Methods

Aim
The aim of this study was to investigate the association with mortality from combination versus monotherapy in a patient cohort with septic shock.

Null hypothesis
There is no difference in mortality between those receiving monotherapy compared to combination therapy.

Sample size estimation
A sample size of 1200 was estimated to provide 80% power to detect an 8% difference in mortality between groups, assuming equal numbers in each treatment group and baseline mortality of 40%, and with an alpha of 0.05 using a chi-square test.

Setting
The intensive care units (ICUs) of Royal Brisbane and Women’s Hospital (RBWH), Australia, and Rigshospitalet (RH), Denmark.

Original inclusion and exclusion criteria
1. Time period: 01.01.2009 – 31.12.2013
2. Inotropic use (= shock): Noradrenaline, Adrenaline, Dobutamine, Dopamine, Vasopressin, Metaraminol for > 24 hours
3. Antibiotic use > 72 hours
Exclusion criteria:
1. Age < 18 years

Study design
Retrospective observational study using data from electronical medical records.

Primary endpoint
In-hospital mortality

Data extraction
The following raw data (form, unit) was extracted from the electronic medical record

- age (integer, years)
- gender (categorical)
- hospital (categorical, Bris vs CPH)
- APACHE II score (integer, points)
- chronic health evaluation (categorical yes/no for 5 categories: respiratory, cardiac, liver, renal, immunosuppression).

For each antibiotic episode, the following were recorded:
- name (string)
- route (categorical, IV or PO)
- prescribed daily dosage (integer)
- start and cessation time (date, hour, min)

The following outcomes are available:
- in-hospital mortality (categorical)
- ICU mortality (categorical)

Additional data:
- Hospital and ICU admission and discharge dates (date, hour, min)

Data Management

1. **Categorizing antibiotics.** Antibiotics were categorized by their names into 21 categories. At RH, low doses of erythromycin were often given orally in order to stimulate the bowel. This was considered a separate category, not relevant for the analysis. The categories used were:
   1. beta-lactams
   2. aminoglycosides
   3. quinolones
   4. macrolides
   5. polypeptides
   6. glycopeptides
   7. oxazolidinones
   8. lincosamides
   9. sulfonamides
   10. tetracyklins
   11. antifungals
   12. folic acid inhibitors
   13. anti-helminths
   14. nitroimidazoles
   15. rifamycin
   16. sulfones
   17. lipopeptide
   18. fucidic acid
   19. onsoniazide
   20. nitrofurantoin
   21. erythromycin to stimulate bowel (excluded, see above)

2. **Definition of relevant empiric antibiotic combinations.** The following combinations were considered relevant to the study question:
   1. beta-lactam + aminoglycoside
   2. beta-lactam + quinolone
   3. beta-lactam + macrolide
   4. beta-lactam + colistin
   5. beta-lactam + glycopeptide
   6. beta-lactam + oxazolidinone
   7. beta-lactam + lincosamide
   8. carbapenem + metronidazole
   9. piperacillin / tazobactam + metronidazole
   10. lincosamide + metronidazole
   11. beta-lactam + fucidic acid
Carbapenems were defined as imipenem, meropenem, ertapenem or doripenem.

3. **Calculating treatment duration.** For each antibiotic episode, duration (in hours) was derived from start and cessation times. There was a difference in data quality between subjects from RBWH and RH, especially regarding date variables; data quality was much lower from RH. Cases with date inconsistencies (treatment starting before admission, treatment ongoing after discharge, treatment duration longer than hospitalization and negative treatment duration) were checked for obvious data entry errors (such as the one starting quinolone treatment in the year 1900). However, in most cases it was not obvious which date was wrong. Date inconsistencies like these were present in 0.5% of antibiotic episodes. For antibiotic episodes with such inconsistencies, treatment duration was labelled as missing/not applicable in order to be imputed.

4. **Missing data.** In the RH data, there was a number of antibiotic episodes for which cessation time was missing. This was due to a computerized system where the clinician had to terminate the antibiotic in a certain way in order to achieve a cessation time in the system. No starting times were missing. In total, 10.8% antibiotic episodes had a missing cessation time. In addition, three patients had a missing hospital length of stay.

