THE LANCET
Respiratory Medicine

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Smit JM, Van Der Zee PA, Stoof SCM, et al. Predicting benefit from adjuvant therapy with corticosteroids in community-acquired pneumonia: a data-driven analysis of randomised trials. *Lancet Respir Med* 2025; published online Jan 29. https://doi.org/10.1016/S2213-2600(24)00405-3.

Supplementary appendix:

Predicting benefit from adjuvant therapy with corticosteroids in community-acquired pneumonia: a data-driven analysis of randomized trials.

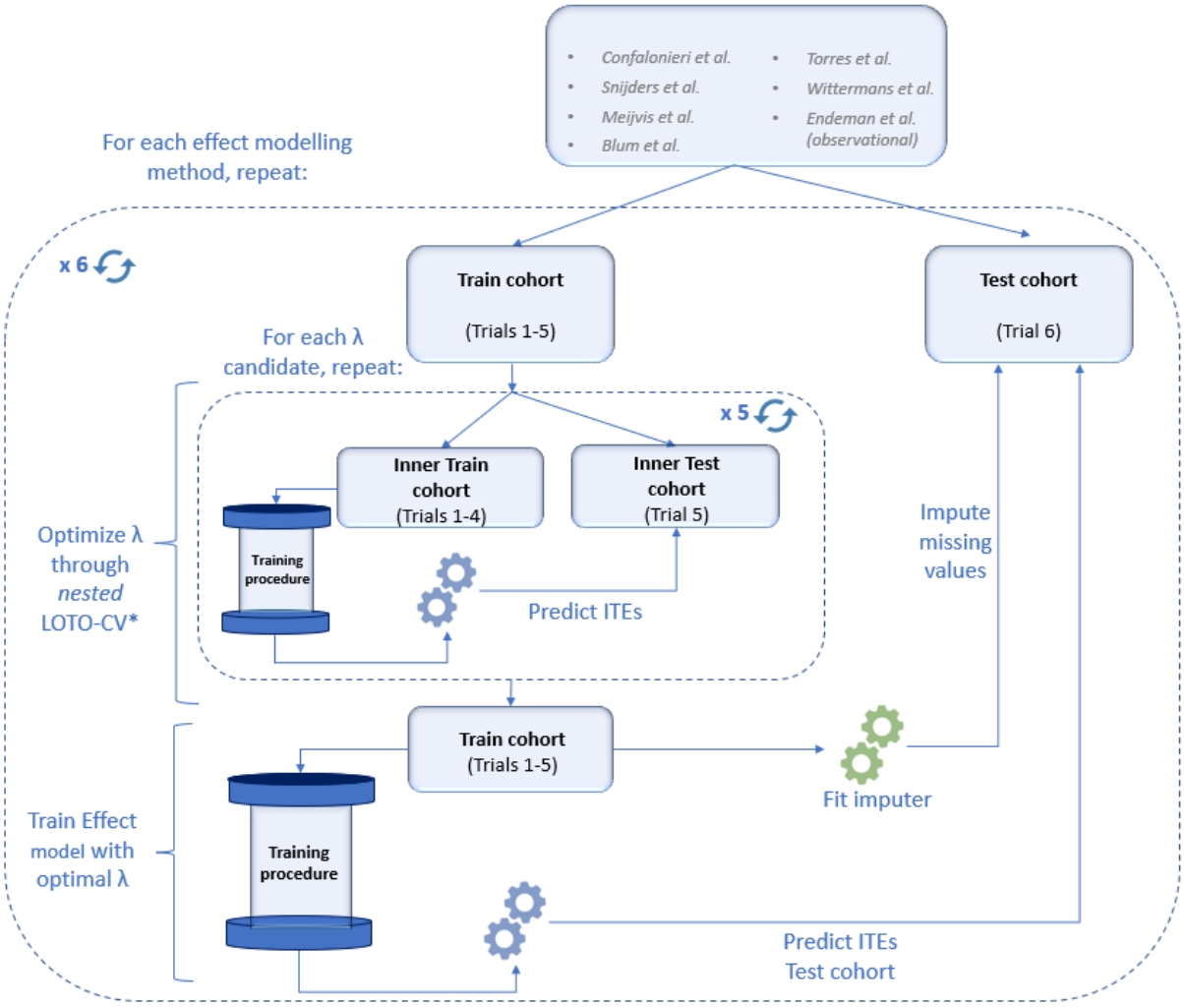
J.M. Smit[1,2], P.A. Van Der Zee[1,3], S.C.M. Stoof[1], M.E. Van Genderen[1], D. Snijders[4], W. G. Boersma[5], P. Confalonieri[6], F. Salton[6], M. Confalonieri[6], M-C. Shih[7], G.U. Meduri[8] , P.-F. Dequin[9], A. Le Gouge[10], M. Lloyd[11], H. Karunajeewa[12], G. Bartminski[12], S. Fernández-Serrano[13], G. Suárez-Cuartín[14], D. van Klaveren[15,16], M. Briel[17,18], C.M. Schönenberger [17], E.W. Steyerberg[19], D.A.M.P.J. Gommers[1], H.I. Bax[20], W J. W. Bos[21,22], E.M.W. Van De Garde[23], E. Wittermans[21], J.C. Grutters[24], C.A. Blum[25,26], M. Christ-Crain[26], A. Torres[27], A. Motos[28,29], M.J.T. Reinders[2], J. Van Bommel[1], J.H. Krijthe[2], H. Endeman[1]

1 Department of Intensive Care, Erasmus MC - University Medical Center Rotterdam, Rotterdam, The Netherlands
2 Pattern Recognition & Bioinformatics group, Delft University of Technology, Delft, The Netherlands
3 Department of Pulmonary Medicine, Erasmus MC - University Medical Center Rotterdam, Rotterdam The Netherlands
4 Department of Pulmonary Medicine, Spaarne Gasthuis, Haarlem, The Netherlands
5 Department of Pulmonary Medicine, Noordwest Hospital, Alkmaar, The Netherlands
6 Department of Pulmonary Medicine, University Hospital of Cattinara, Trieste, Italy
7 Department of Veterans Affairs, Cooperative Studies Program Coordinating Center, Palo Alto, CA, USA
8 Department of Pharmaceutical Sciences, University of Tennessee Health Science Center, Memphis, USA
9 Medecine intensive reanimation, Chru Hôpitaux De Tours, Hospital Bretonneau, Tours, France
10 INSERM CIC1415, CHRU de Tours, Tours, France.
11Centre for Medicine Use and Safety, Monash University, Melbourne, Australia.
12 Department of Medicine, The Western Health Chronic Disease Alliance and the University of Melbourne, Melbourne, Australia
13 Pulmonary Medicine, Badalona Serveis Assistencials S.A., Badalona, Spain
14 Pulmonary Medicine, Bellvitge University Hospital, L'Hospitalet de Llobregat, Spain
15 Predictive Analytics and Comparative Effectiveness Center, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, USA
16 Department of Public Health, Erasmus University Medical Center, Rotterdam, the Netherlands
17 CLEAR Methods Center, Division of Clinical Epidemiology, Department of Clinical Research, University Hospital Basel, University of Basel, Basel, Switzerland
18 Department of Health Research Methodology, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada
19 Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, Netherlands
20 Department of Medical Microbiology and Infectious Diseases and the Department of Internal Medicine, Section of Infectious Diseases , Erasmus University Medical Center, Rotterdam, The Netherlands
21 Department of Internal Medicine, St. Antonius Hospital, Nieuwegein, Netherlands
22 Department of Internal Medicine, Leiden University Medical Center, Leiden, Netherlands
23 Department of Pharmacy, St. Antonius Hospital, Utrecht/Nieuwegein, Netherlands
24 Department of Pulmonary Medicine, St. Antonius Hospital, Nieuwegein, Netherlands
25 Hormonpraxis Aarau, Aarau, Switzerland
26 Division of Endocrinology, Diabetes and Metabolism, Department of Clinical Research, Universitätsspital Basel, Basel, Switzerland
27 Department of Pulmonology, Hospital Clinic. University of Barcelona, Barcelona, Spain
28 Centro de Investigación Biomédica En Red-Enfermedades Respiratorias (CIBERES), Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Universitat de Barcelona, Barcelona, Spain.
29 Center for Research in Transplantation and Translational Immunology, UMR 1064, Nantes Université, Inserm, CHU Nantes, Nantes, France.

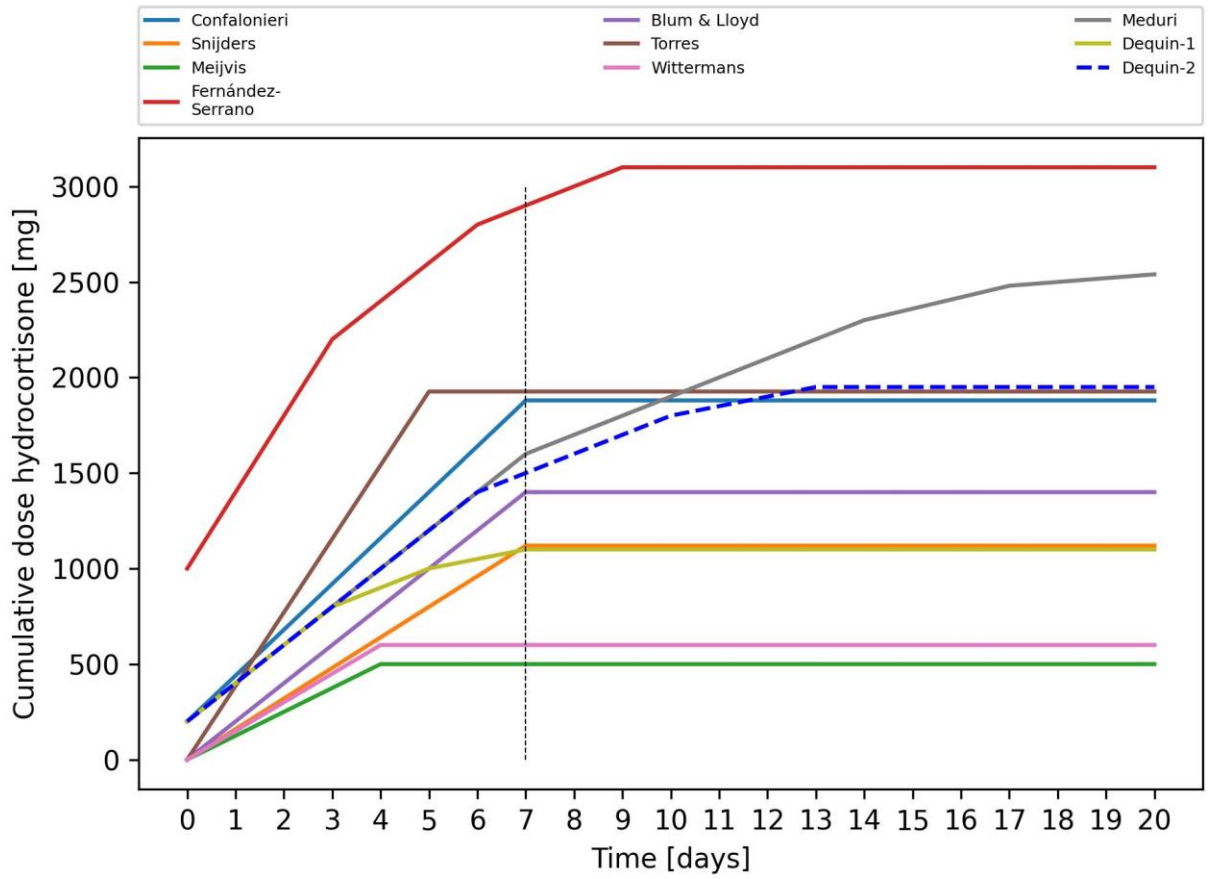
Outline

Appendix Part 1: Supplementary Figures and Tables

Appendix Figure S1: Schematic overview of the 'leave-one-trial-out' (LOTO) cross-validation procedure for method selection.*The training procedure is described in more detail in Appendix Part 5 (p 41). * The *nested* 'leave-one-trial-out' cross-validation procedure is the exact same as the procedure described in Appendix Part 5, but then with 5 of the 6 trials which form the train cohort in the fold of the outer LOTO cross validation.

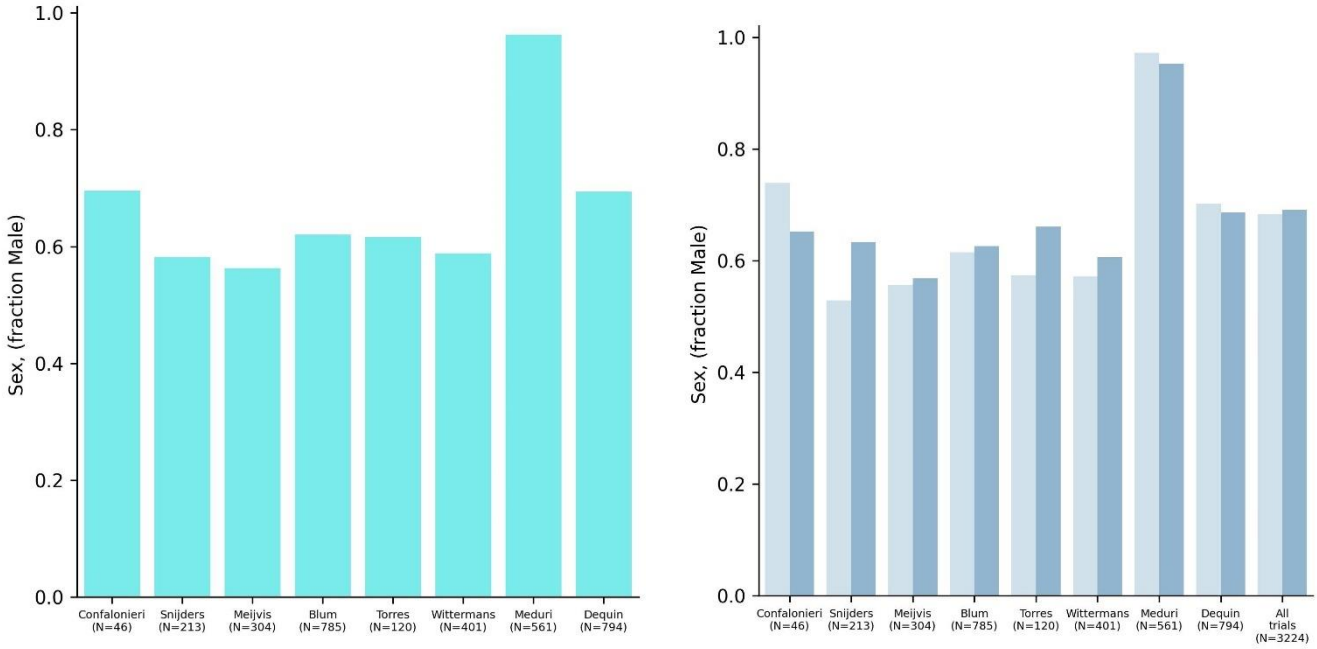


Appendix Figure S2: Cumulative dose of corticosteroids for each study. All doses were transformed into equivalent quantities of hydrocortisone (in mg), using Clincalc's Corticosteroid Conversion Calculator.(1) To calculate the cumulative dose in the treatment regime of *Torres et al*(2)., which assigned patients in the treatment arm to 0.5 mg/kg per 12 hours of methylprednisolone, we assumed an average weight of 84 kg for male patients and 65.9 kg for female patients.(3)

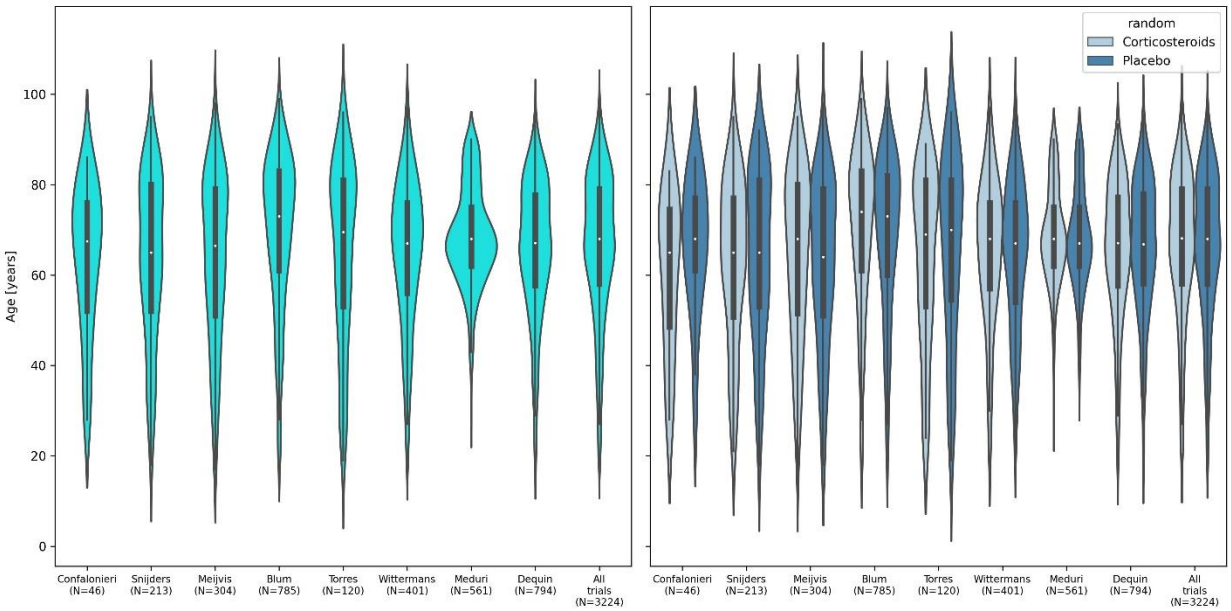


Appendix Figure S3: Violin plots representing the distributions of included variables among the patients from the different included trials and the observational study(4) (left panel) and the distributions split for treatment arms for all included trials (right panel). The x-axis specifies the number of patients per distribution (which could be smaller than the study size due to missingness). In some trials, a variable was completely missing and therefore no distribution is plotted. Distributions of the placebo and corticosteroid arms were compared using a Fisher exact test for categorical variables and a two-sample t test for continuous variables, without adjusting for multiple testing. Significant differences between the distributions (ie, P<0.05) are marked with an asterisk (*).

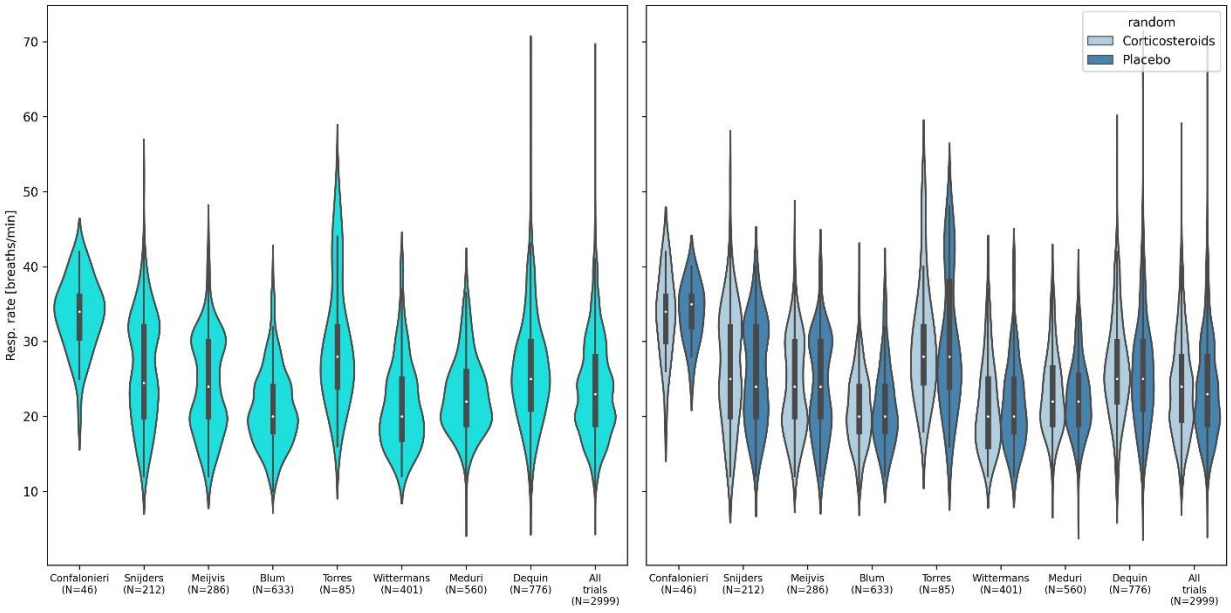
(a) Sex



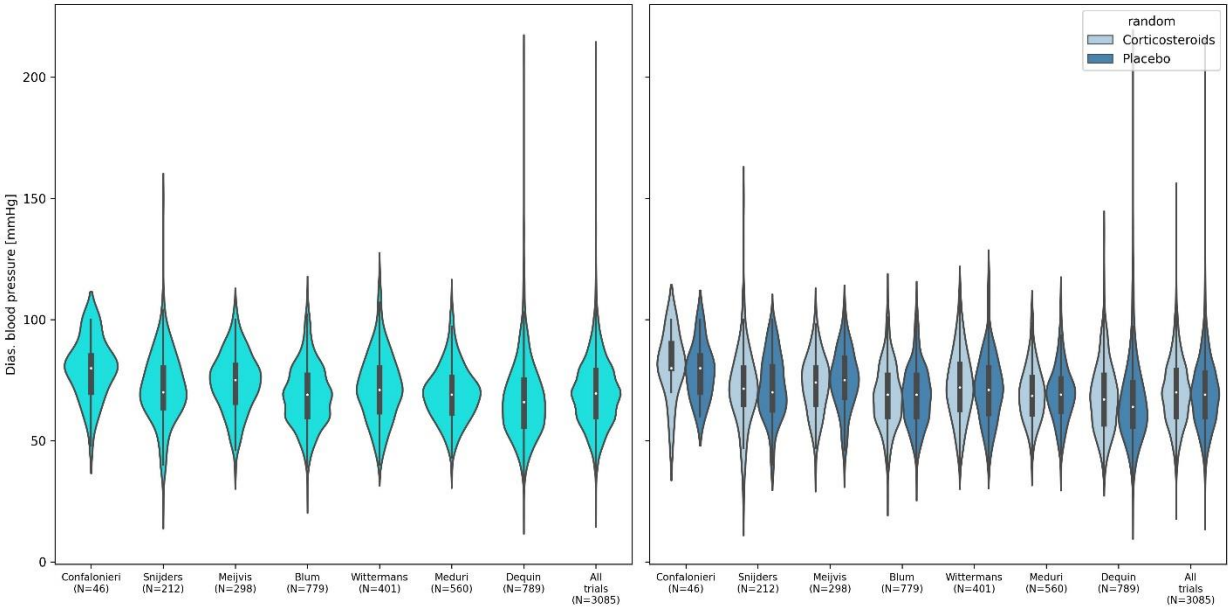
(b) Age



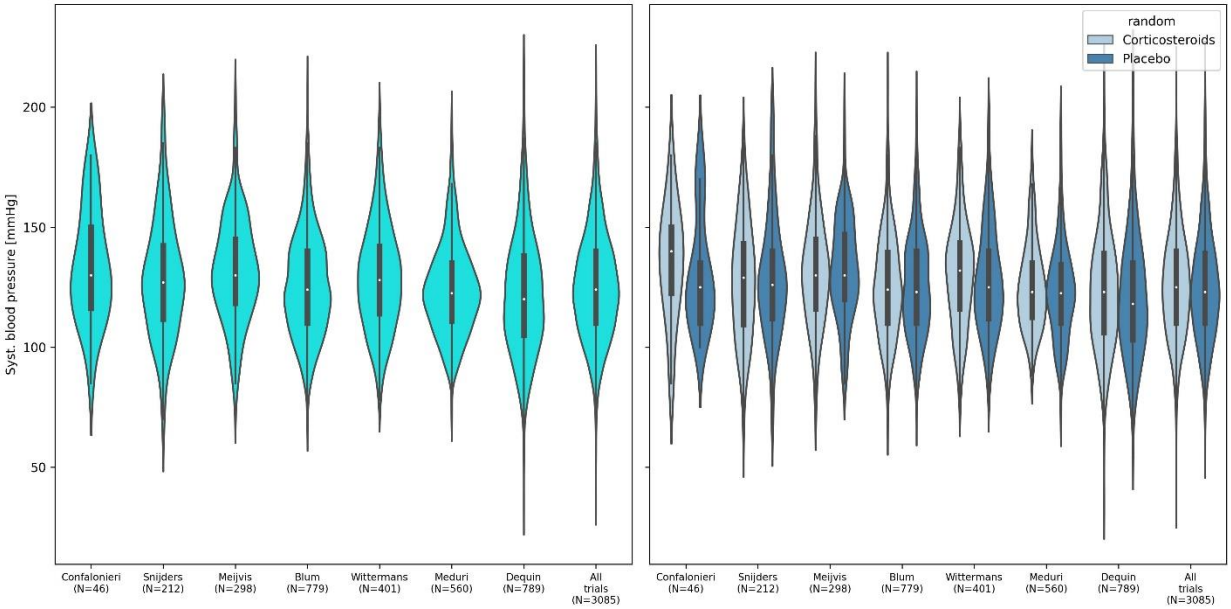
(c) Respiratory rate



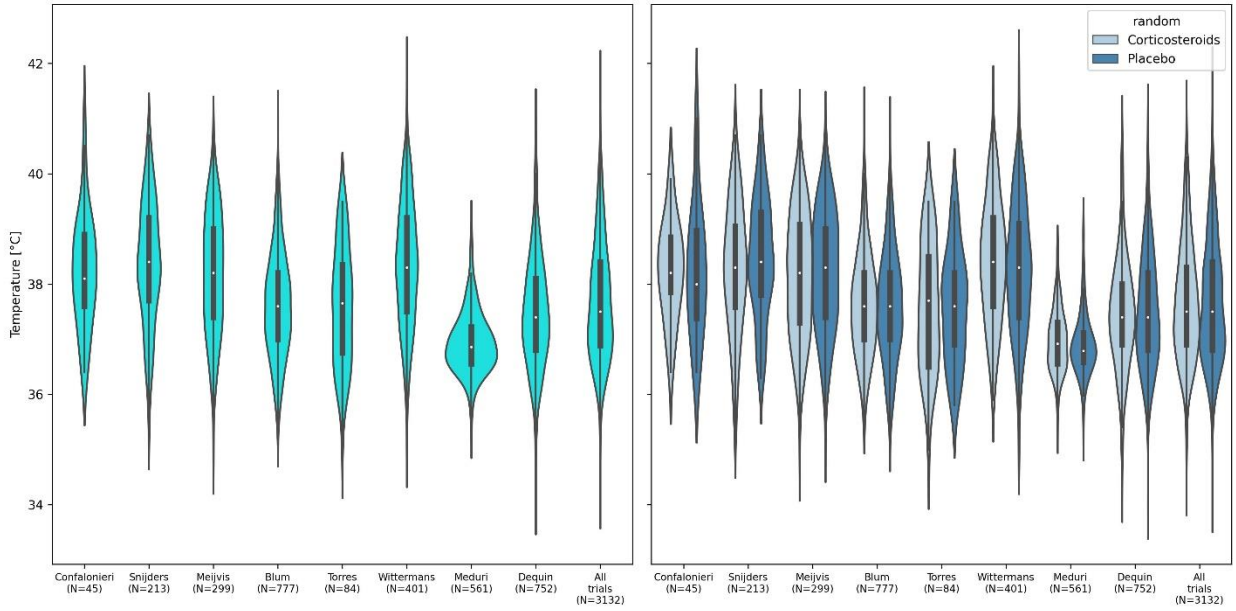
(d) Diastolic blood pressure



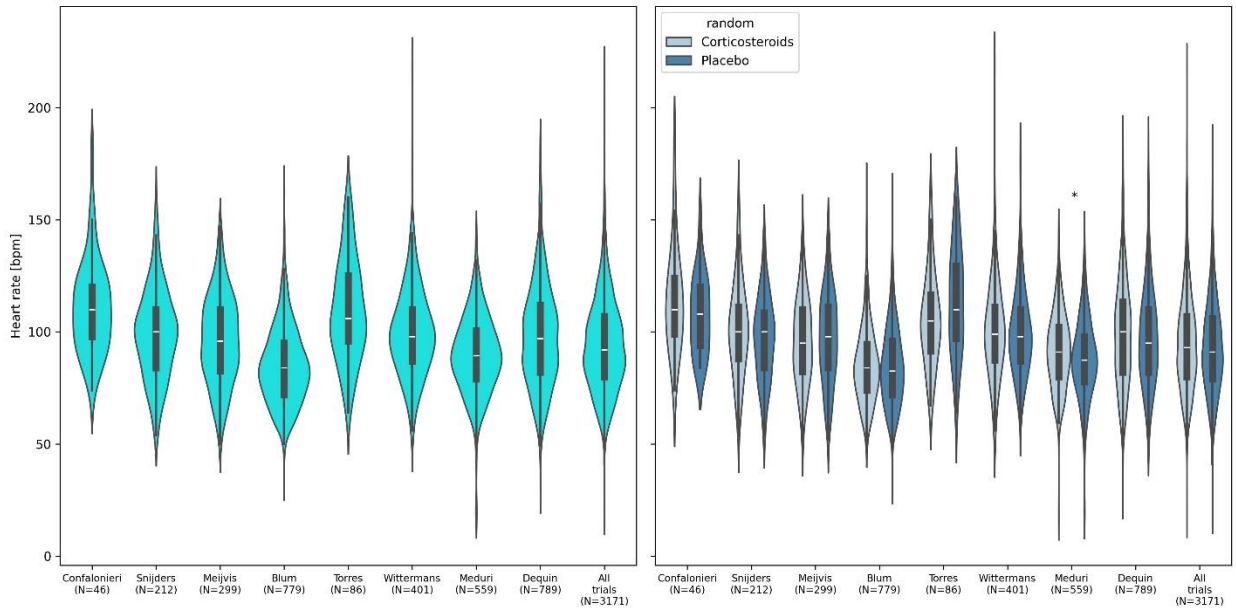
(e) Systolic blood pressure



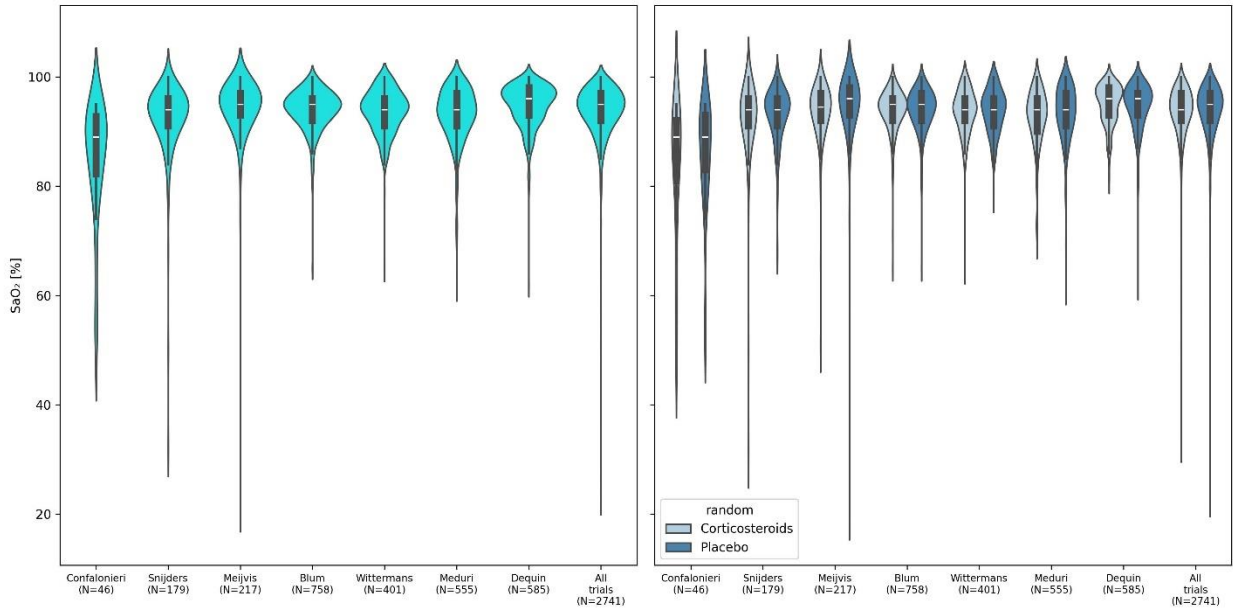
(f) Temperature



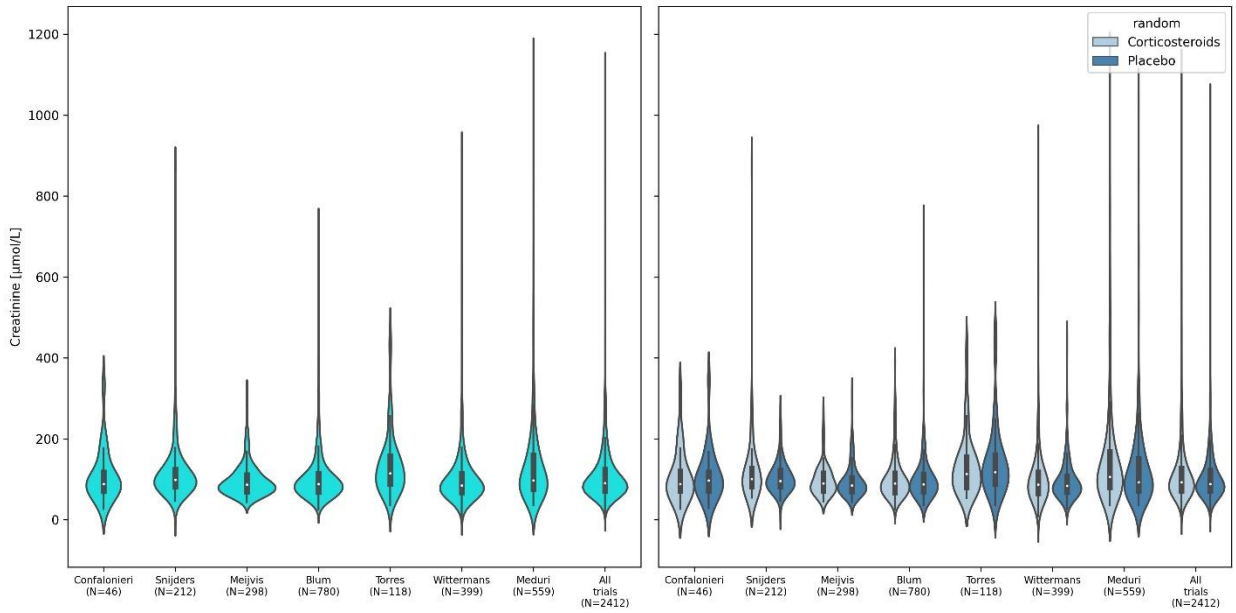
(g) Heart rate



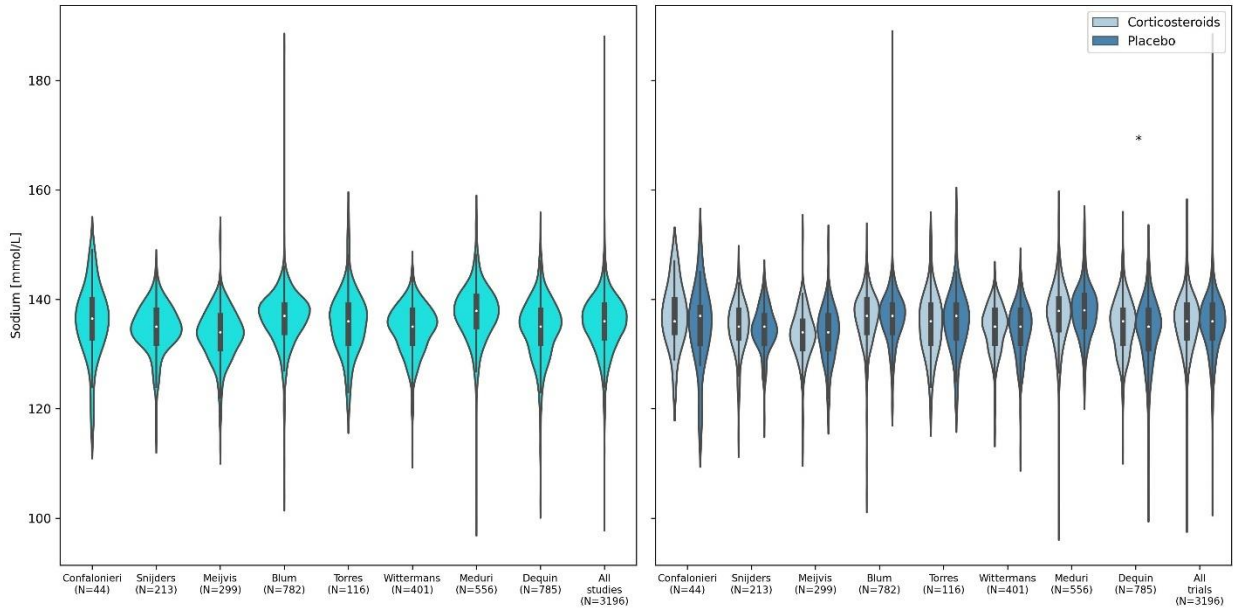
(h) Oxygen saturation (SaO_2)



