Predicting benefit from adjuvant therapy with corticosteroids in community-acquired pneumonia

a data-driven analysis of randomised trials

Smit, Jim M.; Van Der Zee, Philip A.; Snijders, Dominic; Boersma, Wim G.; Confalonieri, Paola; Salton, Francesco; Gommers, Diederik A.M.P.J.; Reinders, Marcel J.T.; Krijthe, Jesse H.; More Authors

DOI
10.1016/S2213-2600(24)00405-3

Publication date
2025

Document Version
Final published version

Published in
The Lancet Respiratory Medicine

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

Jim M Smit, Philip A Van Der Zee, Sara C M Stoof, Michel E Van Genderen, Dominic Snijders, Wim G Boersma, Paola Confalonieri, Francesco Salton, Marco Confalonieri, Mei-Chiung Shih, Gianfranco U Meduri, Pierre-François Dequin, Amélie Le Gouge, Melanie Lloyd, Harin Karunajeewa, Grzegorz Bartminski, Silvia Fernández-Serrano, Guillermo Suárez-Cuartín, David van Klaveren, Matthias Briel, Christof M Schönenberger, Ewout W Steyerberg, Diederik A M P J Gommers, Hannelore I Bax, Wilem Jan W Bos, Ewoudt M W van de Garde, Esther Wittermans, Jan C Grutters, Claudine A Blum, Mirjam Christ-Crain, Antoni Torres, Ana Motos, Marcel J T Reinders, Jasper Van Bommel, Jesse H Krijthe, Henrik Endeman

Summary

Background Despite several randomised controlled trials (RCTs) on the use of adjuvant treatment with corticosteroids in patients with community-acquired pneumonia (CAP), the effect of this intervention on mortality remains controversial. We aimed to evaluate heterogeneity of treatment effect (HTE) of adjuvant treatment with corticosteroids on 30-day mortality in patients with CAP.

Methods In this individual patient data meta-analysis, we included RCTs published before July 1, 2024, comparing adjuvant treatment with corticosteroids versus placebo in patients hospitalised with CAP. The primary endpoint was 30-day all-cause mortality, collected across all trials, and analyses followed the intention-to-treat principle. We analysed HTE using risk and effect modelling. For risk modelling, patients were classified as having less severe or severe CAP based on the pneumonia severity index (PSI), comparing PSI class I–III versus class IV–V. For effect modelling, we trained a corticosteroid-effect model on six trials and externally validated it using data from two trials, received after model preregistration. This model classified patients into two groups: no predicted benefit and predicted benefit from adjuvant treatment with corticosteroids. The literature search was registered on PROSPERO, CRD42022380746.

Findings We included eight RCTs with 3224 patients. Across all eight trials, 246 (7·6%) patients died within 30 days (106 [6·6%] of 1618 in the corticosteroid group *vs* 140 [8·7%] of 1606 in the placebo group; odds ratio [OR] 0·72 [95% CI 0·56–0·94], p=0·017). The corticosteroid-effect model, which selected C-reactive protein (CRP), showed significant HTE during external validation in the two most recent trials. In these trials, 154 (11·4%) of 1355 patients died within 30 days (88 [13·1%] of 671 in the placebo group *vs* 66 [9·6%] of 684 in the corticosteroid group; OR 0·71 [95% CI 0·50–0·99], p=0·044). Among patients predicted to have no benefit (CRP ≤204 mg/L, n=725), no significant effect was observed (OR 0·98 [95% CI 0·63–1·50]), whereas for those with predicted benefit (CRP >204 mg/L, n=630), 39 (13·0%) of 301 patients died in the placebo group compared with 20 (6·1%) of 329 in the corticosteroid group (0·43 [0·25–0·76], $p_{interaction}$=0·026). No significant HTE was found between less severe CAP (PSI class I–III, n=229) and severe CAP (PSI class IV–V, n=1126). Corticosteroid therapy significantly increased hyperglycaemia risk (44 [12·8%] of 344 in the placebo group *vs* 84 [24·8%] of 339 in the corticosteroid group; OR 2·50 [95% CI 1·63–3·83], p<0·0001) and hospital re-admission risk (30 [3·7%] of 814 in the placebo group *vs* 57 [7·0%] of 819 in the corticosteroid group; 1·95 [1·24–3·07], p=0·0038).

Interpretation Overall, adjuvant therapy with corticosteroids significantly reduces 30-day mortality in patients hospitalised with CAP. The treatment effect varied significantly among subgroups based on CRP concentrations, with a substantial mortality reduction observed only in patients with high baseline CRP.

Funding None.

Introduction

Community-acquired pneumonia (CAP) is a major cause of hospitalisation, with high mortality.[1] Although the exact cause is often unidentified, it can be caused by diverse pathogens, including viruses, bacteria, and fungi. This diversity complicates treatment. Adjuvant treatment with corticosteroids could reduce the excessive inflammatory response, which is associated with higher mortality.[2] However, despite evidence from several randomised controlled trials (RCTs) and a meta-analysis suggesting overall survival benefit, their routine use remains debated. Recently published treatment guidelines offer conflicting perspectives.[3–5] The concept of heterogeneity of treatment effect (HTE) has been described as understanding how

Lancet Respir Med 2025

Published **Online** January 29, 2025 https://doi.org/10.1016/S2213-2600(24)00405-3

See **Online/Comment** https://doi.org/10.1016/S2213-2600(24)00418-1

Department of Intensive Care (J M Smit MSc, P A Van Der Zee PhD, S C M Stoof PhD, M E Van Genderen PhD, Prof D A M P J Gommers PhD, J Van Bommel PhD, H Endeman PhD), **Department of Pulmonary Medicine** (P A Van Der Zee), **Department of Public Health** (D van Klaveren PhD), **Department of Medical Microbiology and Infectious Diseases, and Department of Internal Medicine, Section of Infectious Diseases** (H I Bax PhD), **Erasmus MC—University Medical Center Rotterdam, Rotterdam, Netherlands; Pattern Recognition & Bioinformatics Group, Delft University of Technology, Delft, Netherlands** (J M Smit, Prof M J T Reinders PhD, J H Krijthe PhD); **Department of Pulmonary Medicine, Spaarne Gasthuis, Haarlem, Netherlands** (D Snijders PhD); **Department of Pulmonary Medicine, Noordwest Hospital, Alkmaar, Netherlands** (W G Boersma PhD); **Department of Pulmonary Medicine, University Hospital of Cattinara, Trieste, Italy** (P Confalonieri MSc, F Salton MSc, Prof M Confalonieri PhD); **Department of Veterans Affairs, Cooperative Studies Program Coordinating Center, Palo Alto, CA, USA** (M-C Shih PhD); **Department of Pharmaceutical Sciences, University of Tennessee Health Science**