5. **Imputation of antibiotic treatment duration.** In total, 11.4% of antibiotic episodes had a missing or inconsistent date and thus a missing duration. All missing cases were from RH. Given the reasons for missing data above, these were considered missing at random (MAR). These were imputed 10 times with a Multiple Imputation with Chained Equations (MICE) approach using predictive mean matching as the imputation method and 10 iterations to converge. In the imputation, the predictor matrix consisted of the variables age, gender, hospital, APACHE II score, chronic health evaluation, antibiotic category, treatment duration, in-hospital mortality and hospital length of stay. A datafile in the long format was created.

6. **Simultaneous starting time.** For each antibiotic episode, the difference in hours between patient admission and the start of that episode was established. This was done in the long dataset. If two (or more) antibiotics were started within 12 hours of each other, they were considered to have a simultaneous starting time, which was one of the prerequisites for combination therapy.

7. **Determining which combination.** If two antibiotic episodes had a starting time within 12 hours of each other, as defined above, they were screened for appropriate combinations. For example, a betalactam and a quinolone that started simultaneously would be considered a possible combination, but a quinolone and a glycopeptide would not, according to the list above.

8. **Duration of combination (overlap).** For two antibiotics with a simultaneous start and within the prespecified categories, the overlapping duration was determined. The overlapping duration was defined as the first cessation time (i.e. the last start time). In the imputed dataset, this was possible for all antibiotic episodes. For complete case analysis several overlapping durations were considered missing. If one of the antibiotics had a missing cessation time, the overlapping duration was considered missing (Fig. S1).
This patient was given two antibiotics from different classes. They were initiated within 12 hours of each other. The beta-lactam was given for 46 hours but the cessation time and duration of the quinolone was unknown. Most likely, this represents a valid combination as the median duration of quinolone therapy at RH was much longer than the 12 hours needed to fulfill the overlapping duration criterion. However, at this point this combination could not be considered as either valid or invalid without imputation. The imputation used data regarding center (RH), antibiotic category (quinolone), treatment duration for non-missing cases etc. to make 10 imputations of duration, which were then pooled in the analysis.

9. **Determining the number of valid combinations per patient.** A valid combination could be identified as two antibiotics from prespecified combinations started within 12 hours of each other and with an overlapping duration of at least 12 hours. The combination beta-lactam + aminoglycoside was exempted from the overlapping duration criteria, as aminoglycosides are often given as a single dose. The number of valid combinations per patient was calculated.

10. **Adding the main exposure variable, any combination (yes/no).** All patients with at least one valid combination were considered as receiving combination therapy. In the imputed datasets, all patients could be classified. For the complete case analysis, those who had only combinations with missing duration were classified as missing.

11. **Adding more variables.** For each combination, the following types of variables, beta-lactam + aminoglycoside (integer), any beta-lactam+ aminoglycoside (yes/no) and duration of beta-lactam + aminoglycoside etc. were constructed. An overall variable for number of combinations and hours of combination treatment was also derived.

12. **Back transforming long dataset.** The long dataset with imputed variables, as well as variables derived from imputations, were back transformed to a .mids object, in order to enable pooled estimates from statistical modelling (see R script).

**Step by step statistical analysis**
The main outcome (in-hospital mortality) will be the dependent variable in a logistic regression analysis, using combination treatment as the main predictor, controlling for the other baseline variables.

1. **Determining outcome.** The frequency of the main outcome (in-hospital mortality) will be determined. Any cases with missing outcome will be discarded.
2. Crude analysis. Separate bivariate logistic regressions will be carried out for all covariates. This will be done one-by-one, on the original scaling.

3. Variable transformations. To relax the assumption of linearity, the continuous covariates (age and APACHE II score) will be modeled as 4-knot restricted cubic splines.

4. Fitting the multivariate model. The imputations and transformations will be used to fit the multivariate model.

5. Interactions. For the combination therapy variable, a global interaction test will be performed. If this global test is significant, specific interactions will be pursued.