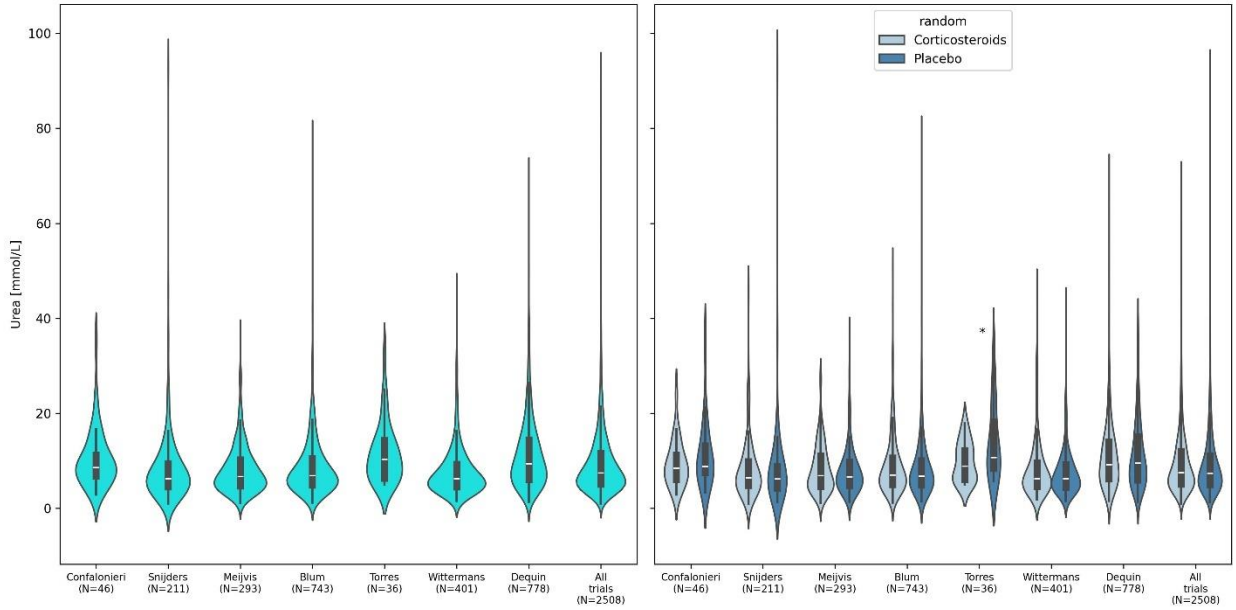
(i) Creatinine



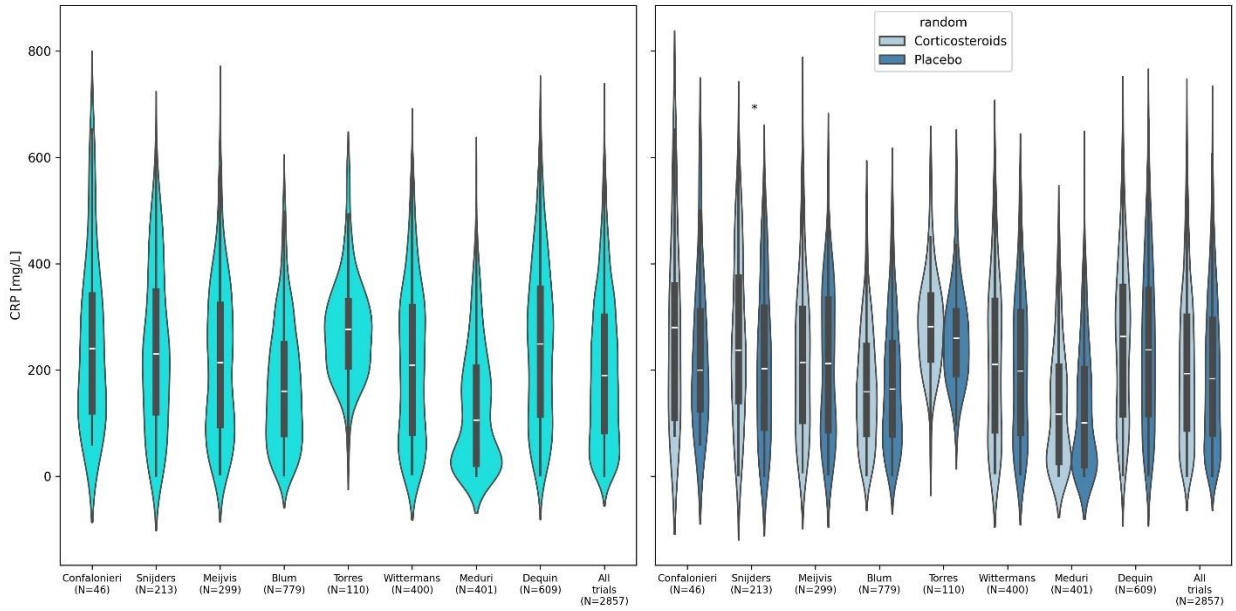
(j) Sodium



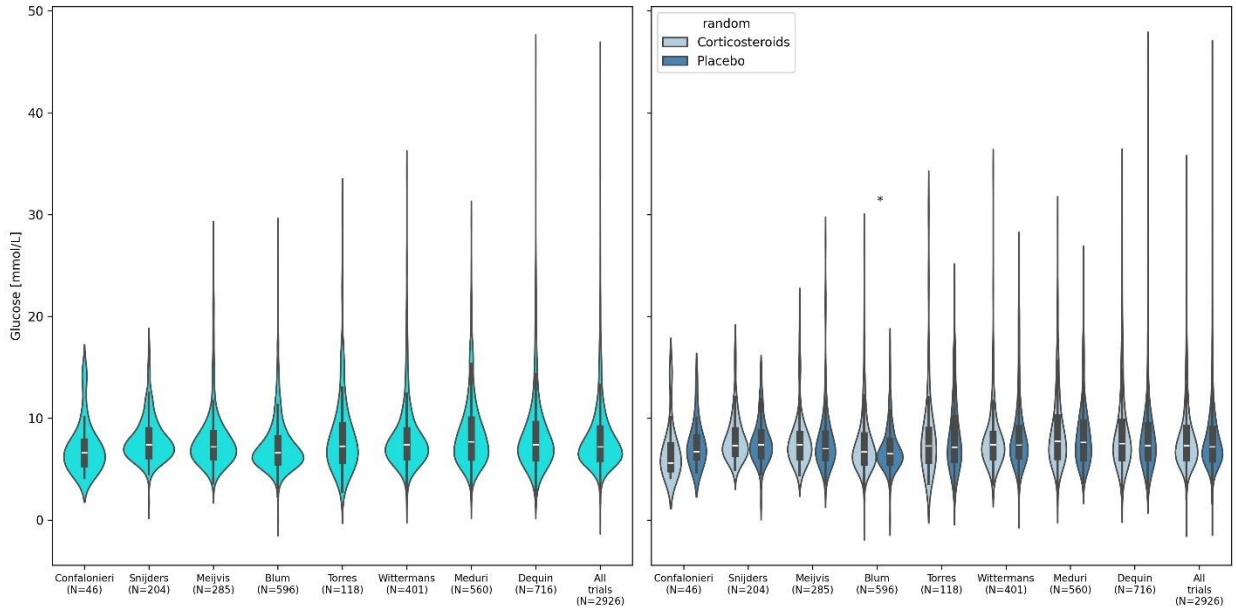
(k) Urea



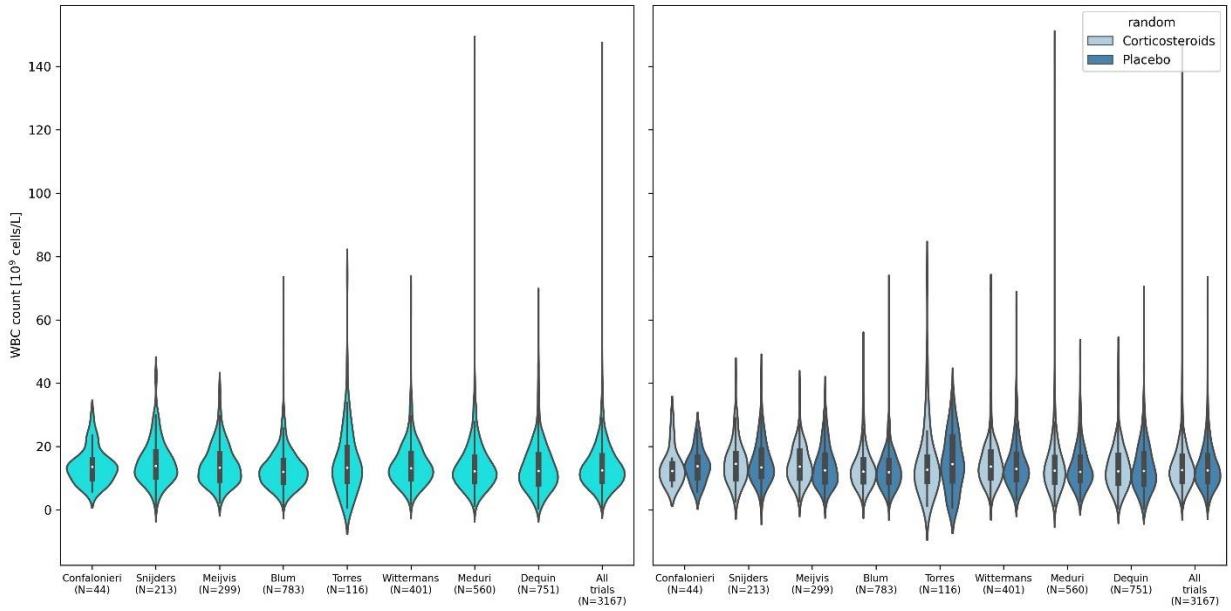
(l) C-reactive protein



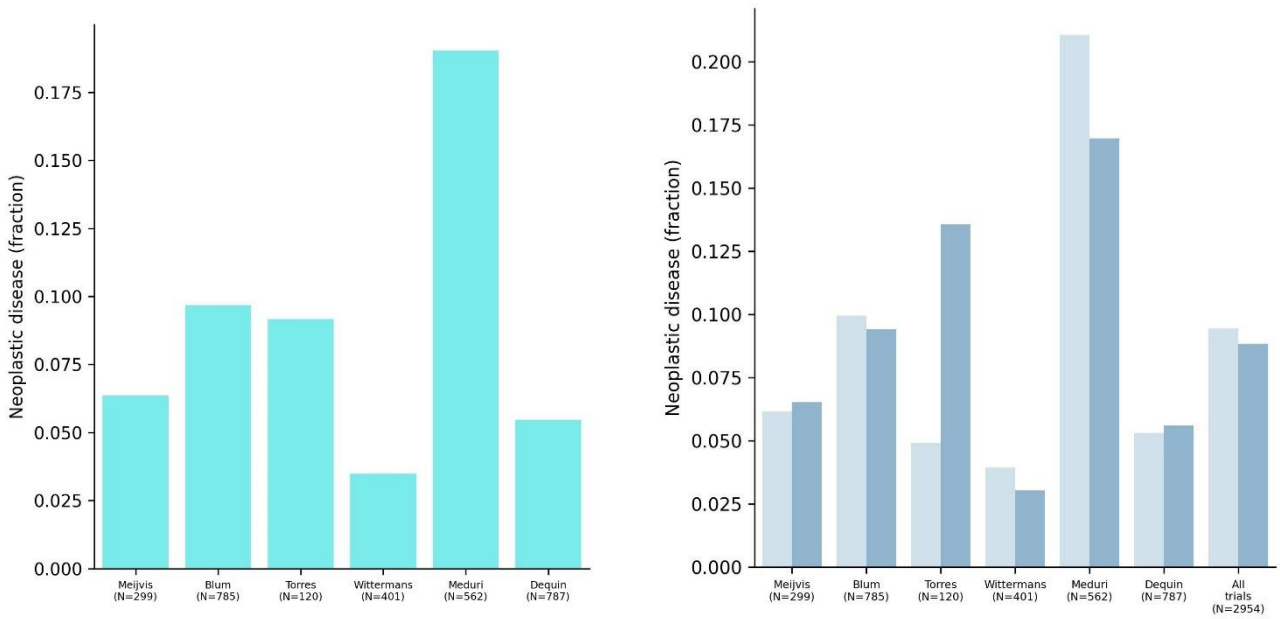
(m) Glucose



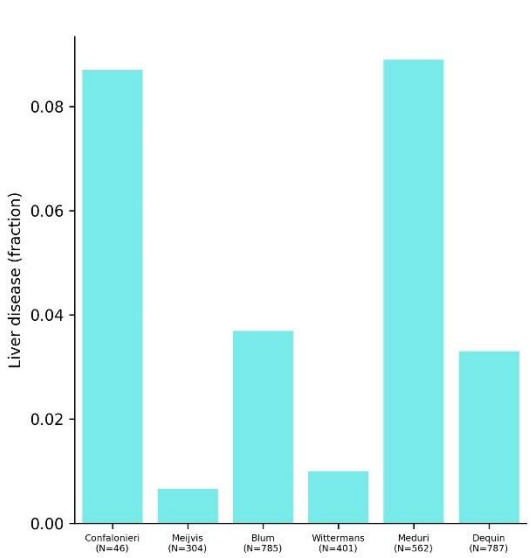
(n) White cell count

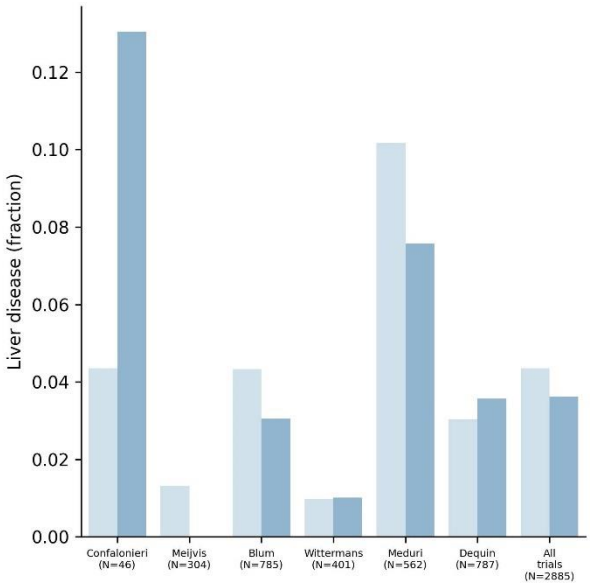


(o) Neoplastic disease

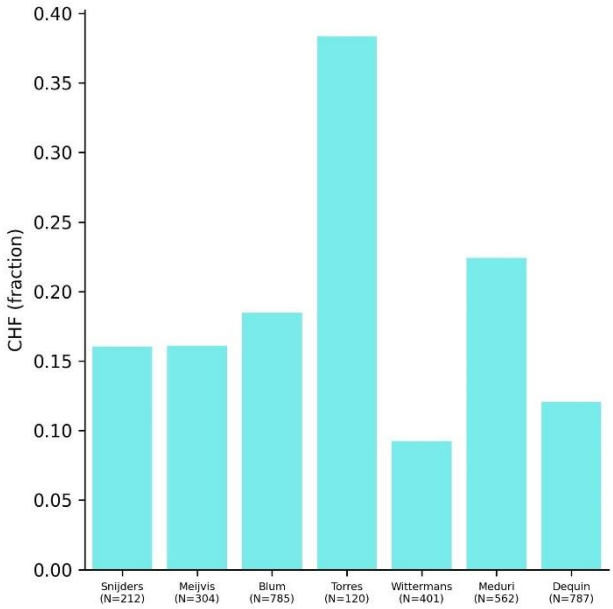


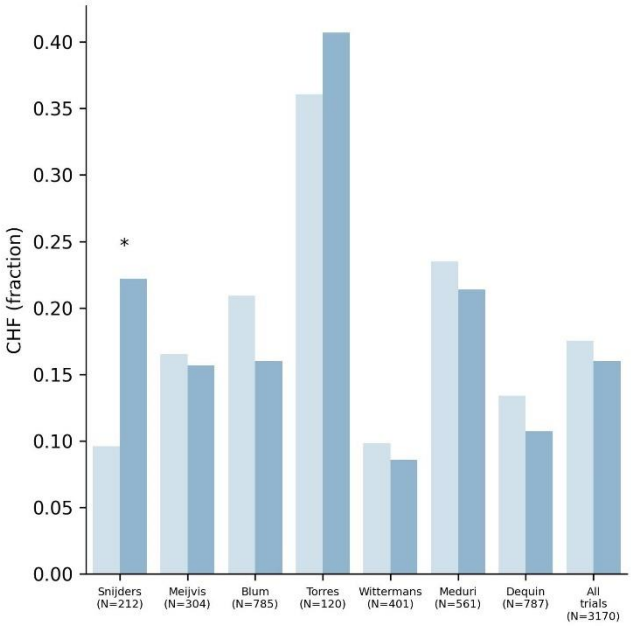
(p) Liver disease



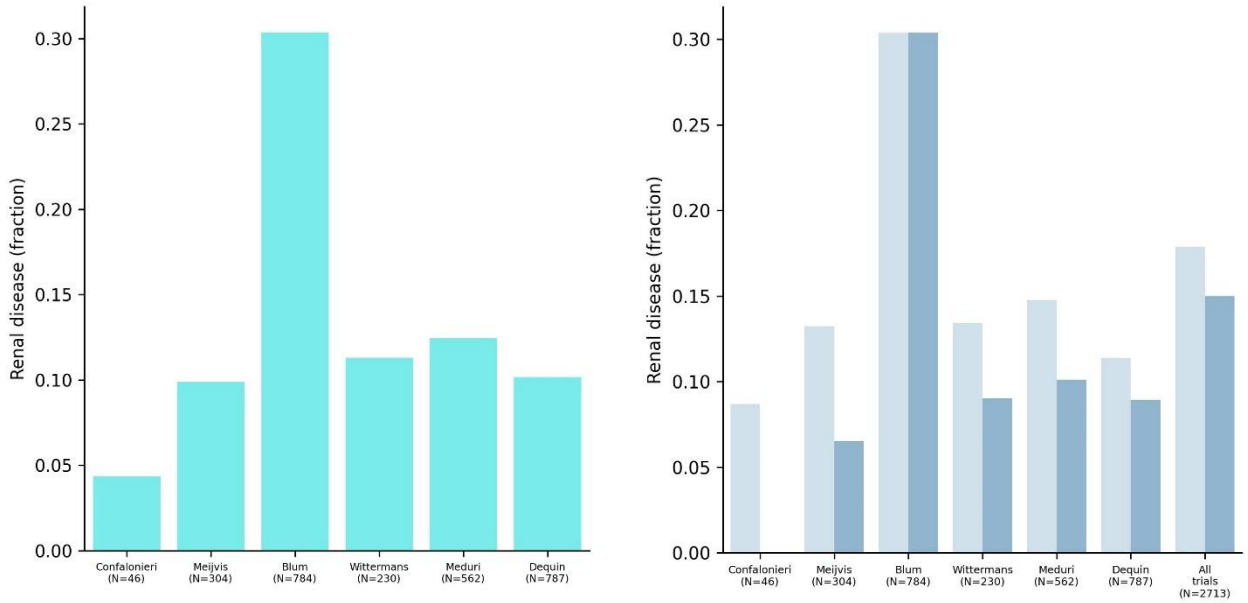


(q) Congestive heart failure

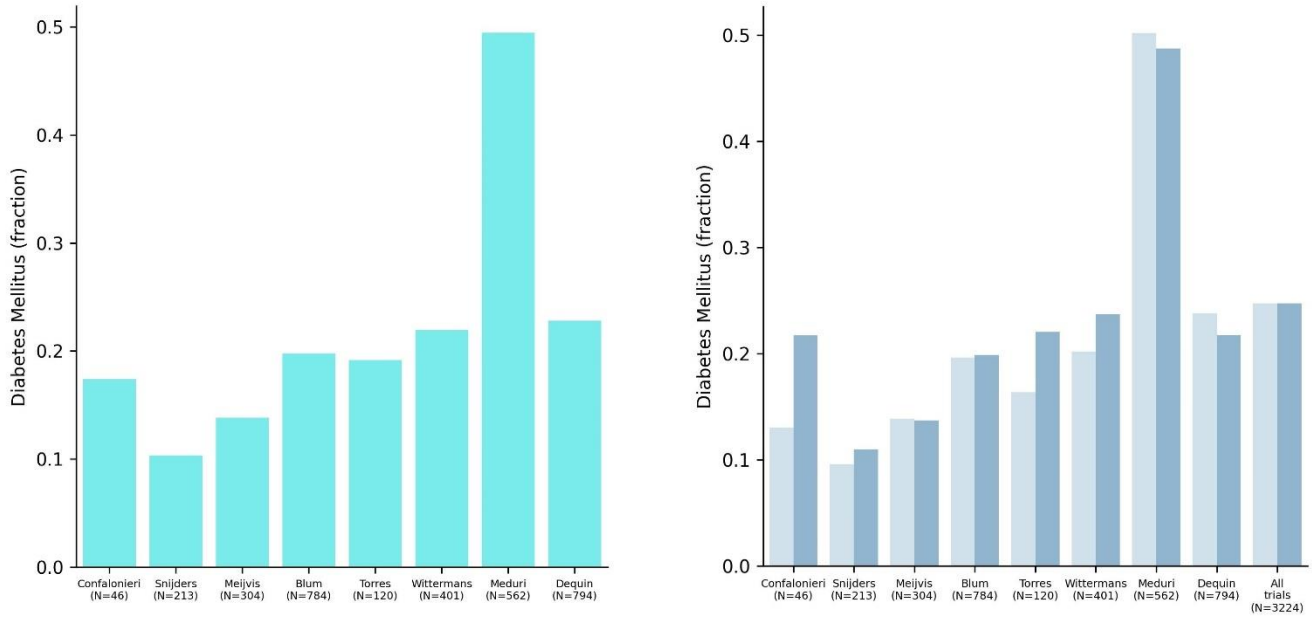




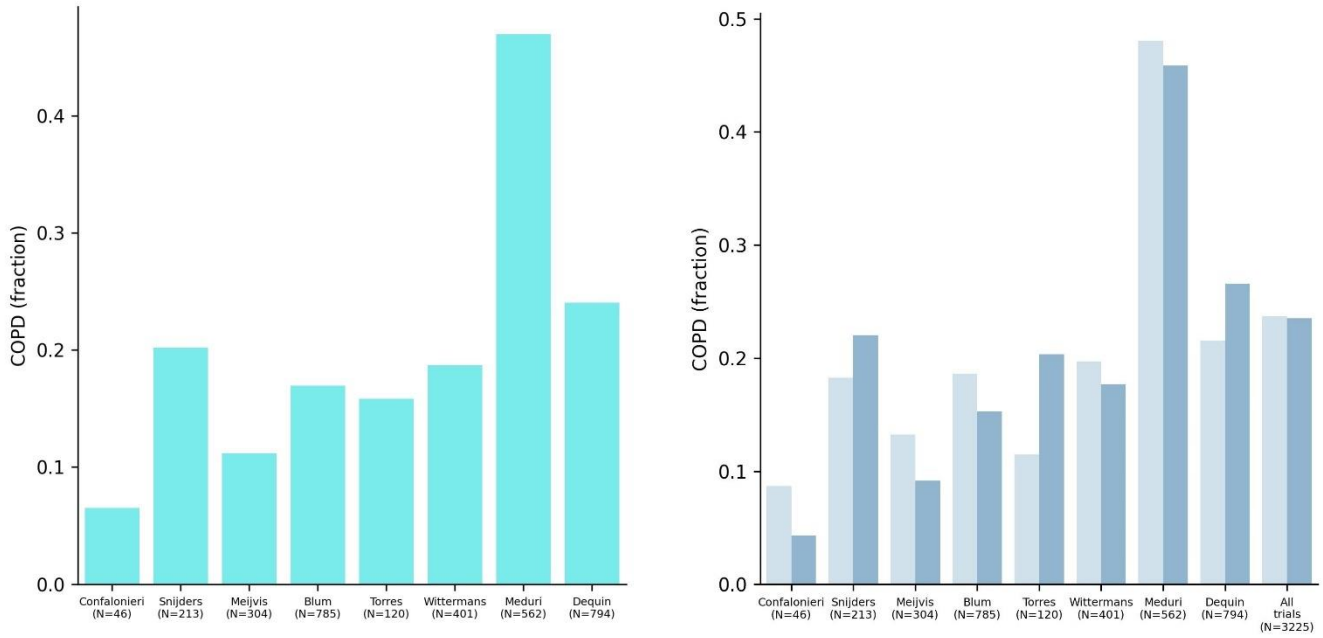
(r) Renal disease



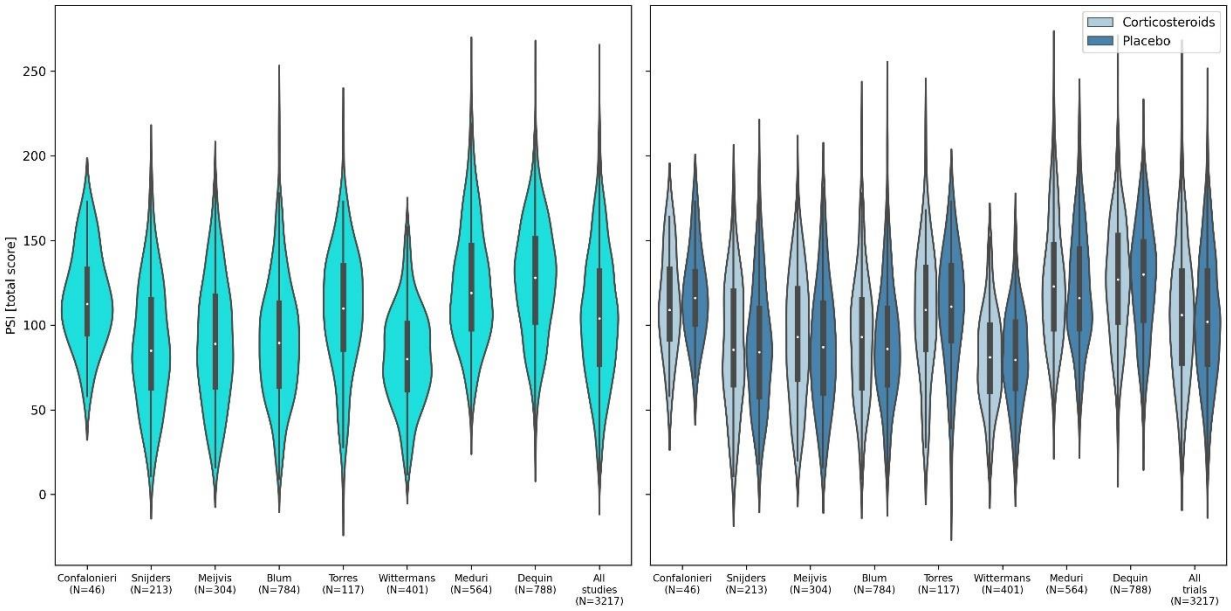
(s) Diabetes Mellitus



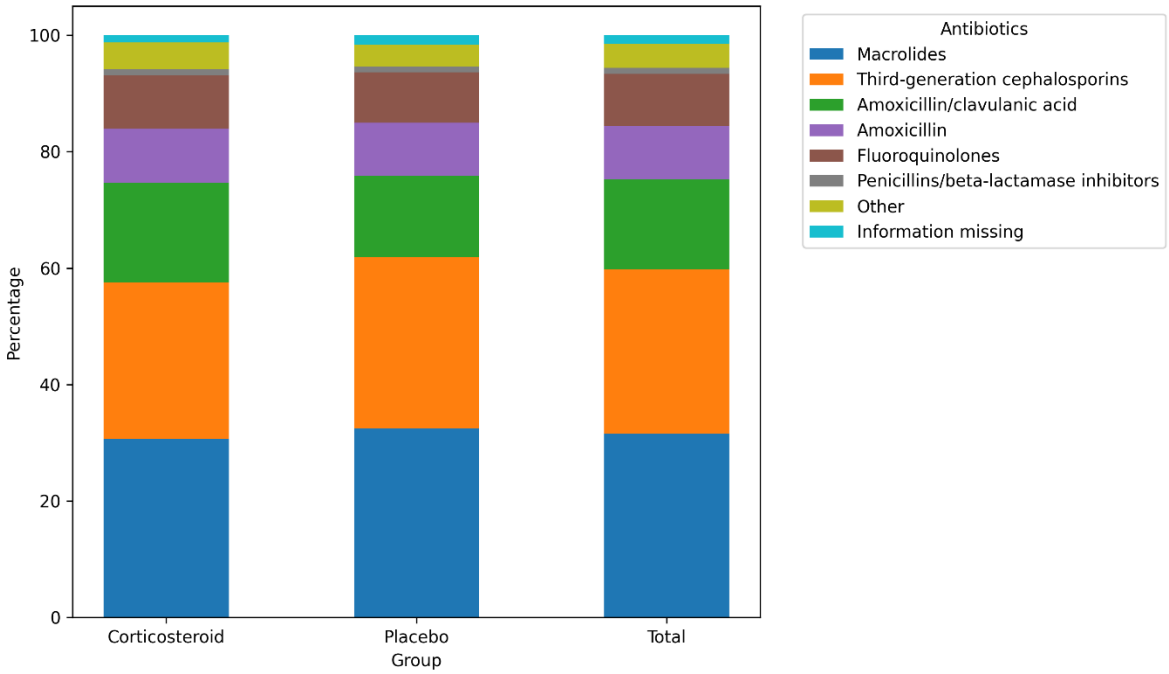
(t) Chronic obstructive pulmonary disease



(u) PSI

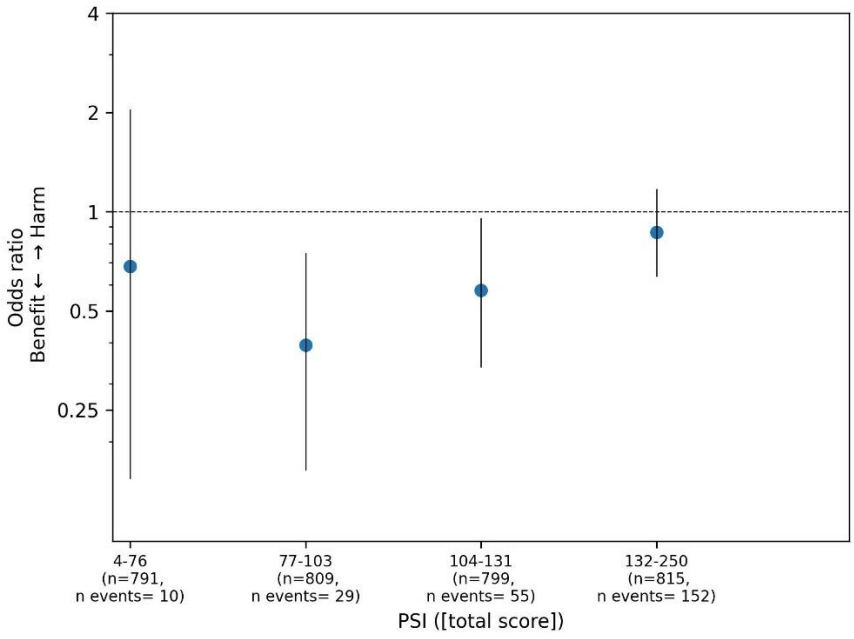


Appendix Figure S4: Stacked bar charts presenting initial antimicrobial treatment incidence. Data are in n (%). Figure is based on the patients from 315 the four trials (2,11–13) from whom we obtained data regarding antimicrobial treatment.

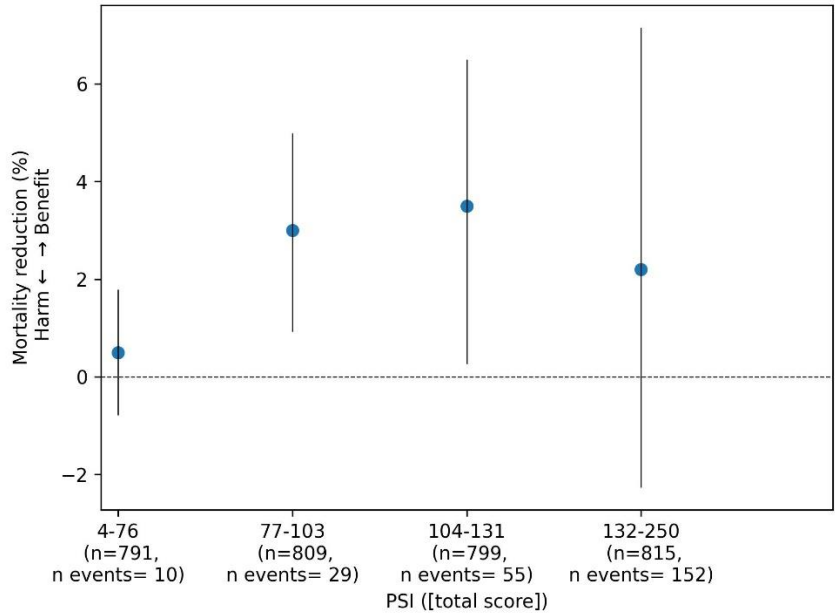


Appendix Figure S5: Heterogeneity of treatment effect on the (a) relative and (b) absolute scale for different PSI score quartiles. Analysis based on all patients (ie, train and test cohort combined), excluding patients with missing values for Pneumonia Severity Index (PSI).

(a) Relative scale (odds ratio)

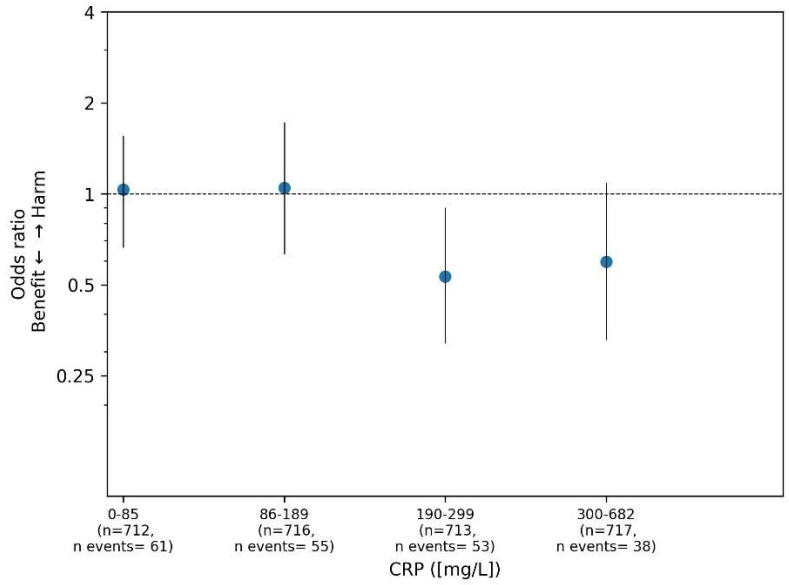


(b) Absolute scale (mortality risk reduction)

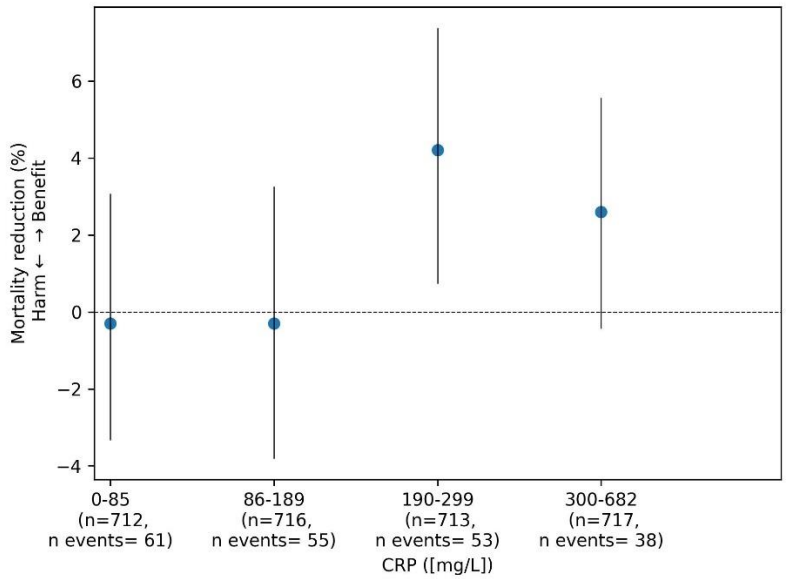


Appendix Figure S6: Heterogeneity of treatment effect on the (a) relative and (b) absolute scale for different C-reactive protein quartiles. Analysis based on all patients (ie, train and test cohort combined), excluding patients with missing values for C-reactive protein.

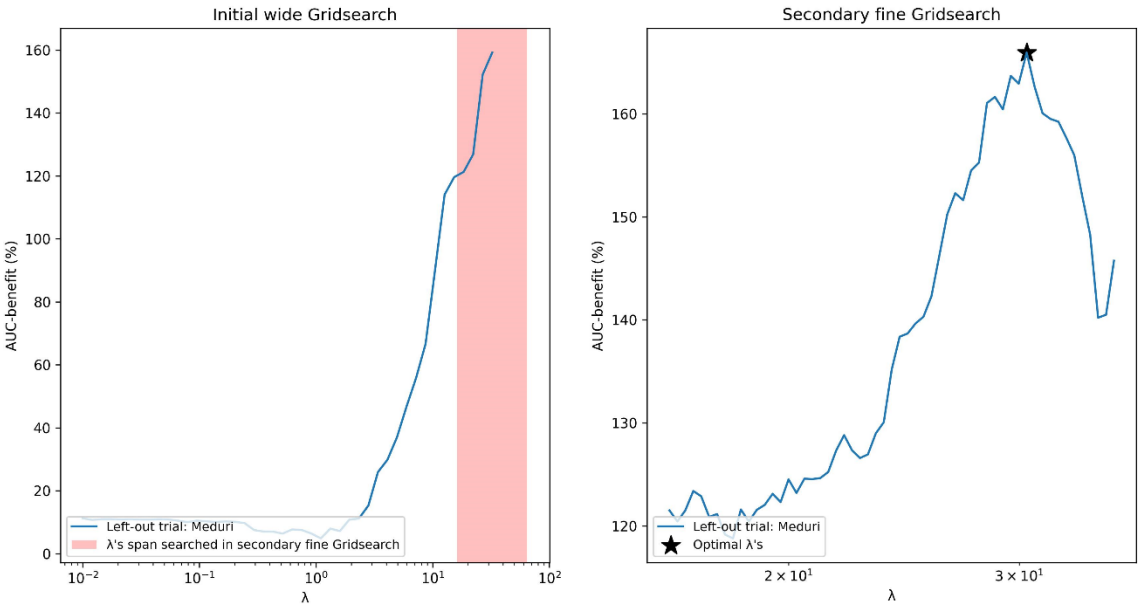
(a) Relative scale (odds ratio). The bars represent the 95% confidence intervals.



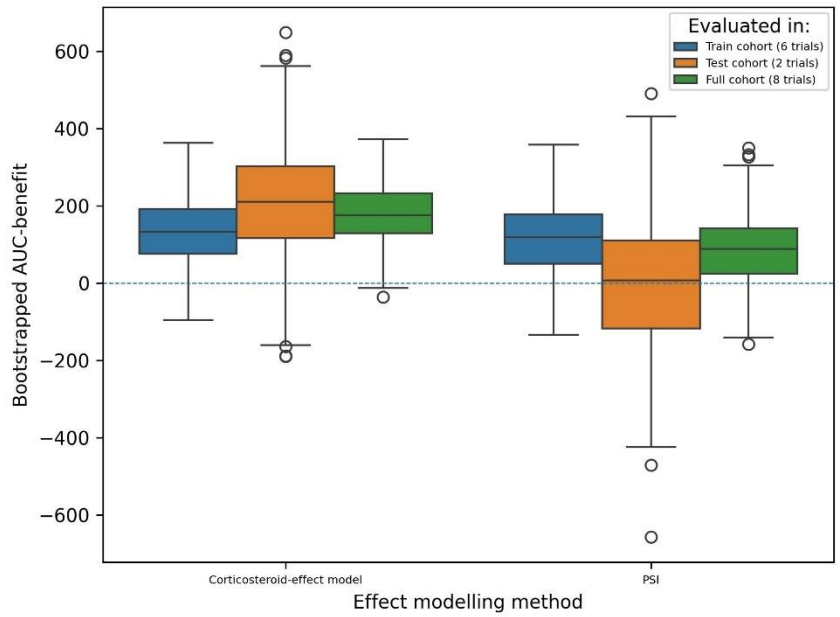
(b) Absolute scale (mortality risk reduction). The bars represent the 95% confidence intervals.



Appendix Figure S7: Results of the initial wide and second fine grid search for the Lasso penalty strength (λ) optimization.

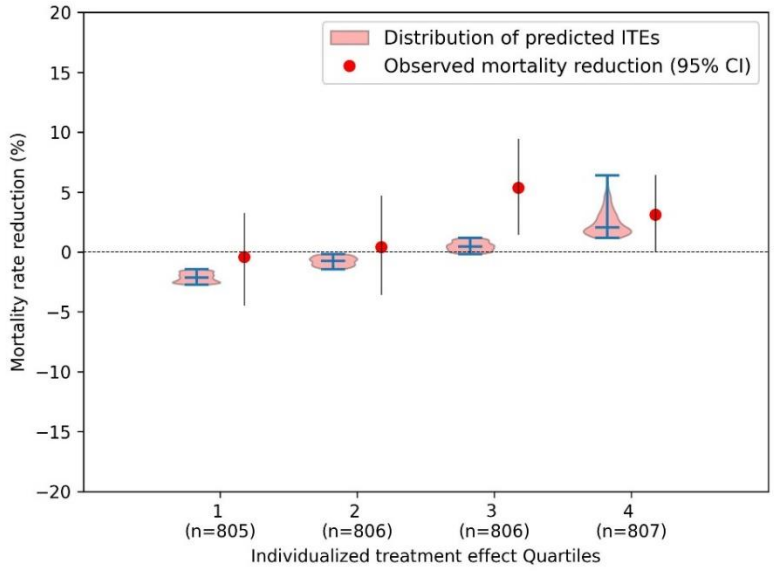


Appendix Figure S8: Discriminative performance of the corticosteroid-effect model (ie, the Tian method) and the PSI in the train cohort (ie, 'apparent validation' and in the test cohort (ie, external validation). The AUC-benefits resulting from 500 bootstrap samples are plotted using boxplots.

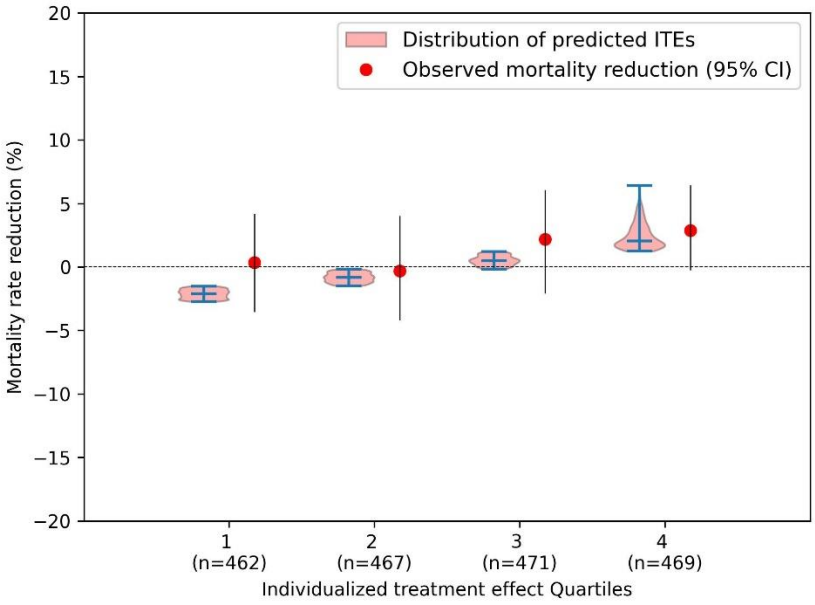


Appendix Figure S9: Calibration for benefit results for the corticosteroid-effect model in the full cohort (ie, all eight included trials combined) and in the train cohort (ie, six trials, ie, 'apparent validation'). For four patient groups based on ascending ITE quartiles, the ITE distributions are using violin plots, next to the observed mortality reductions in each quartile.