Center, Memphis, TN, USA (Prof G U Meduri PhD); Medecine intensive reanimation (P-F Dequin PhD) and INSERM CIC1415 (A Le Gouge MSc), Chru Hôpitaux De Tours, Hospital Bretonneau, Tours, France; Centre for Medicine Use and Safety, Monash University, Melbourne, VIC, Australia (M Lloyd PhD); Department of Medicine, The Western Health Chronic Disease Alliance and the University of Melbourne, Melbourne, VIC, Australia (H Karunajeewa PhD, G Bartminski MD); Pulmonary Medicine, Badalona Serveis Assistencials, Badalona, Spain (S Fernández-Serrano PhD); Pulmonary Medicine, Bellvitge University Hospital, L'Hospitalet de Llobregat, Barcelona, Spain (G Suárez-Cuartín PhD); Predictive Analytics and Comparative Effectiveness Center, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA, USA (D van Klaveren); CLEAR Methods Center, Division of Clinical Epidemiology, Department of Clinical Research, University Hospital Basel, University of Basel, Basel, Switzerland (M Briel PhD, C M Schönenberger MSc); Department of Health Research Methodology, Evidence, and Impact, McMaster University, Hamilton, ON, Canada (M Briel); Department of Biomedical Data Sciences (Prof E W Steyerberg PhD) and Department of Internal Medicine (Prof W J W Bos PhD); Leiden University Medical Center, Leiden, Netherlands; Department of Internal Medicine (Prof W J W Bos, E Wittermans PhD), Department of Pulmonary Medicine (J C Grutters PhD), and Department of Pharmacy (E M W van de Garde PhD), St Antonius Hospital, Nieuwegein, Netherlands; Hormonpraxis Aarau, Aarau, Switzerland (C A Blum PhD), Division of Endocrinology, Diabetes and Metabolism, Department of Clinical Research, Universitätsspital Basel, Basel, Switzerland (C A Blum, M Christ-Crain PhD); Department of Pulmonology, Hospital Clinic, University of Barcelona, Barcelona, Spain (Prof A Torres PhD); Centro de

Research in context

Evidence before this study

Routine use of corticosteroids for patients hospitalised with community-acquired pneumonia (CAP) remains controversial, with two recently published treatment guidelines presenting conflicting perspectives. The most recent guideline recommends corticosteroid treatment in patients with severe CAP. However, severe CAP does not have a unified definition, and the hypothesis that patients with severe CAP benefit more from corticosteroids is based on comparisons between, rather than within, trials, which is problematic because each trial includes a mixture of CAP severities. We updated the literature search from the previous individual patient data meta-analysis by searching MEDLINE, Embase, and the Cochrane Central Register of Controlled Trials (CENTRAL) for randomised controlled trials (RCTs) comparing corticosteroids versus placebo in patients hospitalised with CAP, covering the period from July 1, 2017, to July 1, 2024, using medical subject headings related to "pneumonia" and "corticosteroid".

Added value of this study

Subsequently, we conducted a data-driven individual patient data meta-analysis, including eight RCTs with 3224 patients hospitalised with CAP. Importantly, the collection of individual patient data allowed us to analyse heterogeneity of treatment effects (HTE) by comparing patients within trials rather than between them. In addition to validating the earlier hypothesised HTE between severe and less severe CAP, we developed and externally validated a corticosteroid-effect model to predict the benefit from adjuvant corticosteroid therapy, which selected baseline C-reactive protein (CRP) as an important effect modifier. Our results show a significant reduction in overall mortality from adjuvant corticosteroid therapy in all patients hospitalised with CAP from the eight included RCTs. This treatment effect varied significantly between subgroups identified by the corticosteroid-effect model: patients with high baseline CRP concentrations had a survival benefit, whereas those with low baseline CRP concentrations did not. No significant HTE was found between less severe versus severe CAP (defined as pneumonia severity index class I–III *vs* IV–V). Additionally, we found a significant, overall increase in hospital re-admissions and incidence of hyperglycaemia associated with corticosteroid treatment.

Implications of all the available evidence

Contrary to current guidelines, which do not mention CRP in their recommendations for corticosteroid treatment, our results suggest that corticosteroids should be considered for patients with increased baseline CRP concentrations. Furthermore, as we found a significant overall increase in hospital re-admission and hyperglycaemia incidence due to corticosteroid treatment, a threshold-based CRP decision for treatment might be preferred. The proposed CRP threshold based on our externally validated corticosteroid-effect model was 204 mg/L, but the optimal threshold might be in a wider range around this value.

treatment effect can vary across patients.[6] A widespread hypothesis is that patients with more severe pneumonia derive greater benefits from corticosteroids.[4,7] However, severe CAP does not have a unified definition. Moreover, even with a unified definition for severe CAP, such a hypothesised HTE cannot be confirmed through an aggregate data meta-analysis, because this type of analysis relies on differences between rather than within RCTs.[8] An individual patient data meta-analysis enables examination of HTE within RCTs, with subgroups based on individual baseline characteristics.[9] Conventional individual patient data meta-analyses are based on subgroup analyses of one variable at a time, which are limited by low statistical power, multiple testing, and the inability to integrate multiple patient characteristics simultaneously. By contrast, in our data-driven individual patient data meta-analysis, we performed a predictive HTE analysis,[6] which addresses these limitations by providing predictions of individualised (ie, based on one or more patient characteristics) treatment effects[10] of corticosteroids in patients hospitalised with CAP.

Methods

Search strategy and selection criteria

Study reporting adhered to the PRISMA-IPD guidelines (appendix p 40),[11] and the literature search was registered on PROSPERO (CRD42022380746).[12] We published our preliminary results on the preprint platform medRxiv[13] with the purpose of preregistering our trained model and external validation plan before receiving the data needed for the external validation.