6. Internal validation. The model will be validated through 1000 bootstrap resamples to determine the amount of overfitting and obtain optimism-corrected performance estimates for overall performance measures including c statistic, Brier score and R².

7. Goodness-of-fit. This was evaluated using a Hosmer-Lemeshow test with 20 quantiles as well as through a calibration plot using 1000 bootstrap resamples.

8. Sensitivity analysis. A complete case analysis was performed using the steps outlined above and is presented below.

**Calibration plot**

![Calibration plot](image)

**Figure S2.** A calibration plot for the main multivariable model from the manuscript showed good overall calibration, but less so in the higher predicted values where cases were few.
Subgroup analysis – different antibiotic combinations

Four groups of antibiotic combinations according to their adjudicated main intended treatment benefit:

**Combination therapy of gram-negative bacteria:**
- Beta-lactam + Aminoglycoside
- Beta-lactam + Quinolone
- Beta-lactam + Colistine

**Combination therapy of anaerobes:**
- Carbapenem + Metronidazole
- Piperacillin / Tazobactam + Metronidazole
- Lincosamide + Metronidazole

**Gram positive double coverage and extended spectrum:**
- Beta-lactam + Lincosamide
- Beta-lactam + Glycopeptide
- Beta-lactam + Oxazolidinone
- Beta-lactam + Fucidic acid.

**Extended spectrum for atypical pathogens**
- Beta-lactam + Quinolone
- Beta-lactam + Macrolide

Table S1. Subgroup analysis of different antibiotic combinations.

| Subgroups                      | Coef. | S.E  | Wald z | p-value | Adjusted OR (95% CI) |
|--------------------------------|-------|------|--------|---------|----------------------|
| Gram neg. combination         | -0.15 | 0.17 | -0.86  | 0.39    | 0.86 (0.62 – 1.21)   |
| Anaerobe combination          | -0.25 | 0.29 | -0.87  | 0.38    | 0.78 (0.44 – 1.37)   |
| Gram pos. combination         | -0.16 | 0.16 | -1.02  | 0.31    | 0.85 (0.62 – 1.16)   |
| Atypical combination          | -0.29 | 0.19 | -1.52  | 0.13    | 0.75 (0.52 – 1.09)   |

Analyses of the separate combinations versus no combination. All analyses are adjusted for age, sex, center, APACHE II score, cardiac disease, pulmonary disease, hepatic disease, renal disease and immunosuppression. coef. = beta coefficient, S.E = standard error, OR = Odds ratio, CI = confidence interval.

Sensitivity analysis – complete case analysis

Patients with missing data that hindered strict classification regarding their combination therapy status were excluded from complete case analysis. In total, 142 patients of 1667 (8.5%) could not be classified. As stated before, these were all from RH. Of the 1525 complete cases, 844 (55%) were given combination therapy and 681 (45%) were not. When treatment duration was imputed, 110 of the 142 were considered to be in the combination group and 32 in the no combination group in the full dataset. The baseline characteristics for complete cases was similar to that in the full dataset (Table S2).

Antibiotics complete cases vs full dataset
At RBWH (where all cases were complete), 404 of 1006 (40%) were treated with a combination, equivalent to in the manuscript. At RH, 440 (85%) of the 519 complete cases
were given a combination, compared to 550 of 661 (83%) in the full dataset. The number and duration of antibiotics at RH for complete cases vs. the full dataset are displayed in Table S3.

Regression analyses
The separate, univariate crude analysis against hospital mortality was largely similar for complete cases and the full dataset as seen in Table S4. In the multiple analysis, only minor differences were seen in the models when complete cases and the full dataset were analyzed (Fig. S3). The complete cases model had an $R^2$ of 0.20 and a Brier score of 0.18, with optimism-corrected estimates of 0.18 and 0.18 respectively, after internal validation, representing a slope of 0.93.

Interpretation of sensitivity analysis
With the data quality issues from RH, we deemed it necessary to impute durations for antibiotic combinations. These imputations did not alter the results in a significant way.