(a) Calibration for benefit results in full cohort (ie, train and test cohorts combined)

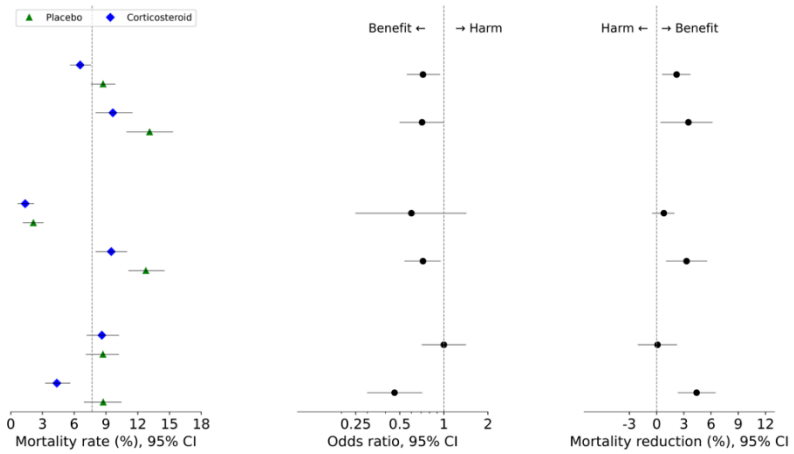


(b) Calibration for benefit results in train cohort (ie, ‘apparent validation’)



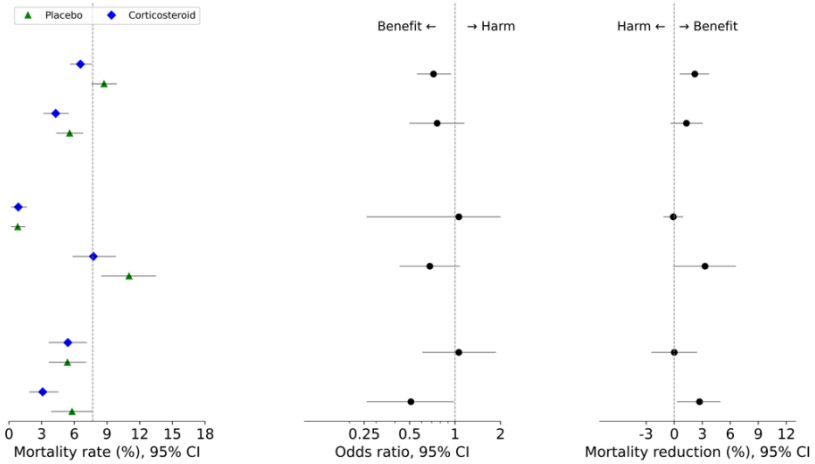
Appendix Figure S10: Results of the validation of the **corticosteroid-effect model and the PSI regarding 30-day mortality** in the **full cohort** (ie, eight trials, train and test cohort combined). Heterogeneity of Treatment Effect (HTE) on the relative, odds ratio scale and the absolute, mortality risk scale. For the relative scale, we added the P value for interaction and for the absolute scale, we added the size of the difference between treatment effects of the subgroups indicated with arrows. OR=odds ratio, NNT=number of patients needed to treat.

	Placebo, Mortality rate, n (%)	Corticosteroid, Mortality rate, n (%)	OR (95% CI)	Mortality reduction, % (95% CI)	NNT	P value for interaction
Full cohort (8 trials, n=3,224)	140/1,606 (8.7)	106/1,618 (6.6)	0.72 (0.56 to 0.94)	2.2% (0.6 to 3.7)	46	
Test cohort (2 trials, n=1,355)	88/671 (13.1)	66/684 (9.6)	0.71 (0.5 to 0.99)	3.5% (0.5 to 6.1)	28	
Full cohort subgroups by PSI						0.77
Less severe CAP (PSI Class I-III, n=1,197)	13/611 (2.1)	8/586 (1.4)	0.6 (0.25 to 1.42)	0.8% (-0.5 to 1.9)	131	
Severe CAP (PSI Class IV-V, n=2,027)	127/995 (12.8)	98/1,032 (9.5)	0.72 (0.54 to 0.95)	3.3% (1.1 to 5.5)	30	
Full cohort sub-groups by cortico-steroid-effect model						0.0054
Predicted no benefit (CRP ≤ 204, n=1,709)	76/873 (8.7)	72/836 (8.6)	1.0 (0.71 to 1.41)	0.1% (-2.0 to 2.2)	1,073	
Predicted benefit (CRP > 204, n=1,515)	64/733 (8.7)	34/782 (4.3)	0.46 (0.3 to 0.71)	4.4% (2.4 to 6.5)	22	



Appendix Figure S11: Results of the validation of the **corticosteroid-effect model and the PSI regarding 30-day mortality** in the **train cohort (six trials)**. Heterogeneity of Treatment Effect (HTE) on the relative, odds ratio scale and the absolute, mortality risk scale. For the relative scale, we added the P value for interaction and for the absolute scale, we added the size of the difference between treatment effects of the subgroups indicated with arrows. OR=odds ratio, NNT=number of patients needed to treat.

	Placebo, Mortality rate, n (%)	Corticosteroid, Mortality rate, n (%)	OR (95% CI)	Mortality reduction, % (95% CI)	NNT	P value for interaction
Full cohort (8 trials, n=3,224)	140/1,606 (8.7)	106/1,618 (6.6)	0.72 (0.56 to 0.94)	2.2% (0.6 to 3.7)	46	
Train cohort (6 trials, n=1,869)	52/935 (5.6)	40/934 (4.3)	0.76 (0.5 to 1.15)	1.3% (-0.4 to 3.0)	78	
Train cohort subgroups by PSI						0.55
Less severe CAP (PSI Class I-III, n=968)	4/499 (0.8)	4/469 (0.9)	1.06 (0.26 to 4.28)	-0.1% (-1.1 to 0.9)	-1,950	
Severe CAP (PSI Class IV-V, n=901)	48/436 (11.0)	36/465 (7.7)	0.68 (0.43 to 1.07)	3.3% (-0.1 to 6.6)	30	
Train cohort sub-groups by cortico-steroid-effect model						0.088
Predicted no benefit (CRP ≤ 204, n=984)	27/503 (5.4)	26/481 (5.4)	1.06 (0.61 to 1.86)	0.0% (-2.4 to 2.4)	-2,658	
Predicted benefit (CRP > 204, n=885)	25/432 (5.8)	14/453 (3.1)	0.51 (0.26 to 0.98)	2.7% (0.3 to 4.9)	37	



Appendix Table S1: The Pneumonia Severity Index (PSI), as published in 1997 in the New England Journal of Medicine.(5) A total point score for a given patient is obtained by summing the patient's age in years (age minus 10 for women) and the points for each applicable characteristic.

Characteristic	**Points Assigned**
Demographic factor	
Age	
Men	Age (years)
Women	Age (years) -10
Nursing home resident	+10
Coexisting illnesses	
Neoplastic disease	+30
Liver disease	+20
Congestive heart failure	+10
Cerebrovascular disease	+10
Renal disease	+10
Physical-examination findings	
Altered mental status	+20
Respiratory rate ≥ 30 breaths/min	+20
Systolic blood pressure <90 mmHg	+20
Temperature < 35°C or ≥ 40°C	+15
Heart rate ≥ 125 bpm	+10
Laboratory and radiographic findings	
Arterial pH <7.35	+30
Blood urea nitrogen ≥ 30 mg/dL	+20
Sodium <130 mmol//L	+20
Glucose ≥ 250 mg/dL	+10
Hematocrit <30%	+10
PaO_2 <60 mmHg	+10
Pleural effusion	+10

Appendix Table S2: The CURB-65 score, as published in 2003 in Thorax.(6) A total point score for a given patient is obtained by summing points.

*defined as a Mental Test Score of 8 or less, or new disorientation in person, place or time

Criterion	**Points Assigned**
Confusion*	1
Urea > 7 mmol/	1
Respiratory rate ≥ 30/min	1
Blood pressure (SBP <90 mm Hg or DBP ≤60 mm Hg)	1
Age ≥ 65 years	1

Appendix Table S3: The R implementations (using the lme4 package(7)) for the linear mixed-effects logistic regression models (LMMs) used to estimate the marginal and conditional ORs, as well as to perform the interaction test. The term "subgroup_mean" denotes the mean of the subgroup variable in each trial, and the term "subgroup_centered" denotes the subgroup variable centered about the trial-specific mean of the subgroup variable in each trial.

Model to ..	**R Implementation**
… calculate marginal odds ratio	formula <- "mortality ~ T + (1 | trial)" lmm <- glmer(formula, data = data, family = binomial)
… calculate conditional odds ratio	formula <- "mortality ~ PSI + age + T + (1 | trial)" lmm <- glmer(formula, data = data, family = binomial)
… test the interaction between patient subgroups and treatment	formula <- "mortality ~ T + subgroup + T:subgroup + (1 | trial)" lmm <- glmer(formula, data = data, family = binomial)
… test the interaction between steroid type/dose and treatment, adjusting for the subgroups identified by the corticosteroid-effect model (see Appendix part 10, Tables S48-49)	formula <- "mortality ~ T + steroid_type/dose + subgroup + T:steroid_type/dose + T:subgroup + (1 | trial)" lmm <- glmer(formula, data = data, family = binomial)
… test the interaction between subgroups and treatment, disentangling within-study and across-study information (8,9) (see also Appendix Part 10, Table S35)	formula <- "mortality ~ T + subgroup + T:subgroup_mean + T:subgroup_centered + (1 | trial)" lmm <- glmer(formula, data = data, family = binomial)
… test the interaction between PSI class subgroups and treatment, "PSI_class" is a categorical variable representing the PSI class, encoded ordinally (ie, Class I-II = 1, Class III = 2, Class IV = 3, Class V = 4; see Appendix part 10, Tables S38-43, Figure S28)	formula <- "mortality ~ T + PSI_class + T: PSI_class + (1 | trial)" lmm <- glmer(formula, data = data, family = binomial)

Appendix Table S4: Baseline characteristics of the 1,869 patients **in the train cohort (ie, six trials)**. Data are n (%) or median (IQR). *PSI values are missing for 0·2%, CURB-65 scores for 16.7%, and information regarding initial need for IMV for 55.9% of the patients, therefore the total numbers in the severity groups do not add up to total number of patients in treatment arms.

	Corticosteroid group (N=934)	**Placebo group (N=935)**	**Missings (% corticosteroid group, % placebo group)**
Demographics			
Female sex	386 (41·3)	359 (38·4)	(0·0, 0·0)
Age, (years)	70·0 (56·2-80·0)	69·0 (55·0-80·0)	(0·0, 0·0)
Clinical parameters			
Resp· rate, (breaths/min)	22·0 (18·0-27·0)	22·0 (18·0-28·0)	(12·1, 10·1)
Dias· blood pressure, (mmHg)	70·0 (61·0-80·0)	70·0 (61·0-80·0)	(7·3, 7·0)
Syst· blood pressure, (mmHg)	128·0 (112·0-142·0)	126·0 (112·0-141·0)	(7·3, 7·0)
Temperature, (°C)	37·9 (37·2-38·7)	38·0 (37·2-38·7)	(2·7, 2·7)
Heart rate, (bpm)	92·0 (80·0-106·0)	92·0 (79·0-106·0)	(2·6, 2·4)
SpO_2, (%)	94·0 (92·0-96·0)	95·0 (92·0-97·0)	(14·7, 14·0)
Laboratory values			
Creatinine, (µmol/L)	90·0 (70·7-120·2)	89·0 (71·8-117·0)	(1·1, 0·7)
Sodium, (mmol/L)	136·0 (133·0-139·0)	136·0 (133·0-138·0)	(1·0, 0·5)
Urea, (mmol/L)	6·8 (4·8-10·6)	6·7 (4·7-9·8)	(7·7, 7·2)
CRP, (mg/L)	196·0 (98·0-300·0)	188·1 (87·4-292·9)	(1·3, 1·1)
Glucose, (mmol/L)	7·1 (6·0-8·5)	6·9 (6·0-8·4)	(11·2, 12·2)
WBC count, (10^9 cells/L)	12·8 (9·4-17·1)	12·7 (9·1-16·9)	(0·9, 0·6)
Comorbidities			
Neoplastic disease	114 (12·2)	113 (12·1)	(14·1, 14·1)
Liver disease	47 (5·0)	42 (4·5)	(17·7, 18·0)
Congestive heart failure	231 (24·7)	206 (22·0)	(2·5, 2·6)
Renal disease	192 (20·6	174 (18·6)	(26·7, 27·4)
Diabetes mellitus	310 (33·2)	306 (32·7)	(0·0, 0·1)
COPD	290 (31·0)	281 (30·1	(0·0, 0·0)
Baseline disease severity indicators*			
PSI			
Total score	90·0 (64·0-115·0)	87·0 (65·0-111·0)	(0·1, 0·2)
Class I	124 (13·3)	114 (12·2)	-
Class II	167 (17·9)	159 (17·0)	-
Class III	177 (19·0)	224 (24·0)	-
Class IV	335 (35·9)	319 (34·1)	-
Class V	130 (13·9)	117 (12·5)	-
CURB-65			
Total score	1·0 (0·0-2·0)	1·0 (0·0-2·0)	(17·9, 15·6)
Score 0-2	690 (73·9)	727 (77·8)	-
Score 3-5	77 (8·2)	62 (6·6)	-
Other			
Initial ICU admission	94 (10·1)	91 (9·7)	(0·0, 0·0)
Initial need for IMV	2 (0·2)	3 (0·3)	(55·6, 56·1)

Appendix Table S5: Baseline characteristics of the 1,869 patients **in the test cohort**. Data are n (%) or median (IQR). *PSI values are missing for 0·6%, CURB-65 scores for 44.9%, and information regarding initial ICU admission and initial need for IMV for 42.3% of the patients, therefore the total numbers in the severity groups do not add up to total number of patients in treatment arms.

	Corticosteroid group (N=697)	**Placebo group (N=682)**	**Missings (% corticosteroid group, % placebo group)**
Demographics			
Female sex	127 (18·2)	137 (20·1)	(0·0, 0·1)
Age, (years)	67·6 (60·0-77·0)	67·0 (60·0-77·4)	(0·0, 0·1)
Clinical parameters			
Resp· rate, (breaths/min)	24·5 (20·0-29·0)	24·0 (20·0-28·0)	(1·3, 1·8)
Dias· blood pressure, (mmHg)	68·0 (59·0-76·5)	66·5 (57·0-75·0)	(0·4, 0·7)
Syst· blood pressure, (mmHg)	123·0 (109·0-138·0)	120·0 (105·5-135·0)	(0·4, 0·7)
Temperature, (°C)	37·1 (36·7-37·7)	37·0 (36·6-37·7)	(3·3, 3·1)
Heart rate, (bpm)	94·0 (81·5-109·0)	91·0 (79·5-106·0)	(0·6, 0·7)
SpO_2, (%)	94·0 (92·0-97·0)	95·0 (92·0-97·0)	(15·8, 15·7)
Laboratory values			
Creatinine, (µmol/L)	106·1 (79·6-168·0)	92·8 (70·7-150·3)	(57·7, 58·2)
Sodium, (mmol/L)	136·6 (133·0-139·3)	136·0 (133·0-139·0)	(0·9, 1·0)
Urea, (mmol/L)	9·2 (6·2-14·0)	9·5 (6·0-15·0)	(43·8, 43·3)
CRP, (mg/L)	187·5 (78·0-311·0)	173·0 (63·1-299·0)	(27·4, 23·8)
Glucose, (mmol/L)	7·6 (6·2-9·8)	7·4 (6·1-9·4)	(5·6, 6·2)
WBC count, (10^9 cells/L)	12·3 (8·7-17·0)	12·1 (8·6-17·6)	(3·6, 3·1)
Comorbidities			
Neoplastic disease	76 (10·9)	74 (10·9)	(2·0, 1·8)
Liver disease	37 (5·3)	39 (5·7)	(2·0, 1·8)
Congestive heart failure	125 (17·9)	96 (14·1)	(2·0, 1·8)
Renal disease	80 (11·5)	70 (10·3)	(2·0, 1·8)
Diabetes mellitus	243 (34·9)	216 (31·7)	(1·4, 1·3)
COPD	215 (30·8)	240 (35·2)	(1·4, 1·3)
Baseline disease severity indicators*			
PSI			
Total score	125·0 (100·0-151·0)	125·0 (101·0-148·0)	(0·6, 0·7)
Class I	7 (1·0)	5 (0·7)	-
Class II	28 (4·0)	31 (4·5)	-
Class III	86 (12·3)	76 (11·1)	-
Class IV	272 (39·0)	259 (38·0)	-
Class V	300 (43·0)	306 (44·9)	-
CURB-65			
Total score	1·0 (1·0-2·0)	1·0 (1·0-2·0)	(45·1, 44·7)
Score 0-2	355 (50·9)	341 (50·0)	-
Score 3-5	28 (4·0)	36 (5·3)	-
Other			
Initial ICU admission	400 (57·4)	395 (57·9)	(42·6, 42·1)
Initial need for IMV	92 (13·2)	85 (12·5)	(42·6, 42·1)

Appendix Table S6: Pathogen incidence. Data are in n (%). Percentages could add up to more than 100%, as for some patients, multiple pathogens were identified. Tabel is based on the patients from the seven trials (2,10–15) from whom we obtained data regarding aetiology.

	Corticosteroid group (N=1,330)	Placebo group (N=1,333)	All (N=2,663)
No pathogen identified	707 (53)	726 (54)	1,433 (53)
Bacterial	494 (37)	466 (34)	960 (36)
Streptococcus pneumoniae	246 (18)	262 (19)	508 (19)
Legionella pneumophila	64 (4)	46 (3)	110 (4)
Staphylococcus aureus	36 (2)	23 (1)	59 (2)
Mycoplasma pneumoniae	26 (1)	26 (1)	52 (1)
Other bacteria	128 (9)	119 (8)	247 (9)
Viral	149 (11)	136 (10)	285 (11)
Influenza A/B	90 (6)	68 (5)	158 (7)
Other virus	55 (4)	54 (4)	109 (4)
Information missing	27 (2)	40 (3)	67 (2)

Appendix Table S7: Initial antimicrobial treatment incidence. Data are in n (%). Tabel is based on the patients from the four trials (2,10,11,15) from whom we obtained data regarding antimicrobial treatment.

	Corticosteroid group (N=956)	Placebo group (N=956)	All (N=1,912)
Macrolides	293 (30)	310 (32)	603 (31.5)
Third-generation cephalosporins	257 (26)	282 (29)	539 (28.2)
Amoxicillin/clavulanic acid	163 (17)	133 (13)	296 (15.5)
Amoxicillin	89 (9)	87 (9)	176 (9.2)
Fluoroquinolones	88 (9)	83 (8)	171 (8.9)
Penicillins/beta-lactamase inhibitors	10 (1)	9 (0)	19 (1)
Other	44 (4)	36 (3)	80 (4.2)
Information missing	12 (1)	16 (1)	28 (1.5)

Appendix Table S8: Overall treatment effect and heterogeneity in treatment effect of adjuvant therapy with corticosteroids among the subgroups identified by the PSI and corticosteroid-effect model for patients included in **the trial by Meduri et al**.(16). OR=odds ratio, NNT=number of patients needed to treat.

	30-day mortality rate, n (%)		**OR (95% CI)**	**Mortality reduction, % (95% CI)**	**NNT**	**P value for interaction**
	Placebo	**Corticosteroid**				
Overall						
(n=561)	39/276 (14.1)	39/285 (13.7)	0.96 (0.65 to 1.43)	0.4% (-4.3 to 5.0)	224	
Subgroups by PSI						P = 0.34
Class I-III (n=99)	5/46 (10.9)	3/53 (5.7)	0.49 (0.0 to 1.91)	5.2% (-4.1 to 13.3)	19	
Class IV-V (n=462)	34/230 (14.8)	36/232 (15.5)	1.06 (0.7 to 1.64)	-0.7% (-6.2 to 4.4)	-136	
Subgroups by corticosteroid-effect model						P = 0.11
Predicted no benefit (n=396)	26/198 (13.1)	31/198 (15.7)	1.23 (0.76 to 2.0)	-2.5% (-8.4 to 3.6)	-39	
Predicted benefitt (n=165)	13/78 (16.7)	8/87 (9.2)	0.51 (0.22 to 1.12)	7.5% (-1.2 to 15.3)	13	

Appendix Table S9: Overall treatment effect and heterogeneity in treatment effect of adjuvant therapy with corticosteroids among the subgroups identified by the PSI and corticosteroid-effect model for patients included in **the trial by Dequin et al**.(15). OR=odds ratio, NNT=number of patients needed to treat.

	30-day mortality rate, n (%)		**OR (95% CI)**	**Mortality reduction, % (95% CI)**	**NNT**	**P value for interaction**
	Placebo	**Corticosteroid**				
Overall						
(n=794)	49/395 (12.4)	27/399 (6.8)	0.51 (0.32 to 0.81)	5.6% (1.9 to 9.3)	17	
Subgroups by PSI						P = 0.28
Class I-III (n=130)	4/66 (6.1)	1/64 (1.6)	0.25 (0.0 to 1.41)	4.5% (-1.1 to 10.2)	22	
Class IV-V (n=664)	45/329 (13.7)	26/335 (7.8)	0.53 (0.33 to 0.85)	5.9% (1.6 to 10.3)	16	
Subgroups by corticosteroid-effect model						P = 0.51
Predicted no benefit (n=329)	23/172 (13.4)	15/157 (9.6)	0.68 (0.38 to 1.24)	3.8% (-2.2 to 9.6)	26	
Predicted benefit (n=465)	26/223 (11.7)	12/242 (5.0)	0.40 (0.2 to 0.72)	6.7% (2.4 to 11.1)	14	

Appendix Table S10: Overall treatment effects in each of the included trials. OR=odds ratio, NNT=number of patients needed to treat.

	30-day mortality rate, n (%)		Marginal OR (95% CI)	Conditional OR (95% CI)	Mortality reduction, % (95% CI)	NNT
	Placebo	**Corticosteroid**				
Confalonieri et al. (n=46)	8/23 (34.8)	0/23 (0.0)	-	-	34.8% (20.0 to 52.4)	2
Snijders et al. (n=213)	6/109 (5.5)	6/104 (5.8)	1.05 (0.33 to 3.37)	1.16 (0.33; 4.06)	-0.3% (-5.8 to 4.7)	-377
Meijvis et al. (n=304)	9/153 (5.9)	9/151 (6.0)	1.01 (0.39 to 2.63)	0.74 (0.26; 2.1)	-0.1% (-4.3 to 4.3)	-1283
Blum et al. (n=785)	13/393 (3.3)	15/392 (3.8)	1.16 (0.55 to 2.48)	0.91 (0.4; 2.04)	-0.5% (-2.7 to 1.4)	-192
Torres et al. (n=120)	9/59 (15.3)	6/61 (9.8)	0.61 (0.20 to 1.82)	0.51 (0.14; 1.8)	5.4% (-4.7 to 15.3)	18
Wittermans et al. (n=401)	7/198 (3.5)	4/203 (2.0)	0.55 (0.16 to 1.90)	0.49 (0.14; 1.77)	1.6% (-0.9 to 4.6)	63
Meduri et al. (n=562)	40/277 (14.4)	39/285 (13.7)	0.94 (0.58 to 1.51)	0.88 (0.54; 1.45)	0.8% (-4.2 to 5.4)	132
Dequin et al. (n=794)	49/395 (12.4)	27/399 (6.8)	0.51 (0.31 to 0.84)	0.5 (0.3; 0.83)	5.6% (1.9 to 9.3)	17

Appendix Table S11: Marginal versus conditional odds ratios. Conditional odds ratios are conditional on the risk factors age and pneumonia severity index (PSI), and the implementation in R is given in Table S3. OR=odds ratio

	30-day mortality rate, n (%)		Marginal OR (95% CI)	Conditional OR (95% CI)
	Placebo	**Corticosteroid**		
Overall				
(n=1,355)	88/671 (13.1)	66/684 (9.6)	0.71 (0.50 to 0.99)	0.67 (0.48 to 0.96)
Subgroups by PSI				
Class I-III (n=229)	9/112 (8.0)	4/117 (3.4)	0.40 (0.12 to 1.36)	0.40 (0.12; 1.34)
Class IV-V (n=1,126)	79/559 (14.1)	62/567 (10.9)	0.75 (0.52 to 1.06)	0.71 (0.49; 1.03)
Subgroups by corticosteroid-effect model				
Predicted no benefit (n=725)	49/370 (13.2)	46/355 (13.0)	0.98 (0.63 to 1.50)	0.89 (0.57; 1.40)
Predicted benefit (n=630)	39/301 (13.0)	20/329 (6.1)	0.43 (0.25 to 0.76)	0.44 (0.25; 0.80)

Appendix Table S12: Overall effect of adjuvant therapy with corticosteroids on binary secondary outcomes. Analysis is based on the patients from whom we obtained data regarding the corresponding trials. *The minus sign denotes risk increase (ie, harm), rather than reduction (ie, benefit). ** Three(10,11,13) of the four trials which included for the readmission outcome, reported readmissions within 30 days after study enrolment (ie, '30-day readmission'), whereas one trial(12) reported readmissions within 30 days after hospital discharge.

	Outcome rate, n (%)		**OR (95% CI)**	**Risk reduction, % (95% CI)***	**NNT***	**P value**
	Placebo	Corticosteroid				
90-day mortality, (n=1,745, from four trials (2,11,14,15)*)*	94/870 (10.8)	70/875 (8.0)	0.71 (0.51 to 0.99)	2.8% (0.4 to 5.2)	35	0.042
28-day IMV, (n=1,568, who did not require IMV at baseline, from four trials, (2,11,14,15)*)*	120/785 (15.3)	82/783 (10.5)	0.59 (0.42 to 0.82)	4.8% (2.1 to 7.5)	20	0.0019
28-day vasopressors, (n=1,625, who did not require vasopressors at baseline, from three trials (11,14,15)*)*	154/811 (19.0)	98/814 (12.0)	0.54 (0.40 to 0.72)	6.9% (4.0 to 9.7)	14	<0.0001
Hospital readmission, (n=1,633, from four trials (10–13)*)***	30/814 (3.7)	57/819 (7.0)	1.95 (1.24 to 3.07)	-3.3% (-5.3 to -1.5)	-30	0.0038
30-day hospital readmission, (n=1,334, from three trials (10,11,13)*)*	23/661 (3.5)	50/673 (7.4)	2.22 (1.34 to 3.68)	-3.9% (-6.0 to -2.1)	-25	0.0020

Appendix Table S13: Overall effect of adjuvant therapy with corticosteroids on length-of-stay secondary outcomes. Analysis is based on the patients from whom we obtained data regarding the corresponding trials. *P value calculated through Kruskal-Wallis test for difference, using the 'kruskal' function from the Scipy library in Python.(17).

	Median length of stay, IQR (days)		**Reduction in median length of stay in days (95% CI)**	**P value***
	Placebo	Corticosteroid		
Hospital stay (n=1,831, from six trials (2,10–14)*)*	7.0 (4.5 ; 11.0)	6.0 (4.0 ; 9.0)	1.0 (0.5 to 1.0)	P= 0.0002
*Hospital stay, **excluding** patients who deceased within 30 days (n=1,756, from six trials* (2,10–14)*)*	7.0 (4.5 ; 11.0)	6.0 (4.0 ; 9.0)	1.0 (0.5 to 1.0)	P= 0.0002
ICU stay (n=930, from four trials (2,11,14,15)*)*	7.0 (4.0 ; 12.0)	5.0 (3.0 ; 9.0)	2.0 (0.0 to 2.0)	P= 0.0009
*ICU stay, **excluding** patients who deceased within 30 days (n=838, from four trials* (2,11,14,15)*)*	6.0 (4.0 ; 11.0)	5.0 (3.0 ; 9.0)	1.0 (0.0 to 2.0)	P = 0.0020

Appendix Table S14: Overall effect of adjuvant therapy with corticosteroids on adverse events compatible with corticosteroid use. *The minus sign denotes risk increase (ie, harm), rather than reduction (ie, benefit).

	Adverse event rate, n (%)		**OR (95% CI)**	**Risk reduction, % (95% CI)***	**NNT***	**P value**
	Placebo	Corticosteroid				
Hyperglycaemia, (n=683, from four trials (2,10,12,14)*)*	44/344 (12.8)	84/339 (24.8)	2.50 (1.63 to 3.83)	-12.0% (-17.0 to -6.9)	-8	< 0.0001
Hospital-acquired infection, (n=2,650, from seven trials (2,10–15)*)*	172/1320 (13.0)	159/1330 (12.0)	0.88 (0.63 to 1.22)	0.9% (-1.6 to 3.9)	92	0.44
Gastro-intestinal bleeding, (n=1,958, from five trials (2,10,11,14,15)*)*	17/979 (1.7)	16/979 (1.6)	0.93 (0.47 to 1.85)	0.1% (-0.8 to 1.0)	979	0.85

Appendix Table S15: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **90-day mortality**. Analysis is based on the patients from the four trials [2,11,14,15] from whom we obtained data regarding 90-day mortality. *The minus sign denotes risk increase (ie, harm), rather than reduction (ie, benefit).

	90-day mortality rate, n (%)		**OR (95% CI)**	**90-day mortality rate reduction, % (95% CI)***	**NNT**	**P for interaction**
	Placebo	**Corticosteroid**				
Subgroups by PSI						P = 0.60
Class I-III (n=572)	8/294 (2.7)	7/278 (2.5)	0.92 (0.33 to 2.56)	0.2% (-2.2 to 2.2)	492	
Class IV-V (n=1,173)	86/576 (14.9)	63/597 (10.6)	0.68 (0.48 to 0.96)	4.4% (1.2 to 7.6)	22	
Subgroups by corticosteroid-effect model						P = 0.07
Predicted no benefit (n=875)	48/443 (10.8)	43/432 (10.0)	0.96 (0.62 to 1.49)	0.9% (-2.6 to 4.3)	113	
Predicted benefit (n=870)	46/427 (10.8)	27/443 (6.1)	0.52 (0.32 to 0.86)	4.7% (1.6 to 7.8)	21	

Appendix Table S16: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **initiation of invasive mechanical ventilation by day 28 (28-day IMV)**. Analysis is based on the patients from the four trials [2,11,14,15] from whom we obtained data regarding 28-day IMV, who did not require IMV at baseline. *The minus sign denotes risk increase (ie, harm), rather than reduction (ie, benefit).

	28-day IMV rate, n (%)		**OR (95% CI)**	**28-day IMV rate reduction, % (95% CI)***	**NNT***	**P for interaction**
	Placebo	**Corticosteroid**				
Subgroups by PSI						P = 0.18
Class I-III (n=553)	12/284 (4.2)	14/269 (5.2)	0.98 (0.41 to 2.34)	-1.0% (-4.0 to 1.7)	-102	
Class IV-V (n=1,015)	108/501 (21.6)	68/514 (13.2)	0.54 (0.38 to 0.78)	8.3% (4.6 to 12.1)	12	
Subgroups by corticosteroid-effect model						P = 0.16
Predicted no benefit (n=777)	44/392 (11.2)	33/385 (8.6)	0.80 (0.46 to 1.37)	2.7% (-1.0 to 6.1)	37	
Predicted benefit (n=791)	76/393 (19.3)	49/398 (12.3)	0.50 (0.33 to 0.76)	7.0% (2.6 to 11.0)	14	

Appendix Table S17: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **initiation of vasopressors by day 28 (28-day vasopressors)**. Analysis is based on the patients from the three trials (11,14,15) from whom we obtained data regarding 28-day vasopressors, who did not require vasopressors at baseline.

	28-day vasopressor rate, n (%)		OR (95% CI)	28-day vasopressor rate reduction, % (95% CI)	NNT	P for interaction
	Placebo	**Corticosteroid**				
Subgroups by PSI						P = 0.34
Class I-III (n=538)	12/278 (4.3)	4/260 (1.5)	0.28 (0.08 to 0.93)	2.8% (0.5 to 5.2)	35	
Class IV-V (n=1,087)	142/533 (26.6)	94/554 (17.0)	0.55 (0.40 to 0.75)	9.7% (5.7 to 13.6)	10	
Subgroups by corticosteroid-effect model						P = 0.53
Predicted no benefit (n=846)	76/426 (17.8)	41/420 (9.8)	0.48 (0.31 to 0.75)	8.1% (4.3 to 12.0)	12	
Predicted benefit (n=779)	78/385 (20.3)	57/394 (14.5)	0.60 (0.41 to 0.87)	5.8% (1.3 to 10.3)	17	

Appendix Table S18: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **hospital readmission**. Analysis is based on the patients from the four trials (10–13) from whom we obtained data regarding hospital readmission. *The minus sign denotes risk increase (ie, harm), rather than reduction (ie, benefit).

	Readmission rate, n (%)		OR (95% CI)	Readmission rate reduction, % (95% CI)*	NNT*	P for interaction
	Placebo	**Corticosteroid**				
Subgroups by PSI						P = 0.07
Class I-III (n=903)	11/468 (2.4)	30/435 (6.9)	3.07 (1.52 to 6.22)	-4.5% (-7.1 to -2.4)	-21	
Class IV-V (n=730)	19/346 (5.5)	27/384 (7.0)	1.30 (0.71 to 2.39)	-1.5% (-4.5 to 1.4)	-64	
Subgroups by corticosteroid-effect model						P = 0.52
Predicted no benefit (n=894)	18/454 (4.0)	29/440 (6.6)	1.71 (0.93 to 3.12)	-2.6% (-5.2 to -0.2)	-38	
Predicted benefit (n=739)	12/360 (3.3)	28/379 (7.4)	2.31 (1.16 to 4.62)	-4.1% (-7.1 to -1.7)	-24	

Appendix Table S19: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **30-day hospital readmission**. Analysis is based on the patients from the three trials (10,11,13) from whom we obtained data regarding 30-day hospital readmission. *The minus sign denotes risk increase (ie, harm), rather than reduction (ie, benefit).

	Readmission rate, n (%)		**OR (95% CI)**	**Readmission rate reduction, % (95% CI)***	**NNT***	**P for interaction**
	Placebo	**Corticosteroid**				
Subgroups by PSI						P = 0.28
Class I-III (n=743)	10/379 (2.6)	27/364 (7.4)	2.96 (1.41 to 6.20)	-4.8% (-7.2 to -2.1)	-20	
Class IV-V (n=591)	13/282 (4.6)	23/309 (7.4)	1.67 (0.83 to 3.36)	-2.8% (-6.2 to 0.3)	-35	
Subgroups by corticosteroid-effect model						P = 0.96
Predicted no benefit (n=750)	14/379 (3.7)	29/371 (7.8)	2.21 (1.15 to 4.25)	-4.1% (-6.9 to -1.3)	-24	
Predicted benefit (n=584)	9/282 (3.2)	21/302 (7.0)	2.26 (1.02 to 5.02)	-3.8% (-6.7 to -0.6)	-26	

Appendix Table S20: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **median length of hospital stay**. Analysis is based on the patients from six trials (2,10–14) from whom we obtained data regarding length of hospital stay.

	Median length of hospital stay, IQR (days)		**Reduction in median length of hospital stay in days (95% CI)**
	Placebo	**Corticosteroid**	
Subgroups by PSI			
Class I-III (n=958)	6.0 (4.0 ; 8.5)	5.0 (3.5 ; 7.0)	1.0 (0.0 to 1.0)
Class IV-V (n=873)	9.0 (6.0 ; 14.0)	7.5 (5.0 ; 12.0)	1.5 (1.0 to 3.0)
Subgroups by corticosteroid-effect model			
Predicted no benefit (n=969)	7.0 (4.0 ; 10.0)	6.0 (4.0 ; 9.0)	1.0 (0.5 to 1.0)
Predicted benefit (n=862)	7.5 (5.0 ; 12.0)	6.5 (4.5 ; 10.0)	1.0 (0.0 to 2.0)

Appendix Table S21: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **median length of ICU stay**. Analysis is based on the patients from four trials (2,11,14,15) from whom we obtained data regarding length of ICU stay, who were admitted to the ICU during their hospitalization. *The minus sign denotes length of stay increase (ie, harm), rather than reduction (ie, benefit).

	Median length of ICU stay, IQR (days)		**Reduction in median length of ICU stay in days (95% CI)***
	Placebo	**Corticosteroid**	
Subgroups by PSI			
Class I-III (n=166)	4.5 (3.0 ; 9.0)	5.0 (3.0 ; 7.8)	-0.5 (-1.0 to 1.0)
Class IV-V (n=764)	7.0 (4.0 ; 13.0)	6.0 (3.0 ; 10.0)	1.0 (0.0 to 2.5)
Subgroups by corticosteroid-effect model			
Predicted no benefit (n=374)	6.0 (3.75 ; 10.0)	6.0 (4.0 ; 9.0)	0.0 (-1.0 to 1.0)
Predicted benefit (n=556)	7.0 (4.0 ; 14.0)	5.0 (3.0 ; 9.0)	2.0 (1.0 to 3.0)

Appendix Table S22: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **hyperglycaemia**. Analysis is based on the patients from the four trials (2,10,12,14) from whom we obtained data regarding hyperglycaemia. *The minus sign denotes risk increase (ie, harm), rather than reduction (ie, benefit).

	Hyperglyceamia rate, n (%)		**OR (95% CI)**	**Hyperglyceamia rate reduction, % (95% CI)***	**NNT***	**P for interaction**
	Placebo	**Corticosteroid**				
Subgroups by PSI						P = 0.15
Class I-III (n=323)	16/170 (9.4)	35/153 (22.9)	3.94 (1.95 to 7.94)	-13.5% (-19.9 to -6.6)	-7	
Class IV-V (n=360)	28/174 (16.1)	49/186 (26.3)	1.83 (1.06 to 3.15)	-10.3% (-17.2 to -3.4)	-9	
Subgroups by corticosteroid-effect model						P = 0.70
Predicted no benefit (n=291)	20/159 (12.6)	37/132 (28.0)	2.79 (1.47 to 5.31)	-15.5% (-23.0 to -7.7)	-6	
Predicted benefit (n=392)	24/185 (13.0)	47/207 (22.7)	2.30 (1.30 to 4.06)	-9.7% (-16.0 to -3.2)	-10	

Appendix Table S23: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **hospital-acquired infections**. Analysis is based on the patients from the seven trials (2,10–15) from whom we obtained data regarding hospital-acquired infections. *The minus sign denotes risk increase (ie, harm), rather than reduction (ie, benefit).

	Hospital-acquired infection rate, n (%)		**OR (95% CI)**	**Hospital-acquired infection rate reduction, % (95% CI)***	**NNT**	**P for interaction**
	Placebo	**Corticosteroid**				
Subgroups by PSI						P = 0.33
Class I-III (n=1,088)	77/558 (13.8)	62/530 (11.7)	0.54 (0.24 to 1.21)	2.1% (-0.7 to 5.1)	47	
Class IV-V (n=1,562)	95/762 (12.5)	97/800 (12.1)	0.96 (0.66 to 1.39)	0.3% (-2.4 to 3.1)	292	
Subgroups by corticosteroid-effect model						P = 0.09
Predicted no benefit (n=1,307)	82/671 (12.2)	72/636 (11.3)	1.20 (0.74 to 1.94)	0.9% (-1.6 to 3.9)	111	
Predicted benefit (n=1,343)	90/649 (13.9)	87/694 (12.5)	0.67 (0.43 to 1.06)	1.3% (-1.7 to 4.4)	75	

Appendix Table S24: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **gastrointestinal bleedings**. Analysis is based on the patients from the five trials (2,10,11,14,15) from whom we obtained data regarding hyperglycaemia. *The minus sign denotes risk increase (ie, harm), rather than reduction (ie, benefit).

	Gastrointestinal bleeding rate, n (%)		**OR (95% CI)**	**Gastrointestinal bleeding rate reduction, % (95% CI)***	**NNT***	**P for interaction**
	Placebo	**Corticosteroid**				
Subgroups by PSI						-
Class I-III (n=691)	0/357 (0.0)	0/334 (0.0)	-	0.0%	-	
Class IV-V (n=1,267)	17/622 (2.7)	16/645 (2.5)	0.91 (0.45 to 1.81)	0.3% (-1.2 to 1.7)	396	
Subgroups by corticosteroid-effect model						P = 0.99
Predicted harm group (n=970)	8/498 (1.6)	7/472 (1.5)	0.96 (0.34 to 2.67)	0.1% (-1.2 to 1.4)	810	
Predicted benefit group (n=988)	9/481 (1.9)	9/507 (1.8)	0.95 (0.37 to 2.41)	0.1% (-1.3 to 1.5)	1042	

Appendix Table S25: Overview of maximum time between presentation at the hospital, and the measurement of the baseline C-reactive Protein (CRP), for the eight included trials.

Reference, year	**Maximum time between hospital presentation and measurement of baseline CRP**
Confalonieri, 2005	6 hours (for all patients)
Snijders, 2010	8 hours (for all patients)
Meijvis, 2011	24 hours (for all patients)
Blum, 2015	24 hours (for all patients)
Torres, 2015	24 hours (for 116/120 patients, 97%)
Wittermans, 2021	24 hours (for all patients)
Meduri, 2022	*unknown*
Dequin, 2023	24 hours (for 452/794 patients, 57%), 36 hours (for 660/794 patients, 83%), 48 hours (for 705/794 patients, 89%),

Appendix Table S26: Overview of the used definitions for the adverse outcome hyperglycaemia, for the four trials (2,10,12,14) from whom we obtained data regarding hyperglycaemia.

Reference, year	**Used definition for hyperglycaemia**
Confalonieri, 2005	No protocolized definition used.
Snijders, 2010	No protocolized definition used.
Meijvis, 2011	Non-fasting blood glucose > 11 mmol/L
Torres, 2015	No protocolized definition used.

