characteristics were not contacted for additional information.

3. Estimation of 28-day mortality from other mortality measures

For the 20 trials that did not report 28-day mortality, we estimated 28-day mortality based on reported hospital, ICU or 90-day mortality using linear regression with data from trials that reported both 28-day and another mortality measure.

eTable 1. Equations used to estimate 28-day mortality.

eFigure 2. Regression estimation plots of 28-day mortality from (A) hospital mortality, (B) ICU mortality and (C) 90-day mortality.

4. Mean mortality rates and heterogeneity in subcategories of trials.

Mean mortality and heterogeneity parameters were estimated using a weighted random-effects model with mortality on the log-odds scale. Mortality was back-transformed from log-odds to percentages for clarity. The value of τ represents the estimated between-trial standard deviation in mortality on the logodds scale. The 95% confidence intervals indicate the precision of the estimated mean, while the 95% prediction intervals represents the estimated between-trial variability in mortality rates after adjusting for random chance and sample size (equal to the mean mortality rate ± 1.96 τ on the log-odds scale). There was significant heterogeneity in all subcategories of trials. * The median sample size separating the 50% largest from smallest trials was 36 patients.

5. Regression goodness-of-fit statistics.

eTable 3. Univariate regression goodness-of-fit statistics.

The model with the smallest AIC (Akaike Information Criterion, which penalized additional regressors) was chosen. For equal AIC's, the linear model was preferred.

6. Plots of the association between baseline variables and mortality

eFigure 3. Associations between control group mortality and selected population characteristics. Circle sizes are proportional to the trial sample sizes. The lines represent significant linear associations estimated using inverse variance weighted random-effects model with mortality on the log-odds scale. Morality rates are back-transformed to percentages for clarity. APACHE, Acute Physiology and Chronic Health Evaluation; SAPS, Simplified Acute Physiology Score; SOFA, Sequential Organ Failure Assessment score; MAP, mean arterial pressure

7. Results from multivariate linear model

To estimate the proportion of heterogeneity that can be explained by a combination of population characteristics, we constructed a weighted multivariate random-effects linear model with multiple imputation to account for missing values.

Population variables that were reported by at least 25% of the included trials with a univariate regression $R^2 \geq 0.10$ were candidate regressors: The proportion of patients with chronic heart disease, mean SAPS-II score, mean SOFA score, the proportion of patients on mechanical ventilation, the proportion of patients on vasopressors, mean serum creatinine, the amount of fluids before randomization, the proportion of Gram-positive infections and the proportion of culture-negative infections.

Multiple imputation (generating 20 datasets) with predictive mean matching was used to estimate missing observations from the candidate variables. In addition to the candidate regressors above, population variables that were reported by more than 50% of trials were included in the multiple imputation datasets to strengthen the validity of the imputation: The year of publication, mean age, the proportion of male patients, mean APACHE II score, mean mean arterial pressure, mean heart rate, mean serum lactate, and the proportions of patients with respiratory, abdominal and urological infections. These variables were included in the imputation but not in the multivariate regression analysis.

Candidate variables that were eliminated stepwise from the model (for P≥0.05) were, in order of elimination: The amount of fluids before randomization (P=0.779), the proportion of patients on vasopressors (P=0.528), the proportion of Gram-negative cultures (P=0.712), the proportion of culturenegative infections ($P=0.366$), the proportion of patients with chronic heart disease ($P=0.132$), the proportion of Gram-positive cultures (P=0.126), mean SAPS-II score (P=0.102).

The final model estimating the effect of the independent variables on control-group 28-day mortality (expressed on the log-odds scale) was:

eTable 4. Results from multivariate random-effects linear model

Mixed effects model ($k=65$; Tau² estimator: Maximum Likelihood) Tau² = 0.2955447; Tau = 0.5436402;

Test for Residual Heterogeneity: $QE(df = 61) = 685.7083$, p-val < .0001

 $R^2 = 0.4144$, residual $I^2 = 81.9\%$

Estimated function: *Logodds(mortality rate) = -3.341 + 0.103 × [mean SOFA score] +*

0.979 × [% mechanical ventilation] + 0.007 × [mean serum creatinine, µmol/L]

8. Results from regression tree analysis.

We constructed a regression tree using the Recursive Partitioning and Regression Tree (RPART) routines implemented in R (1, 2).