Table S2. Baseline characteristics for complete cases vs full dataset

| Variable                  | Complete cases | Full dataset |
|---------------------------|----------------|--------------|
|                           | No combination | Combination therapy (n = 681) | No combination | Combination therapy (n = 713) |
| Age, mean (SD)            | 55(18)         | 59(15)       | 56 (18)        | 60 (15)                      |
| Sex, male                 | 408 (60%)      | 400 (47%)    | 420 (59%)      | 438 (46%)                    |
| Hospital, RH              | 79 (12%)       | 440 (52%)    | 111 (16%)      | 550 (58%)                    |
| APACHE II score, mean (SD)| 23 (8)         | 26 (8)       | 23 (8)         | 26 (8)                       |
| Respiratory disease       | 29 (4%)        | 50 (6%)      | 32 (4%)        | 55 (6%)                      |
| Cardiac disease           | 31 (5%)        | 48 (6%)      | 33 (5%)        | 58 (6%)                      |
| Liver disease             | 29 (4%)        | 59 (7%)      | 31 (4%)        | 70 (7%)                      |
| Renal disease             | 28 (4%)        | 54 (6%)      | 30 (4%)        | 60 (6%)                      |
| Immunosuppression         | 90 (13%)       | 206 (24%)    | 93 (13%)       | 233 (24%)                    |
| ICU stay median (IQR), days| 10 (7-16)     | 13 (7-22)    | 10 (6-16)      | 12 (7-21)                    |

Continuous variables are described with mean (SD) or median (IQR), categorical variables with counts (percentages of column total. RH = Rigshospitalet.

Table S3. Antibiotics at RH, complete cases vs full dataset

| Antibiotic combination | Complete cases | Full dataset |
|------------------------|----------------|--------------|
|                        | Count | Duration (IQR) | Count | Duration (IQR) |
| Beta-lactam + quinolone| 380   | 144 (72-219)   | 479   | 120 (64-216)   |
| Beta-lactam + glycopeptide| 95     | 158 (82-240)   | 122   | 122 (63-216)   |
| Carbapenem + nitroimizadole| 192   | 159 (96-235)   | 234   | 144 (72-233)   |
| Beta-lactam + aminoglycoside| 0     | NA            | 0     | NA             |
| Beta-lactam + lincomamide| 78     | 120 (72-214)   | 110   | 96 (58-168)    |
| Beta-lactam + macrolide  | 6     | 115 (86-156)   | 8     | 110 (72-120)   |
| Beta-lactam + oxazolidinone| 16   | 67 (36-137)    | 23    | 72 (37-168)    |
| Beta-lactam + colistin   | 19    | 178 (144-350)  | 30    | 168 (72-312)   |
| Lincomamide + nitroimizadole| 26   | 144 (64-192)   | 28    | 144 (55-192)   |
| Beta-lactam + fucidic acid| 18    | 80 (31-168)    | 19    | 96 (26-168)    |
| Piperacillin/tazobactam + nitroimizadole| 1     | 65 (65-65)    | 1     | 65 (65-65)    |
| All combinations         | 831   | 144 (72-221)   | 1054  | 120 (59-216)   |

Note: individual patients can have more than one combination. Expressed as median (interquartile range). RH = Rigshospitalet.
Table S4. Crude analysis of mortality, complete cases vs full dataset