Appendix Part 2: Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data Checklist

PRISMA-IPD Section/topic	Item No	Checklist item	Reported on page
Title			
Title	1	Identify the report as a systematic review and meta-analysis of individual participant data.	1
Abstract			
Structured summary	2	Provide a structured summary including as applicable:	5-6
		Background: state research question and main objectives, with information on participants, interventions, comparators and outcomes.	
		Methods: report eligibility criteria; data sources including dates of last bibliographic search or elicitation, noting that IPD were sought; methods of assessing risk of bias.	
		Results: provide number and type of studies and participants identified and number (%) obtained; summary effect estimates for main outcomes (benefits and harms) with confidence intervals and measures of statistical heterogeneity. Describe the direction and size of summary effects in terms meaningful to those who would put findings into practice.	
		Discussion: state main strengths and limitations of the evidence, general interpretation of the results and any important implications.	
		Other: report primary funding source, registration number and registry name for the systematic review and IPD meta-analysis.	
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known.	7
Objectives	4	Provide an explicit statement of the questions being addressed with reference, as applicable, to participants, interventions, comparisons, outcomes and study design (PICOS). Include any hypotheses that relate to particular types of participant-level subgroups.	7
Methods			
Protocol and registration	5	Indicate if a protocol exists and where it can be accessed. If available, provide registration information including registration number and registry name. Provide publication details, if applicable.	8
Eligibility criteria	6	Specify inclusion and exclusion criteria including those relating to participants, interventions, comparisons, outcomes, study design and characteristics (e.g. years when conducted, required minimum follow-up). Note whether these were applied at the study or individual level i.e. whether eligible participants were included (and ineligible participants excluded) from a study that included a wider population than specified by the review inclusion criteria. The rationale for criteria should be stated.	8
Identifying studies - information sources	7	Describe all methods of identifying published and unpublished studies including, as applicable: which bibliographic databases were searched with dates of coverage; details of any hand searching including of conference proceedings; use of study registers and agency or company databases; contact with the original research team and experts in the field; open adverts and surveys. Give the date of last search or elicitation.	8, appendix part 3

Identifying studies - search	8	Present the full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	appendix part 3
Study selection processes	9	State the process for determining which studies were eligible for inclusion.	appendix part 3
Data collection processes	10	Describe how IPD were requested, collected and managed, including any processes for querying and confirming data with investigators. If IPD were not sought from any eligible study, the reason for this should be stated (for each such study). If applicable, describe how any studies for which IPD were not available were dealt with. This should include whether, how and what aggregate data were sought or extracted from study reports and publications (such as extracting data independently in duplicate) and any processes for obtaining and confirming these data with investigators.	8, appendix part 3
Data items	11	Describe how the information and variables to be collected were chosen. List and define all study level and participant level data that were sought, including baseline and follow-up information. If applicable, describe methods of standardising or translating variables within the IPD datasets to ensure common scales or measurements across studies.	8
IPD integrity	A1	Describe what aspects of IPD were subject to data checking (such as sequence generation, data consistency and completeness, baseline imbalance) and how this was done.	8
Risk of bias assessment in individual studies.	12	Describe methods used to assess risk of bias in the individual studies and whether this was applied separately for each outcome. If applicable, describe how findings of IPD checking were used to inform the assessment. Report if and how risk of bias assessment was used in any data synthesis.	8, appendix part 3
Specification of outcomes and effect measures	13	State all treatment comparisons of interests. State all outcomes addressed and define them in detail. State whether they were pre-specified for the review and, if applicable, whether they were primary/main or secondary/additional outcomes. Give the principal measures of effect (such as risk ratio, hazard ratio, difference in means) used for each outcome.	8-9
Synthesis methods	14	Describe the meta-analysis methods used to synthesise IPD. Specify any statistical methods and models used. Issues should include (but are not restricted to): • Use of a one-stage or two-stage approach. • How effect estimates were generated separately within each study and combined across studies (where applicable). • Specification of one-stage models (where applicable) including how clustering of patients within studies was accounted for. • Use of fixed or random effects models and any other model assumptions, such as proportional hazards. • How (summary) survival curves were generated (where applicable). • Methods for quantifying statistical heterogeneity (such as I^2 and τ^2). • How studies providing IPD and not providing IPD were analysed together (where applicable). • How missing data within the IPD were dealt with (where applicable).	9-11
Exploration of variation in effects	A2	If applicable, describe any methods used to explore variation in effects by study or participant level characteristics (such as estimation of interactions between effect and covariates). State all participant-level characteristics that were analysed as potential effect modifiers, and whether these were pre-specified.	9-11

Risk of bias across studies	15	Specify any assessment of risk of bias relating to the accumulated body of evidence, including any pertaining to not obtaining IPD for particular studies, outcomes or other variables.	12, appendix part 10
Additional analyses	16	Describe methods of any additional analyses, including sensitivity analyses. State which of these were pre-specified.	12, appendix part 10
Results			
Study selection and IPD obtained	17	Give numbers of studies screened, assessed for eligibility, and included in the systematic review with reasons for exclusions at each stage. Indicate the number of studies and participants for which IPD were sought and for which IPD were obtained. For those studies where IPD were not available, give the numbers of studies and participants for which aggregate data were available. Report reasons for non-availability of IPD. Include a flow diagram.	13, appendix part 3
Study characteristics	18	For each study, present information on key study and participant characteristics (such as description of interventions, numbers of participants, demographic data, unavailability of outcomes, funding source, and if applicable duration of follow-up). Provide (main) citations for each study. Where applicable, also report similar study characteristics for any studies not providing IPD.	13
IPD integrity	A3	Report any important issues identified in checking IPD or state that there were none.	13
Risk of bias within studies	19	Present data on risk of bias assessments. If applicable, describe whether data checking led to the up-weighting or down-weighting of these assessments. Consider how any potential bias impacts on the robustness of meta-analysis conclusions.	-
Results of individual studies	20	For each comparison and for each main outcome (benefit or harm), for each individual study report the number of eligible participants for which data were obtained and show simple summary data for each intervention group (including, where applicable, the number of events), effect estimates and confidence intervals. These may be tabulated or included on a forest plot.	13-14
Results of syntheses	21	Present summary effects for each meta-analysis undertaken, including confidence intervals and measures of statistical heterogeneity. State whether the analysis was pre-specified, and report the numbers of studies and participants and, where applicable, the number of events on which it is based.	13-14
		When exploring variation in effects due to patient or study characteristics, present summary interaction estimates for each characteristic examined, including confidence intervals and measures of statistical heterogeneity. State whether the analysis was pre-specified. State whether any interaction is consistent across trials.	
		Provide a description of the direction and size of effect in terms meaningful to those who would put findings into practice.	
Risk of bias across studies	22	Present results of any assessment of risk of bias relating to the accumulated body of evidence, including any pertaining to the availability and representativeness of available studies, outcomes or other variables.	16, appendix part 10
Additional analyses	23	Give results of any additional analyses (e.g. sensitivity analyses). If applicable, this should also include any analyses that incorporate aggregate data for studies that do not have IPD. If applicable, summarise the main meta-analysis results following the inclusion or exclusion of studies for which IPD were not available.	16, appendix part 10
Discussion			

Summary of evidence	24	Summarise the main findings, including the strength of evidence for each main outcome.	17
Strengths and limitations	25	Discuss any important strengths and limitations of the evidence including the benefits of access to IPD and any limitations arising from IPD that were not available.	18, 20-21
Conclusions	26	Provide a general interpretation of the findings in the context of other evidence.	19
Implications	A4	Consider relevance to key groups (such as policy makers, service providers and service users). Consider implications for future research.	19
Funding			
Funding	27	Describe sources of funding and other support (such as supply of IPD), and the role in the systematic review of those providing such support.	-

413

Appendix Part 3: Systematic literature search

Methods

Randomized controlled trials (RCTs) eligible for this study compared placebo with low-dose oral or intravenous corticosteroid therapy as adjunctive therapy in community-acquired pneumonia (CAP) patients. We excluded studies with pseudo randomization or with treatment combinations that did not allow investigation of an independent corticosteroid effect. We updated the systematic search by Briel and colleagues, which identified eligible studies up to July 2017.(18) As such, we electronically searched Medline, Embase, and the Cochrane Controlled Trials Registry from July 2017 to July 2024 using medical subject headings based on the terms 'pneumonia' and 'corticosteroid'. Table S27 contains the detailed search strategies. Two reviewers (JS and PvdZ) independently assessed trial eligibility based on title and abstracts, full-texts, and further information from investigators if needed. From all eligible trials, individual patient data (IPD), including demographic, clinical, and laboratory data, were requested by the authors. The data were verified against the reported results and inconsistencies were resolved with the corresponding authors. The risk of bias (ROB) arising from the randomization process, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result in included trials was assessed independently by two reviewers (JS and PvdZ), using the updated version of the Cochrane ROB assessment tool.(19)

Results

From the literature search, we identified 10 eligible trials (Figure S12). We contacted the corresponding authors of all eligible trials. The authors of Nafae et al.(20) did not respond, the authors of Sabry et al.(21) responded that the dataset was lost. The authors of Meduri et al.(16) and Dequin et al.(15) were requested to wait with sharing the IPD until the corticosteroid-effect model was published as a pre-print.(22) Five studies were judged as having overall low ROB, while for the remaining studies, some concerns were raised (Table S28). Concerns were raised for bias arising from the randomization process, the selection of the reported result or both.

442 Appendix Table S27: Search query per database.

443

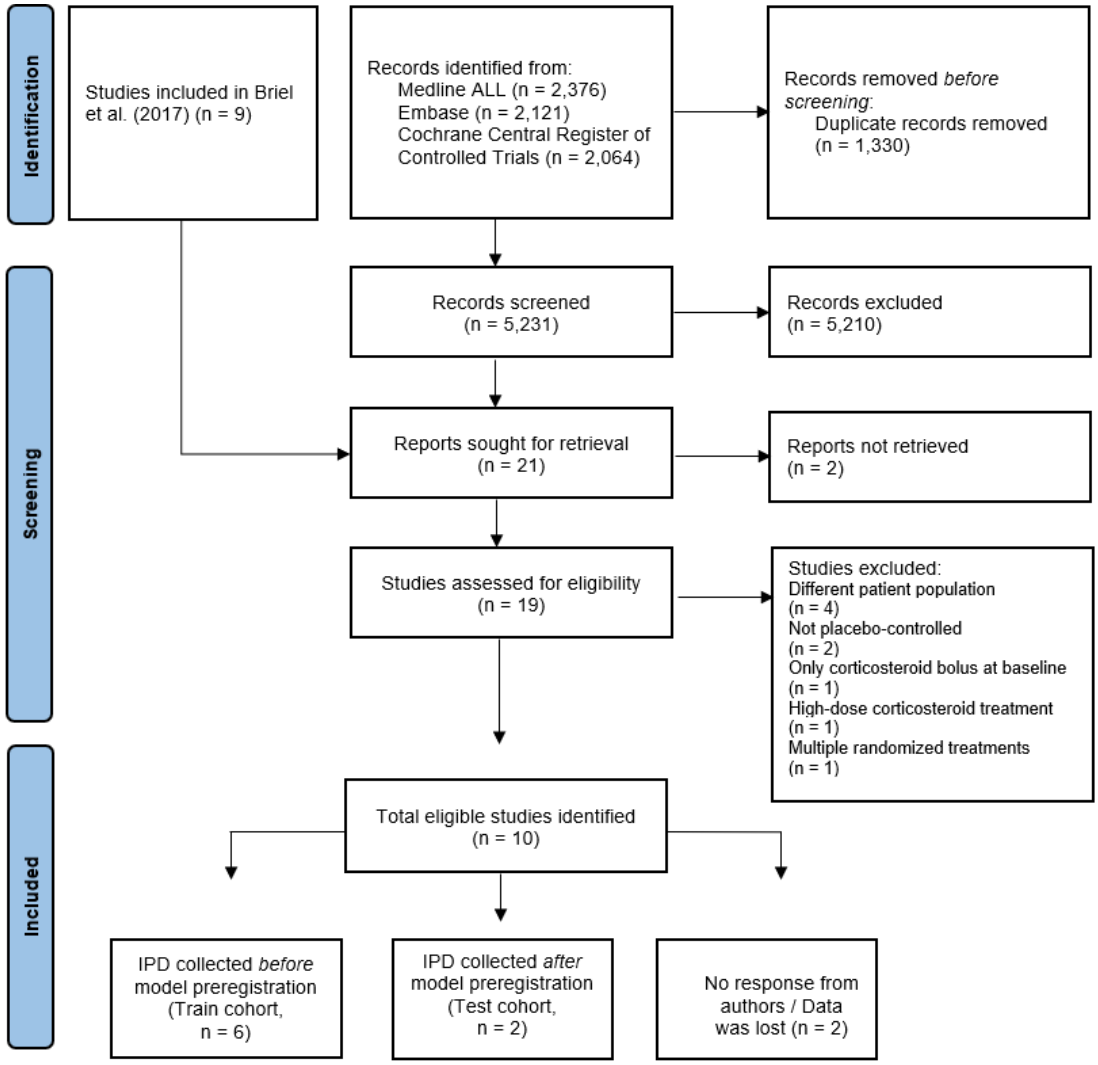
Database	Records after duplicates removed	Search Query
Medline ALL through Ovid	1870	(exp Pneumonia / OR (cap OR hap OR pneumon*).ab,ti.) AND (exp Steroids / OR exp Adrenal Cortex Hormones / OR (prednison* OR prednisolon* OR methylprednisolon* OR betamethason* OR dexamethason* OR triamcinolone OR hydrocortison* OR alclometason* OR algeston* OR amcinonid* OR amelometason* OR beclometason* OR budesonid* OR butixocort* OR chloroprednison* OR ciclesonid* OR ciprocinonid* OR clobetasol* OR clobetason* OR clocortolon* OR cloprednol* OR cortivazol* OR deflazacort* OR diflorason* OR diflucortolon* OR difluprednat* OR domoprednat* OR drocinonid* OR dutimelan* OR etiprednol-dicloacetat* OR fluclorolon* OR fludrocortison* OR fludroxycortid* OR flumetason* OR flumoxonid* OR flunisolid* OR fluocinolon* OR fluocinonid* OR fluocortin* OR fluocortolon* OR fluorometholon* OR flupredniden* OR fluprednisolon* OR fluticason* OR formocortal* OR halcinonid* OR halometason* OR halopredon* OR hydrocortison* OR icometasone-enbutat* OR isoflupredon* OR itrocinonid* OR locicortolone-dicibat* OR lorinden-a* OR lorinden-t* OR loteprednol* OR mazipredon* OR medryson* OR meprednison* OR mometasone-furoat* OR nicocortonid* OR nivacortol* OR oropivalon* OR paramethason* OR prednisolon* OR prednison* OR pregnenolon* OR procinonid* OR promestrien* OR resocortol* OR rimexolon* OR rofleponid* OR ticabeson* OR timobeson* OR tipredan* OR tixocortol* OR triamcinolon* OR ulobetasol-propionat* OR uniderm* OR vamorolon* OR zoticason* OR steroid* OR corticosteroid* OR Adrenal-Cortex-Hormone* OR glucocorticoid* OR hydroxycorticosteroid*).ab,ti.) AND (Exp Controlled clinical trial/ OR "Double-Blind Method"/ OR "Single-Blind Method"/ OR "Random Allocation"/ OR (random* OR factorial* OR crossover* OR cross over* OR placebo* OR ((doubl* OR singl*) ADJ blind*) OR assign* OR allocat* OR volunteer* OR trial OR groups).ab,ti,kf.) NOT (exp Animals/ NOT Humans/) NOT ((exp child/ OR exp infant/ OR pediatrics/ OR adolescent/) NOT exp adult/) AND 2017:2030.(sa_year).
Embase through Embase.com	1180	(pneumonia/exp OR (cap OR hap OR pneumon*):Ab,ti) AND ('steroid'/de OR 'corticosteroid'/exp OR (prednison* OR prednisolon* OR methylprednisolon* OR betamethason* OR dexamethason* OR triamcinolone OR hydrocortison* OR alclometason* OR algeston* OR amcinonid* OR amelometason* OR beclometason* OR budesonid* OR butixocort* OR chloroprednison* OR ciclesonid* OR ciprocinonid* OR clobetasol* OR clobetason* OR clocortolon* OR cloprednol* OR cortivazol* OR deflazacort* OR diflorason* OR diflucortolon* OR difluprednat* OR domoprednat* OR drocinonid* OR dutimelan* OR etiprednol-dicloacetat* OR fluclorolon* OR fludrocortison* OR fludroxycortid* OR flumetason* OR flumoxonid* OR flunisolid* OR fluocinolon* OR fluocinonid* OR fluocortin* OR fluocortolon* OR fluorometholon* OR flupredniden* OR fluprednisolon* OR fluticason* OR formocortal* OR halcinonid* OR halometason* OR halopredon* OR hydrocortison* OR icometasone-enbutat* OR isoflupredon* OR itrocinonid* OR locicortolone-dicibat* OR lorinden-a* OR lorinden-t* OR loteprednol* OR mazipredon* OR medryson* OR meprednison* OR mometasone-furoat* OR nicocortonid* OR nivacortol* OR oropivalon* OR paramethason* OR prednisolon* OR prednison* OR pregnenolon* OR procinonid* OR promestrien* OR resocortol* OR rimexolon* OR rofleponid* OR ticabeson* OR timobeson* OR tipredan* OR tixocortol* OR triamcinolon* OR ulobetasol-propionat* OR uniderm* OR vamorolon* OR zoticason* OR steroid* OR corticosteroid* OR Adrenal-Cortex-Hormone* OR glucocorticoid* OR hydroxycorticosteroid*):ab,ti) AND ('randomised controlled trial'/exp OR 'single blind procedure'/exp OR 'double blind procedure'/exp OR 'crossover procedure'/exp OR (random* OR placebo* OR factorial* OR crossover* OR 'cross-over' OR 'cross over' OR assign* OR allocat* OR volunteer* OR ((singl* OR doubl*) NEAR/2 (blind* OR mask*))):ab,ti) NOT [conference abstract]/lim NOT ([animals]/lim NOT [humans]/lim) NOT (juvenile/exp NOT adult/exp) AND [2017-07-01]/sd
Cochrane Central Register of Controlled Trials through Wiley	894	((cap OR hap OR pneumon*):Ab,ti) AND ((prednison* OR prednisolon* OR methylprednisolon* OR betamethason* OR dexamethason* OR triamcinolone OR hydrocortison* OR alclometason* OR algeston* OR amcinonid* OR amelometason* OR beclometason* OR budesonid* OR butixocort* OR chloroprednison* OR ciclesonid* OR

		ciprocinonid* OR clobetasol* OR clobetason* OR clocortolon* OR cloprednol* OR cortivazol* OR deflazacort* OR diflorason* OR diflucortolon* OR difluprednat* OR domoprednat* OR drocinonid* OR dutimelan* OR etiprednol-dicloacetat* OR fluclorolon* OR fludrocortison* OR fludroxycortid* OR flumetason* OR flumoxonid* OR flunisolid* OR fluocinolon* OR fluocinonid* OR fluocortin* OR fluocortolon* OR fluorometholon* OR flupredniden* OR fluprednisolon* OR fluticason* OR formocortal* OR halcinonid* OR halometason* OR halopredon* OR hydrocortison* OR icometasone-enbutat* OR isoflupredon* OR itrocinonid* OR locicortolone-dicibat* OR lorinden-a* OR lorinden-t* OR loteprednol* OR mazipredon* OR medryson* OR meprednison* OR mometasone-furoat* OR nicocortonid* OR nivacortol* OR oropivalon* OR paramethason* OR prednisolon* OR prednison* OR pregnenolon* OR procinonid* OR promestrien* OR resocortol* OR rimexolon* OR rofleponid* OR ticabeson* OR timobeson* OR tipredan* OR tixocortol* OR triamcinolon* OR ulobetasol-propionat* OR uniderm* OR vamorolon* OR zoticason* OR steroid* OR corticosteroid* OR Adrenal-Cortex-Hormone* OR glucocorticoid* OR hydroxycorticosteroid*):ab,ti)

Appendix Table S28: Results of risk of bias assessment for each eligible study. The updated version of the Cochrane ROB assessment tool(19) was used to assess bias arising from the randomization process (R), deviations from intended interventions (D), missing outcome data (Mi), measurement of the outcome (Me), and selection of the reported result (S). Overall ROB was judged to be low if ROB was judged to be low in all domains. Overall ROB (O) was judged as 'some concerns' if some concerns were raised in at least one domain, but no domain was judged as high ROB.

First author, year (reference)	**R**	**D**	**Mi**	**Me**	**S**	**O**
Confalonieri, 2005	SOME CONCERNS	LOW RISK	LOW RISK	LOW RISK	SOME CONCERNS	SOME CONCERNS
Snijders, 2010	LOW RISK	LOW RISK	LOW RISK	LOW RISK	SOME CONCERNS	SOME CONCERNS
Meijvis, 2011	LOW RISK	LOW RISK	LOW RISK	LOW RISK	LOW RISK	LOW RISK
Sabry, 2011	SOME CONCERNS	LOW RISK	LOW RISK	LOW RISK	SOME CONCERNS	SOME CONCERNS
Nafae, 2013	SOME CONCERNS	LOW RISK	LOW RISK	LOW RISK	SOME CONCERNS	SOME CONCERNS
Torres, 2015	LOW RISK	LOW RISK	LOW RISK	LOW RISK	LOW RISK	LOW RISK
Blum, 2015	LOW RISK	LOW RISK	LOW RISK	LOW RISK	LOW RISK	LOW RISK
Wittermans, 2021	LOW RISK	LOW RISK	LOW RISK	LOW RISK	LOW RISK	LOW RISK
Meduri, 2022	LOW RISK	LOW RISK	LOW RISK	LOW RISK	SOME CONCERNS	SOME CONCERNS
Dequin, 2023	LOW RISK	LOW RISK	LOW RISK	LOW RISK	LOW RISK	LOW RISK

Appendix Figure S12: Flow diagram of study selection resulting from the systematic literature search.



Appendix Part 4: Implementation of the LASSO penalty

We used Statsmodels' 'fit_regularized' function to implement penalized logistic regression,(23) which minimizes the following loss function:

$$\frac{LL}{2 * n_{train}} + \lambda\left((1 - L1_{wt})\frac{||x||_2^2}{2} + L1_{wt}||x||_1\right)$$

where LL represents the logistic loss, n_{train} the number of obervations (ie, patients) used to train the model, λ the penalization strength, $L1_{wt}$ the fraction of the penalty given to the L1 penalty term (ie, $L1_{wt}$=0 results in a Ridge fit, $L1_{wt}$=1 results in a LASSO fit), and $||x||_2$ and $||x||_1$ the L2 and L1 norms, respectively.

In this particular implementation, the penalization strength is influenced by the training data's size (n_{train}). In our study, we employed a leave-one-trial-out cross-validation, which resulted in variations in the size of the training set across different folds. For instance, the training set was notably smaller in the cross-validation fold where the study conducted by Blum et al.,(11) which accounted for 42% of all patients included in our study, formed the test cohort. To enhance the robustness of the penalization approach against changes in training set size across the cross-validation folds, we modified the loss function as follows:

$$\frac{LL}{2 * n_{train}} + \frac{\lambda}{n_{train}}\left((1 - L1_{wt})\frac{||x||_2^2}{2} + L1_{wt}||x||_1\right)$$

We implemented this by dividing λ of the 'fit_regularized' function (in the Statsmodels' implementation this argument is called 'alpha') by the size of the train set.

Appendix Part 5: Detailed description of the corticosteroid-effect model training

The training procedure of the corticosteroid-effect model comprised multiple steps, as visualized in Figure S13.

Step 1: A priori variable selection

Step one involved including variables based on availability: those available for at least two-thirds (ie, 67%) of patients in both train and test cohorts were included.

Step 2: Data imputation and normalization

In step two, missing values were imputed and data normalization was performed. We addressed missing values by the K-Nearest-Neighbour (KNN) imputation algorithm. This algorithm imputes missing values using values from the five nearest neighbours (i.e., the shortest Euclidean distance regarding the remaining variables) that have a value for that variable, averaging these uniformly. For binary variables, after averaging, we mapped values < 0.5 to 0 and values ≥ 0.5 to 1. To accomplish this, we first normalized all variables in the train cohort and the data from the observational study,(4) using centering (ie, making the data zero-mean) and standard scaling (ie, making the data unit variance). We fitted the imputer algorithm using the combined data of the train cohort with the observational study,(4) and used it to fill in missing values in both the train and test cohorts. Subsequently, we transformed the imputed datasets back to their original scale. Lastly, we normalized the imputed train and test cohorts once again by centering and scaling each variable (ie, both continuous and binary variables) based on its standard deviation, ensuring that all variables in the training data are zero-mean and have unit variance before these are used for model training.

Step 3: Encoding of the treatment variable

As proposed by Tian et al.,(24) in step three we encoded the treatment variable as placebo = -1 and corticosteroid treatment = 1.

Step 4: Addition of the treatment variable and variable-treatment interaction terms

Step four involved creating interaction terms by multiplying the included variables with the encoded treatment variable and adding them, together with the encoded treatment variable, to the logistic regression model.

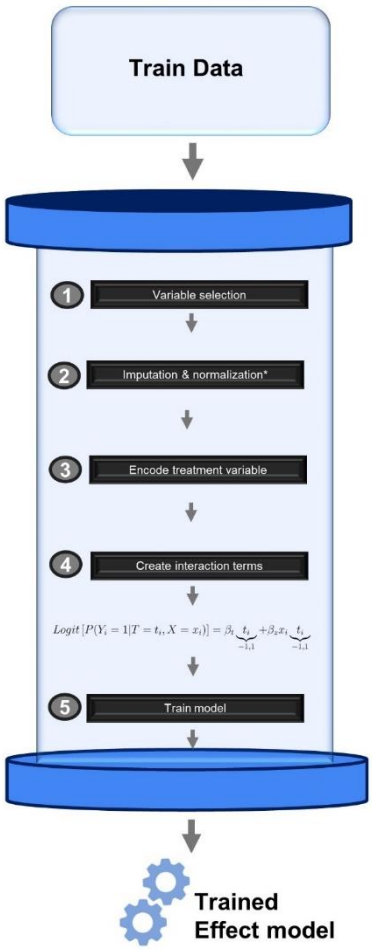
Step 5: Model training

In step 5, before training the Lasso regression model, the penalty strength (λ) was optimized through a 'leave-one-trial-out' cross-validation within the train cohort, selecting the λ that yielded the best cross-validated discrimination for benefit (see appendix part 6 for a detailed description). Then this optimal λ is used to train the penalized logistic regression model using all data from the train cohort, penalizing both treatment variable and interaction terms.

Step 6: Predict individualized treatment effects (ITEs) in the test cohort

Finally, in step six, this trained model is used to predict ITEs for patients in the test cohort.

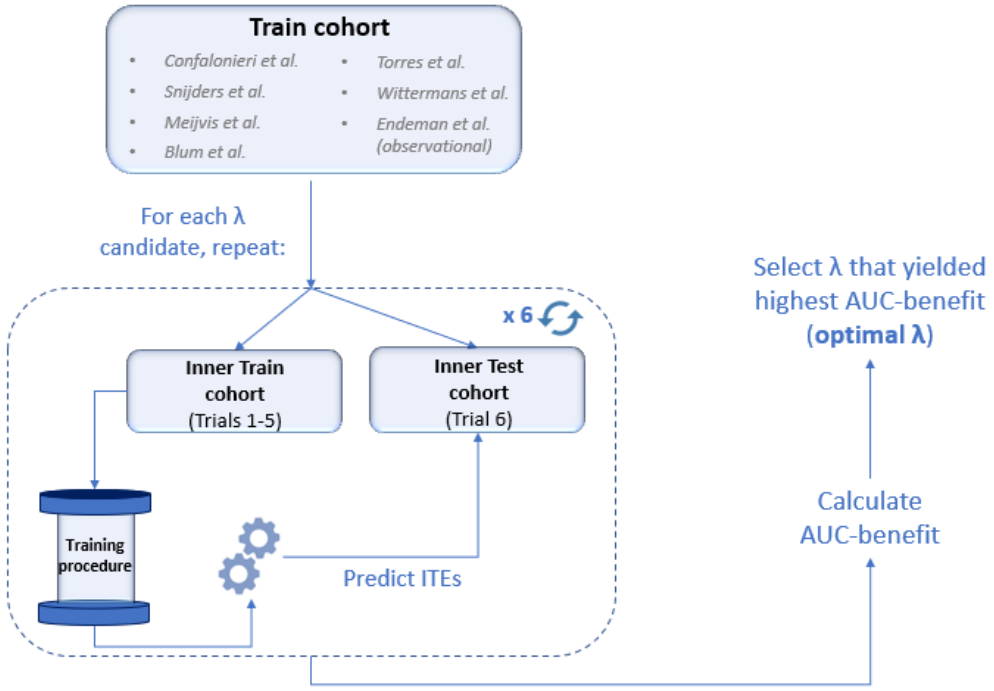
Appendix Figure S13: Visualization of the corticosteroid-effect model training procedure.



Appendix Part 6: Detailed description of the penalty strength (λ) optimization

The penalty strength (λ) was optimized using an exhaustive grid search, where the performance of each λ candidate was evaluated through a 'leave-one-trial-out' (LOTO) cross-validation (Figure S14). Here, in six iterations, combined IPD from five trials formed the 'inner train cohort', and the held out trial the 'inner test cohort'. Now, the modelling steps as described in Appendix Part 5 are repeated using the inner train cohort and the candidate λ, whereafter the model is trained and ITEs are predicted for the patients in the inner test cohort. Candidate λs resulting in zero weights for interaction terms and the treatment variable (ie, resulting in zero ITEs only) in at least one of the folds, were not considered. The predicted ITEs from the six iterations were then combined and from these predictions, we took 1000 bootstrap samples. For each bootstrap sample, we calculated the AUC-benefit, and the λ that yielded the highest median AUC-benefit (ie, the optimal λ) was selected. The first grid search used a default wide grid, and after the wide grid search, the optimal λ was used to define the center point of a finer grid for the second grid search (see Table S29), and the whole LOTO-CV procedure is repeated using this finer grid.

Appendix Figure S14: Schematic overview of the 'leave-one-trial-out' (LOTO) cross-validation procedure for penalty strength (λ) optimization.



Appendix Table S29: The (default) wide and fine grid spaces used in the grid searches. All grids were created evenly spaced on a logarithmic scale. The variable 'center' is defined as $\log^{10}(\lambda_{opt})$, where λ_{opt} is the optimal λ found in the first, wide grid search.

Grid search	Searched grid	N steps
Wide (default)	10^{-2} to 10^{2}	50
Fine	$10^{(center - 0.3)}$ to $10^{(center + 0.3)}$	100

Appendix Part 7: Definition of the 'Area under the Δ-benefit curve' (AUC-benefit)

The AUC-benefit is closely related to the (area under the) 'Qini' or 'Uplift' curve, as the Δ-benefit curve is a special case of the Qini/Uplift curve where treated and untreated patients are ranked jointly and the volumes are expressed in relative numbers (ie, percentiles).(25)

It involves considering different ITE thresholds to divide patients into two groups: a predicted harm group (where ITE ≤ threshold) and a predicted benefit group (where ITE > threshold). Both groups are further divided into those who received corticosteroids (G_1 and G_3) and those who received placebo (G_2 and G_4, Figure S15).

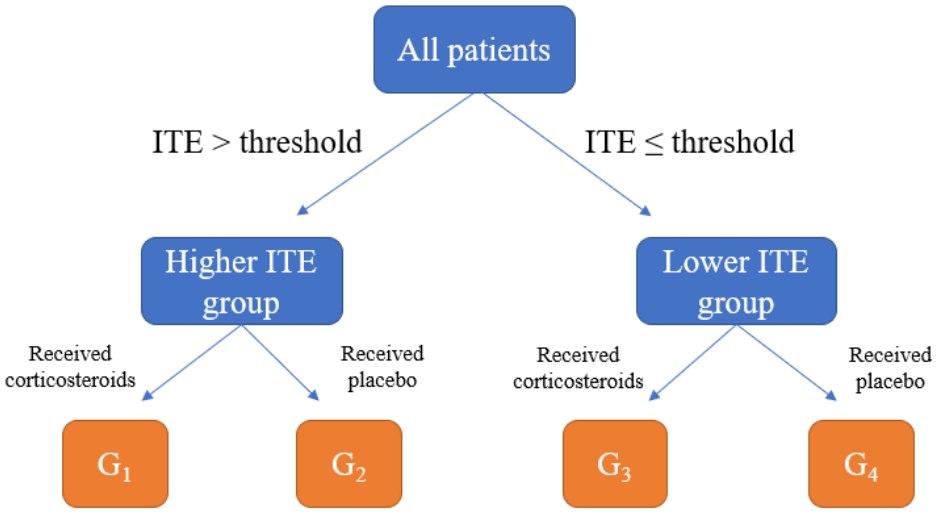
The Δ-benefit is defined as follows:

$$\Delta_{benefit} = \left[\frac{\sum_{i \subseteq G_2} y_i}{n_2} - \frac{\sum_{i \subseteq G_1} y_i}{n_1}\right] - \left[\frac{\sum_{i \subseteq G_4} y_i}{n_4} - \frac{\sum_{i \subseteq G_3} y_i}{n_3}\right]$$

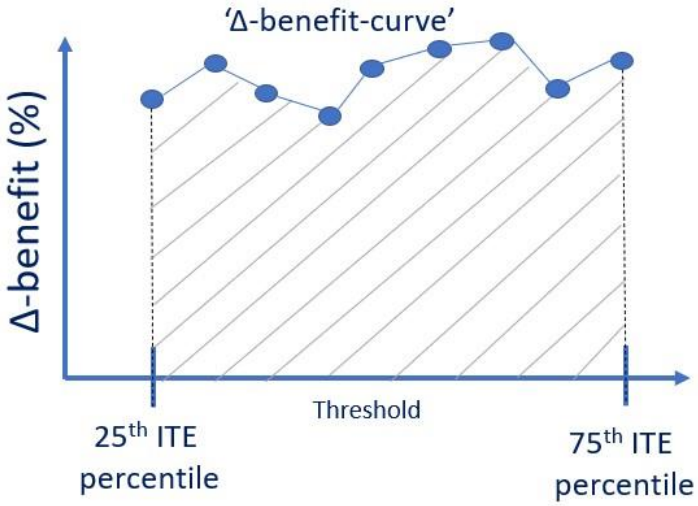
where i indexes the patient, y_i equals 1 in case 30-day mortality and 0 otherwise, and $n_{1\text{-}4}$ denote the number of patients in $G_{1\text{-}4}$.

The Δ-benefit is calculated considering a range of ten thresholds, starting with a threshold at the 25th percentile, and increase the percentiles in ten equal steps until the 75th percentile of the full ITE distribution. The calculated Δ-benefits for the different thresholds forms the 'Δ-benefit-curve', and the area under the Δ-benefit-curve (AUC-benefit) is calculated as the trapezoidal area under this curve (Figure S16). We used Sklearn's 'metrics.auc' function to calculate the AUC-benefit.

Appendix Figure S15: Schematic overview of patient grouping according to a certain ITE threshold.



Appendix Figure S16: Schematic overview of the area under the benefit-curve (AUC-benefit).



Appendix Part 8: Method Selection

Methods

Before obtaining IPD of the test cohort, we selected our effect modelling method (ie, the Tian method(24)) among alternative penalized regression methods. We explored different modelling choices, as we experimented with different modelling choices regarding:

- the **penalty** type (ie, Lasso or Ridge)
- the inclusion of **main effects** in the logistic regression model
- the inclusion of an **intercept term** in the logistic regression model
- the **encoding** of the **treatment variable** (ie, {0, 1} vs {-1, 1}

yielding 16 unique effect modelling methods (Table S30). Please note that 'effect-8' with LASSO penalization corresponds to the Tian method. We evaluated the discrimination for benefit in terms of AUC-benefit of each method through an 'internal-external validation' within the train cohort (see Figure S1 in appendix part 1). The predicted ITEs from the six iterations were then combined and from these predictions, we took 1000 bootstrap samples. For each bootstrap sample, we calculated the AUC-benefit, and the λ that yielded the highest median AUC-benefit (ie, the optimal λ) was selected.

Results

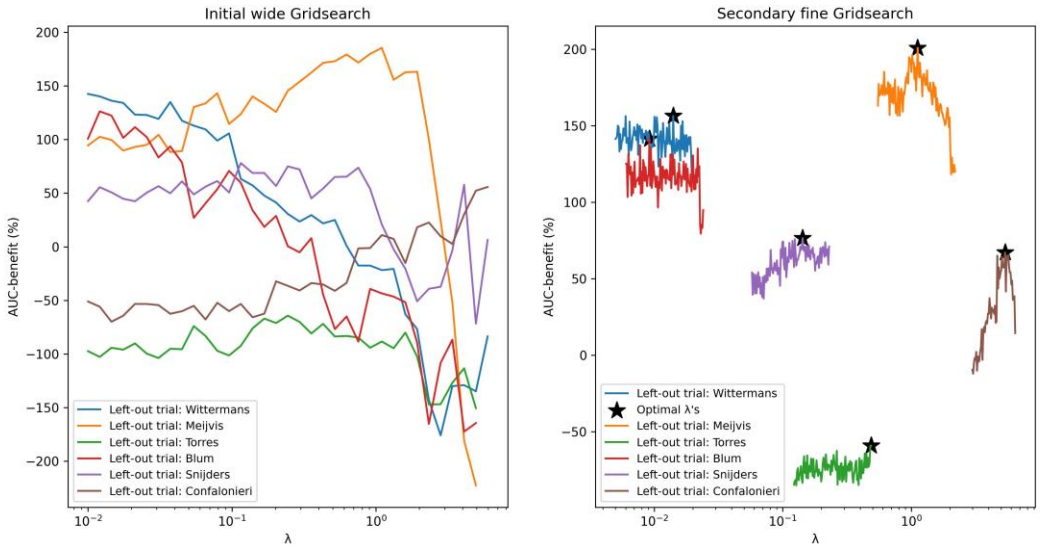
The results of the grid searches for all effect modelling procedures are visualized in Figure S18. The resulting weights of the trained models for the different modelling procedures are visualized in Figures S18. The boostrapped AUC-benefits for each method are visualized in Figure S19. The effect-8 model using a Lasso penalty resulted in the highest median AUC-benefit.

Appendix Table S30: Description of the different modelling procedures, where i indexes the patients, Y is the mortality, T is the treatment (ie, corticosteroids or placebo), X is the set of included variables, β_0 is the intercept term, β_t is the coefficient for the treatment variable, β_m includes the coefficients for the main effects (x_i) and β_z includes the coefficients for the treatment-variable interaction terms ($x_i\, t_i$) for an individual patient.

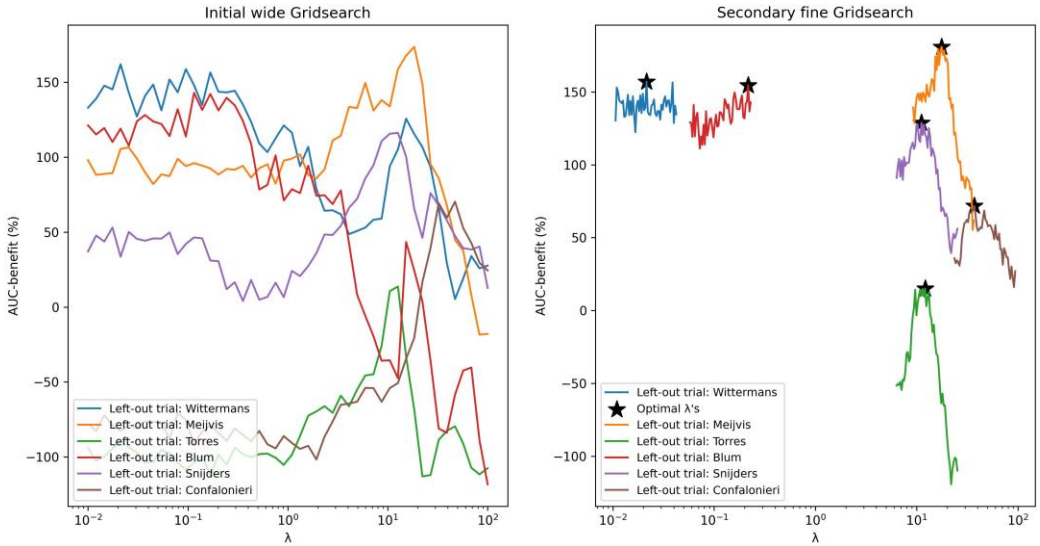
Model name	**Main effects**	**Intercept term**	**Encoding treatment variable**	**Formula**
Effect-1	✓	✓	{0, 1}	$Logit\left[P(Y_i = 1 \mid T = t_i, X = x_i)\right] = \beta_0 + \beta_t \underbrace{t_i}_{0,1} + \beta_m x_i + \beta_z x_i \underbrace{t_i}_{0,1}$
Effect-2	✓	✓	{-1, 1}	$Logit\left[P(Y_i = 1 \mid T = t_i, X = x_i)\right] = \beta_0 + \beta_t \underbrace{t_i}_{-1,1} + \beta_m x_i + \beta_z x_i \underbrace{t_i}_{-1,1}$
Effect-3	✓	×	{0, 1}	$Logit\left[P(Y_i = 1 \mid T = t_i, X = x_i)\right] = \beta_t \underbrace{t_i}_{0,1} + \beta_m x_i + \beta_z x_i \underbrace{t_i}_{0,1}$
Effect-4	✓	×	{-1, 1}	$Logit\left[P(Y_i = 1 \mid T = t_i, X = x_i)\right] = \beta_t \underbrace{t_i}_{-1,1} + \beta_m x_i + \beta_z x_i \underbrace{t_i}_{-1,1}$
Effect-5	×	✓	{0, 1}	$Logit\left[P(Y_i = 1 \mid T = t_i, X = x_i)\right] = \beta_0 + \beta_t \underbrace{t_i}_{0,1} + \beta_z x_i \underbrace{t_i}_{0,1}$
Effect-6	×	✓	{-1, 1}	$Logit\left[P(Y_i = 1 \mid T = t_i, X = x_i)\right] = \beta_0 + \beta_t \underbrace{t_i}_{-1,1} + \beta_z x_i \underbrace{t_i}_{-1,1}$
Effect-7	×	×	{0, 1}	$Logit\left[P(Y_i = 1 \mid T = t_i, X = x_i)\right] = \beta_t \underbrace{t_i}_{0,1} + \beta_z x_i \underbrace{t_i}_{0,1}$
Effect-8 (Tian)	×	×	{-1, 1}	$Logit\left[P(Y_i = 1 \mid T = t_i, X = x_i)\right] = \beta_t \underbrace{t_i}_{-1,1} + \beta_z x_i \underbrace{t_i}_{-1,1}$

Appendix Figure S17: Results of the initial (wide) and (secondary) fine grid searches for λ optimalization in each LOTO-CV fold, resulting from the different variations of the modelling procedures without additional dichotomized variables.