We extended the systematic search for RCTs comparing adjuvant therapy with corticosteroids with placebo in patients hospitalised with CAP performed in the previous individual patient data meta-analysis on this topic,[7] updating it to July 1, 2024 (appendix p 44). Inclusion criteria generally included radiographic consolidations and multiple CAP symptoms (eg, cough, sputum, and fever), with some trial variations. Two reviewers (JMS and PAvdZ) independently screened articles and assessed risk of bias using the RoB2 tool.[14] Ten eligible trials[15–24] were identified (figure 1A). We initially obtained individual patient data from six trials,[15–17,20–22] and, after preregistering our trained model and external validation plan as a preprint on medRxiv,[13] we obtained individual patient data from the two latest trials[23,24] (figure 1A). Authors of remaining trials[18,19] either did not respond or reported unavailable individual patient data. Initially collected individual patient data included age, sex, clinical parameters, baseline laboratory values, comorbidities, and the pneumonia severity index (PSI; appendix p 22).[25] Post hoc (ie, not prespecified in the medRxiv submission), we

collected the CURB-65 score (appendix p 22),[26] and individual patient data regarding initial admission to an intensive care unit (ICU), initial need for invasive mechanical ventilation, microbiological aetiology, and initial antimicrobial treatment. Individual patient data from an additional observational CAP study[27] were also collected, but only to improve the imputation of missing data in trial datasets. Individual patient data were checked for consistency with the original publications, and remaining issues were resolved with the corresponding authors. We did not collect data regarding ethnicity or race. All included trials had ethical approval from local institutional Medical Ethics Committees and recruited patients who provided written informed consent, or a waiver for consent was granted. The primary study endpoint was 30-day all-cause mortality, which was gathered from all included trials (and specifically requested from investigators whose RCTs did not originally report this outcome). Post hoc, we analysed the effect of corticosteroids on several secondary endpoints (ie, 90-day mortality, length of hospital and ICU stay, 28-day need for invasive mechanical ventilation and vasopressors, and hospital re-admission) and adverse events (ie, hyperglycaemia, hospital-acquired infections, and gastrointestinal bleeding). All patients were analysed in the study group to which they were randomly assigned (intention-to-treat principle). We excluded patients with unobserved primary endpoints and with implausible reported baseline C-reactive protein (CRP; >1000 mg/L).

Predictive HTE analysis

We conducted a predictive HTE analysis following the framework proposed by Kent and colleagues,[6] who distinguish two categories of predictive HTE approaches: risk modelling, in which a multivariable model predicts the mortality risk and is applied to disaggregate patients within RCTs to define risk-based variation in benefit from corticosteroids, and effect modelling, in which a model is trained on RCT data by incorporating a variable for treatment assignment, and interactions between treatment and baseline covariates are modelled (figure 1B). We explored both risk and effect modelling strategies. For risk modelling, rather than building a new risk model, we used the PSI,[25] a well established risk model for adverse outcomes in CAP. For effect modelling, we trained a new model, which is henceforth referred to as the corticosteroid-effect model. As effect modelling requires a large amount of data and is prone to overfitting,[28] and given the small number of patients and (mortality) events in our setting, we initially limited ourselves to effect modelling methods based on penalised regression. Specifically, we used a method proposed by Tian and colleagues,[29] in which the individualised treatment effect is modelled directly through least absolute shrinkage and selection operator (LASSO) logistic regression, consisting of only the treatment variable and covariate–treatment interaction terms (appendix p 48).

Corticosteroid-effect model training and external validation

The training procedure of the corticosteroid-effect model comprised multiple steps: a priori variable selection based on availability, data imputation using the K-nearest-neighbour imputation algorithm (trained using patients from the train cohort and the observational study[27]), data normalisation, training of the LASSO regression model with a tuned penalty strength (λ) in the train cohort, and prediction of individualised treatment effects for the test cohort, defined as the predicted probability of 30-day mortality with placebo treatment, minus the predicted probability with corticosteroid treatment (figure 1B). Each step is described in more detail in the appendix (p 49). Before training the LASSO regression model, λ was optimised through a leave-one-trial-out cross-validation within the train cohort, selecting the λ that yielded the best cross-validated discrimination for benefit (appendix p 51).[30] Discrimination for benefit represents the model's ability to rank patients based on the benefit they would derive from adjuvant treatment with corticosteroids. To quantify this, we proposed the area under the Δ-benefit curve (AUC-benefit), which summarises the difference in observed mortality rate reduction between subgroups identified by use of various thresholds for the predicted individualised treatment effects (appendix p 53).

The initially collected individual patient data from six trials[15–17,20–22] formed the train cohort, whereas individual patient data from the two most recent trials[23,24] formed the test cohort for external validation (figure 1A). Crucially, we requested the authors of the two latest trials to provide individual patient data after the trained corticosteroid-effect model and evaluation plan were published on medRxiv,[13] ensuring that the individual patient data of these trials could not influence the model training. After receiving the individual patient data of the test cohort, we imputed missing baseline covariates based on the train cohort, and predicted individualised treatment effects for the patients in the test cohort using the trained corticosteroid-effect model. As the PSI was developed before any of the included trials were conducted,[25] its performance in all the included trials can be interpreted as an external validation.