The dependent variable in the model was 28-day mortality. The recursive partitioning algorithm selected among all inclusion criteria and baseline characteristics the most informative variable, which was then 'split' at the value that best differentiates low- versus high mortality. The algorithm then selected the most informative variable for each of the two resulting subgroups, and split it again. When a splitting variable was missing for a specific trial, a surrogate variable (the variable most closely correlated to the splitting variable) was used.

The resulting regression tree is shown in eFigure 4. The cumulative proportion of variability in 28-day mortality explained the regression tree is (R^2) is 0.43 (eFigure 5, left panel). The cross-validated relative error decreased to below the the root (split 0) value, which indicates that the tree is likely not overfitted. (eFigure 5, right panel).

The algorithm output with primary splits and surrogate splits for missing values can be found at the end on this section.

eFigure 4. Regression tree for the prediction of 28-day mortality rate. The values between each terminal branch are the mean mortality rates for the n trials in the terminal branch. The splitting variables are: mean age (split at 64.8 years); the proportion of patients with a respiratory infection (split at 54.5%); the proportion of patients on mechanical ventilation (split at 74.3%); and the proportion of male patients (splits at 63.8% and 53.8%). The terminal numbers represent the mean mortality rate (and n number of trials) for that tree branch.

eFigure 5. The cumulative proportion of variability in 28-day mortality explained at each split (R², left panel) and the cross-validated relative error at each split (right panel). The R^2 for the whole tree is 0.42. The cross-validated relative error decreases to below the the root (split 0) value, which indicates that the tree is likely not overfitted.

R output:

```
rpart(formula = f, data = d, weights = sqrt(mort.cg.n), control = list(minsplit = 10))
n= 65
Node number 1: 64 observations, complexity param=0.1677901
  mean=0.354528, MSE=0.02436645 
  left son=2 (42 obs) right son=3 (22 obs)
  Primary splits:
    age.m.cg < 64.85 to the left, improve=0.1770292, (1 missing)
  Surrogate splits:
   male.p.cg < 0.678044 to the left, agree=0.747, adj=0.066, (1 split)
Node number 2: 42 observations, complexity param=0.09663411
  mean=0.3160983, MSE=0.01539218 
  left son=4 (12 obs) right son=5 (30 obs)
  Primary splits:
    mechvent.p.cg < 0.7434334 to the left, improve=0.1826555, (19 missing)
  Surrogate splits:
   population.incl.hyperlactatemia < 0.5 to the right, agree=0.862, adj=0.618, (19 split)
Node number 3: 22 observations, complexity param=0.1184911
  mean=0.4609157, MSE=0.03380377 
  left son=6 (13 obs) right son=7 (9 obs)
  Primary splits:
    oriresp.p.cg < 0.5447349 to the left, improve=0.32381720, (4 missing)
  Surrogate splits:
   oriabd.p.cg \leq 0.1540171 to the right, agree=0.741, adj=0.378, (2 split)
    year < 2013.5 to the right, agree=0.679, adj=0.229, (2 split)
Node number 4: 12 observations
```
Node number 5: 30 observations, complexity param=0.04253843 mean=0.3575663, MSE=0.01591277 left son=10 (20 obs) right son=11 (10 obs) Primary splits: male.p.cg < 0.6381818 to the left, improve=0.09969404, (1 missing) Surrogate splits: population.incl.mechanical.ventilation < 0.5 to the left, agree=0.772, adj=0.186, (1 split) Node number 6: 13 observations mean=0.3814903, MSE=0.0120705 Node number 7: 9 observations mean=0.5979004, MSE=0.0416422 Node number 10: 20 observations, complexity param=0.04253843 mean=0.3317222, MSE=0.0139477 left son=20 (13 obs) right son=21 (7 obs) Primary splits: male.p.cg < 0.537751 to the right, improve=0.23640110, (1 missing) Surrogate splits: year < 2014.5 to the left, agree=0.791, adj=0.013, (1 split) Node number 11: 10 observations mean=0.4258359, MSE=0.01467858 Node number 20: 13 observations mean=0.3013633, MSE=0.004979949

Node number 21: 7 observations mean=0.4484868, MSE=0.0312601

References

- 1. Atkinson EJ, Therneau TM. An introduction to recursive partitioning using the RPART routines. *Rochester Mayo Found* 2017;
- 2. Therneau T, Atkinson B, Ripley B. rpart: Recursive Partitioning and Regression Trees. *R Packag version 41-11* 2017;

9. List of excluded trials

This list contains the search results that were deemed possibly relevant (evaluated in full text) but which were excluded for the reasons stated below.