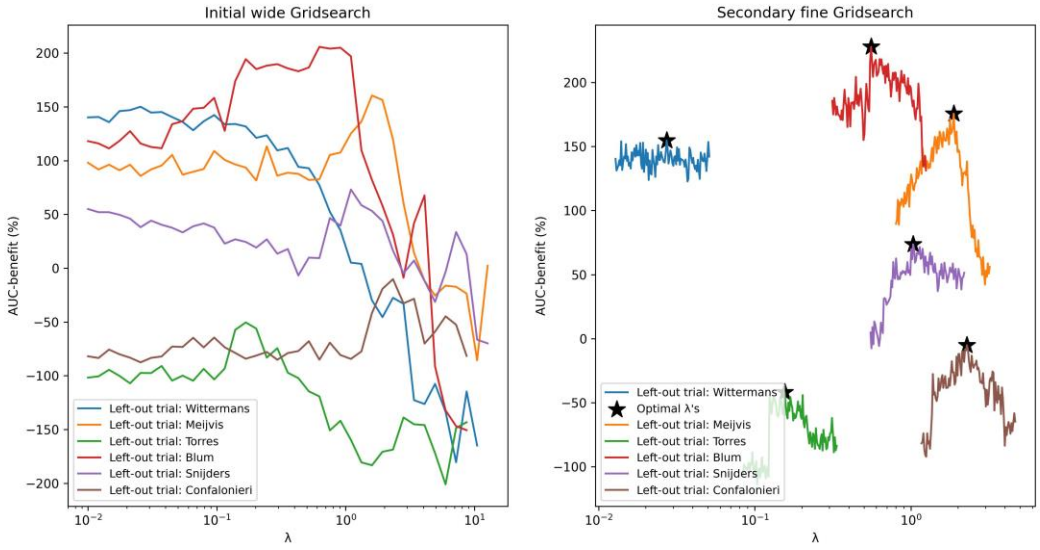
(a) Effect-1, Lasso



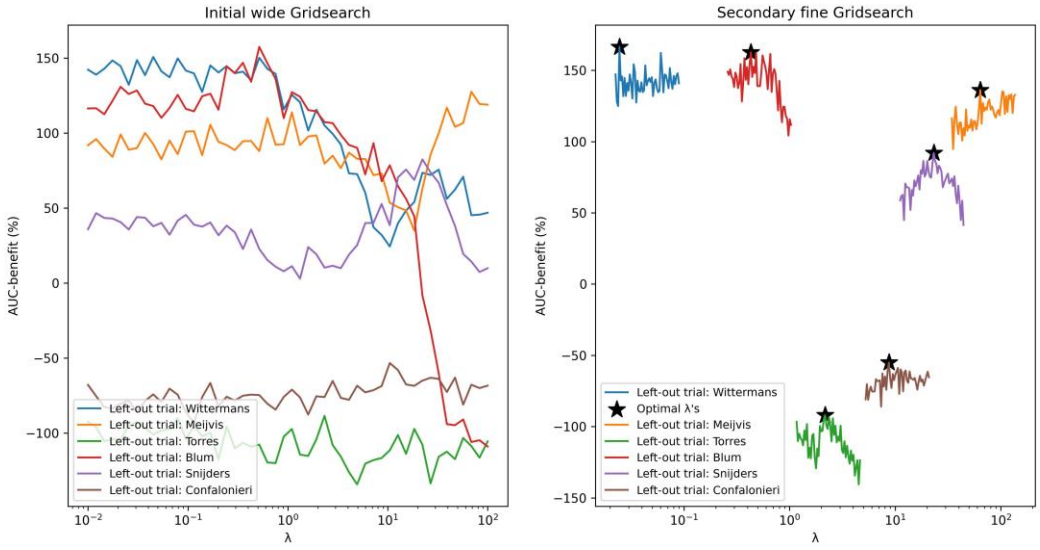
(b) Effect-1, Ridge



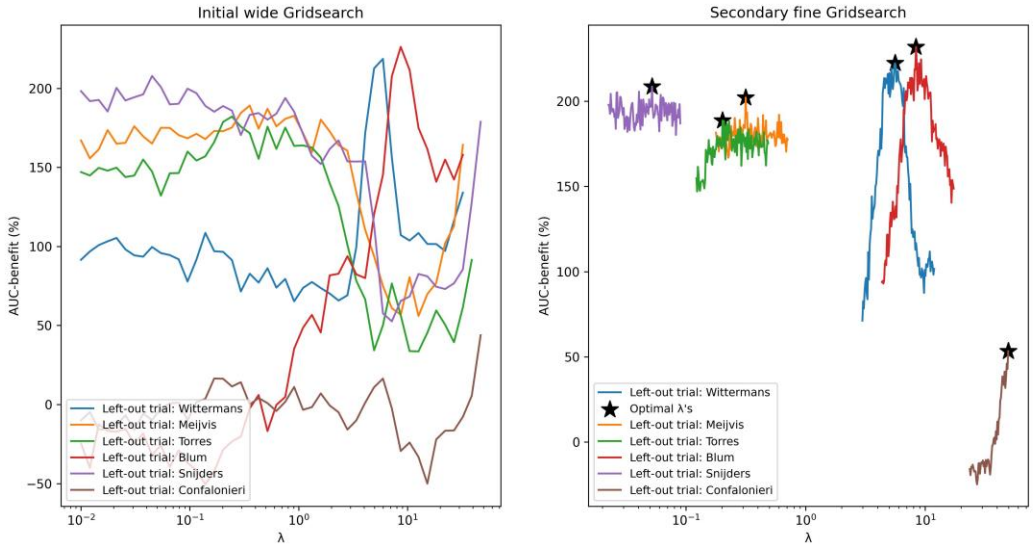
(c) Effect-2, Lasso



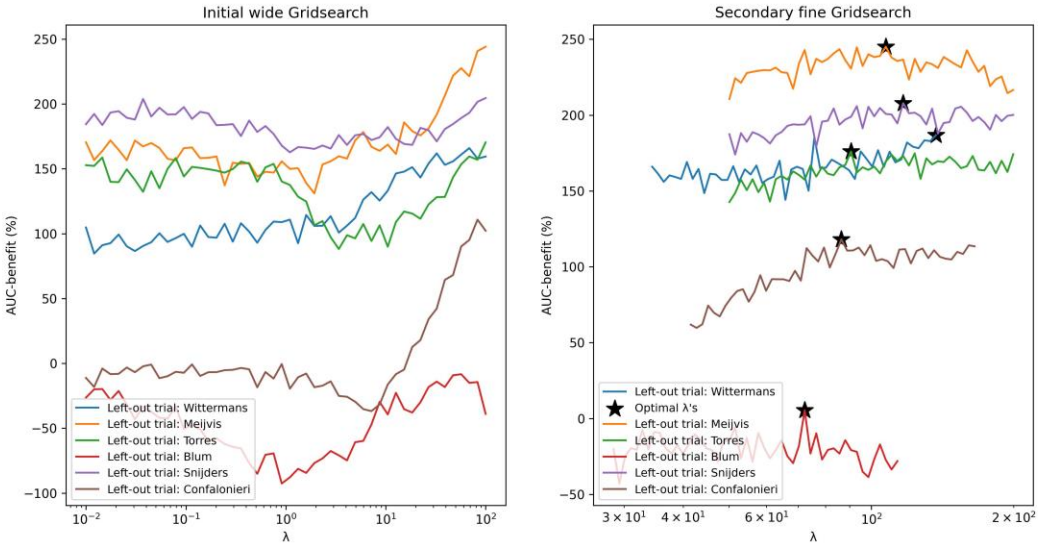
(d) Effect-2, Ridge



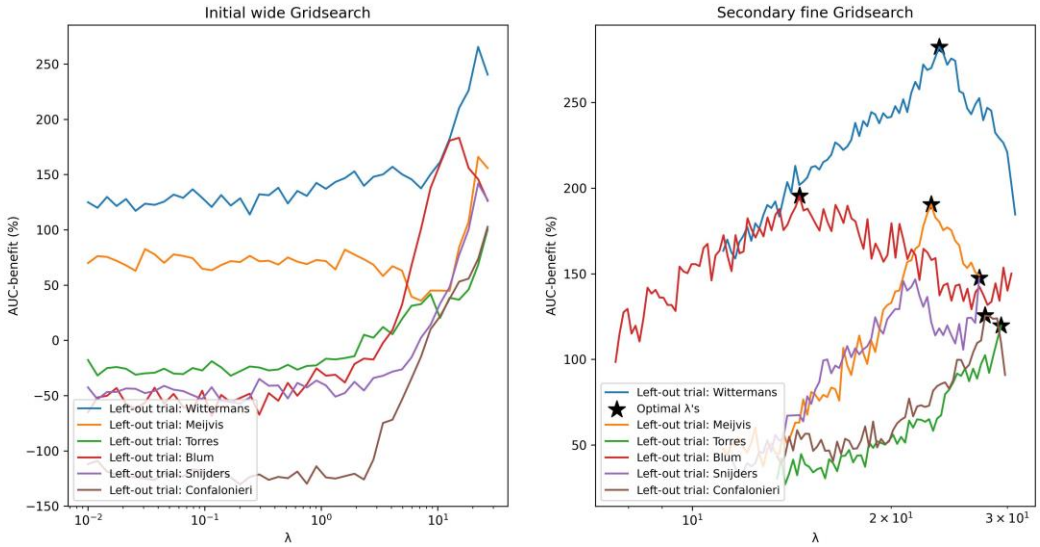
(e) Effect-3, Lasso



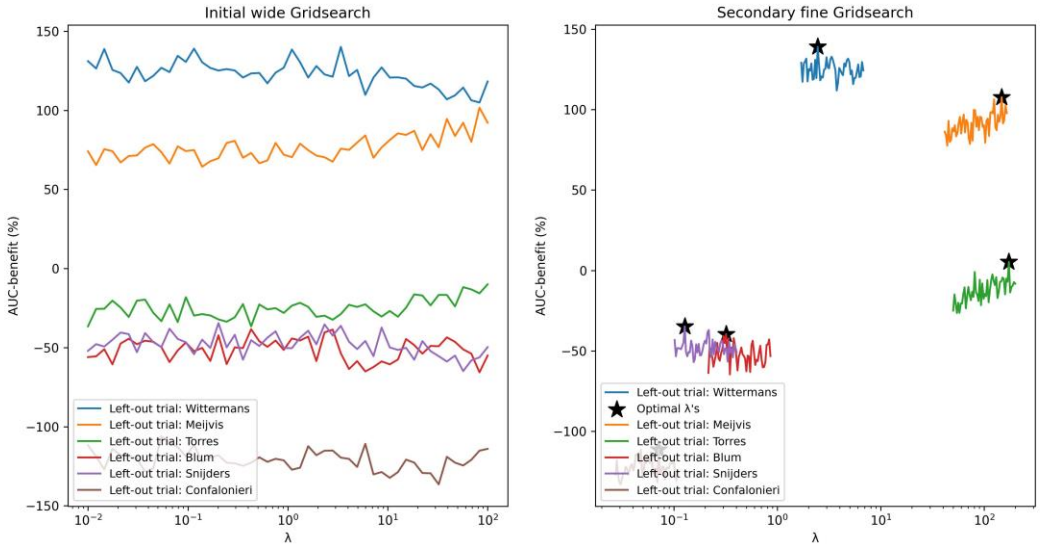
(f) Effect-3, Ridge



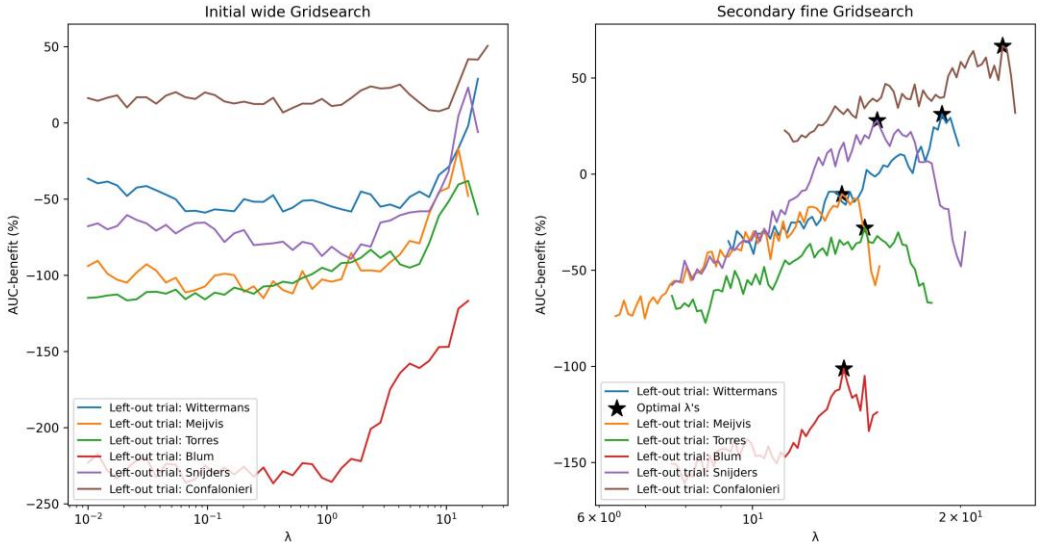
(g) Effect-4, Lasso



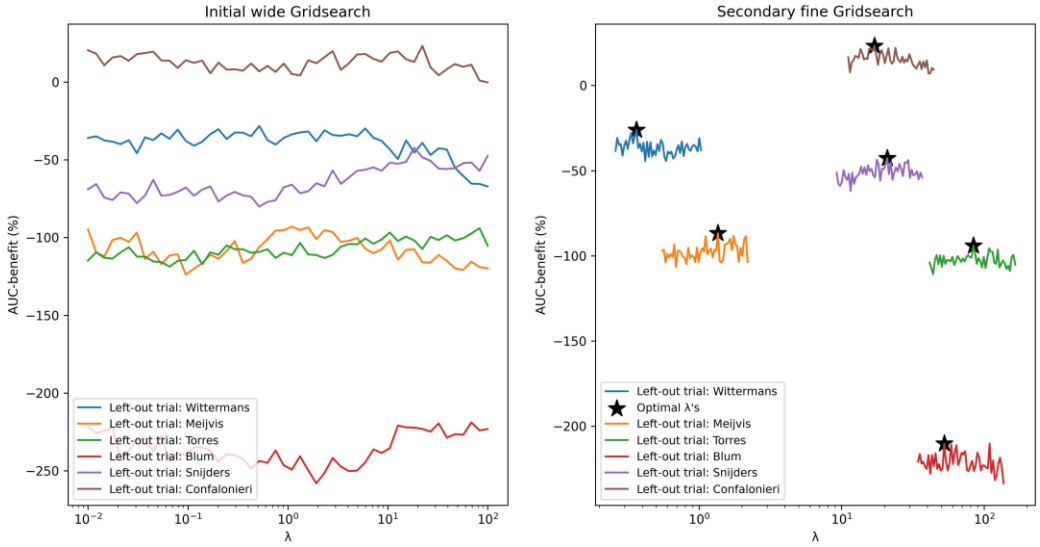
(h) Effect-4, Ridge



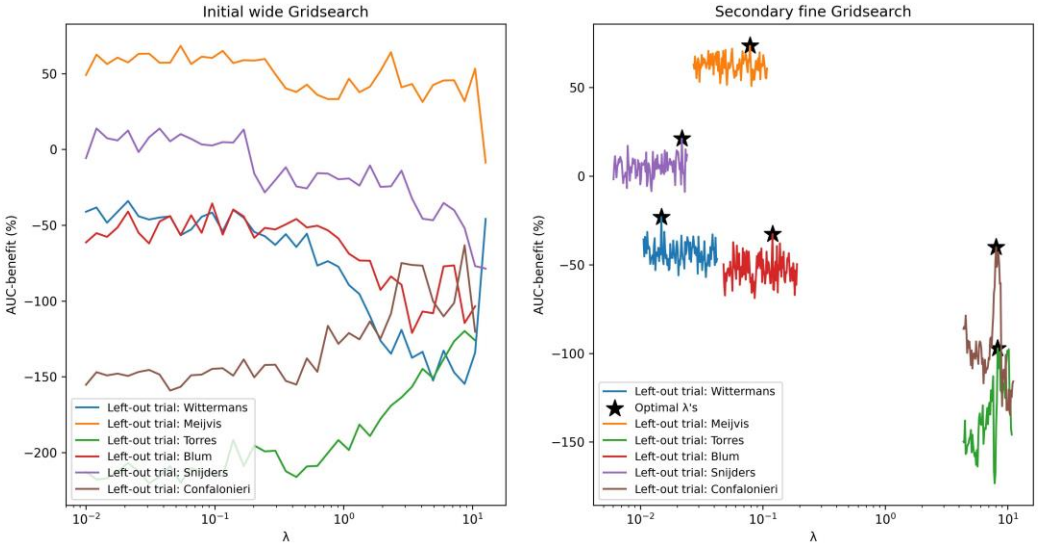
(i) Effect-5, Lasso



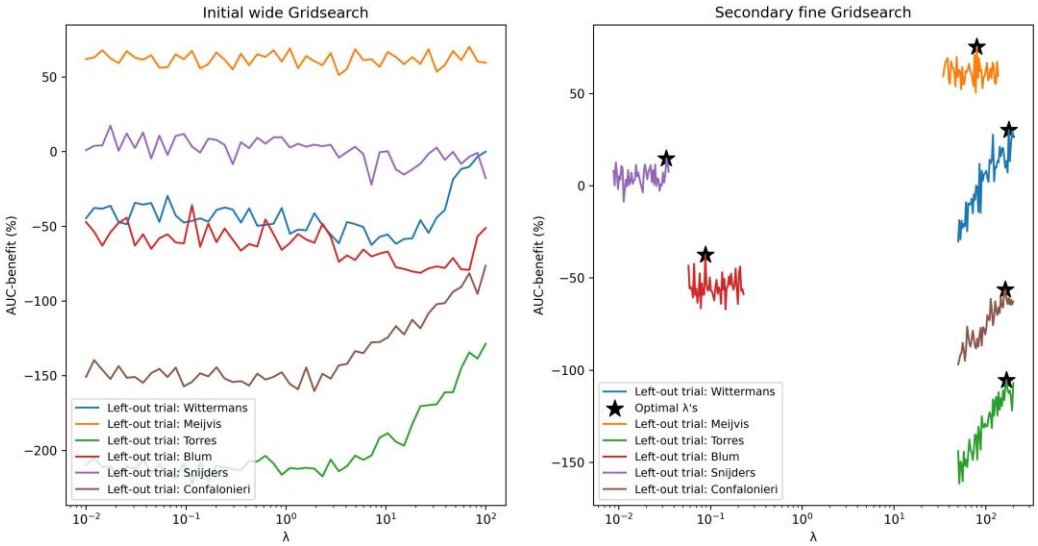
(j) Effect-5, Ridge



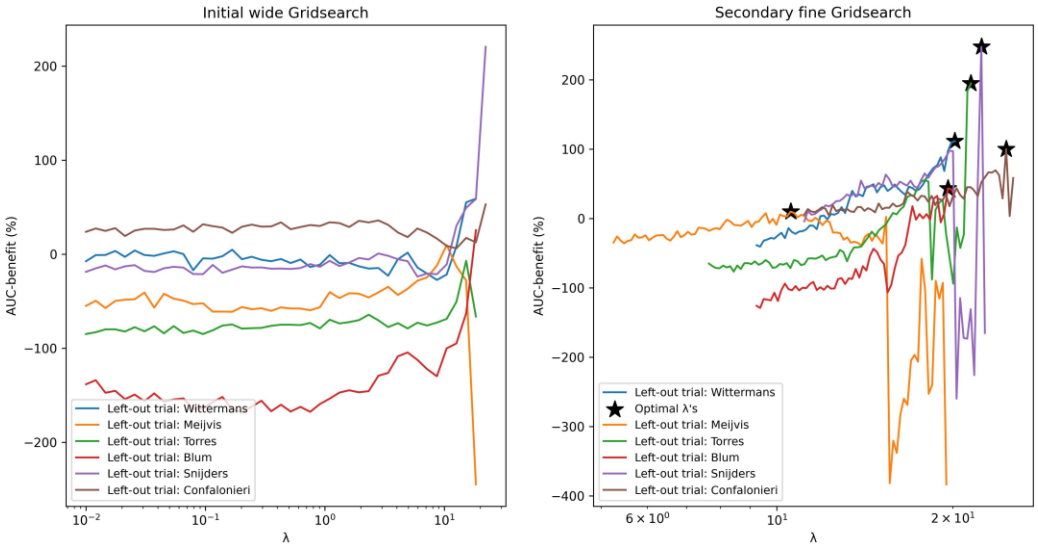
(k) Effect-6, Lasso



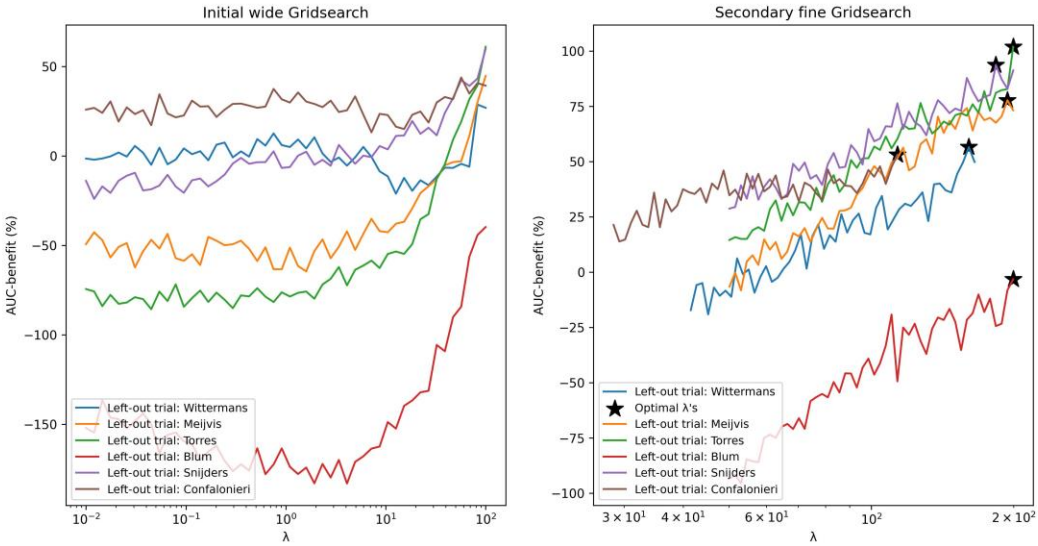
(l) Effect-6, Ridge



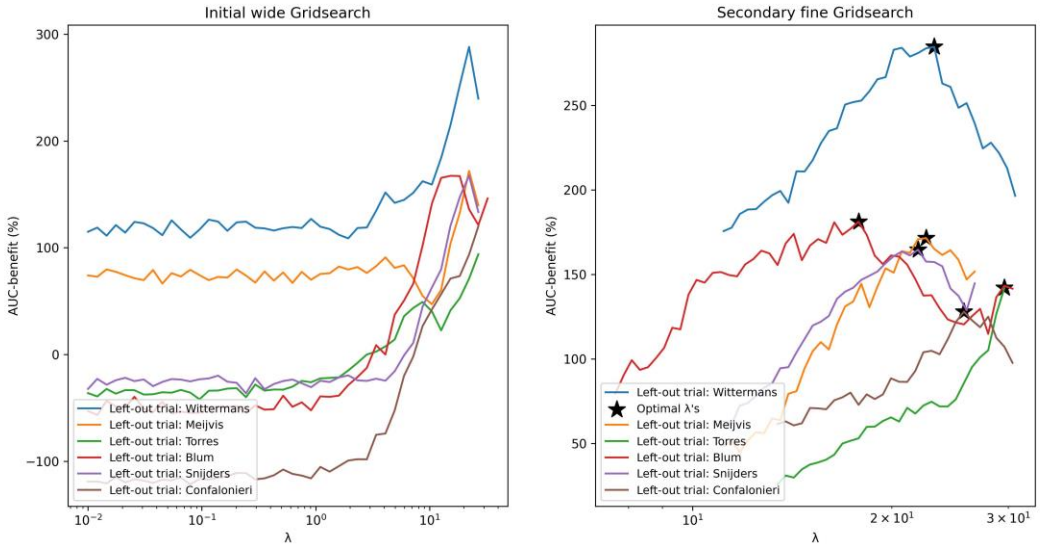
(m) Effect-7, Lasso



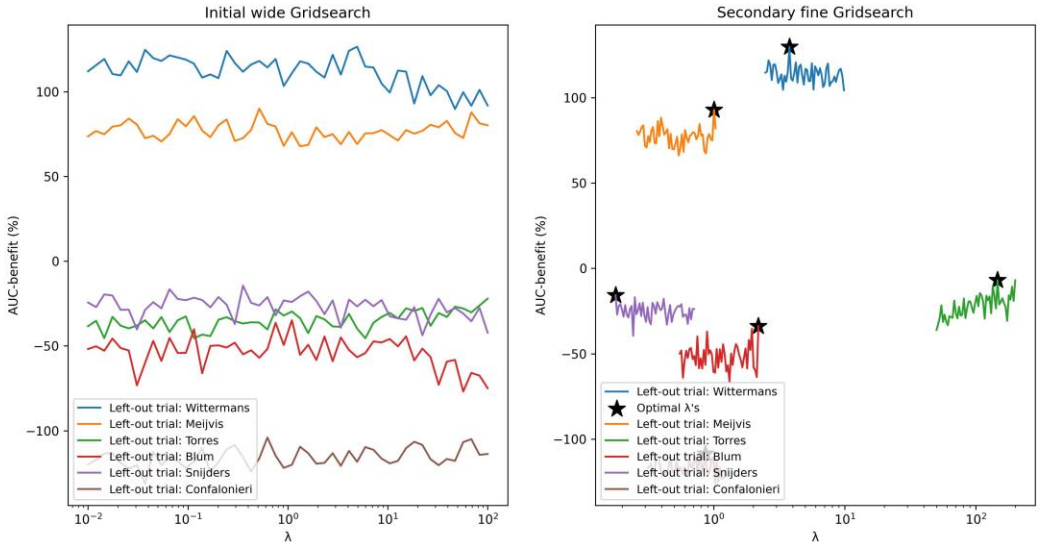
(n) Effect-7, Ridge



(o) Effect-8, Lasso

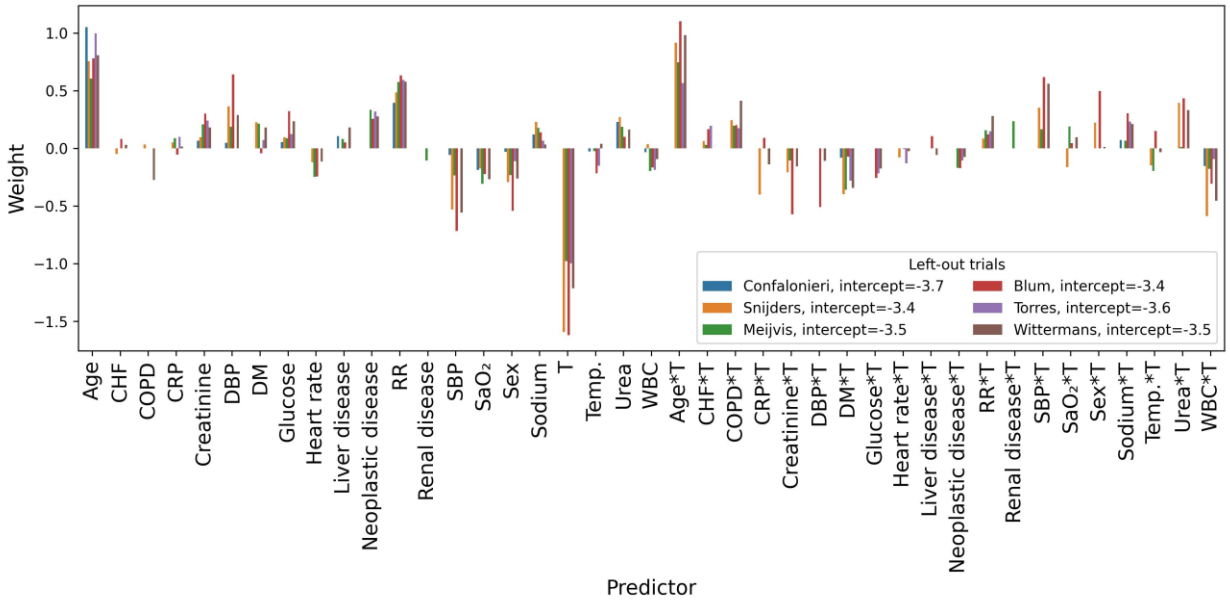


(p) Effect-8, Ridge

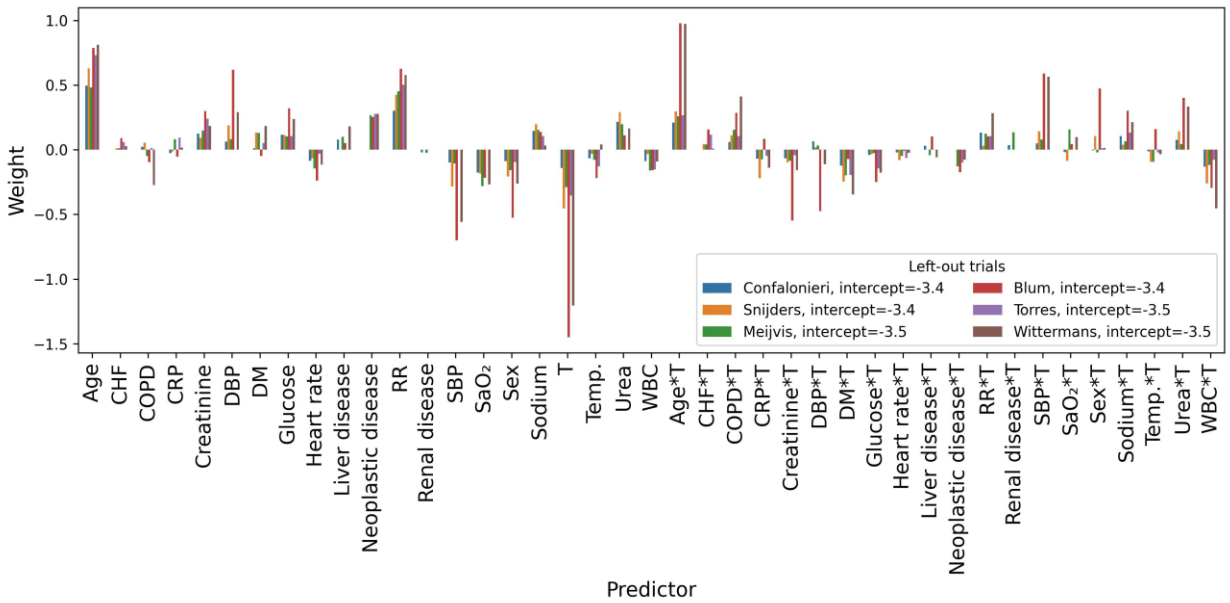


Appendix Figure S18: Bar charts of all non-zero weights of the fitted logistic regression models in each LOTO-CV fold, resulting from the different variations of modelling procedures without additional dichotomized variables. RR=respiratory rate, DBP=Diastolic blood pressure, SBP=Systolic blood pressure, Temp.=Body temperature, CRP=C-reactive protein, WBC=White cell count, T=Treatment variable