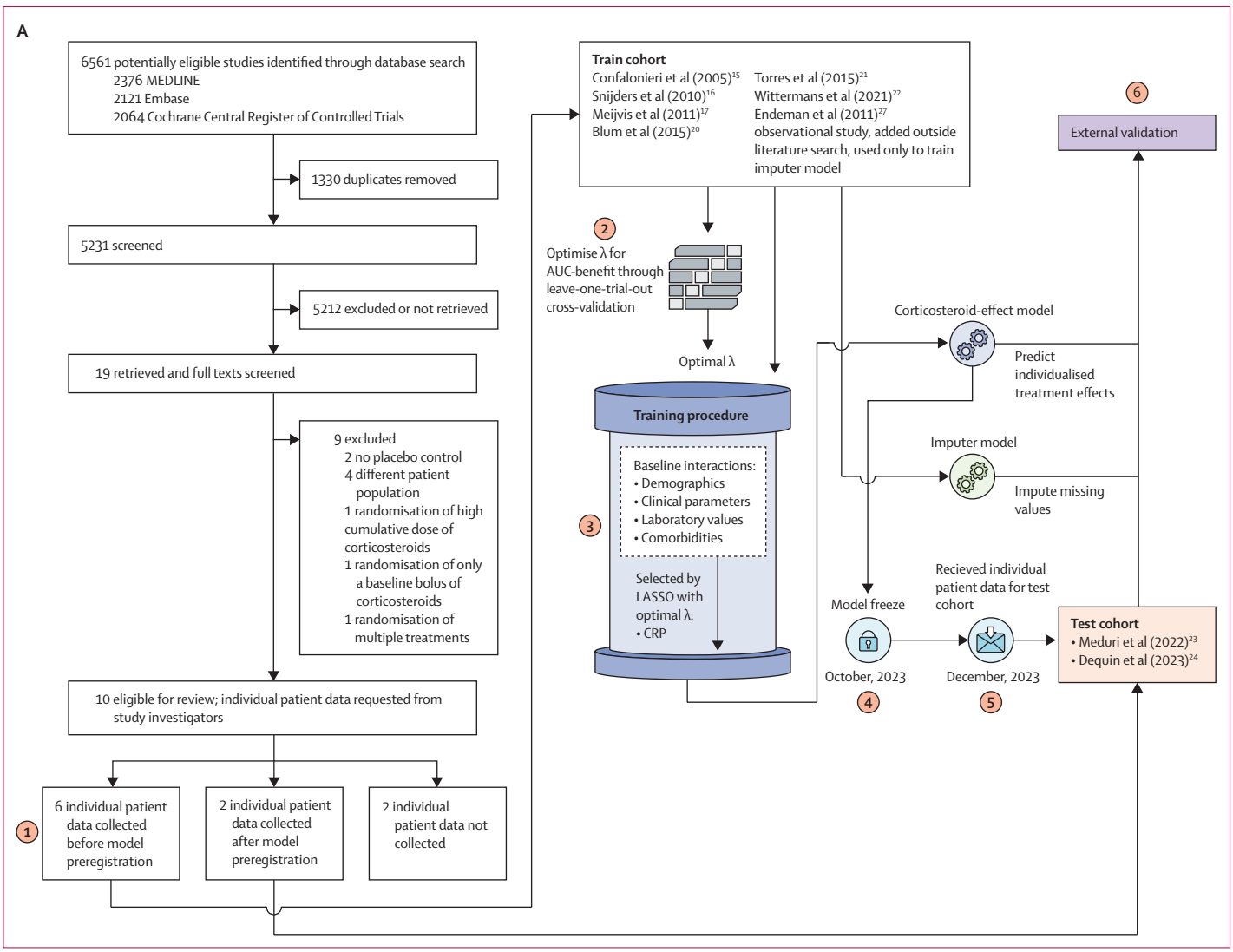
Model evaluation

As prespecified on medRxiv, we assessed discrimination and calibration for benefit for both the PSI and the corticosteroid-effect model, and tested for HTE among the subgroups identified by these models, in the test cohort (ie, the external validation). Discrimination for benefit was evaluated using the AUC-benefit, and the calibration for benefit by dividing patients into four groups based on ascending predicted individualised treatment effect quartiles and plotting the predictions next to the observed mortality reductions. To test for HTE, we first assumed a decision threshold: ie, the

Investigación Biomédica En Red-Enfermedades Respiratorias (CIBERES), Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Universitat de Barcelona, Barcelona, Spain (A Motos PhD); **Center for Research in Transplantation and Translational Immunology, UMR 1064, Nantes Université, INSERM, CHU Nantes, Nantes, France** (A Motos)

Correspondence to:
Mr Jim M Smit, Department of Intensive Care, Erasmus MC—University Medical Center Rotterdam, 3015 GD Rotterdam, Netherlands
j.smit@erasmusmc.nl

See **Online** for appendix



(Figure 1 continues on next page)

value above which treatment of patients is considered worthwhile.[30] For the corticosteroid-effect model, we assumed treatment to be worthwhile if any reduction in mortality risk was predicted (regardless of the magnitude), dividing patients into a subgroup of those who would have been advised for treatment (ie, individualised treatment effect >0; the predicted benefit subgroup) or a subgroup of those who would have been advised against treatment (individualised treatment effect ≤0; the predicted no benefit subgroup; figure 1B). The PSI, however, predicts mortality risk, rather than treatment effect. To test the hypothesis that patients with severe CAP benefit more from corticosteroids, we chose a decision threshold of 90 for the PSI (as prespecified on medRxiv[13]), which splits patients into PSI class I–III (ie, less severe CAP) versus PSI class IV–V (ie, severe CAP; figure 1B). For the resulting subgroups, we estimated relative treatment effects in terms of odds ratios (ORs) with 95% CIs in a one-stage approach[31] using a linear mixed-effects (logistic regression) model, including the trial as a random intercept to account for between-trial variability. Addressing the non-collapsibility of the OR measure, we (post hoc) also calculated ORs conditional on two strong and widely available prognostic factors for the primary endpoint (ie, age and PSI), through direct adjustment[32] (implementations in R version 4.2.1; appendix p 23) in the test cohort. Finally, we tested for HTE through an interaction test by adding the subgroup variables in turn to the mixed-effects model as a main effect and as an interaction term with the treatment variable. Additionally, we presented for each subgroup the observed mortality rates in the treatment groups, observed treatment effects in terms of mortality reduction (95% CIs around mortality reductions were calculated through bootstrapping using 1000 bootstrap