| Variable                  | Complete cases |               | p value | Full dataset |               | p value |
|---------------------------|----------------|---------------|---------|--------------|---------------|---------|
|                           | OR (95%CI)     | OR (95%CI)    |         |              |               |         |
| APACHE II score (points)  | 1.09 (1.08 – 1.11) | 1.09 (1.07 – 1.10) | <0.001 | <0.001       |               |         |
| Immunosuppression         | 2.96 (2.28 – 3.85) | 3.03 (2.35 – 3.89) | <0.001 | <0.001       |               |         |
| Age (years)               | 1.03 (1.02 – 1.04) | 1.03 (1.02 – 1.04) | <0.001 | <0.001       |               |         |
| Hospital, RH              | 2.07 (1.65 – 2.60) | 1.85 (1.49 – 2.29) | <0.001 | <0.001       |               |         |
| Liver disease             | 2.70 (1.75 – 4.18) | 2.35 (1.56 – 3.52) | <0.001 | <0.001       |               |         |
| Renal disease             | 2.43 (1.55 – 3.81) | 2.36 (1.54 – 3.62) | <0.001 | <0.001       |               |         |
| Cardiac disease           | 2.36 (1.49 – 3.72) | 2.10 (1.37 – 3.22) | <0.001 | <0.001       |               |         |
| Respiratory disease       | 1.60 (1.00 – 2.54) | 1.78 (1.15 – 2.77) | 0.05   | 0.01         |               |         |
| Combination therapy       | 1.35 (1.08 – 1.69) | 1.33 (1.07 – 1.66) | 0.01   | 0.01         |               |         |
| Sex, male                 | 0.86 (0.69 – 1.07) | 0.86 (0.69 – 1.05) | 0.16   | 0.16         |               |         |

CI = confidence interval, OR = Odds Ratio. RH = Rigshospitalet. All covariates are presented on their original scale, coded as indicated.
Figure S3 Main effects of multiple model displaying log odds for the predictors, including non-linear effects, sorted horizontally by decreasing association with the outcome in-hospital mortality. Complete cases (top panel) and full dataset (bottom panel).
Robustness analysis – duration of combination therapy

To test the robustness of the results, different lengths of overlapping treatment duration for the antibiotic combinations, i.e. 12 hours, 24 hours, 36 hours, 48 hours and 72 hours, were analysed. These were tested in a dichotomous way (i.e. combination >24 hours vs. combination <24 hours etc.), as a factor (12 hours vs. no combination, 24 hours vs. no combination etc.) and in an ordinal way. Combination therapy was not associated with outcome in any of these analyses (Table S5).

**Table S5.** Robustness analyses – overlapping duration

| Duration of combination | Coef. | S.E  | Wald z | p-value | Adjusted OR (95% CI) |
|-------------------------|-------|------|--------|---------|----------------------|
| 12+ h vs. none          | -0.12 | 0.13 | -0.91  | 0.36    | 0.88 (0.68 – 1.15)  |
| 24+ h vs. none          | -0.06 | 0.13 | -0.51  | 0.61    | 0.93 (0.72 – 1.21)  |
| 36+ h vs. none          | -0.01 | 0.14 | -0.07  | 0.94    | 0.99 (0.76 – 1.29)  |
| 48+ h vs. none          | 0.03  | 0.14 | 0.25   | 0.80    | 1.03 (0.79 – 1.35)  |
| 72+ h vs. none          | 0.03  | 0.14 | 0.23   | 0.82    | 1.03 (0.79 – 1.34)  |
| 12 h – 24 h vs. none    | -0.28 | 0.32 | -0.89  | 0.37    | 0.76 (0.41 – 1.40)  |
| 24 h – 36 h vs. none    | -0.31 | 0.30 | -1.03  | 0.30    | 0.73 (0.40 – 1.33)  |
| 36 h – 48 h vs. none    | -0.29 | 0.35 | -0.82  | 0.41    | 0.75 (0.38 – 1.49)  |
| 48 h – 72 h vs. none    | -0.06 | 0.28 | -0.20  | 0.84    | 0.95 (0.55 – 1.64)  |
| 72 h+ vs. none          | -0.05 | 0.15 | -0.34  | 0.73    | 0.95 (0.75 – 1.28)  |
| 48+ h vs. 12-48 h       | 0.26  | 0.20 | 1.28   | 0.20    | 1.30 (0.87 – 1.93)  |

Duration of combination vs no combination. All analyses are controlled for age, sex, center, APACHE II score, cardiac disease, pulmonary disease, hepatic disease, renal disease and immunosuppression. Coef. = beta coefficient, S.E = standard error, OR = Odds ratio, CI = confidence interval.