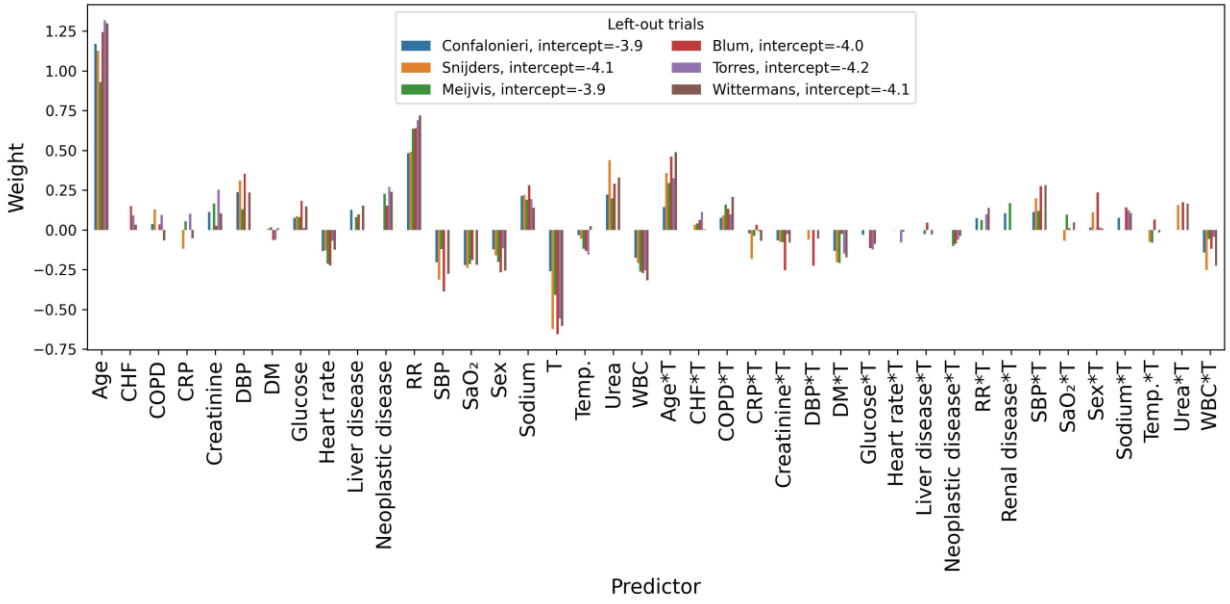
(a) Effect-1, Lasso



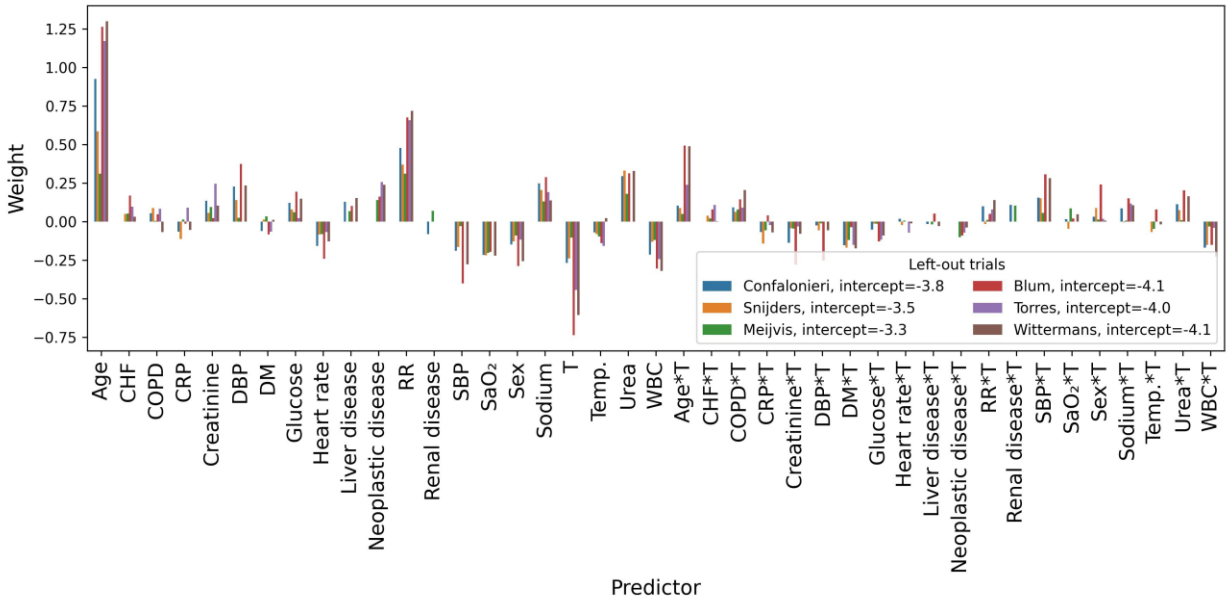
(b) Effect-1, Ridge



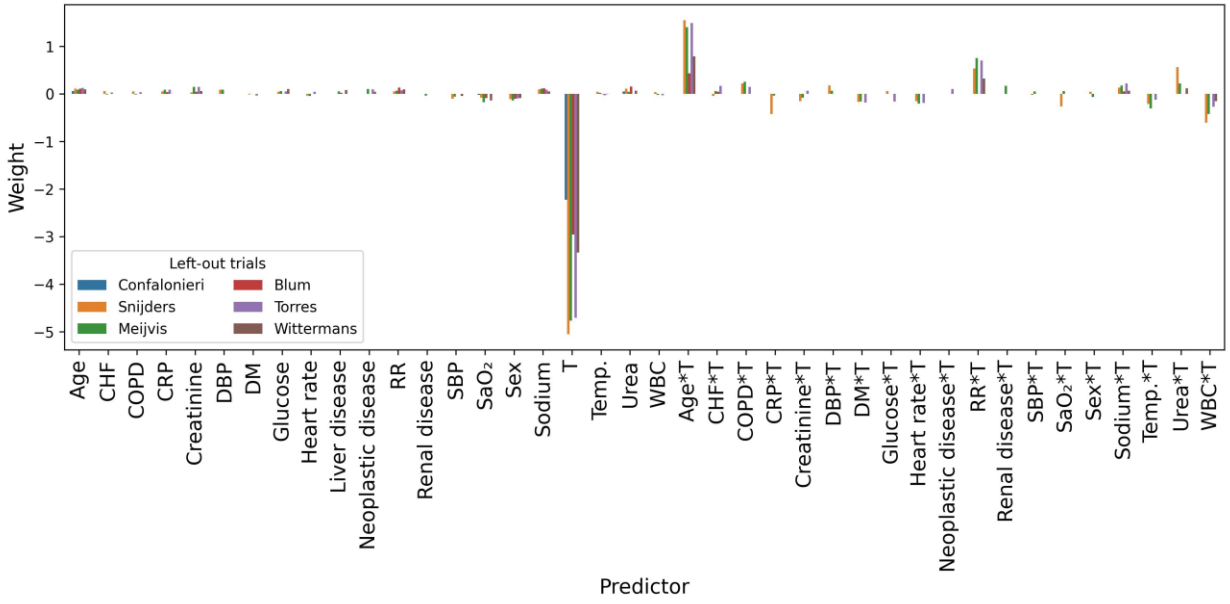
(c) Effect-2, Lasso



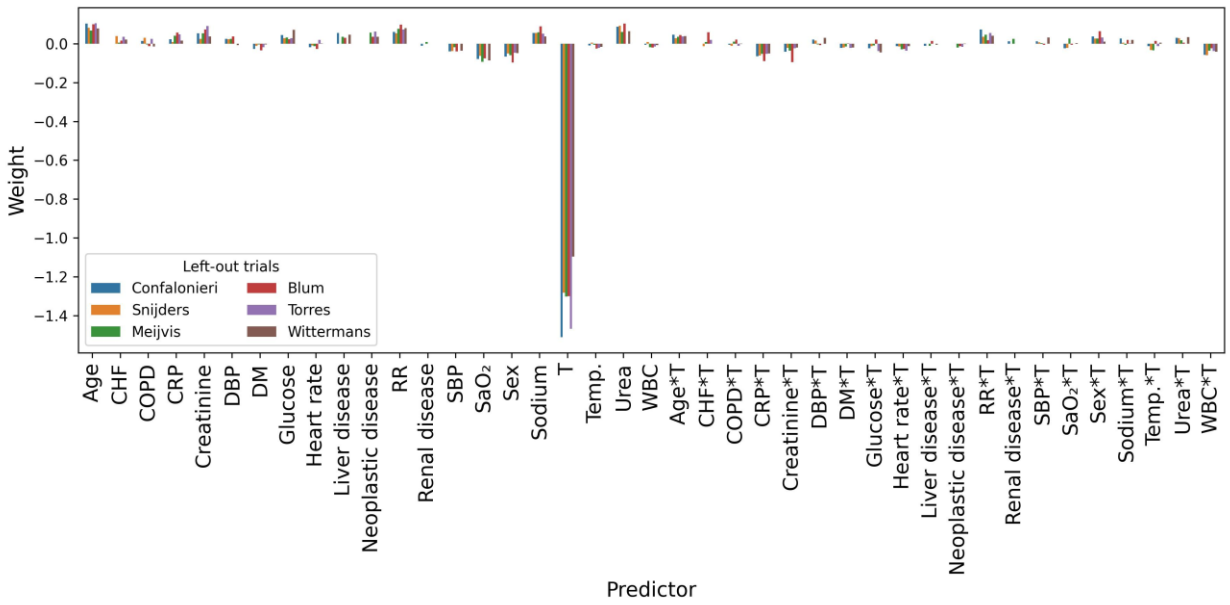
(d) Effect-2, Ridge



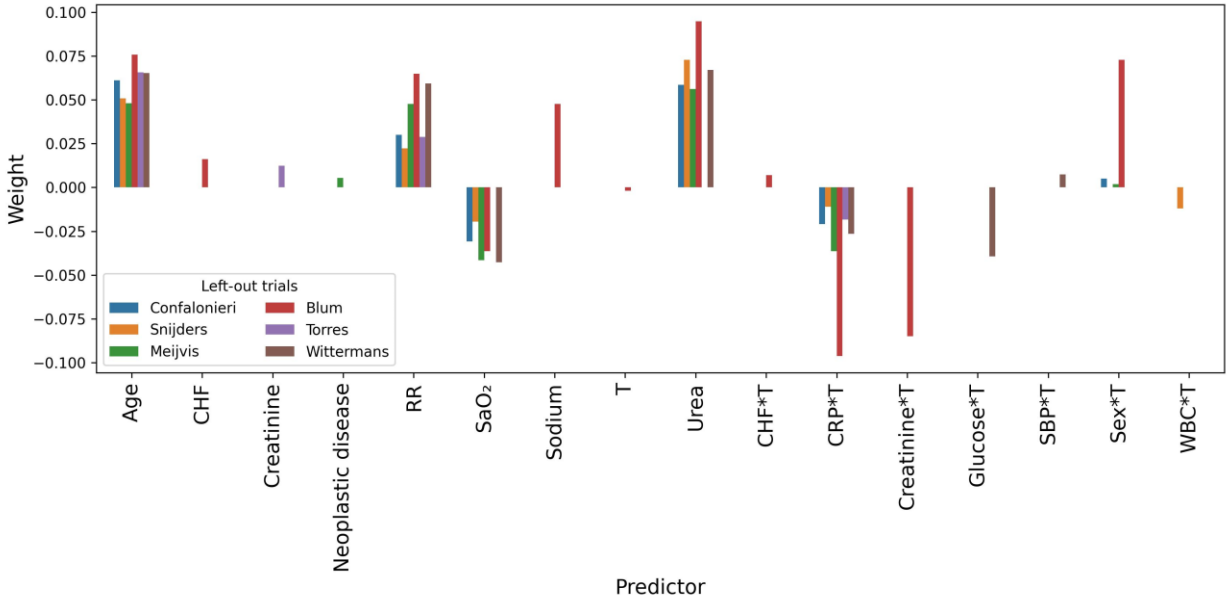
(e) Effect-3, Lasso



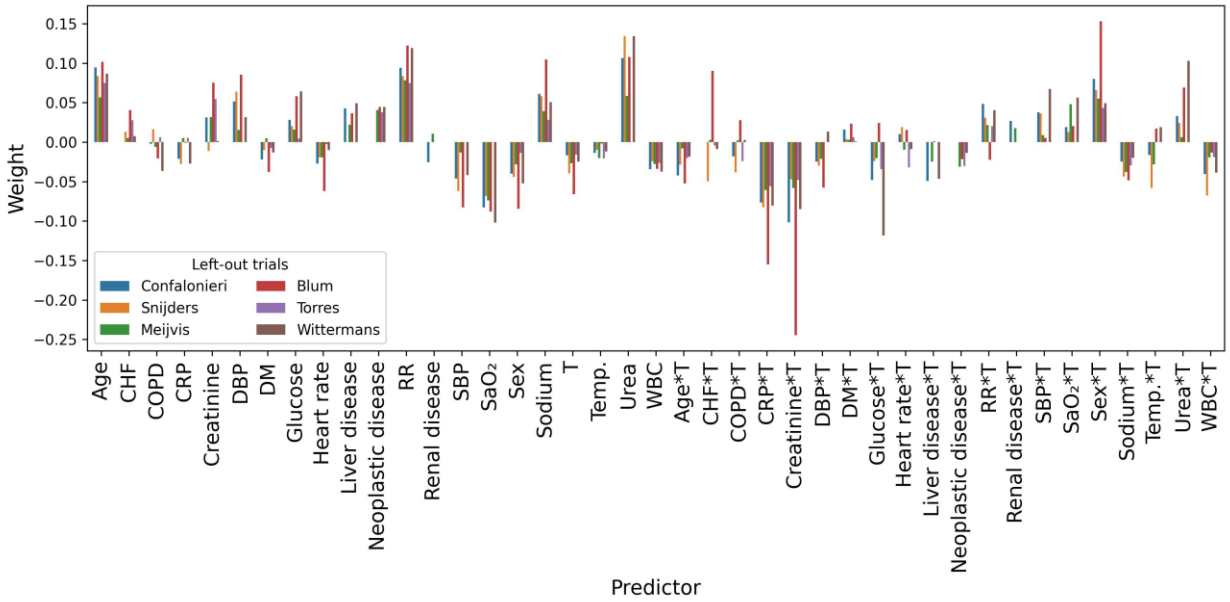
(f) Effect-3, Ridge



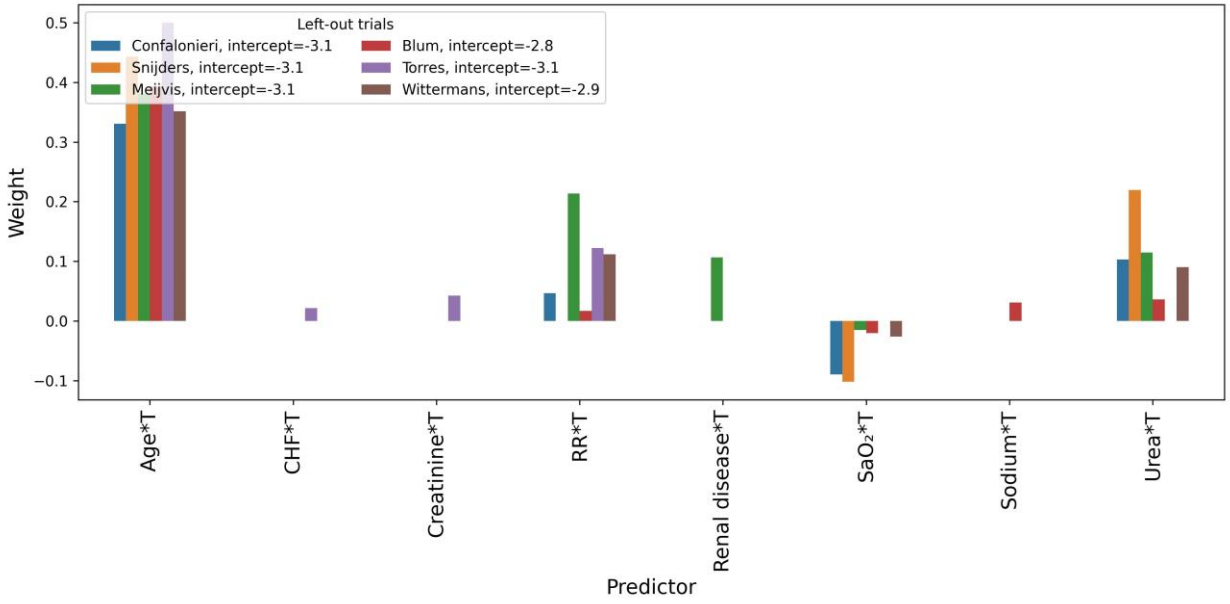
(g) Effect-4, Lasso



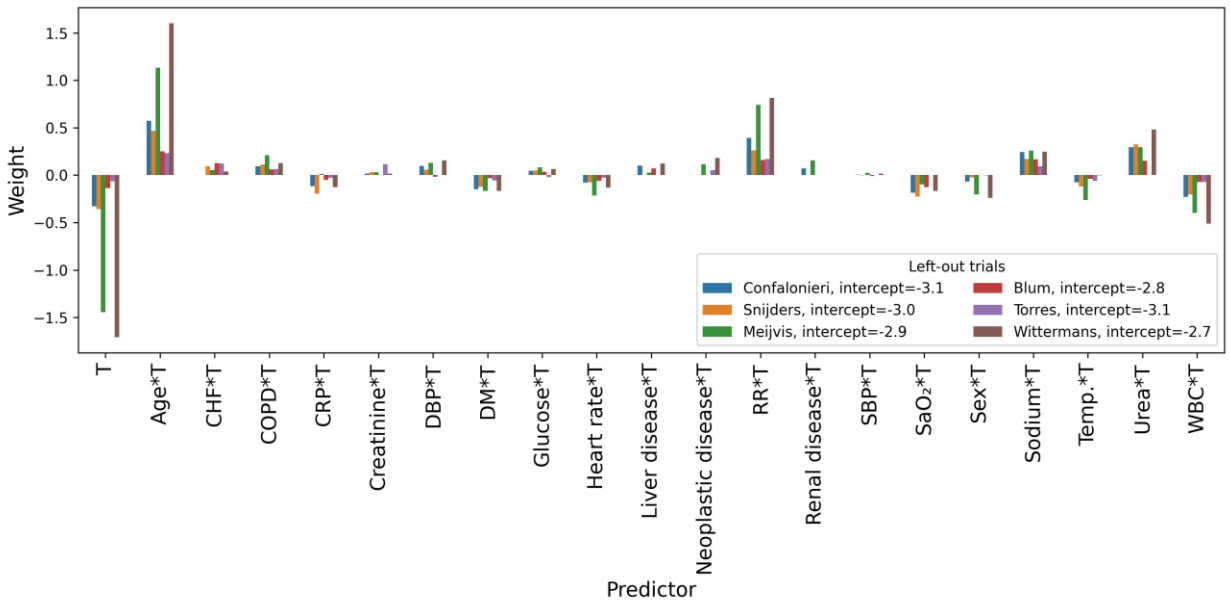
(h) Effect-4, Ridge



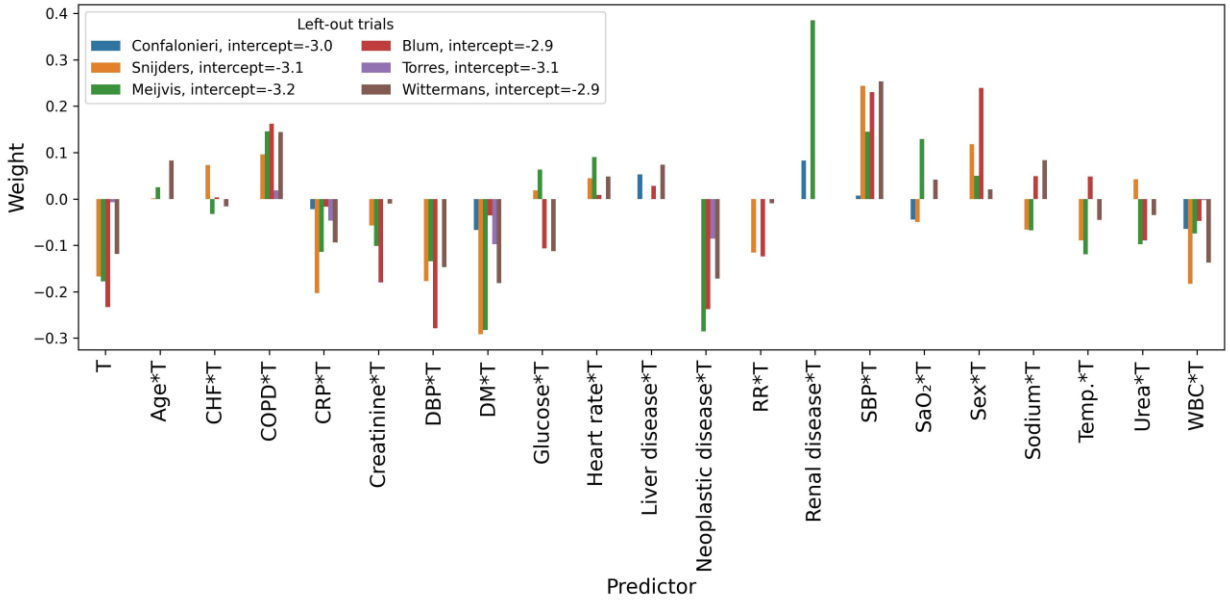
(i) Effect-5, Lasso



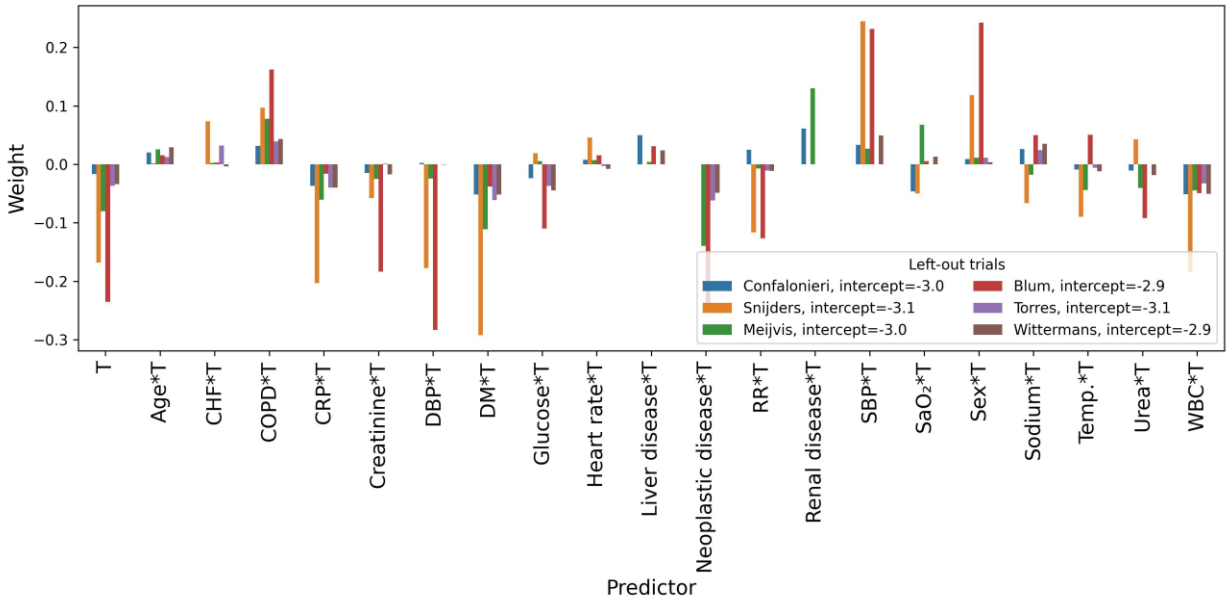
(j) Effect-5, Ridge



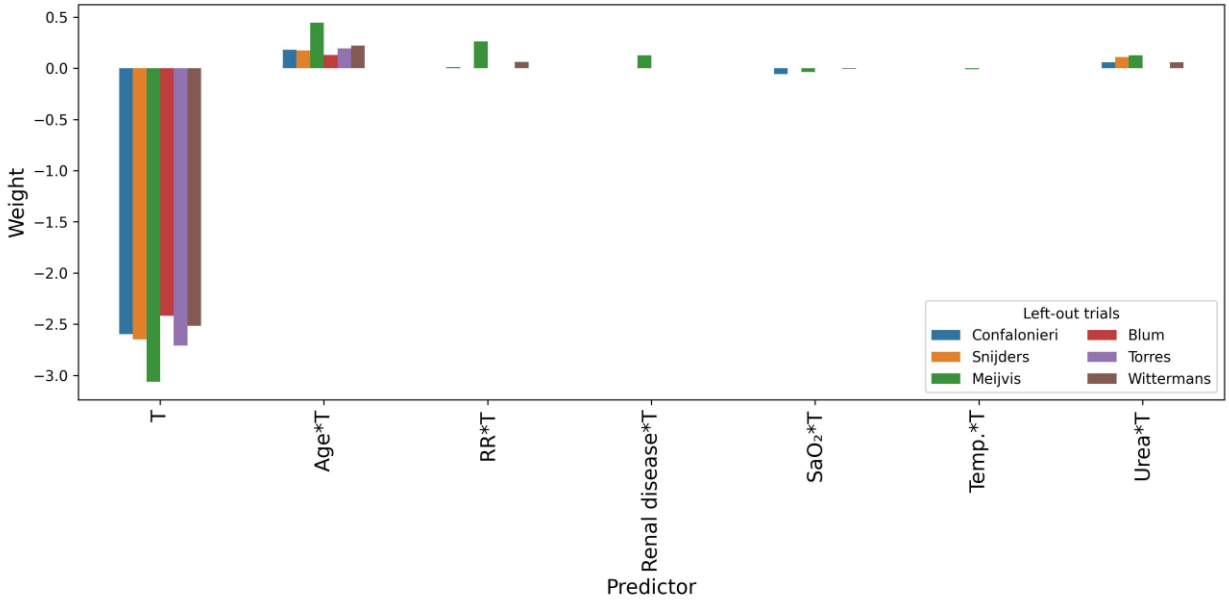
(k) Effect-6, Lasso



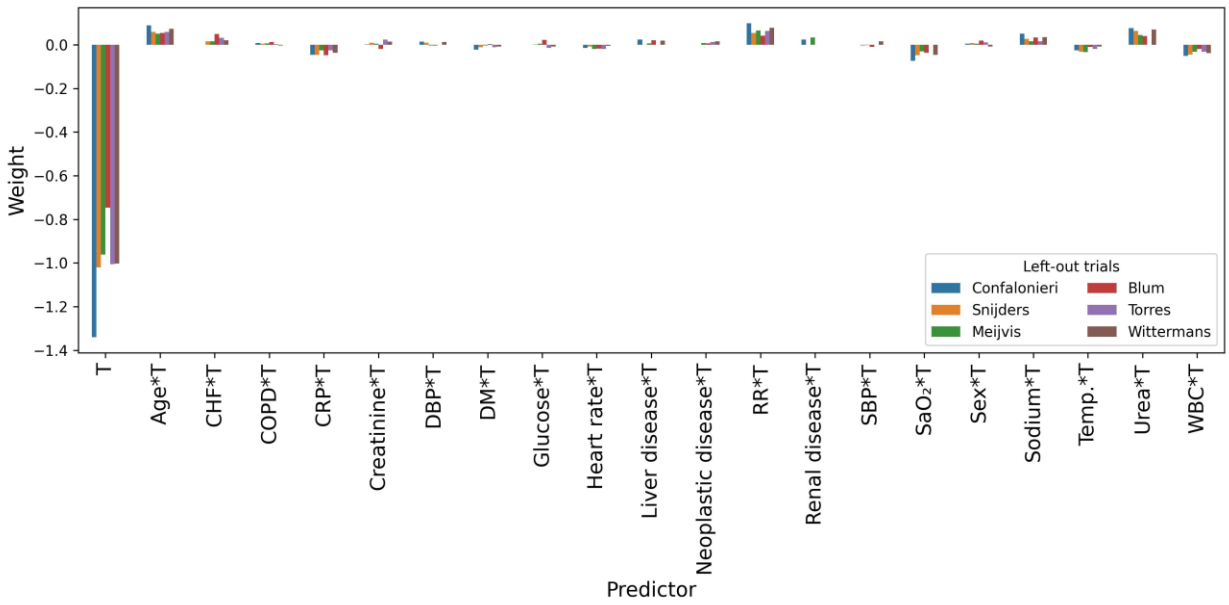
(l) Effect-6, Ridge



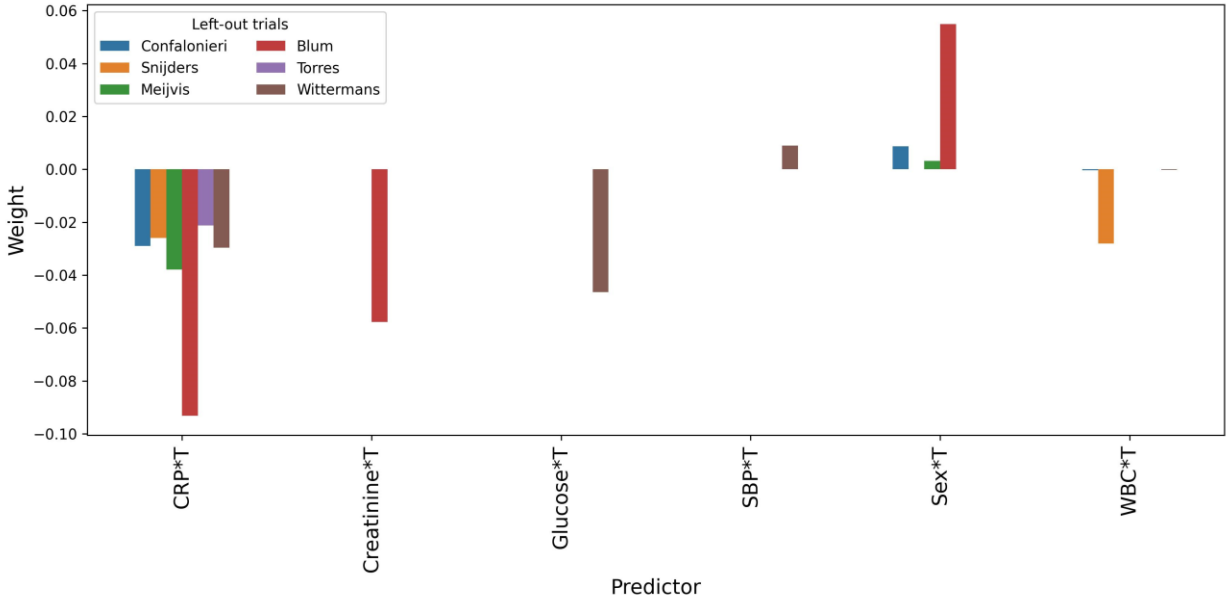
(m) Effect-7, Lasso



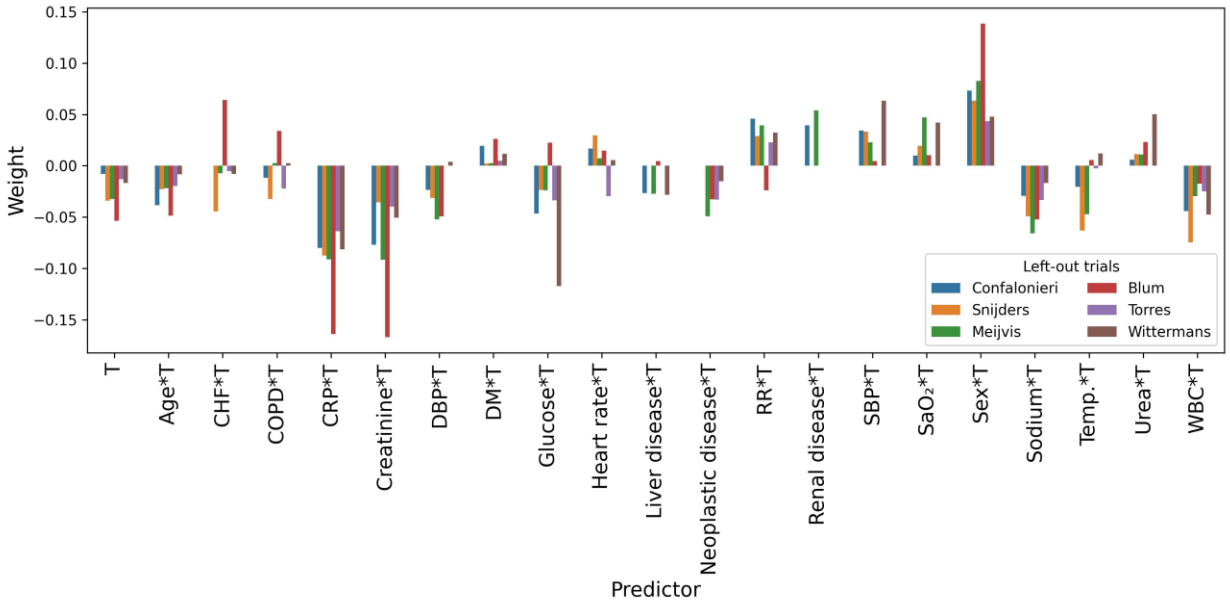
(n) Effect-7, Ridge



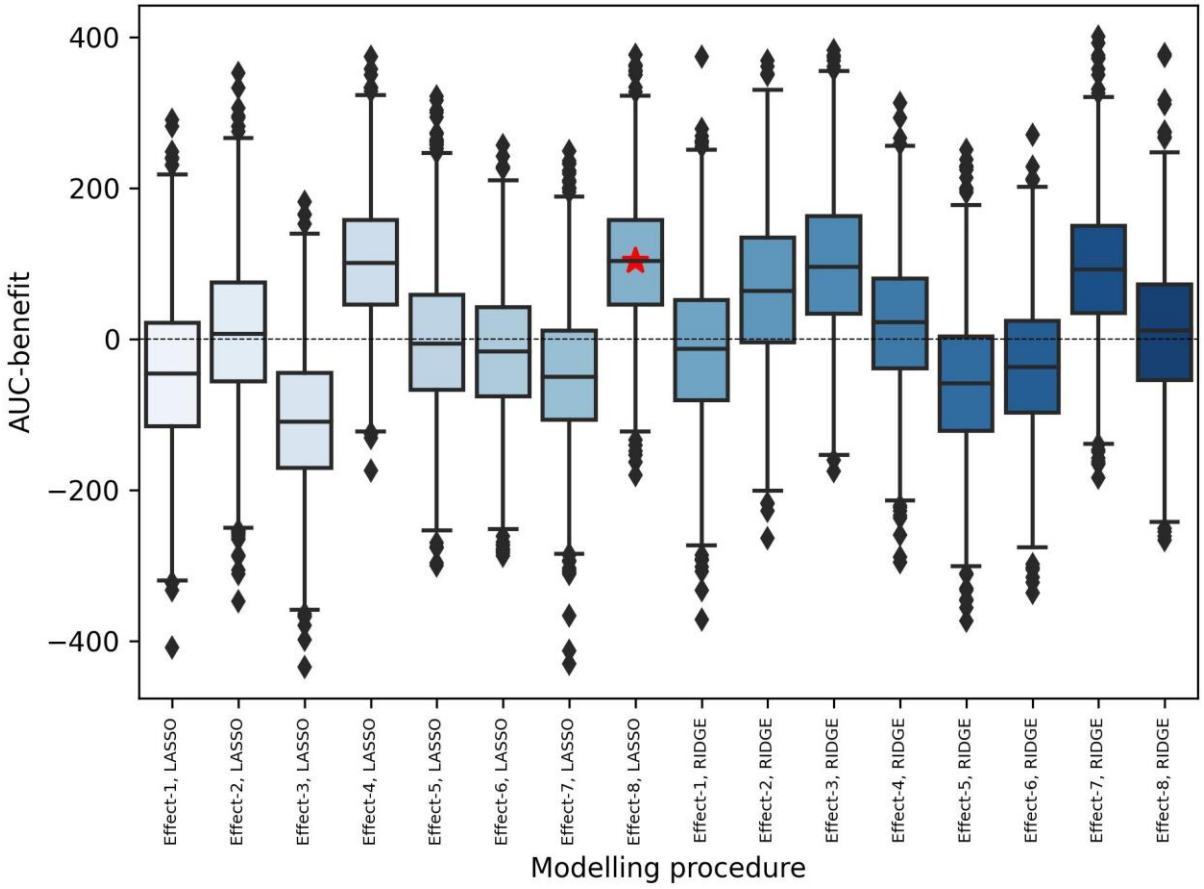
(o) Effect-8, Lasso



(p) Effect-8, Ridge



Appendix Figure S19: Discriminative performances of the different variations of modelling procedures without additional dichotomized variables, in terms of AUC-benefits (boxplots were created with 1000 bootstrap samples).



Appendix Part 9: Non-linear effect modelling

Methods

As the Tian method(24) only allows for linear modelling of treatment-covariate interactions, we additionally explored the performance of non-linear effect modelling methods. Each of the non-linear effect modelling methods is described in detail below.

To allow for modelling non-linear effects through the Tian method, we introduced extra, non-linear terms. Specifically, after imputation, each continuous variable was split based on the median value within the training data and added to the model. For instance, in addition to heart rate as a continuous variable, we also included a binary variable encoded as '1 if heart rate > 100 bpm (ie, the median) , 0 otherwise'. The resulting model following this method was pre-specified in our preliminary results, and we refer to it as ‘Non-linear Tian’.

Post-hoc, we repeated the model training and external validation using more flexible causal machine learning methods, ie, the causal forest (26,27), X-learner (28), and R-learner(29). The causal forest(26,27) is tree-based model that iteratively builds a combination of decision trees for estimating heterogeneous treatment effects that produces predicted values of the unit-level conditional average treatment effects (CATEs) rather than predicted values of the outcome variable, as in the traditional random forest. We applied this method using the EconML(30) ‘grf.CausalForest’ function. The X-learner(28) is a type of meta-algorithm or ‘meta-learner’, which decompose ITE (or CATE) estimation into several sub-regression or sub-classification problems that can be solved with any regression or supervised machine learning method.(28,31) The X-learner(28) uses each observation in the training set in an ‘X’–like shape, as it uses the observed outcomes to estimate the un-observed individual treatment effects, and then estimates the ITE function in a second step as if the ITEs were observed. The R-learner(29) is another two-step meta-learner, which uses the Robinson transformation.(32) We implemented both X- and R-learner with the XGBoost algorithm,(33) which is a tree-based model that builds a collection of decision trees with advanced regularization to reduce overfitting, using the ‘xgboost’ library in Python.

For each method, we optimized hyperparameters in the same fashion as for the linear Tian method, ie, using a grid search and select the options yielding the higest cross-validated AUC-benefit in a leave-one-trial-out cross-validation using the train cohort (see Appendix Part 5). For the non-linear Tian method, we optimized the Lasso

penalty strength (λ). For the Causal Forest, X- and R-learner, we used a single grid search, looping through all possible combinations of a set of method-specific hyper-parameters. The searched grids for each method are outlined in Table S31, and yielded 756, 90 and 90 unique hyperparameter combinations for the causal forest, X-learner and R-learner, respectively.

For each method, we evaluated the discrimination and calibration for benefit in the external validation. To evaluate the discrimination for benefit, we took 1000 bootstrap samples of the test-cohort and calculated the AUC-benefit in each bootstrap sample. We evaluated the calibration for benefit as described in the main text.

We benchmarked discrimination for benefit results to the performance in 'apparent validation'(34), ie, the performance of the non-linear effect models in the data it was trained on (the train cohort). Additionally, we added the discrimination for benefit results of the linear Tian method (ie, the corticosteroid-effect model presented in the main text) for comparison.

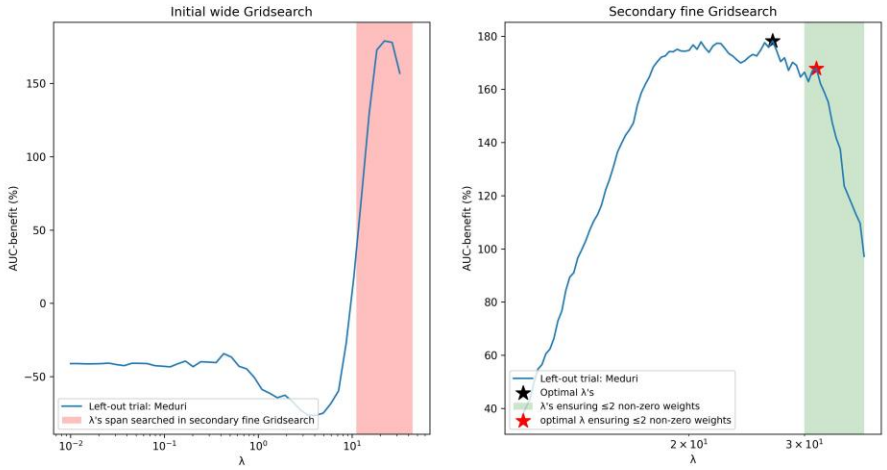
Results

Figure S20 shows the results of the λ optimization through the wide and fine grid search for the non-linear Tian method. Model training using this optimized λ resulted in a model with four non-zero weights: for the interaction terms with CRP, dichotomized glucose (ie, glucose > 7 mmol/L), creatinine and sex (Table S32). As models that require fewer variables are preferred in clinical practise, we repeated the λ optimization with an extra constraint. Namely, instead of selecting the optimal λ, we selected the optimal λ among λs which resulted in a final model with *maximally two* non-zero weights (Figure S20). Model training using this λ resulted in a model with two non-zero weights for the interactions with CRP and dichotomized glucose (ie, glucose > 7 mmol/L; Table S33).

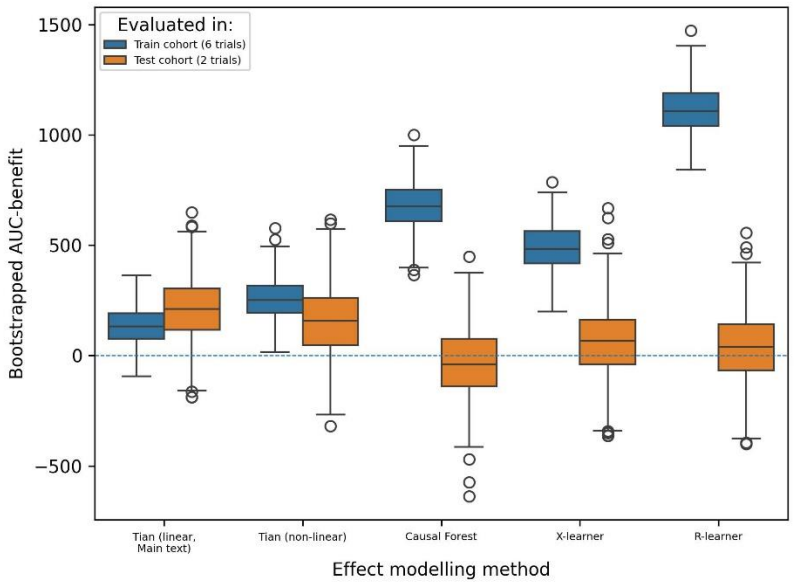
Figure S21 shows the bootstrapped AUC-benefit results of each method, compared to the linear Tian method. None of the methods showed an advantage over the linear Tian methods, and for the more flexible machine learning methods, the median AUC-benefit were even close to zero (not outperforming randomly generated ITEs). Notably, each of the non-linear methods yielded lower AUC-benefits in the external validation as compared to the apparent validation, and this difference was more extreme for the more flexible machine learning methods, suggesting that

these methods are more prone for overfitting than the linear Tian method. Figure S22 shows the calibration for benefit results for each of the non-linear method.

Appendix Figure S20: Results of the initial wide and second fine grid search for λ optimization in the procedure with the introduction of extra dichotomized variables (black star) and the same procedure with the constraint to select the optimal λ that results in maximally two non-zero weights (red star).



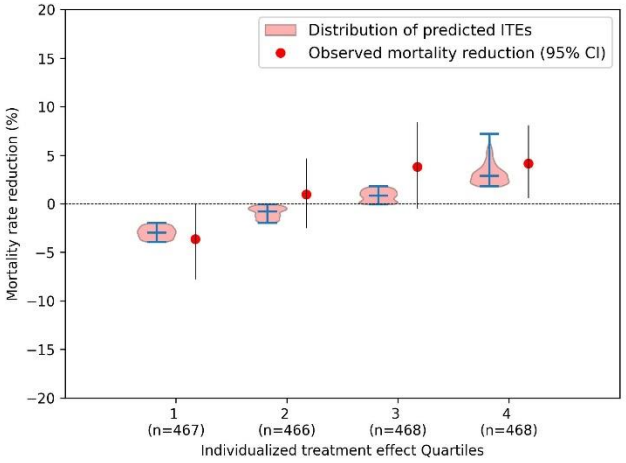
Appendix Figure S21: Discrimination for benefit performance of the different effect modelling methods in the train cohort (ie, 'apparent validation') and in the test cohort (ie, external validation). The AUC-benefits resulting from 500 bootstrap samples are plotted using boxplots.



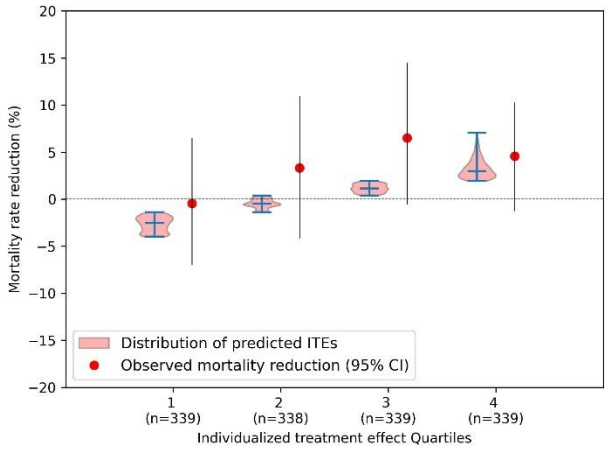
Appendix Figure S22: Calibration for benefit performance of the different effect modelling methods in the train cohort (ie, 'apparent validation' and in the test cohort (ie, external validation).

(a) Non-linear Tian

(i) Evaluated in train cohort (apparent validation)

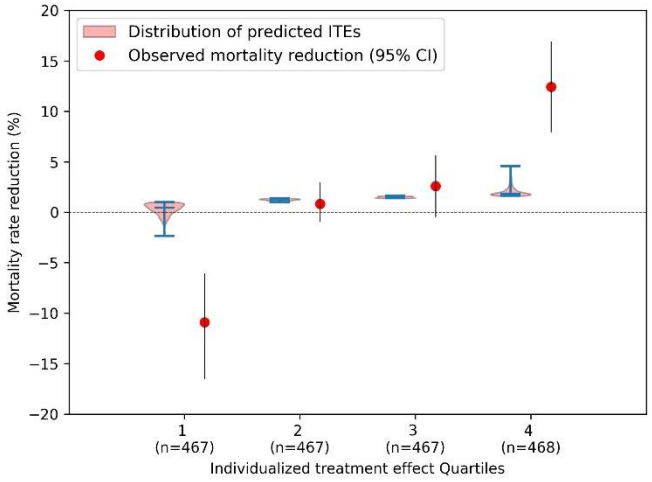


(ii) Evaluated in test cohort (external validation)

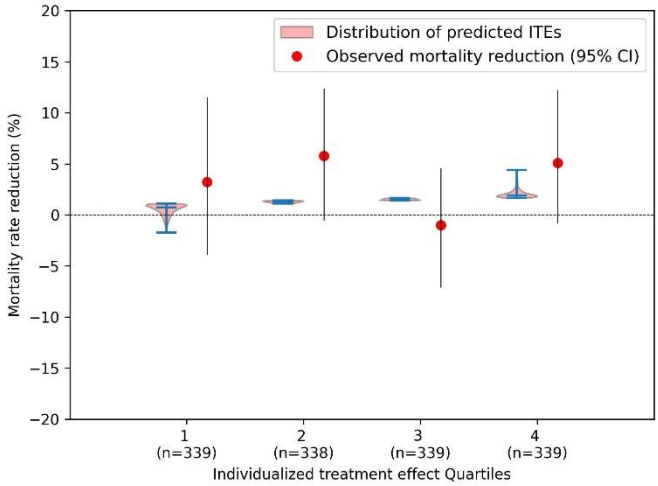


(b) Causal forest

(i) Evaluated in train cohort (apparent validation)



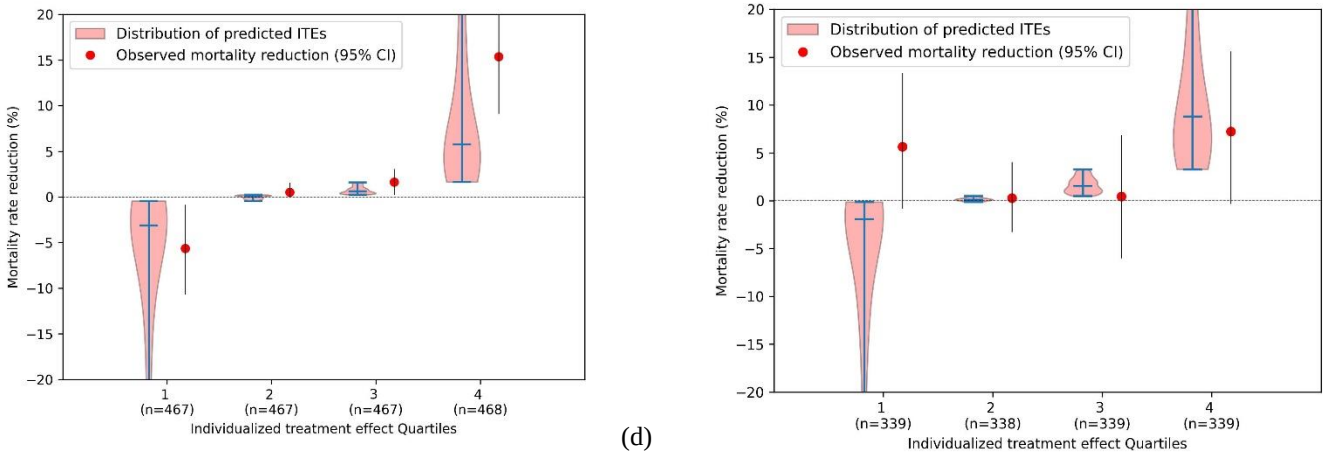
(ii) Evaluated in test cohort (external validation)



(c) X-learner

(i) Evaluated in train cohort (apparent validation)

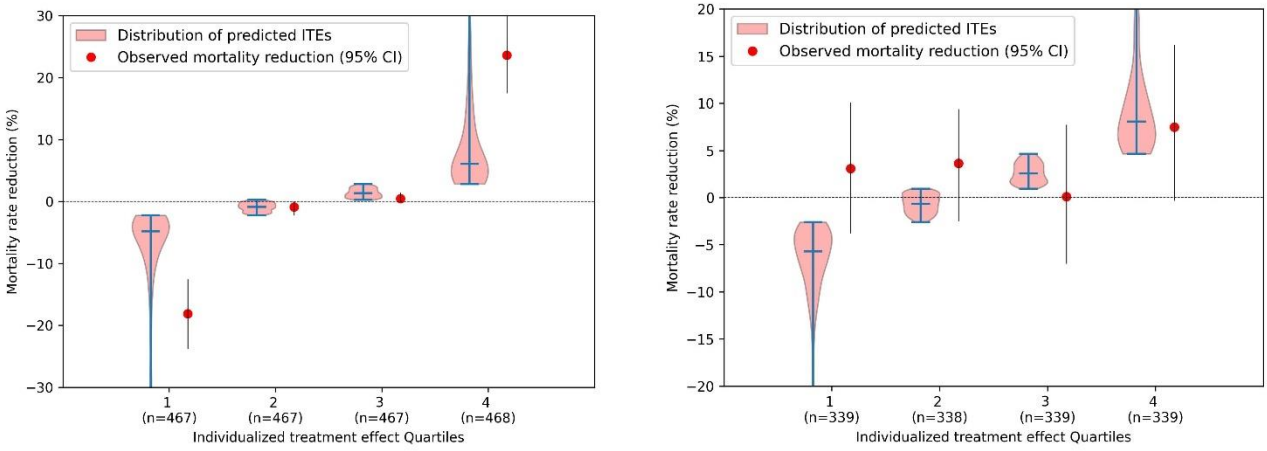
(ii) Evaluated in test cohort (external validation)



(d) R-learner

(i) Evaluated in train cohort (apparent validation)

(ii) Evaluated in test cohort (external validation)



Appendix Table S31: Hyperparameter grids searched for the different non-linear effect modelling methods.