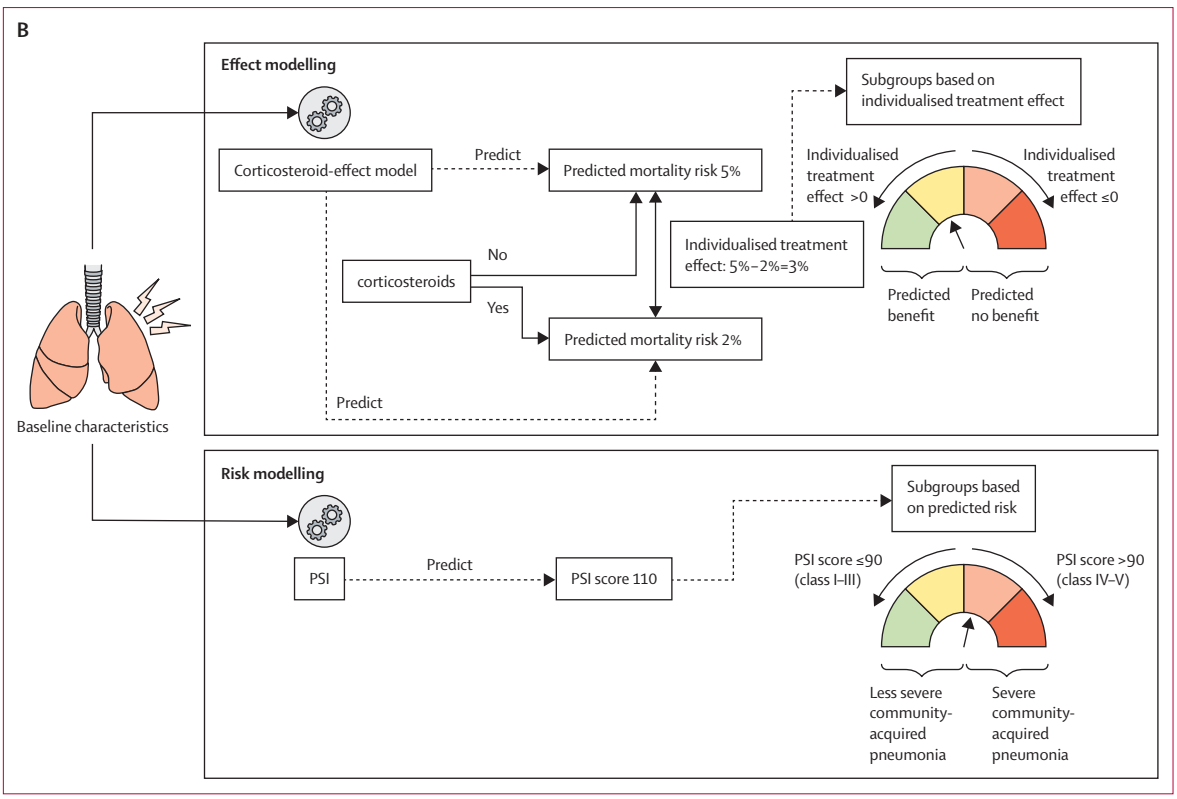


Figure 1: **Study selection and design**
(A) Schematic overview of literature search, model training, and external validation: first, eligible trials were identified through a systematic literature search; second, λ was optimised using leave-one-trial-out cross-validation; third, the corticosteroid-effect model was trained with optimised λ; fourth, the model was preregistered on medRxiv;[13] fifth, individual patient data were obtained from test cohort trials; and sixth, the corticosteroid-effect model was externally validated in the test cohort. (B) Schematic overview of risk and effect modelling. λ=penalty strength. AUC-benefit=area under the Δ-benefit curve. CRP=C-reactive protein. LASSO=least absolute shrinkage and selection operator. PSI=pneumonia severity index.

samples), and the number needed to treat (NNT). Additionally, we tested for HTE and reported observed mortality rates, treatment effects, and NNTs among the subgroups identified by both PSI and corticosteroid-effect model in the individual trials that made up the test cohort, the train cohort, and the full cohort (ie, all eight trials in the train and test cohorts combined). As the corticosteroid-effect model was trained using the train cohort (constituting the majority of the full cohort), the observed HTE among the subgroups identified by the corticosteroid-effect model in the train cohort and full cohort is at risk of overfitting.

To evaluate the overall effect of adjuvant corticosteroid therapy on the primary endpoint in all patients hospitalised with CAP, regardless of subgroups, we estimated relative treatment effects using the same one-stage approach[31] and reported observed mortality rates, treatment effects, and NNTs based on the full cohort (ie, all eight included trials). Additionally, we assessed the overall effects of corticosteroids in the train and test cohorts separately, and in each individual trial.

We applied the same approach to evaluate overall corticosteroid effects, as well as HTE between subgroups identified by the PSI and corticosteroid-effect model, on secondary endpoints and adverse events, using all trials for which individual patient data were available for each specific endpoint or adverse event. However, because logistic regression is not suited for testing the significance of corticosteroid effects on length of stay outcomes, we used the Kruskal–Wallis test (using Scipy's kruskal function[33]) for the length of hospital and ICU stay outcomes. All p values were two sided, and values less than 0·05 were deemed to be significant.

Method selection and non-linear effect modelling

Before obtaining individual patient data of the test cohort, we selected our effect modelling method of choice (ie, the Tian method[29]) among alternative penalised regression strategies (appendix p 55). Since our implementation of the Tian method[29] only supported linear effect modelling, we also evaluated the performance of more flexible approaches that allow for non-linearities.

As prespecified,[13] we trained a modified Tian model on the train cohort with dichotomised interaction terms for each continuous variable to capture non-linearities. Additionally, we post hoc trained models using methods capable of handling non-linear relationships and higher-order interactions (ie, a causal forest,[34] X-learner,[35] and

	Country	Modelling cohort	Patients in ITT analysis	30-day mortality rate			Age, years	Sex		PSI score	Corticosteroid, route, dose, and duration	Cumulative dose on day 7, in hydrocortisone equivalents, mg
				Overall	Corticosteroid group	Placebo group		Female	Male			
Confalonieri et al (2005)[15]	Italy	Train cohort	46	8/46 (17·4%)	0/23	8/23 (34·8%)	67 (52–76)	14/46 (30·4%)	32/46 (69·6%)	113 (95–133)	Hydrocortisone (intravenous), 200 mg bolus followed by 240 mg daily for 7 days	1880
Snijders et al (2010)[16]	Netherlands	Train cohort	213	12/213 (5·6%)	6/104 (5·8%)	6/109 (5·5%)	65 (52–80)	89/213 (41·8%)	124/213 (58·2%)	85 (63–115)	Prednisolone (intravenous or oral), 40 mg daily for 7 days	1120
Meijvis et al (2011)[17]	Netherlands	Train cohort	304	18/304 (5·9%)	9/151 (6·0%)	9/153 (5·9%)	67 (51–79)	133/304 (43·8%)	171/304 (56·3%)	89 (64–117)	Dexamethasone (intravenous), 5 mg daily for 4 days	500
Blum et al (2015)[20]	Switzerland	Train cohort	785	28/785 (3·6%)	15/392 (3·8%)	13/393 (3·3%)	73 (61–83)	298/785 (38·0%)	487/785 (62·0%)	89 (64–113)	Prednisone (oral), 50 mg daily for 7 days	1400
Torres et al (2015)[21]	Spain	Train cohort	120	15/120 (12·5%)	6/61 (9·8%)	9/59 (15·3%)	70 (53–81)	46/120 (38·3%)	74/120 (61·7%)	110 (86–135)	Methylprednisolone (intravenous), 0·5 mg/kg twice daily for 5 days	1927*
Wittermans et al (2021)[22]	Netherlands	Train cohort	401	11/401 (2·7%)	4/203 (2·0%)	7/198 (3·5%)	67 (56–76)	165/401 (41·1%)	236/401 (58·9%)	80 (62–101)	Dexamethasone (oral), 6 mg daily for 4 days	600
Meduri et al (2022)[23]	USA	Test cohort	584	80/584 (13·7%)	40/297 (13·5%)	40/287 (13·9%)	68 (62–75)	22/584 (3·8%)	562/584 (96·2%)	119 (98–147)	Methylprednisolone (intravenous), 40 mg bolus followed by a tapering dose (40 mg to 4 mg daily) for 20 days	1600
Dequin et al (2023)[24]	France	Test cohort	795	76/795 (9·6%)	27/400 (6·8%)	49/395 (12·4%)	67 (58–78)	243/795 (30·6%)	552/795 (69·4%)	128 (102–151)	Hydrocortisone (intravenous), tapering dose (200 mg to 0 mg daily) for 8 days or 14 days	1100 or 1500†

Data are n/N (%) or median (IQR) unless otherwise stated. Percentages might not sum to 100 as a result of rounding. The distributions of age, sex, and PSI scores in the separate treatment groups of each individual trial are shown in the appendix (p 5). ITT=intention-to-treat. PSI=pneumonia severity index. *To calculate the cumulative dose in the treatment regimen of Torres and colleagues' study,[21] which administered 0·5 mg/kg of methylprednisolone every 12 h, we assumed average weights of 84·0 kg for males and 65·9 kg for females due to the unavailability of individual patient weight data. †In the trial by Dequin and colleagues,[24] on day 4, the medical team used predefined criteria to decide whether to administer hydrocortisone for a total of 8 days or 14 days, with different doses on days 5, 6, and 7.

Table 1: **Characteristics of the eligible randomised controlled trials**

R-learner[36]) and evaluated them on the test cohort (appendix p 74).

Sensitivity analyses

We conducted two prespecified sensitivity analyses to explore effect modification for individual covariates, and to validate our corticosteroid-effect model in two other, ineligible trials.[37,38] Additionally, we conducted several post-hoc sensitivity analyses, including assessing bias from missingness in the primary endpoint,[39] and investigating risk of aggregation bias.[31] We examined the impact of data imputation, the robustness of our analysis with different imputation methods, and the influence of excluding patients with high missingness in baseline characteristics. We examined HTE among separate PSI classes, among severity classes based on the CURB-65 score,[26] and between patients who did or did not require initial ICU admission or invasive mechanical ventilation. We investigated HTE based on microbiological aetiology, and evaluated outcomes for 30-day mortality and hospital-acquired infections, considering variations in corticosteroid types, doses, and treatment timing. We examined the performance of the corticosteroid-effect model among patient subgroups based on microbiological aetiology. Finally, we re-assessed the overall effects of corticosteroids on length of stay outcomes, excluding patients with 30-day mortality, and on re-admission, only considering 30-day re-admissions (appendix p 81).

Role of the funding source

There was no funding source for this study.

Results

Eight of ten eligible trials provided individual participant data, including 3248 patients (table 1). We excluded 24 patients from the test cohort in external validation: 20 had missing 30-day mortality data, three had CRP higher than 1000 mg/L (appendix p 112), and one had more than 80% missing baseline data. No patients were excluded in the train cohort. Overall 30-day mortality ranged from 4–17% among the individual trials: 246 (7·6%) of 3224 patients died across all trials. Corticosteroid doses varied, with a day 7 hydrocortisone-equivalent dose of 500–2000 mg (appendix p 4), and treatment started within 12–96 h of hospital admission. Baseline characteristics were similar between corticosteroid and placebo groups, with similar missingness across treatment groups (table 2), as well as in train and test cohorts (appendix pp 24–25). Baseline characteristic distributions per study and per treatment group showed some small but significant imbalances between treatment groups in individual trials (appendix p 5). We obtained individual patient data on microbiological aetiology from seven trials,[15–17,20–22,24] finding no pathogen in 1433 (53·8%) of 2663 patients, whereas 960 (36·0%) tested positive

for bacterial and 285 (10·7%) for viral agents (appendix p 26). Data on antimicrobial treatments came from four trials, with macrolides, third-generation cephalosporins, amoxicillin plus clavulanic acid, and amoxicillin used in 1614 (84·4%) of 1912 patients (appendix pp 16, 26). Plots of relative and absolute treatment effects by PSI and CRP quartiles showed a non-linear relationship between CRP and treatment effect (appendix pp 16–18).

A priori, we selected 20 baseline variables consisting of demographics, clinical parameters, laboratory values, and comorbidities (table 2). The rate of missingness per variable was generally low (ie, missing for 25% of the patients or less), and was similar between the corticosteroid and placebo groups. The rate of missingness per patient was also generally low: 1689 (90·4%) of 1869 patients in the train cohort and 1332 (98·3%) of 1355 patients in the test cohort had low (ie, ≤20%) missingness among baseline characteristics (appendix p 110). After optimising λ and training the corticosteroid-effect model (appendix p 18), LASSO penalisation selected only the interaction term with CRP out of the 20 interaction terms based on each baseline variable.