Effect modelling technique	**Hyperparameter**	**Grid searched**
Causal Forest	**Criterion** - The function to measure the quality of a split	[‘mse’, ‘het’]
	max_depth – The maximum depth of the tree. If ‘None’, then nodes are expanded until all leaves are pure or until all leaves contain less than min_samples_split samples.	[1, 2, 3, 5, 7, None]
	Min_samples_split – The minimum number of samples required to split an internal node	[5, 10, 20]
	min_samples_leaf – The minimum number of samples required to be at a leaf node.	[2, 5, 10]
	max_features The number of features to consider when looking for the best split:	[1, 2, 3, ‘auto’, ‘sqrt’, None]
XGBoost (used for implementation of the R- and X-learner)	**max_depth** – Maximum tree depth for base learners.	[3, 5, 7, 11, 15]
	grow_policy – Tree growing policy. 0: favor splitting at nodes closest to the node, i.e. grow depth-wise. 1: favor splitting at nodes with highest loss change.	['depthwise', 'lossguide']
	learning_rate – Boosting learning rate (xgb’s “eta”)	[0.1, 0.01, 0.001]
	subsample – Subsample ratio of the training instances.	[0.5, 0.7, 1]

Appendix Table S32: Values of non-zero weights of the model resulting from the procedure with the introduction of extra dichotomized variables. CRP=C-reactive protein, T=treatment variable.

Variable	weight
*Sex*T*	0.00633
*Creatinine*T*	-0.00072
*CRP*T*	-0.03922
*Glucose > 7 mmol/L *T*	-0.04154

Appendix Table S33: Values of non-zero weights of the model resulting from the procedure with the introduction of extra dichotomized variables and the constraint to select the optimal λ that results in maximally two non-zero weights.

Variable	Weight
*CRP*T*	-0.03099
*Glucose > 7 mmol/L *T*	-0.03255

Appendix Part 10: Sensitivity analyses

Methods:

Pre-specified sensitivity analyses

Two additional prespecified sensitivity analyses were conducted. First, to explore whether the variables, for which the interactions were selected by the LASSO operator in the training of the corticosteroid-effect model and the non-linear effect model using the Tian-method (described in Appendix Part 9), act as individual relative effect modifiers, we tested for HTE using the same one-stage approach through mixed-effects modelling, but now, rather than the subgroup variables, adding the selected variables as continuous variables in turn to the models as a fixed effect and as an interaction term with the treatment variable (see Table S3 in Appendix part 1). Here, we only included patients with non-missing values for the tested variables. Second, we validated our model on two other trials(35,36) that we considered ineligible. We considered the trial by Fernandez-Serrano et al.(35) ineligible due to a high corticosteroid dose and the trial by Lloyd et al.(36) due to randomization of multiple treatments, including corticosteroids. Despite their ineligibility, we examined their potential impact on the results by validating the corticosteroid-effect model with these trials forming the test cohort.

Post-hoc sensitivity analyses

Twelve more sensitivity analyses were added after receiving the IPD of the test cohort.

First, 20 patients (1.5%) in the test cohort had missing data regarding mortality and were therefore excluded in the primary analysis. If missingness of the outcome data is related to prognostic factors at baseline as well as treatment group, their exclusion will create a baseline imbalance in prognosis leading to biased effect estimates.(37) Therefore, we examined the relationship of missing outcome with treatment group and baseline prognostic factors. Second, the results as shown in the main text used the following one-stage approach(8) to test the interaction between the subgroups identified by the corticosteroid-effect model and the treatment:

$$Logit\left[P(Y_i = 1)\right] = \alpha_j + \beta_1[t_{ij}] + \beta_2[x_{ij}] + \underbrace{\gamma_a[x_{ij}t_{ij}]}_{amalgated}$$

$$\alpha_j \sim N(\alpha, \tau^2)$$

in which a_j is the random intercept of the j^{th} trial (following a $N(0, \tau^2)$ distribution, where τ^2 is the residual between-trial heterogeneity), t_{ij} the treatment (0=placebo, 1=corticosteroids), x_{ij} the subgroup variable (ie, 0=predicted no benefit, 1=predicted benefit) and $x_{ij}t_{ij}$ the interaction term between the treatment and subgroup. This approach may allow between-trial information to contribute toward the summary interaction estimate, in combination with within-trial information, which may lead to *aggregation bias* (also known as ecological bias).(8,9) In other words, the estimated interaction (γ_a) is an amalgamation of within-trial and between-trial information.

Therefore, as proposed by Riley and colleagues,(8) we disentangled within-trial and between-study information in the one-stage model by centring the subgroup covariate about the trial-specific means ($\bar{x}_j$), and including the trial-specific mean as an additional adjustment term (ie, $\bar{x}_j t_{ij}$) to explain between-trial heterogeneity (see Table S3 in appendix part 1 for the R implementation):

$$Logit\left[P(Y_i = 1)\right] = \alpha_j + \beta_1[t_{ij}] + \beta_2[x_{ij}] + \underbrace{\gamma_b[\bar{x}_j t_{ij}]}_{between} + \underbrace{\gamma_w[(x_{ij} - \bar{x}_j)t_{ij}]}_{within}$$

$$\alpha_j \sim N(\alpha, \tau^2)$$

where γ_b is the additional term to explain between-study heterogeneity in the overall treatment effect, and γ_w the within-trial interaction estimate (ie, the parameter of interest).

We performed the disentangling of the interaction estimate both for the interaction between the subgroups identified by the corticosteroid-effect model (ie, CRP≤204 vs CRP>204 mg/L), both in the test cohort (ie, the two most recent trials(15,16)) and in the full cohort (ie, all eight included trials), and reported the estimates and corresponding P values of the different interaction estimates (γ_a, γ_w) and the between-trial heterogeneity term (γ_b). Additionally, to examine within-trial interactions, we also calculated the γ_w, and visualized the relative and absolute treatment effects for the subgroups in each individual trial.

Third, to examine the effect of imputation as part of our primary analysis, we repeated the validation of the corticosteroid-effect model only in patients with non-missing values for the prognostic factors used for adjustment as well as for the variables of which the interactions were selected in each of the models (ie, complete case analysis).

Fourth, we examined the robustness of the analysis for the used imputation method, repeating the external validation while varying the 'K' parameter of the KNN imputer, which determines the number of neighboring samples used for imputation. We repeated the external validation varying the K parameter between 3 and 20. Additionally, we repeated the external validation using an alternative imputation method, scikit-learn's 'IterativeImputer'. This imputation method (inspired by R's MICE package) imputes each variable with missing values based on the remaining variables with Bayesian ridge regression in an iterated round-robin fashion. For this imputer, we used default settings.

Fifth, we examined the influence of patients with high (ie, > 20%) missingness among baseline characteristics, by repeating the external validation excluding these patients.

Sixth, we categorized patients into individual PSI classes (combining Class I and II as distinguishment required data not obtained in this study), and compared the effect of corticosteroids on 30-day mortality, and other secondary outcomes, between these classes.

Seventh, in addition to the PSI,(5) we conducted risk modelling using another well-established risk stratification score for CAP, ie, the CURB-65 score.(6) We categorized patients into CURB-65 score 0-2 (indicating 'less severe' CAP) and 3-5 (indicating 'severe' CAP), and compared the effect of corticosteroids on 30-day mortality between the resulting subgroups.

Eighth, we analysed HTE of corticosteroids on 30-day mortality between patients who required initial ICU admission or mechanical ventilation and those who did not.

Nineth, we investigated HTE among patient subgroups based on microbiological aetiology differences, comparing the following subgroups: patients with a bacterial infection (n=960), patients with a Streptococcus pneumoniae infection (the most predominant bacterial agent, n=508), patients with a viral infection (potentially in combination with a bacterial infection, n=285), patients with solely a viral infection (n=202), patients with an influenza infection (potentially in combination with a bacterial infection, n=158), patients with solely an influenza infection (n=114).

Tenth, we assessed HTE on 30-day mortality and hospital-acquired infections, comparing different corticosteroid types, doses, and treatment timing. For corticosteroid type, we analysed four subgroups: hydrocortisone, prednisone/prednisolone (grouped together), methylprednisolone, and dexamethasone. For dose, we compared three

subgroups based on cumulative doses by day 7 (as most studies stopped treatment by study day 7), converted to hydrocortisone equivalents: <1,000 mg, 1,000–1,500 mg, and >1,500 mg. For timing, we compared patients treated within 24 or 48 hours of hospital admission to those treated later.

Eleventh, we examined the performance of the (CRP-based) corticosteroid-effect model across different CAP subtypes, comparing patients without an identified pathogen, those with a bacterial pathogen, those with a Streptococcus pneumoniae infection (ie, a subgroup of the bacterial group), and those with a viral pathogen.

Finally, we assessed the overall effect of corticosteroids on hospital and ICU length of stay, excluding patients with 30-day mortality from the analysis, and assessed the overall effect of corticosteroids on hospital readmissions, we assessed the overall effect of corticosteroids on hospital readmission, only considering trials which used the same follow-up period for this outcome.

Results:

Pre-specified sensitivity analyses

We tested the variables for which the interactions were selected by the LASSO operator in the training of the corticosteroid-effect model and the non-linear effect model using the Tian-method (ie, CRP and dichotomized glucose, see Appendix Part 9) as individual effect modifiers, but found no significant HTE by CRP (P=0.051) or by dichotomized (ie, >7 mmol/L) glucose (P=0.17). In the two ineligible trials,(35,36) the corticosteroid-effect model predicted harm in 664 and benefit in 208 patients, although harm was observed in both groups and was more pronounced in the predicted benefit group (Table S34).

Post-hoc sensitivity analyses

First, the patients with missing outcomes were equally divided between the treatment arms (nine in the corticosteroid group and eleven in the placebo group) and the distributions of age, PSI and respiratory rate were similar between patients for who the outcome was missing and those for who the outcome was available (Figure

S23). Therefore, we assume the risk of bias due to missingness in the primary outcome to be small.

Second, after disentangling within-study and across-study information in the test cohort, the within-trial interaction estimate (γ_w) in the test cohort was less strong as compared to the amalgated interaction (γ_a), with P value of 0.088 compared to 0.0026 (Table S35). The test cohort, however, consists of only two trials, complicating the accurate modelling of the between-trial heterogeneity term (γ_b). The large difference between the γ_b estimated in the test cohort and in the full cohort (ie, -2.68 vs -1.73) suggests that γ_b is overestimated in the test cohort, potentially leading to an underestimation of the within-trial interaction (γ_w). In the full cohort, consisting of 8 trials, γ_b can be estimated more accurately. Here, the difference between the within-trial interaction (γ_w) the amalgated interaction (γ_a) was smaller, with estimates of -0.78 vs -0.70, and P values of 0.0054 vs 0.017, respectively. Moreover, except for the trial by Snijders et al.(10), similar interactions were observed in all individual trials (Figures S24-25), suggesting strong within-trial interaction. Hence, although we cannot rule out aggregation bias for the interaction found in the test cohort due to the small number of trials in this cohort, we consider it very unlikely that the interaction between the subgroups identified by the corticosteroid-effect model (ie, CRP $\leq$ 204 vs CRP > 204 mg/L) is mostly driven by between-trial heterogeneity.

Third, CRP was missing in 26% of the patients in the test cohort, respectively. The complete case analysis resulted in similar point estimates but wider confidence intervals compared to the primary analysis due to the smaller sample size (Table S36).

Fourth, the relative effects in the predicted no benefit and predicted benefit groups, as well as the resulting P values for the interaction tests between these subgroup and the effect or corticosteroids on 30-day mortality, were very similar for the repeated analyses varying the 'K' parameter of the KNN imputer (Figure S26a). Also the analysis using the IterativeImputer resulted in similar estimates and P value (Figure S26b).

Fifth, after excluding 23 out of 1,355 (1.7%) patients with high missingness in baseline characteristics from the test cohort (Figure S27), we observed similar effect estimates and P values (Table S37).

Sixth, neither regarding 30-day mortality (Figure S28), nor for any of the secondary outcomes (Tables S38-43), did we observe significant HTE between individual PSI classes.

Seventh, we obtained IPD regarding CURB-65 scores from six trials,(2,10–13,15) totalling 2,315 patients, with 2,112 and 203 patients with CURB-65 scores 0-2 and 3-5, respectively. In these patients, we observed benefit from corticosteroids in patients with CURB-65 scores 0-2 (indicating 'less severe' CAP), reducing mortality from 7% to 3·9% (OR 0·53, 95% CI 0·36 to 0·78), whereas we observed harm from corticosteroids in patients with CURB-65 scores 3-5 (indicating 'severe' CAP), increasing mortality from 13.3% to 17·1% (OR 1·32, 95% CI 0·60 to 2·89), as reflected in a strong interaction (P=0·033; Table S44).

Eighth, we obtained IPD regarding initial ICU admission and initial need for mechanical ventilation from seven (2,10–15) and four (2,12,13,15) trials, respectively. Here, we did not observe significant HTE between the subgroups (Tables S45-46).

Nineth, we observed no significant heterogeneity of treatment effect between subgroups based on microbiological aetiology differences (Table S47). However, we observed point estimates indicating harm from corticosteroids in patients with viral infections and those with influenza infections. These harmful effects were more pronounced when analysed only for patients without an additional bacterial infection, with the tests for HTE showing a trend towards significance.

Tenth, after adjustment for subgroups observed by the corticosteroid-effect model (see Table S3, appendix part 1), the subgroup treated with hydrocortisone showed a significantly greater benefit compared to patients treated with other corticosteroids types (Table S48), but no HTE among subgroup based on dose (Table S49). We obtained data regarding time between hospital admission and start of corticosteroid treatment for only one trial.(15) Here, we observed significantly greater benefit in the subgroup treated within 24 hours compared to patients treated later than 24 hours (P=0·022), and the subgroup treated after 48 hours even showed substantial mortality increase (ie, harm),

also reflected in a strong interaction (P=0·021; Table S50). We did not observe HTE regarding hospital-acquired infections across subgroups based on used corticosteroid type, dose or timing (Tables S51-53).

Eleventh, different patient subgroups by microbiological aetiology showed notable differences in CRP distributions, with higher CRP values for patients with bacterial infections, compared to those without an identified pathogen or a viral infection (Figure S29). The HTE between the subgroups identified by the corticosteroid-effect model was consistent across patient subgroups by microbiological aetiology, except for the viral infection group. In this subgroup, we found a similar (non-significant) harmful effect in both patient groups, ie, those predicted to show no benefit (CRP ≤ 204 mg/L) and those predicted to benefit (CRP > 204 mg/L; Tables S54-57).

Finally, examining the overall effect of corticosteroids on length of hospital and ICU stay, after excluding patients with 30-day mortality, resulted in similar, significant effects (Table S13, appendix part 1). Among the four trials(10–13) from whom we obtained data regarding hospital readmission, the Meijvis et al. trial(12) studied readmissions within 30 days after *hospital discharge*, compared to hospital readmission within 30 days after study enrolment in the other trials(10,11,13) who included this outcome. We assessed the overall effect of corticosteroids on hospital readmissions, excluding the patients from the Meijvis et al. trial(12), and found a similar, significant effect (Tables S12, appendix part 1).

Appendix Table S34: Heterogeneity in treatment effect of adjuvant therapy with corticosteroids among the subgroups identified by the corticosteroid-effect model in the (ineligible) trials **by Lloyd et al.(36) and Fernandez-Serrano et al.(35).** OR= odds ratio, NNT=number of patients needed to treat.

	30-day mortality rate, n (%)		**OR (95% CI)**	**Mortality reduction, % (95% CI)**	**NNT**
	Placebo	**Corticosteroid**			
Overall					
(n=872)	40/278 (14.4)	40/287 (13.9)	0.88 (0.54; 1.45)	0.5% (-4.4 to 5.5)	221
Subgroups by corticosteroid-effect model					
Predicted harm group (n=664)	34/340 (10.0)	33/324 (10.2)	1.02 (0.68 to 1.55)	-0.2% (-4.0 to 3.5)	-540
Predicted benefit group (n=208)	10/103 (9.7)	14/105 (13.3)	1.43 (0.67 to 3.06)	-3.6% (-10.5 to 4.0)	-27

Appendix Table S35: Effect estimates and P value for the amalgated interaction term (γ_a), the additional term to explain between-study heterogeneity in the overall treatment effect (γ_b) and the (disentangled) within-trial interaction term (γ_w), estimated both in the test cohort and in the full (ie, train and test combined) cohort.

		Test cohort (2 trials, n=1,355)	**Full cohort (8 trials, n=3,224)**
γ_a			
	Estimate	-0.81	-0.78
	P value	0.0026	0.0054
γ_b			
	Estimate	-2.68	-1.73
	P value	0.0042	0.067
γ_w			
	Estimate	-0.63	-0.70
	P value	0.088	0.017

Appendix Table S36: Complete case analysis results: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on 30-day mortality in the **test cohort (external validation) and full cohort** by the corticosteroid-effect model, **excluding patients with missing values for C-reactive protein.** OR= odds ratio, NNT=number of patients needed to treat.

	30-day mortality rate, n (%)		**OR (95% CI)**	**Mortality reduction, % (95% CI)**	**NNT**	**P value for interaction**
	Placebo	**Corticosteroid**				
Overall						
Test cohort (n=1,010)	64/515 (12.4)	51/495 (10.3)	0.81 (0.55; 1.19)	2.1% (-1.1 to 5.3)	47	
Full cohort (n=2,858)	116/1,441 (8.0)	91/1,417 (6.4)	0.78 (0.59; 1.04)	1.6% (0.1 to 3.3)	61	
Subgroups by corticosteroid-effect model, in test cohort						P = 0.10
Predicted harm group (n=551)	37/290 (12.8)	35/261 (13.4)	1.06 (0.65; 1.74)	-0.7% (-5.3 to 4.4)	-153	
Predicted benefit group (n=459)	27/225 (12.0)	16/234 (6.8)	0.54 (0.28; 1.03)	5.2% (0.2 to 10.0)	19	
Subgroups by corticosteroid-effect model, in full cohort						P = 0.018
Predicted harm group (n=1,526)	64/789 (8.1)	61/737 (8.3)	1.06 (0.73; 1.54)	-0.2% (-2.6 to 2.2)	-605	
Predicted benefit group (n=1,332)	52/652 (8.0)	30/680 (4.4)	0.52 (0.33; 0.83)	3.6% (1.2 to 5.7)	28	

Appendix Table S37: Heterogeneity in treatment effect of adjuvant therapy with corticosteroids among the subgroups identified in the in the **test cohort (external validation)** by the corticosteroid-effect model, **excluding 23/1,355 (1.7%) of patients with high missingness among baseline characteristics**. OR= odds ratio, NNT=number of patients needed to treat.

	30-day mortality rate, n (%)		**OR (95% CI)**	**Mortality reduction, % (95% CI)**	**NNT**	**P value for interaction**
	Placebo	**Corticosteroid**				
Overall						
(n=1,332)	88/659 (13.4)	65/673 (9.7)	0.69 (0.49; 0.97)	3.7% (0.8 to 6.5)	27	
Subgroups by corticosteroid-effect model						P = 0.015
Predicted harm group (n=712)	49/365 (13.4)	46/347 (13.3)	0.99 (0.64 to 1.51)	0.2% (-3.9 to 4.0)	594	
Predicted benefit group (n=620)	39/294 (13.3)	19/326 (5.8)	0.40 (0.23 to 0.72)	7.4% (3.6 to 11.3)	13	

Appendix Table S38: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **90-day mortality** among **individual PSI classes**(5). Analysis is based on the patients from the four trials (2,11,14,15) from whom we obtained data regarding 90-day mortality. The P value for interaction is calculated using an interaction test between treatment and PSI class categories, encoding the PSI classes ordinally (see Appendix Table S3, page 22). *The minus sign denotes risk increase (ie, harm), rather than reduction (ie, benefit).

	90-day mortality rate, n (%)		**OR (95% CI)**	**90-day mortality rate reduction, % (95% CI)***	**NNT**	**P for interaction**
	Placebo	**Corticosteroid**				
Subgroups by PSI						P = 0.92
Class I-II (n=295)	3/144 (2.1)	2/151 (1.3)	0.63 (0.10 to 3.38)	0.8% (-1.8 to 3.4)	131	
Class III (n=277)	5/150 (3.3)	5/127 (3.9)	1.19 (0.34 to 4.20)	-0.6% (-4.5 to 2.7)	-165	
Class IV (n=644)	31/309 (10.0)	20/335 (6.0)	0.57 (0.32 to 1.02)	4.1% (1.0 to 7.3)	24	
Class V (n=529)	55/267 (20.6)	43/262 (16.4)	0.76 (0.49 to 1.18)	4.2% (-1.3 to 9.4)	23	

Appendix Table S39: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **initiation of invasive mechanical ventilation by day 28 (28-day IMV)** among **individual PSI classes**(5). Analysis is based on the patients from the four trials (2,11,14,15) from whom we obtained data regarding 28-day IMV, who did not require IMV at baseline. The P value for interaction is calculated using an interaction test between treatment and PSI class categories, encoding the PSI classes ordinally (see Appendix Table S3, page 22). *The minus sign denotes risk increase (ie, harm), rather than reduction (ie, benefit).

	28-day IMV rate, n (%)		**OR (95% CI)**	**28-day IMV rate reduction, % (95% CI)***	**NNT***	**P for interaction**
	Placebo	**Corticosteroid**				
Subgroups by PSI						P = 0.64
Class I-II (n=290)	2/142 (1.4)	5/148 (3.4)	2.70 (0.47 to 15.54)	-2.0% (-5.3 to 0.7)	-50	
Class III (n=263)	10/142 (7.0)	9/121 (7.4)	1.06 (0.46 to 1.46)	-0.4% (-5.8 to 5.0)	-252	
Class IV (n=594)	58/291 (19.9)	30/303 (9.9)	0.41 (0.24 to 0.69)	10.0% (5.4 to 15.1)	9	
Class V (n=421)	50/210 (23.8)	38/211 (18.0)	0.70 (0.42 to 1.14)	5.8% (-0.6 to 12.1)	17	

Appendix Table S40: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **initiation of vasopressors by day 28 (28-day vasopressors)** among **individual PSI classes**(5). Analysis is based on the patients from the three trials (11,14,15) from whom we obtained data regarding 28-day vasopressors, who did not require vasopressors at baseline. The P value for interaction is calculated using an interaction test between treatment and PSI class categories, encoding the PSI classes ordinally (see Appendix Table S3, page 22).

	28-day vasopressor rate, n (%)		**OR (95% CI)**	**28-day vasopressor rate reduction, % (95% CI)**	**NNT**	**P for interaction**
	Placebo	**Corticosteroid**				
Subgroups by PSI						P = 0.61
Class I-II (n=276)	1/135 (0.7)	1/141 (0.7)	0.98 (0.06 to 16.82)	0.0% (-1.5 to 1.6)	3172	
Class III (n=262)	11/143 (7.7)	3/119 (2.5)	0.21 (0.05 to 0.84)	5.2% (0.9 to 9.5)	19	
Class IV (n=596)	61/284 (21.5)	37/312 (11.9)	0.45 (0.28 to 0.72)	9.6% (4.5 to 14.9)	10	
Class V (n=491)	81/249 (32.5)	57/242 (23.6)	0.65 (0.43 to 0.97)	9.0% (1.7 to 15.7)	11	

Appendix Table S41: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **hospital readmission** among **individual PSI classes**(5). Analysis is based on the patients from the four trials (10–13) from whom we obtained data regarding hospital readmission. The P value for interaction is calculated using an interaction test between treatment and PSI class categories, encoding the PSI classes ordinally (see Appendix Table S3, page 22). *The minus sign denotes risk increase (ie, harm), rather than reduction (ie, benefit).

	Readmission rate, n (%)		**OR (95% CI)**	**Readmission rate reduction, % (95% CI)***	**NNT***	**P for interaction**
	Placebo	**Corticosteroid**				
Subgroups by PSI						P = 0.099
Class I-II (n=535)	6/260 (2.3)	15/275 (5.5)	2.44 (0.93 to 6.39)	-3.1% (-6.3 to -0.6)	-31	
Class III (n=368)	5/208 (2.4)	15/160 (9.4)	4.20 (1.49 to 11.82)	-7.0% (-11.1 to -2.9)	-14	
Class IV (n=550)	12/260 (4.6)	20/290 (6.9)	1.53 (0.73 to 3.20)	-2.3% (-5.3 to 1.0)	-43	
Class V (n=180)	7/86 (8.1)	7/94 (7.4)	0.91 (0.30 to 2.70)	0.7% (-6.6 to 7.2)	144	

Appendix Table S42: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **median length of hospital stay** among **individual PSI classes**(5). Analysis is based on the patients from six trials (2,10–14) from whom we obtained data regarding length of hospital stay.

	Median length of hospital stay, IQR (days)		**Reduction in median length of hospital stay in days (95% CI)**
	Placebo	**Corticosteroid**	
Subgroups by PSI			
Class I-II (n=565)	5.0 (3.5 ; 7.5)	5.0 (3.0 ; 6.0)	0.0 (0.0 to 1.0)
Class III (n=393)	7.0 (4.5 ; 9.75)	6.0 (4.5 ; 9.0)	1.0 (0.0 to 1.5)
Class IV (n=637)	8.0 (6.0 ; 13.0)	7.0 (5.0 ; 11.0)	1.0 (0.5 to 2.0)
Class V (n=236)	11.0 (8.0 ; 16.0)	8.25 (6.0 ; 15.75)	2.75 (0.0 to 4.5)

Appendix Table S43: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **median length of ICU stay** among **individual PSI classes**(5). Analysis is based on the patients from four trials (2,11,14,15) from whom we obtained data regarding length of ICU stay, who were admitted to the ICU during their hospitalization. *The minus sign denotes length of stay increase (ie, harm), rather than reduction (ie, benefit).

		Median length of ICU stay, IQR (days)		**Reduction in median length of ICU stay in days (95% CI)***
		Placebo	**Corticosteroid**	
Subgroups by PSI				
	Class I-II (n=57)	4.0 (3.0 ; 5.75)	5.0 (3.0 ; 7.0)	-1.0 (-3.0 to 0.0)
	Class III (n=109)	6.0 (3.0 ; 10.75)	5.0 (3.0 ; 8.0)	1.0 (0.0 to 3.0)
	Class IV (n=352)	7.0 (4.0 ; 13.0)	5.0 (3.0 ; 8.0)	2.0 (1.0 to 4.0)
	Class V (n=412)	7.0 (4.0 ; 13.0)	6.0 (3.0 ; 12.0)	1.0 (-1.0 to 2.0)

Appendix Table S44: Risk modelling results: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on 30-day mortality by the **risk groups identified by the CURB-65 score**. Results are based on the six trials(2,10–13,15) for which we obtained IPD regarding CURB-65 scores. OR= odds ratio, NNT=number of patients needed to treat.

	30-day mortality rate, n (%)		**OR (95% CI)**	**Mortality reduction, % (95% CI)**	**NNT**	**P value for interaction**
	Placebo	**Corticosteroid**				
Overall						
(n=2,315)	88/1,166 (7.5)	59/1,149 (5.1)	0.65 (0.46 to 0.91)	2.4% (0.8 to 4.1)	41	
Subgroups by CURB-65						P = 0.033
CURB-65 score 0-2 (n=2,112)	75/1068 (7.0)	41/1044 (3.9)	0.53 (0.36 to 0.78)	3.1% (1.5 to 4.7)	32	
CURB-65 score 3-5 (n=203)	13/98 (13.3)	18/105 (17.1)	1.32 (0.60 to 2.89)	-3.9% (-11.9 to 4.5)	-25	

Appendix Table S45: Risk modelling results: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on 30-day mortality by the **risk groups based on initial ICU admission**. Results are based on the six trials(2,10–15) for which we obtained IPD regarding baseline ICU admission. OR= odds ratio, NNT=number of patients needed to treat.

	30-day mortality rate, n (%)		OR (95% CI)	Mortality reduction, % (95% CI)	NNT	P value for interaction
	Placebo	**Corticosteroid**				
Overall						
(n=2,663)	101/1,330 (7.6)	67/1,333 (5.0)	0.64 (0.46 to 0.88)	2.6% (1.1 to 4.1)	38	
Subgroups by initial ICU admission						P = 0.10
No (n=1,684)	37/844 (4.4)	32/840 (3.8)	0.85 (0.52 to 1.37)	0.6% (-1.0 to 2.2)	174	
Yes (n=979)	64/486 (13.2)	35/493 (7.1)	0.50 (0.33 to 0.78)	6.1% (3.0 to 9.0)	16	

Appendix Table S46: Risk modelling results: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on 30-day mortality by the **risk groups based on initial need for invasive mechanical ventilation (IMV)**. Results are based on the four trials(2,12,13,15) for which we obtained IPD regarding initial need for mechanical ventilation. OR= odds ratio, NNT=number of patients needed to treat.

	30-day mortality rate, n (%)		OR (95% CI)	Mortality reduction, % (95% CI)	NNT	P value for interaction
	Placebo	**Corticosteroid**				
Overall						
(n=1,619)	74/805 (9.2)	46/814 (5.7)	0.59 (0.40 to 0.86)	3.5% (1.3 to 5.5)	28	
Subgroups by initial need for IMV						P = 0.85
No (n=1,437)	51/661 (7.7)	30/661 (4.5)	0.57 (0.37 to 0.88)	3.4% (1.1 to 5.3)	29	
Yes (n=182)	14/88 (15.9)	10/94 (10.6)	0.63 (0.26 to 1.50)	5.3% (-3.5 to 14.0)	18	

Appendix Table S47: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **30-day mortality** for different **identified pathogens**. Analysis is based on the patients from the seven trials (2,10–15) from whom we obtained data regarding aetiology.

Subgroups by aetiology	**30-day mortality rate, n (%)**		**OR (95% CI)**	**Mortality reduction, % (95% CI)**	**NNT**	**P value for interaction**
	Placebo	**Corticosteroid**				
Overall						
(n=2,596)	97/1303 (7.4)	64/1293 (4.9)	0.65 (0.47 to 0.90)	2.5% (0.9 to 4.2)	40	
Identified pathogen						P = 0.87
No (n=1,433)	50/707 (7.1)	35/726 (4.8)	0.67 (0.43 to 1.05)	2.3% (0.2 to 4.5)	44	
Yes (n=1,163)	47/596 (7.9)	29/567 (5.1)	0.64 (0.40 to 1.03)	2.8% (0.5 to 5.2)	36	
Bacterial						P = 0.34
No (n=1,635)	54/808 (6.7)	43/827 (5.2)	0.75 (0.48 to 1.17)	1.5% (-0.6 to 3.8)	67	
Yes (n=960)	43/494 (8.7)	21/466 (4.5)	0.53 (0.31 to 0.92)	4.2% (1.8 to 6.8)	23	
Streptococcus pneumoniae						P = 0.89
No (n=2,087)	85/1056 (8.0)	55/1031 (5.3)	0.65 (0.45 to 0.92)	2.7% (1.0 to 4.7)	36	
Yes (n=508)	12/246 (4.9)	9/262 (3.4)	0.69 (0.29 to 1.68)	1.4% (-1.6 to 4.3)	69	
Viral						P =0.065
No (n=2,217)	89/1110 (8.0)	52/1107 (4.7)	0.57 (0.40 to 0.81)	3.3% (1.6 to 5.0)	30	
Yes (n=285)	6/149 (4.0)	9/136 (6.6)	1.69 (0.58 to 4.88)	-2.6% (-7.1 to 1.5)	-38	
Viral, without bacterial infection						P = 0.078
No (n=2,300)	91/1158 (7.9)	53/1142 (4.6)	0.58 (0.41 to 0.82)	3.2% (1.5 to 4.8)	31	
Yes (n=202)	4/101 (4.0)	8/101 (7.9)	2.02 (0.59 to 6.92)	-4.0% (-9.4 to 1.0)	-25	

Appendix Table S47, continued.

Subgroups by aetiology	**30-day mortality rate, n (%)**		**OR (95% CI)**	**Mortality reduction, % (95% CI)**	**NNT**	**P value for interaction**
	Placebo	**Corticosteroid**				
Influenza						P = 0.10
No (n=2,224)	80/1110 (7.2)	48/1114 (4.3)	0.58 (0.40 to 0.84)	2.9% (1.4 to 4.6)	34	
Yes (n=158)	6/90 (6.7)	7/68 (10.3)	1.61 (0.51 to 5.02)	-3.6% (-11.5 to 4.2)	-27	
Influenza, without bacterial infection						P = 0.17
No (n=2,268)	82/1140 (7.2)	49/1128 (4.3)	0.59 (0.41 to 0.85)	2.8% (1.3 to 4.5)	35	
Yes (n=114)	4/60 (6.7)	6/54 (11.1)	1.75 (0.47 to 6.57)	-4.4% (-13.4 to 4.2)	-22	

Appendix Table S48: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **30-day mortality** for **different types of used corticosteroids**. *The minus sign denotes that treatment had net harm, rather than benefit. **Interaction test with an extra adjustment for the subgoups identified by the corticosteroid-effect model (ie, predicted no benefit vs predicted benefit group; see Table S3 in appendix part 1).

Subgroups by corticosteroid type	**30-day mortality rate, n (%)**		**OR (95% CI)**	**Mortality reduction, % (95% CI)**	**NNT ***	**P value for interaction**	**Adjusted P value for interaction****
	Placebo	**Corticosteroid**					
Hydrocortisone						P = 0.0090	P = 0.029
No (n=2,384)	83/1188 (7.0)	79/1196 (6.6)	0.93 (0.68 to 1.29)	0.4% (-1.4 to 1.9)	262		
Yes (n=840)	57/418 (13.6)	27/422 (6.4)	0.43 (0.27 to 0.70)	7.2% (3.8 to 10.7)	13		
Prednisone/ Prednisolone						P = 0.13	P = 0.18
No (n=2,226)	121/1104 (11.0)	85/1122 (7.6)	0.66 (0.49 to 0.88)	3.4% (1.3 to 5.5)	29		
Yes (n=998)	19/502 (3.8)	21/496 (4.2)	1.12 (0.60 to 2.12)	-0.4% (-2.7 to 1.5)	-222		
Methyl-prednisolone						P = 0.24	P = 0.41
No (n=2,543)	92/1271 (7.2)	61/1272 (4.8)	0.64 (0.46 to 0.90)	2.4% (0.8 to 4.0)	40		
Yes (n=681)	48/335 (14.3)	45/346 (13.0)	0.89 (0.58 to 1.38)	1.3% (-3.2 to 5.4)	75		
Dexamethasone						P = 0.78	P = 0.72
No (n=2,519)	124/1255 (9.9)	93/1264 (7.4)	0.71 (0.54 to 0.95)	2.5% (0.5 to 4.3)	39		
Yes (n=705)	16/351 (4.6)	13/354 (3.7)	0.80 (0.38 to 1.69)	0.9% (-1.5 to 3.1)	112		

Appendix Table S49: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **30-day mortality** for **different cumulative doses on study day 7**, transformed into equivalent quantities of hydrocortisone in mg. *The minus sign denotes that treatment had net harm, rather than benefit. **Interaction test with an extra adjustment for the subgroups identified by the corticosteroid-effect model (ie, baseline CRP ≤ 204 mg/L vs > 204 mg/L, see Table S3 in appendix part 1).

Subgroups by cumulative dose on study day 7 (hydrocortisone equivalent)	**30-day mortality rate, n (%)**		**OR (95% CI)**	**30-day mortality rate reduction, % (95% CI)***	**NNT***	**P value for interaction**	**Adjusted P value for interaction****
	Placebo	**Corticosteroid**					
<1,000 mg						P = 0.78	P = 0.72
No (n=2,519)	124/1255 (9.9)	93/1264 (7.4)	0.71 (0.54 to 0.95)	2.5% (0.5 to 4.3)	29		
Yes (n=705)	16/351 (4.6)	13/354 (3.7)	0.80 (0.38 to 1.69)	0.9% (-1.5 to 3.1)	112		
1,000 – 1,500 mg						P = 0.69	P = 0.90
No (n=1,432)	72/709 (10.2)	58/723 (8.0)	0.76 (0.53 to 1.10)	2.1% (-0.5 to 4.7)	46		
Yes (n=1,792)	68/897 (7.6)	48/895 (5.4)	0.69 (0.47 to 1.01)	2.2% (0.1 to 4.0)	45		
> 1,500 mg						P = 0.84	P = 0.91
No (n=2,497)	84/1248 (6.7)	61/1249 (4.9)	0.71 (0.50 to 0.99)	1.8% (0.2 to 3.3)	54		
Yes (n=727)	56/358 (15.6)	45/369 (12.2)	0.75 (0.49 to 1.14)	3.4% (-0.9 to 7.6)	29		

Appendix Table S50: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **30-day mortality** for **different times between hospital admission and start of treatment**. Results are based on the trial by Dequin et al.,(15) the only trial for which we obtained IPD regarding time between hospital admission and initiation of treatment. OR= odds ratio, NNT=number of patients needed to treat. *The minus sign denotes that treatment had net harm, rather than benefit.

Subgroups by timing between hospital admission and initiation of corticosteroid treatment	**30-day mortality rate, n (%)**		**OR (95% CI)**	**30-day mortality rate reduction, % (95% CI)***	**NNT**	**P value for interaction**
	Placebo	**Corticosteroid**				
						P = 0.022
< 24 hours *(n=452)*	34/230 (14.8)	11/222 (5.0)	0.30 (0.16 to 0.56)	9.8% (4.9 to 14.5)	10	
≥ 24 hours *(n=342)*	15/165 (9.1)	16/177 (9.0)	0.99 (0.52 to 2.01)	0.1% (-5.3 to 4.9)	1947	
						P = 0.021
< 48 hours *(n=705)*	46/352 (13.1)	20/353 (5.7)	0.4 (0.24 to 0.65)	7.4% (3.5 to 11.2)	13	
≥ 48 hours *(n=89)*	3/43 (7.0)	7/46 (15.2)	2.39 (0.71 to 12.26)	-8.2% (-19.5 to 2.4)	-12	

Appendix Table S51: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **hospital-acquired infections** for different **types of used corticosteroids**. Analysis is based on the patients from the seven trials (2,10–15) from whom we obtained data regarding hospital-acquired infections. *The minus sign denotes that treatment had net harm, rather than benefit.

Subgroups by corticosteroid type	**Hospital-acquired infection rate, n (%)**		**OR (95% CI)**	**Hospital-acquired infection rate reduction, % (95% CI)***	**NNT**	**P value for interaction**
	Placebo	**Corticosteroid**				
Hydrocortisone						P = 0.92
No (n=1,810)	128/902 (14.2)	119/908 (13.1)	0.86 (0.52 to 1.41)	1.1% (-1.2 to 3.8)	92	
Yes (n=840)	44/418 (10.5)	40/422 (9.5)	0.89 (0.57 to 1.40)	1.0% (-2.4 to 4.4)	95	
Prednisone/ Prednisolone						P = 0.73
No (n=1,665)	54/828 (6.5)	50/837 (6.0)	0.91 (0.61 to 1.36)	0.5% (-1.4 to 2.6)	182	
Yes (n=985)	118/492 (24.0)	109/493 (22.1)	0.80 (0.43 to 1.47)	1.9% (-1.8 to 6.5)	53	
Methyl-prednisolone						-
No (n=2,530)	172/1261 (13.6)	157/1269 (12.4)	0.85 (0.61 to 1.19)	1.3% (-0.6 to 3.5)	78	
Yes (n=120)	0/59 (0.0)	2/61 (3.3)	-	-3.3% (-6.9 to 0.0)	-30	
Dexamethasone						P = 0.82
No (n=1,945)	162/969 (16.7)	151/976 (15.5)	0.89 (0.62 to 1.27)	1.2% (-1.3 to 3.8)	80	
Yes (n=705)	10/351 (2.8)	8/354 (2.3)	0.79 (0.31 to 2.02)	0.6% (-1.4 to 2.6)	169	

Appendix Table S52: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **hospital-acquired infections** for **different cumulative doses on study day 7**, transformed into equivalent quantities of hydrocortisone in mg. Analysis is based on the patients from the seven trials (2,10–15) from whom we obtained data regarding hospital-acquired infections. *The minus sign denotes that treatment had net harm, rather than benefit.

Subgroups by cumulative dose on study day 7 (hydrocortisone equivalent)	**Hospital-acquired infection rate, n (%)**		**OR (95% CI)**	**Hospital-acquired infection rate reduction, % (95% CI)***	**NNT***	**P value for interaction**
	Placebo	**Corticosteroid**				
<1,000 mg						P = 0.82
No (n=1,945)	162/969 (16.7)	151/976 (15.5)	0.89 (0.62 to 1.27)	1.2% (-1.3 to 3.8)	80	
Yes (n=705)	10/351 (2.8)	8/354 (2.3)	0.79 (0.31 to 2.02)	0.6% (-1.4 to 2.6)	169	
1,000 – 1,500 mg						P = 0.57
No (n=871)	10/433 (2.3)	11/438 (2.5)	1.10 (0.46 to 2.60)	-0.2% (-1.8 to 1.5)	-495	
Yes (n=1,779)	162/887 (18.3)	148/892 (16.6)	0.84 (0.58 to 1.21)	1.7% (-1.0 to 4.6)	59	
> 1,500 mg						-
No (n=2,484)	172/1238 (13.9)	156/1246 (12.5)	0.83 (0.60 to 1.17)	1.4% (-0.6 to 3.6)	72	
Yes (n=166)	0/82 (0.0)	3/84 (3.6)	-	-3.6% (-6.9 to 0.0)	-28	

Appendix Table S53: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on on **hospital-acquired infections** for **different times between hospital admission and start of treatment**. Results are based on the trial by Dequin et al.,(15) the only trial for which we obtained IPD regarding time between hospital admission and initiation of treatment. OR= odds ratio, NNT=number of patients needed to treat. *The minus sign denotes that treatment had net harm, rather than benefit.

Subgroups by timing between hospital admission and initiation of corticosteroid treatment	**Hospital-acquired infection rate, n (%)**		**OR (95% CI)**	**Hospital-acquired infection rate reduction, % (95% CI)***	**NNT ***	**P value for interaction**
	Placebo	**Corticosteroid**				
						P = 0.25
< 24 hours *(n=452)*	31/230 (13.5)	22/222 (9.9)	0.71 (0.42 to 1.18)	3.6% (-1.7 to 8.8)	28	
≥ 24 hours *(n=342)*	13/165 (7.9)	17/177 (9.6)	1.24 (0.68 to 2.49)	-1.7% (-6.7 to 2.9)	-57	
						P = 0.92
< 48 hours *(n=705)*	4/43 (9.3)	4/46 (8.7)	0.4 (0.24 to 0.65)	1.4% (-2.3 to 5.3)	69	
≥ 48 hours *(n=89)*	40/352 (11.4)	35/353 (9.9)	0.93 (0.2 to 5.4)	0.6% (-10.2 to 11.2)	164	

Appendix Table S54: Heterogeneity in treatment effect of adjuvant therapy with corticosteroids among the subgroups identified by the corticosteroid-effect model **in patients without an identified pathogen (n=1,433)**. OR= odds ratio, NNT=number of patients needed to treat.

	30-day mortality rate, n (%)		OR (95% CI)	Mortality reduction, % (95% CI)	NNT	P value for interaction
	Placebo	**Corticosteroid**				
Overall						
(n=1,433)	50/707 (7.1)	35/726 (4.8)	0.67 (0.43; 1.05)	2.3% (0.3 to 4.3)	44	
Subgroups by corticosteroid-effect model						P = 0.095
Predicted harm group (n=828)	28/412 (6.8)	25/416 (6.0)	0.93 (0.53 to 1.64)	0.8% (-2.0 to 3.9)	127	
Predicted benefit group (n=605)	22/295 (7.5)	10/310 (3.2)	0.41 (0.19 to 0.88)	4.2% (1.3 to 7.2)	23	

Appendix Table S55: Heterogeneity in treatment effect of adjuvant therapy with corticosteroids among the subgroups identified by the corticosteroid-effect model **in patients without a bacterial infection (n=960)**. OR= odds ratio, NNT=number of patients needed to treat.

	30-day mortality rate, n (%)		OR (95% CI)	Mortality reduction, % (95% CI)	NNT	P value for interaction
	Placebo	**Corticosteroid**				
Overall						
(n=960)	43/494 (8.7)	21/466 (4.5)	0.51 (0.30; 0.88)	4.2% (1.6 to 6.9)	23	
Subgroups by corticosteroid-effect model						P = 0.34
Predicted harm group (n=325)	16/181 (8.8)	9/144 (6.2)	0.69 (0.29 to 1.60)	2.6% (-2.1 to 7.1)	38	
Predicted benefit group (n=635)	27/313 (8.6)	12/322 (3.7)	0.42 (0.21 to 0.84)	4.9% (1.6 to 7.7)	20	

Appendix Table S56: Heterogeneity in treatment effect of adjuvant therapy with corticosteroids among the subgroups identified by the corticosteroid-effect model **in patients with a Streptococcus Pneumonia infection (n=508)**. OR= odds ratio, NNT=number of patients needed to treat.

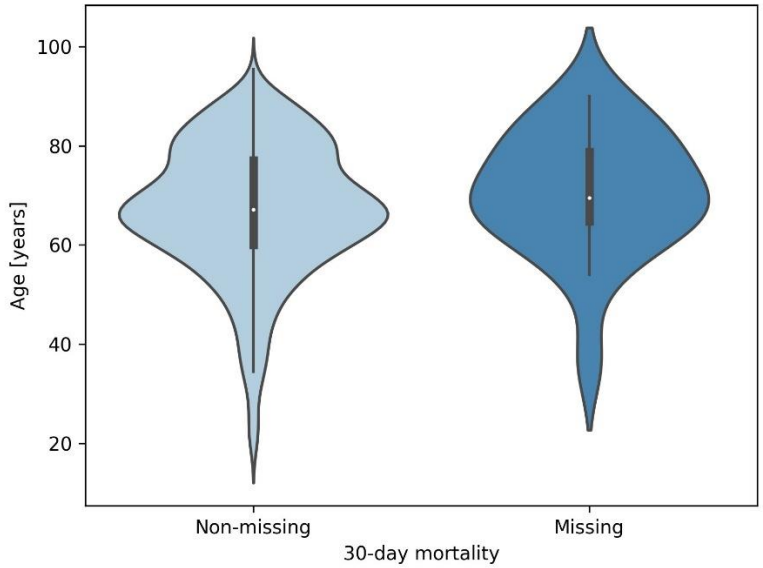
	30-day mortality rate, n (%)		OR (95% CI)	Mortality reduction, % (95% CI)	NNT	P value for interaction
	Placebo	**Corticosteroid**				
Overall						
(n=508)	12/246 (4.9)	9/262 (3.4)	0.69 (0.29; 1.68)	1.4% (-1.6 to 4.4)	69	
Subgroups by corticosteroid-effect model						P = 0.12
Predicted harm group (n=153)	4/87 (4.6)	5/66 (7.6)	1.70 (0.44 to 6.60)	-3.0% (-9.7 to 3.6)	-33	
Predicted benefit group (n=355)	8/159 (5.0)	4/196 (2.0)	0.39 (0.12 to 1.33)	3.0% (-0.2 to 6.2)	33	

Appendix Table S57: Heterogeneity in treatment effect of adjuvant therapy with corticosteroids among the subgroups identified by the corticosteroid-effect model **in patients without a viral infection (n=285)**. OR= odds ratio, NNT=number of patients needed to treat. *Minus sign denotes harm.

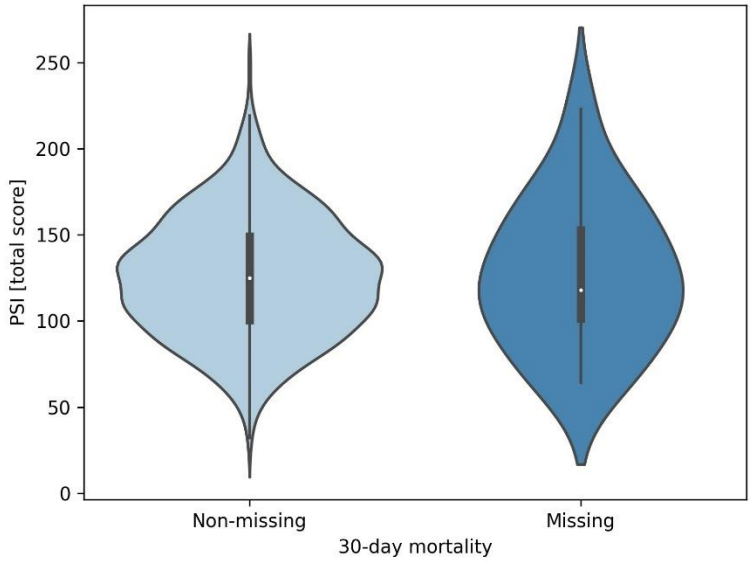
	30-day mortality rate, n (%)		OR (95% CI)	Mortality reduction, % (95% CI)*	NNT*	P value for interaction
	Placebo	**Corticosteroid**				
Overall						
(n=285)	6/149 (4.0)	9/136 (6.6)	1.69 (0.58; 4.88)	-2.6% (-7.1 to 1.5)	-38	
Subgroups by corticosteroid-effect model						P = 0.99
Predicted harm group (n=159)	4/83 (4.8)	6/76 (7.9)	1.69 (0.46 to 6.25)	-3.1% (-9.4 to 3.1)	-32	
Predicted benefit group (n=126)	2/66 (3.0)	3/60 (5.0)	1.68 (0.27 to 10.44)	-2.0% (-7.8 to 3.5)	-50	

Appendix Figure S23: Distribution plots of three important prognostic factors (age, pneumonia severity index [PSI] and respiratory rate) in patients where outcome was missing and not missing.

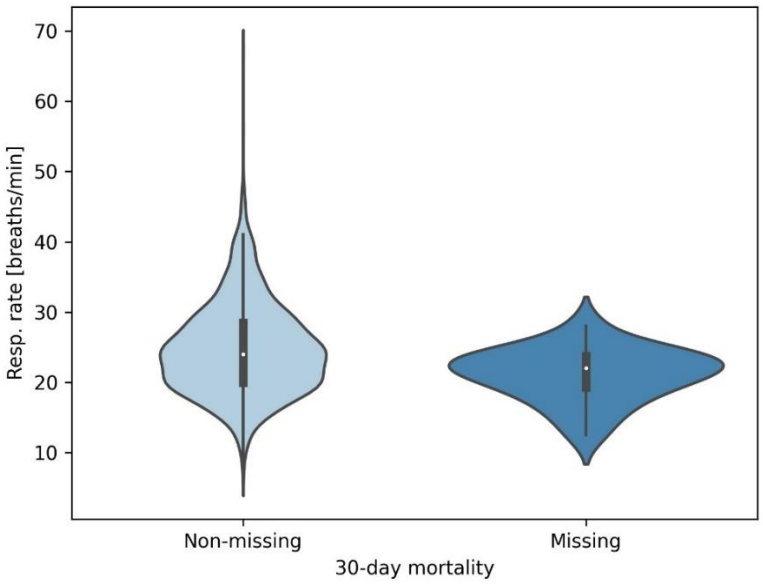
(a) Age



(b) Pneumonia severity index (PSI)

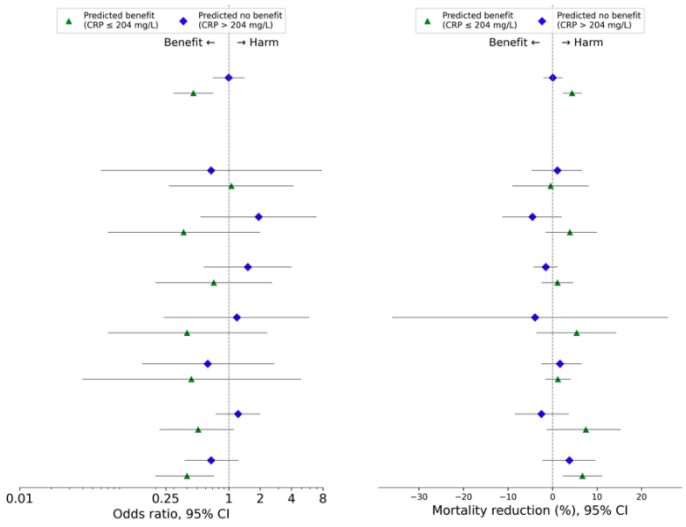


(c) Respiratory rate



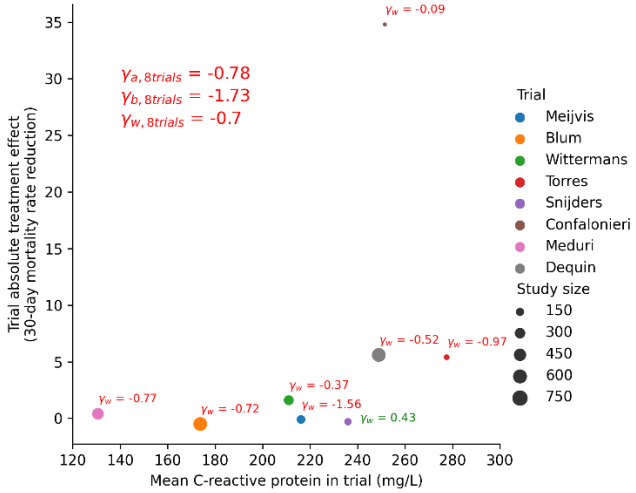
Appendix Figure S24: Heterogeneity of Treatment Effect (HTE) on the relative (odds ratio) and absolute (mortality reduction) scale for the subgroups identified by the corticosteroid-effect model in **each individual trial**. OR=odds ratio. *minus signs denote net harm (ie, mortality increase).

	Predicted benefit, n (%)	Overall OR (95% CI)	OR Predicted no benefit (95% CI)	OR Predicted benefit (95% CI)	Overall mortality reduction*, % (95% CI)	Mort. reduction* Predicted no benefit (95% CI)	Mort. reduction* Predicted benefit (95% CI)
Overall, full cohort	1,515/3,224 (47)	0.72 (0.56 to 0.94)	1.00 (0.71 to 1.41)	0.46 (0.30 to 0.71)	2.2 (0.6 to 3.7)	0.1 (-2.0 to 2.2)	4.4 (2.4 to 6.5)
Confalonieri, (2005)	25/46 (54.3)	-	-	-	34.8 (20 to 52.4)	33.3 (12.5 to 57.1)	36.4 (11.1 to 61.6)
Snijders, (2010)	118/213 (55.4)	1.05 (0.33 to 3.37)	0.68 (0.06 to 7.76)	1.06 (0.27 to 4.16)	-0.3 (-5.8 to 4.7)	1.1 (-4.6 to 6.6)	-0.4 (-9.0 to 8.1)
Meijvis, (2011)	158/304 (52.0)	1.01 (0.39 to 2.63)	1.94 (0.54 to 6.94)	0.37 (0.07 to 1.99)	-0.1 (-4.3 to 4.3)	-4.5 (-11.2 to 2.0)	3.9 (-1.5 to 9.9)
Blum, (2015)	288/785 (36.7)	1.16 (0.55 to 2.48)	1.53 (0.58 to 4.00)	0.72 (0.20 to 2.59)	-0.5 (-2.7 to 1.4)	-1.5 (-4.2 to 1.1)	1.1 (-2.4 to 4.6)
Torres, (2015)	91/120 (75.8)	0.61 (0.20 to 1.82)	1.20 (0.24 to 5.89)	0.40 (0.07 to 2.33)	5.4 (-4.7 to 15.3)	-3.9 (-36.0 to 25.9)	5.4 (-3.5 to 14.3)
Wittermans, (2021)	204/401 (50.9)	0.55 (0.16 to 1.90)	0.63 (0.15 to 2.72)	0.44 (0.04 to 4.92)	1.6 (-0.9 to 4.6)	1.7 (-2.5 to 6.5)	1.2 (-1.6 to 4.0)
Meduri, (2022)	165/561 (29.4)	0.96 (0.65 to 1.43)	1.23 (0.76 to 2.00)	0.51 (0.22 to 1.12)	0.4 (-4.3 to 5.0)	-2.5 (-8.4 to 3.6)	7.5 (-1.2 to 15.3)
Dequin, (2023)	465/794 (58.6)	0.51 (0.32 to 0.81)	0.68 (0.38 to 1.24)	0.40 (0.20 to 0.72)	5.6 (1.9 to 9.3)	3.8 (-2.2 to 9.6)	6.7 (2.4 to 11.1)

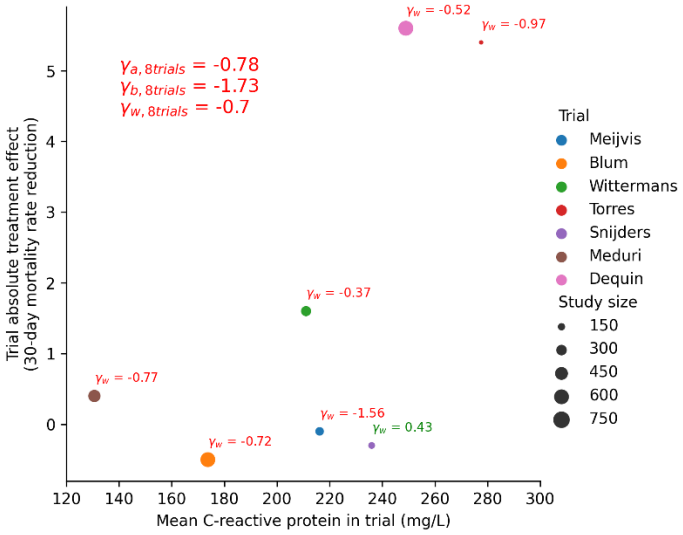


Appendix Figure S25: Scatter plot showing the within-trial interaction terms in each individual trial, and its corresponding overall absolute treatment effect and mean C-reactive proteins. We also added the overall amalgated (γ_a) and within-trial (γ_w) interaction terms, and the between-trial heterogeneity term (γ_b). Each dot represents one trial, and the dot size is proportional to the size of the trial. As the positive, absolute treatment effect of the trial by Confalonieri et al.(14) was much larger than for the other trials (due to its small sample size), we also plotted the same figure without this trial to visualize the remaining 7 trials in more details (figure b). We also plotted the same figure only for the 2 trials which made up the test cohort (figure c).

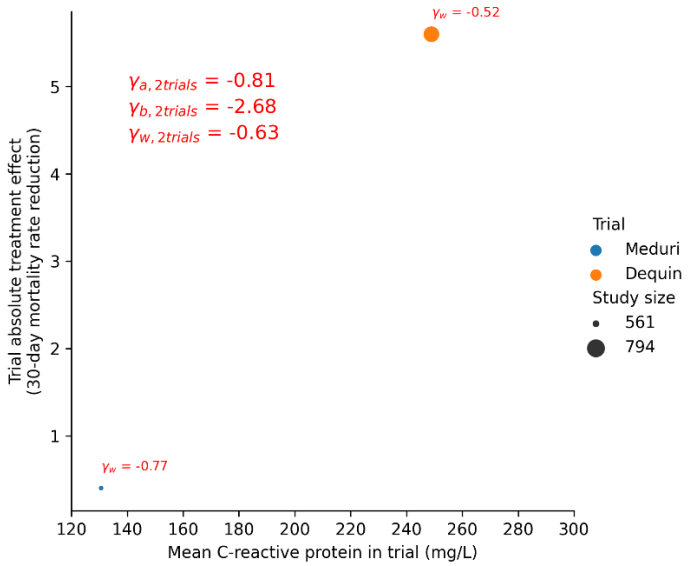
(a) Full cohort (ie, all eight included trials)



(b) Full cohort (ie, all eight included trials), except Confalonieri et al.(14)

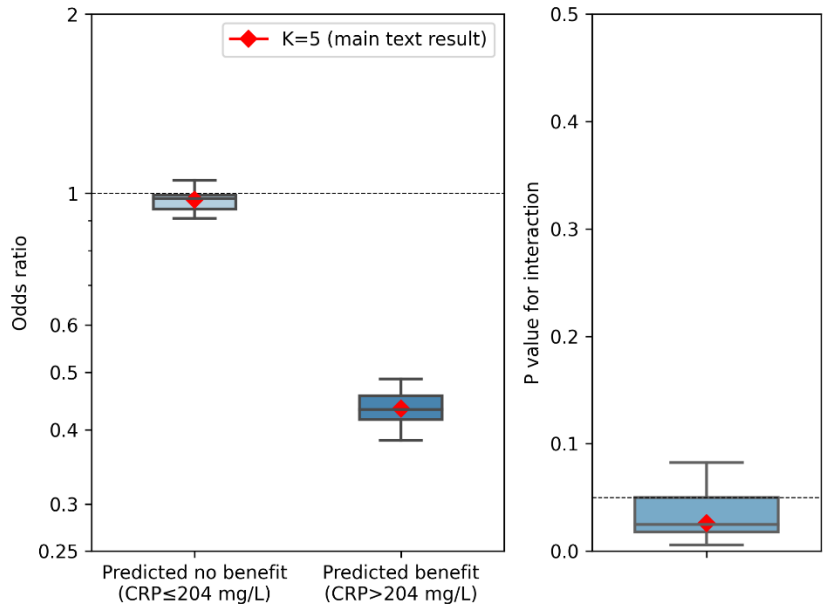


(c) Test cohort (ie, two trials)

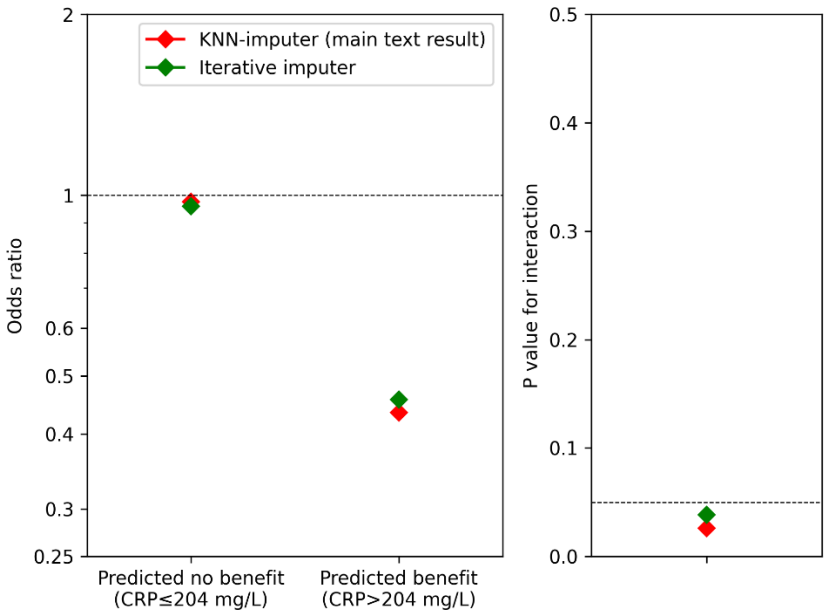


Appendix Figure S26: Relative effects in the patient subgroups identified by the corticosteroid-effect model in the test cohort, as well as the resulting P values for the interaction tests between these subgroup and the effect or corticosteroids on 30-day mortality, for (a) repeated external validations varying the 'K' parameter of the KNN imputer and for (b) the external validation using the IterativeImputer.

(a)



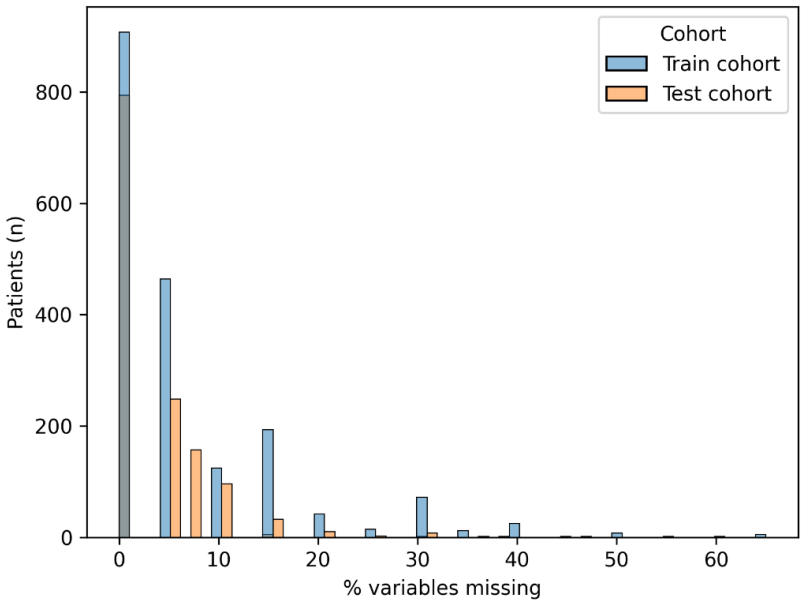
(b)



1243

1244

1245 Appendix Figure S27: Histogram showing the distributions of the percentages of missingness among baseline
1246 characteristics required for external validation, stratified for patients in the train and test cohorts.

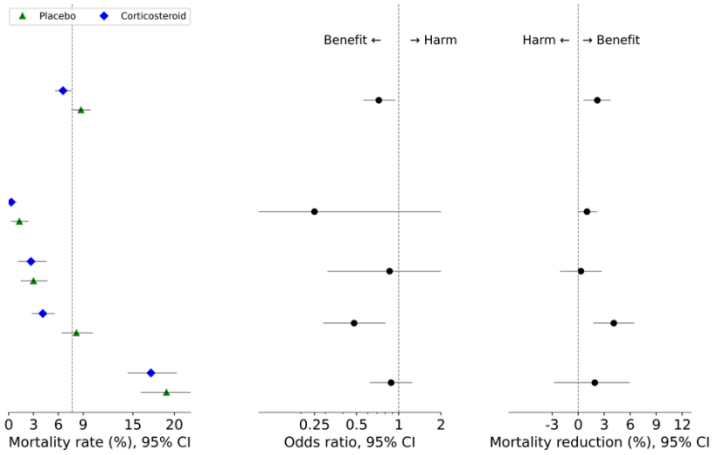


1247

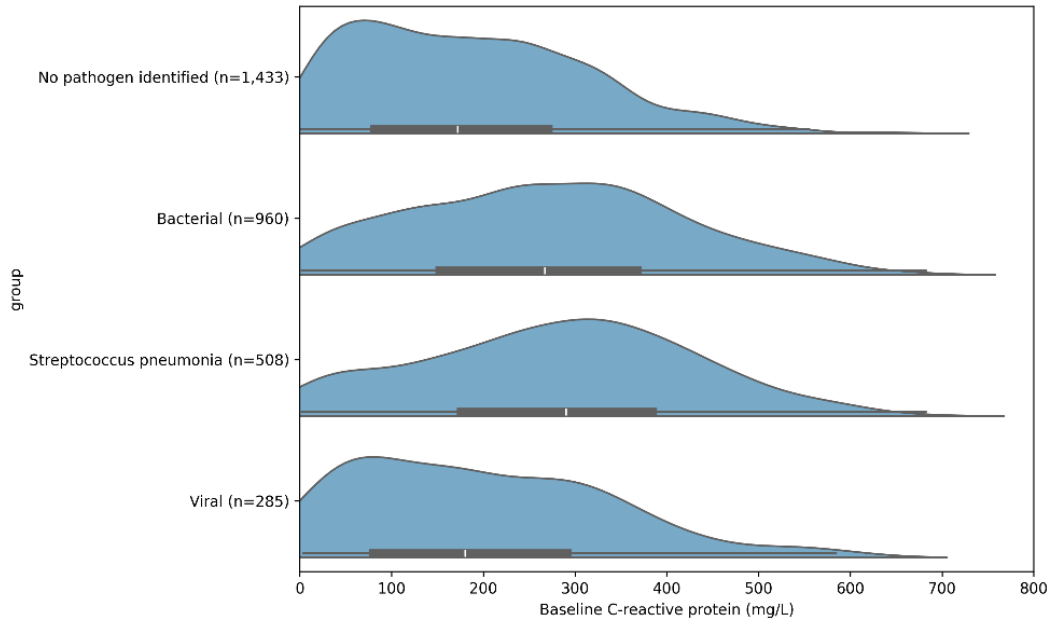
1248

Appendix Figure S28: Heterogeneity of treatment effect of adjuvant therapy with corticosteroids on **30-day mortality** among **individual PSI classes**(5).

	Placebo, Mortality rate, n (%)	Corticosteroid, Mortality rate, n (%)	OR (95% CI)	Mortality reduction, % (95% CI)	NNT	P value for interaction
Full cohort (8 trials, n=3,224)	140/1,606 (8.7)	106/1,618 (6.6)	0.72 (0.56 to 0.94)	2.2% (0.6 to 3.7)	46	
Full cohort subgroups by PSI						0.11
PSI Class I-II, (n=636)	4/311 (1.3)	1/325 (0.3)	0.25 (0.03 to 2.37)	1.0% (0 to 2.1)	102	
PSI Class III, (n=561)	9/300 (3.0)	7/261 (2.7)	0.86 (0.31 to 2.34)	0.3% (-2.1 to 2.7)	314	
PSI Class IV, (n=1,182)	47/575 (8.2)	25/607 (4.1)	0.48 (0.29 to 0.8)	4.1% (1.8 to 6.4)	24	
PSI Class V, (n=845)	80/420 (19.0)	73/425 (17.2)	0.88 (0.62 to 1.24)	1.9% (-2.8 to 5.9)	53	



Appendix Figure S29: Baseline C-reactive protein distributions for patient subgroups based on microbiological aetiologies.



Appendix Part 11: Exclusion of patients with implausible C-reactive protein values

Three patients from the trial by Meduri et al. (16) exhibited baseline CRP levels exceeding 1,000 mg/L, specifically 1,460, 2,220, and 24,930 mg/L. Among the remaining patient data collected in this study, encompassing over 3,500 CAP patients from the other seven included trials, two ineligible trials,(35,36) and the observational dataset,(4) the highest observed CRP value was 568 mg/L. Additionally, literature on extreme CRP values (38,39) has not documented values exceeding 839 mg/L. Therefore, we considered the reported CRP values for these three patients implausible and deemed the associated data unreliable. Consequently, these patients were excluded from the analysis.

Appendix Part 12: Derivation of the C-reactive protein threshold

Assuming a decision threshold (ie, a predicted individualized treatment effect above which treating patients is considered worthwhile) of 0, the corticosteroid-effect model simplifies to a decision trees consisting of one CRP threshold in the absolute scale (ie, in terms of mg/L), respectively, because it consists of only one non-zero weight, where the individualized treatment effect equals 0 for one CRP value.

Derivation of CRP threshold:

The corticosteroid-effect model is represented below:

$$Logit\left[P(Y_i = 1|T = t_i, C = c_i)\right] = w_c c_i \underbrace{t_i}_{-1,1}$$

where i indexes the patients, T represents the treatment variable, C the (standardized) CRP value, and w_c the model's weight for the interaction term with CRP (as presented in Table S58).

To find the CRP value that corresponds with an individualized treatment effect of 0, we equate the models under placebo treatment (ie, t_i = -1) and under corticosteroid treatment (ie, t_i = 1):

$$Logit\left[P(Y_i = 1|T = -1, C = c_i)\right] = Logit\left[P(Y_i = 1|T = 1, C = c_i)\right]$$

which yields:

$$-w_c c_i = w_c c_i$$

$$c_i = 0$$

Hence, for the corticosteroid-effect model, an individualized treatment effect of 0 corresponds with a standardized CRP value of 0 (ie, the mean), which is 204.1 mg/L.

Appendix Table S58: Values of non-zero weights of the corticosteroid-effect model. CRP=C-reactive protein, T=treatment variable.

Variable	weight
*CRP*T*	-0.03564

References

1. Kane S. Corticosteroid Conversion Calculator [Internet]. ClinCalc. [cited 2023 Jan 25]. Available from: https://clincalc.com/corticosteroids
2. Torres A, Sibila O, Ferrer M, Polverino E, Menendez R, Mensa J, et al. Effect of corticosteroids on treatment failure among hospitalized patients with severe community-acquired pneumonia and high inflammatory response: a randomized clinical trial. JAMA. 2015 Feb;313(7):677–86.
3. Average sizes of men and women [Internet]. Available from: https://www.worlddata.info/average-bodyheight.php
4. Endeman H, Meijvis SCA, Rijkers GT, Van Velzen-Blad H, Van Moorsel CHM, Grutters JC, et al. Systemic cytokine response in patients with community-acquired pneumonia. Eur Respir J. 2011;37(6):1431–8.
5. Fine MJ, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, et al. A Prediction Rule to Identify Low-Risk Patients with Community-Acquired Pneumonia. N Engl J Med [Internet]. 1997 Jan 23;336(4):243–50. Available from: https://doi.org/10.1056/NEJM199701233360402
6. Lim WS, van der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. Thorax. 2003 May;58(5):377–82.
7. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. J Stat Softw [Internet]. 2015 Oct 7;67(1 SE-Articles):1–48. Available from: https://www.jstatsoft.org/index.php/jss/article/view/v067i01
8. Riley RD, Debray TPA, Fisher D, Hattle M, Marlin N, Hoogland J, et al. Individual participant data meta-analysis to examine interactions between treatment effect and participant-level covariates: Statistical recommendations for conduct and planning. Stat Med. 2020;39(15):2115–37.
9. Hua H, Burke DL, Crowther MJ, Ensor J, Tudur Smith C, Riley RD. One-stage individual participant data meta-analysis models: estimation of treatment-covariate interactions must avoid ecological bias by separating out within-trial and across-trial information. Stat Med. 2017;36(5):772–89.
10. Snijders D, Daniels JMA, De Graaff CS, Van Der Werf TS, Boersma WG. Efficacy of corticosteroids in community-acquired pneumonia: A randomized double-blinded clinical trial. Am J Respir Crit Care Med.

2010;181(9):975–82.

11. Blum CA, Nigro N, Briel M, Schuetz P, Ullmer E, Suter-Widmer I, et al. Adjunct prednisone therapy for patients with community-acquired pneumonia: A multicentre, double-blind, randomised, placebo-controlled trial. Lancet. 2015;385(9977):1511–8.

12. Meijvis SCA, Hardeman H, Remmelts HHF, Heijligenberg R, Rijkers GT, Van Velzen-Blad H, et al. Dexamethasone and length of hospital stay in patients with community-acquired pneumonia: A randomised, double-blind, placebo-controlled trial. Lancet [Internet]. 2011;377(9782):2023–30. Available from: http://dx.doi.org/10.1016/S0140-6736(11)60607-7

13. Wittermans E, Vestjens SMT, Spoorenberg SMC, Blok WL, Grutters JC, Janssen R, et al. Adjunctive treatment with oral dexamethasone in non-ICU patients hospitalised with community-acquired pneumonia: A randomised clinical trial. Eur Respir J. 2021;58(2):1–10.

14. Confalonieri M, Urbino R, Potena A, Piattella M, Parigi P, Puccio G, et al. Hydrocortisone infusion for severe community-acquired pneumonia: A preliminary randomized study. Am J Respir Crit Care Med. 2005;171(3):242–8.

15. Dequin P-F, Meziani F, Quenot J-P, Kamel T, Ricard J-D, Badie J, et al. Hydrocortisone in Severe Community-Acquired Pneumonia. N Engl J Med. 2023 May;388(21):1931–41.

16. Meduri GU, Shih M-C, Bridges L, Martin TJ, El-Solh A, Seam N, et al. Low-dose methylprednisolone treatment in critically ill patients with severe community-acquired pneumonia. Intensive Care Med. 2022 Aug;48(8):1009–23.

17. kruskal [Internet]. Available from: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.kruskal.html

18. Briel M, Spoorenberg SMC, Snijders D, Torres A, Fernandez-Serrano S, Meduri GU, et al. Corticosteroids in Patients Hospitalized with Community-Acquired Pneumonia: Systematic Review and Individual Patient Data Metaanalysis. Clin Infect Dis. 2018;66(3):346–54.

19. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: A revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:1–8.

20. Nafae RM, Ragab MI, Amany FM, Rashed SB. Adjuvant role of corticosteroids in the treatment of community-acquired pneumonia. Egypt J Chest Dis Tuberc [Internet]. 2013;62(3):439–45. Available from:

http://dx.doi.org/10.1016/j.ejcdt.2013.03.009

21. Sabry NA, Omar EE-D. Corticosteroids and ICU Course of Community Acquired Pneumonia in Egyptian Settings. Pharmacol & Pharm. 2011;02(02):73–81.

22. Smit JM, Zee PA Van Der, Stoof SCM, Genderen ME Van, Snijders D, Boersma WG, et al. Predicting individualized treatment effects of corticosteroids in community-acquired-pneumonia: a data-driven analysis of randomized controlled trials. medRxiv [Internet]. 2023 Jan 1;2023.10.03.23296132. Available from: http://medrxiv.org/content/early/2023/10/03/2023.10.03.23296132.abstract

23. Seabold S, Perktold J. Statsmodels: Econometric and statistical modeling with python. In: 9th Python in Science Conference. 2010.

24. Tian L, Alizadeh AA, Gentles AJ, Tibshirani R. A Simple Method for Estimating Interactions between a Treatment and a Large Number of Covariates. J Am Stat Assoc. 2014 Oct;109(508):1517–32.

25. Devriendt F, Van Belle J, Guns T, Verbeke W. Learning to Rank for Uplift Modeling. IEEE Trans Knowl Data Eng. 2020;

26. Athey S, Tibshirani J, Wager S. Generalized random forests. Ann Stat. 2019;47(2):1179–203.

27. Wager S, Athey S. Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. J Am Stat Assoc [Internet]. 2018;113(523):1228–42. Available from: https://doi.org/10.1080/01621459.2017.1319839

28. Künzel SR, Sekhon JS, Bickel PJ, Yu B. Metalearners for estimating heterogeneous treatment effects using machine learning. Proc Natl Acad Sci U S A. 2019;116(10):4156–65.

29. Nie X, Wager S. Quasi-oracle estimation of heterogeneous treatment effects. Biometrika. 2021;108(2):299–319.

30. Keith Battocchi Eleanor Dillon MHGLPOMOVS. {EconML}: {A Python Package for ML-Based Heterogeneous Treatment Effects Estimation}. 2019.

31. Curth A, van der Schaar M. Nonparametric Estimation of Heterogeneous Treatment Effects: From Theory to Learning Algorithms. Proc Mach Learn Res. 2021;130:1810–8.

32. Robinson PM. Root-N-Consistent Semiparametric Regression. Econometrica [Internet]. 1988 May 7;56(4):931–54. Available from: http://www.jstor.org/stable/1912705

33. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM

SIGKDD International Conference on Knowledge Discovery and Data Mining [Internet]. New York, NY, USA: Association for Computing Machinery; 2016. p. 785–794. (KDD '16). Available from: https://doi.org/10.1145/2939672.2939785

34. Steyerberg EW, Harrell FE. Prediction models need appropriate internal, internal-external, and external validation. J Clin Epidemiol [Internet]. 2016;69:245–7. Available from: http://dx.doi.org/10.1016/j.jclinepi.2015.04.005

35. Fernández-Serrano S, Dorca J, Garcia-Vidal C, Fernández-Sabé N, Carratalà J, Fernández-Agüera A, et al. Effect of corticosteroids on the clinical course of community-acquired pneumonia: A randomized controlled trial. Crit Care. 2011;15(2).

36. Lloyd M, Karahalios A, Janus E, Skinner EH, Haines T, De Silva A, et al. Effectiveness of a Bundled Intervention Including Adjunctive Corticosteroids on Outcomes of Hospitalized Patients with Community-Acquired Pneumonia: A Stepped-Wedge Randomized Clinical Trial. JAMA Intern Med. 2019;179(8):1052–60.

37. Groenwold RHH, Moons KGM, Vandenbroucke JP. Randomized trials with missing outcome data: How to analyze and what to report. Cmaj. 2014;186(15):1153–7.

38. Landry A, Docherty P, Ouellette S, Cartier LJ. Causes and outcomes of markedly elevated C-reactive protein levels. Can Fam Physician. 2017;63(6):e316–23.

39. Vanderschueren S, Deeren D, Knockaert DC, Bobbaers H, Bossuyt X, Peetermans W. Extremely elevated C-reactive protein. Eur J Intern Med. 2006;17(6):430–3.