In the external validation, the corticosteroid-effect model yielded a higher AUC-benefit than the PSI (appendix p 19). In terms of calibration for benefit, the corticosteroid-effect model showed underestimation in the higher prediction regions (figure 2). HTE between the subgroups in the test cohort identified by the corticosteroid-effect model was substantial (figure 3). Treatment benefit was substantial in the predicted benefit subgroup: 20 (6·1%) of 329 patients died within 30 days in the corticosteroid group compared with 39 (13·0%) of 301 in the placebo group (OR 0·43 [95% CI 0·25–0·76]), whereas there was no significant difference in mortality in the predicted no benefit subgroup (0·98 [0·63–1·50]), as reflected in a strong interaction ($p_{interaction}$=0·026). In both trials of the test cohort, we observed HTE between the subgroups identified by the corticosteroid-effect model in the same direction, although it was more pronounced in the trial by Meduri and colleagues[23] (appendix pp 27–28).

In the full cohort (ie, train and test cohort combined), as well as in the train cohort, we observed similar discrimination and calibration for benefit, and similar HTE (appendix pp 19–21).

As the LASSO penalisation selected only the interaction term with (continuous) CRP, each individualised treatment effect predicted by the corticosteroid-effect model corresponds to a specific CRP value, where a 0 prediction (ie, no benefit, no harm) corresponds to a CRP of 204 mg/L (appendix p 113). Post hoc, we examined HTE resulting from CRP thresholds around this value using the full cohort, suggesting similar HTE for thresholds around 200 mg/L (figure 4).

In the test cohort, the relative treatment effect was slightly greater in patients with PSI class I–III than in those with PSI class IV–V, although this difference was

	Corticosteroid group (n=1631)	Placebo group (n=1617)	Missingness (% corticosteroid group, % placebo group)	p value*
Demographics				
Sex	..	..	0·0, 0·1	0·68
Female	513 (31·5%)	496 (30·7%)	..	..
Male	1118 (68·5%)	1121 (69·3%)	..	..
Age, years	68·2 (58·1–79·0)	68·0 (58·0–79·0)	0·0, 0·1	0·77
Clinical parameters				
Respiratory rate, breaths per min	24·0 (19·5–28·0)	23·0 (19·0–28·0)	7·5, 6·5	0·28
Diastolic blood pressure, mm Hg	70·0 (60·0–79·0)	69·0 (60·0–78·0)	4·4, 4·3	0·44
Systolic blood pressure, mm Hg	125·0 (110·0–140·0)	123·0 (110·0–139·0)	4·4, 4·3	0·15
Temperature, °C	37·5 (36·9–38·3)	37·5 (36·8–38·4)	2·9, 2·8	0·96
Heart rate, bpm	92·5 (80·0–107·0)	91·0 (79·0–106·0)	1·7, 1·7	0·07
SpO_2, %	94·0% (92·0–97·0)	95·0% (92·0–97·0)	15·1, 14·7	0·21
Laboratory values				
Creatinine, µmol/L	92·8 (72·0–127·0)	89·0 (71·0–122·2)	25·4, 25·0	0·05
Sodium, mmol/L	136·0 (133·0–139·0)	136·0 (133·0–139·0)	1·2, 0·8	0·20
Urea, mmol/L	7·5 (5·1–12·0)	7·4 (5·0–11·1)	23·1, 22·3	0·85
CRP, mg/L	192·4 (90·0–300·0)	183·1 (80·4–293·9)	12·5, 10·6	0·15
Glucose, mmol/L	7·3 (6·1–9·0)	7·2 (6·0–8·9)	8·8, 9·6	0·08
White blood cell count, 10^9 cells per L	12·6 (9·1–17·0)	12·5 (9·0–17·1)	2·1, 1·7	0·32
Comorbidities				
Neoplastic disease	135 (8·3%)	135 (8·3%)	9·0, 8·9	0·95
Liver disease	59 (3·6%)	56 (3·5%)	11·0, 11·1	0·85
Congestive heart failure	284 (17·4%)	248 (15·3%)	2·3, 2·2	0·12
Renal disease	237 (14·5%)	209 (12·9%)	16·1, 16·6	0·19
Diabetes	405 (24·8%)	392 (24·2%)	0·6, 0·6	0·71
COPD	376 (23·1%)	386 (23·9%)	0·6, 0·6	0·59
Baseline disease severity indicators				
PSI				
Total score	106·0 (78·0–132·0)	103·0 (77·0–132·0)	0·3, 0·4	0·22
Class I	131 (8·0%)	120 (7·5%)	..	..
Class II	195 (12·0%)	190 (11·8%)	..	..
Class III	263 (16·1%)	299 (18·6%)	..	..
Class IV	607 (37·2%)	578 (35·7%)	..	..
Class V	430 (26·4%)	423 (26·2%)	..	..
CURB-65†				
Total score	1·0 (0·0–2·0)	1·0 (0·0–2·0)	29·5, 27·9	0·78
Score 0–2	1045 (64·1%)	1068 (66·0%)	..	..
Score 3–5	105 (6·4%)	98 (6·1%)	..	..
Other				
Initial ICU admission‡	494 (30·3%)	486 (30·1%)	18·2, 17·7	0·91
Initial need for invasive mechanical ventilation§	94 (5·8%)	88 (5·4%)	50·0, 50·2	0·70

Data are n (%) or median (IQR). COPD=chronic obstructive pulmonary disease. CRP=C-reactive protein. ICU=intensive care unit. PSI=pneumonia severity index. *Distributions of the placebo and corticosteroid groups were compared using a Fisher exact test for categorical variables and a two-sample *t* test for continuous variables, without adjusting for multiple testing. †We obtained individual patient data regarding CURB-65 scores from six trials.[16,17,20–22,24] ‡We obtained individual patient data regarding initial ICU admission from seven trials.[15–17,20–22,24] §We obtained individual patient data regarding initial need for invasive mechanical ventilation from four trials.[17,21,22,24]

Table 2: **Patient characteristics measured at baseline of the 3248 patients from the eight included randomised controlled trials**

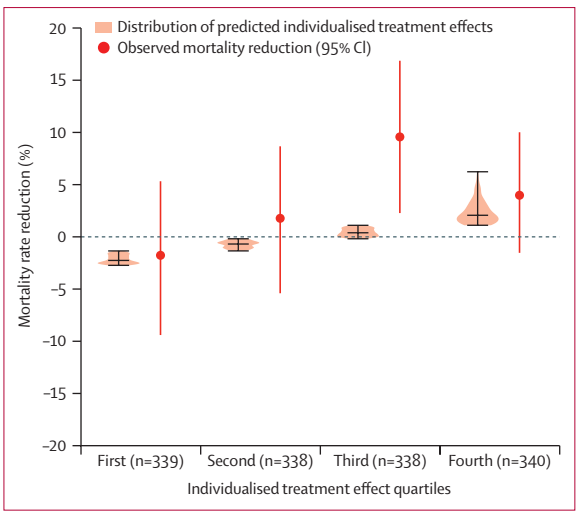


Figure 2: **Calibration for benefit results for the corticosteroid-effect model for external validation in the test cohort**
For four patient groups based on ascending quartiles of the predicted individualised treatment effects , the predicted individualised treatment effect distributions are visualised using violin plots (medians with IQRs) next to the observed mortality reductions (error bars indicate 95% CIs) in each quartile.

not significant (OR 0·40 [95% CI 0·12 to 1·36] *vs* 0·75 [0·52 to 1·06], $p_{interaction}$=0·33). We observed similar results in the individual trials that made up the test cohort (appendix pp 27–28). In the full cohort, there was no significant difference, but the point estimate of relative benefit was smaller in PSI class IV–V compared with PSI class I–III (OR 0·72 [95% CI 0·54 to 0·95] *vs* 0·60 [0·25 to 1·42], $p_{interaction}$=0·77), whereas the point estimate of absolute benefit was greater, with a mortality rate reduction of 3·3% (95% CI 1·1 to 5·5) in PSI class IV–V compared with 0·8% (–0·5 to 1·9) in PSI class I–III (appendix p 21). This phenomenon is known as risk magnification. In the train cohort, both relative and absolute benefit were greater in PSI class IV–V than in PSI class I–III (appendix p 21).

In the full cohort (ie, all patients hospitalised with CAP), adjuvant therapy with corticosteroids significantly reduced 30-day mortality, with 106 (6·6%) of 1618 patients having died at 30 days in the corticosteroid group compared with 140 (8·7%) of 1606 in the placebo group (OR 0·72 [95% CI 0·56–0·94], p=0·017; figure 3). This overall effect was similar in both the test cohort (0·71 [0·50–0·99], p=0·044; figure 3) and train cohort (0·76 [0·50–1·15], p=0·20; appendix p 21), but showed large variety among the individual trials (appendix p 29). We observed small differences between marginal ORs and ORs conditional on prognostic factors (appendix p 30).

The non-linear Tian method with extra, dichotomised terms selected the interaction terms with both (continuous) CRP and dichotomised glucose (split at 7 mmol/L) as predictors for benefit. However, when evaluated using the test cohort, it did not outperform the (linear) Tian method in terms of AUC benefit or calibration for benefit, whereas the more flexible causal machine learning methods showed even worse performances (appendix p 76).

We obtained individual patient data regarding 90-day mortality, length of ICU stay, and 28-day need for invasive mechanical ventilation from four trials,[15,20,21,24] 28-day need for vasopressors from three trials,[15,20,24] hospital re-admission from four trials,[16,17,20,22] length of hospital stay from six trials,[15–17,20–22] hyperglycaemia from four trials,[15–17,21] hospital-acquired infections from six trials,[15–17,20,21,24] gastrointestinal bleeding from five trials,[15,17,20,21,24] CURB-65 scores from six trials,[16,17,20–22,24] initial ICU admission from seven trials,[15–17,20–22,24] initial invasive mechanical ventilation from four trials,[17,21,22,24] and microbiological aetiology from seven trials.[15–17,20–22,24]

Adjuvant therapy with corticosteroids significantly reduced 90-day mortality by 2·8% (95% CI 0·4 to 5·2; OR 0·73 [95% CI 0·51 to 0·99], p=0·042), 28-day need for invasive mechanical ventilation by 4·8% (2·1 to 7·5; 0·59 [0·42 to 0·82], p=0·0019), 28-day need for vasopressors by 6·9% (4·0 to 9·7; 0·54 [0·40 to 0·72], p<0·0001), and median length of hospital stay from 7 days to 6 days (Kruskal–Wallis test p=0·0002) and ICU stay from 7 days to 5 days (Kruskal–Wallis test p=0·0009; appendix pp 31–32). However, it significantly increased hospital re-admission by 3·3% (95% CI –5·3 to –1·5; OR 1·95 [95% CI 1·24 to 3·07], p=0·0038; appendix p 31) and hyperglycaemia by 12·0% (–17·0 to –6·9; 2·25 [1·62 to 3·17], p<0·0001; appendix p 32). Corticosteroid therapy significantly increased hyperglycaemia risk (44 [12·8%] of 344 in the placebo group *vs* 84 [24·8%] of 339 in the corticosteroid group; 2·50 [1·63–3·83], p<0·0001) and hospital re-admission risk (30 [3·7%] of 814 in the placebo group *vs* 57 [7·0%] of 819 in the corticosteroid group; 1·95 [1·24–3·07], p=0·0038). Hospital-acquired infections occurred in 331 (12·5%) of 2650 patients and gastrointestinal bleeding in 33 (1·7%) of 1958, but corticosteroids had no significant effect on either outcome. We observed no significant HTE on secondary endpoints or adverse events between the subgroups identified by the PSI and the corticosteroid-effect model (appendix pp 33–38).

Evaluating the corticosteroid-effect model in the ineligible trials,[37,38] we observed no benefit in both predicted benefit and predicted no benefit subgroups. No significant effect modifiers were found, although (continuous) CRP interaction approached significance. We judged risk of bias due to missingness in the primary outcome to be small, but adjustment for aggregation bias reduced the strength of the interaction between the corticosteroid-effect model subgroups and the treatment. The primary analysis without imputation showed similar results, but wider CIs due to lower sample size. Different imputation methods, as well as the removal of patients with high missingness among baseline characteristics, yielded similar results. Unlike the PSI, CURB-65 suggested more benefit in less severe CAP. No significant

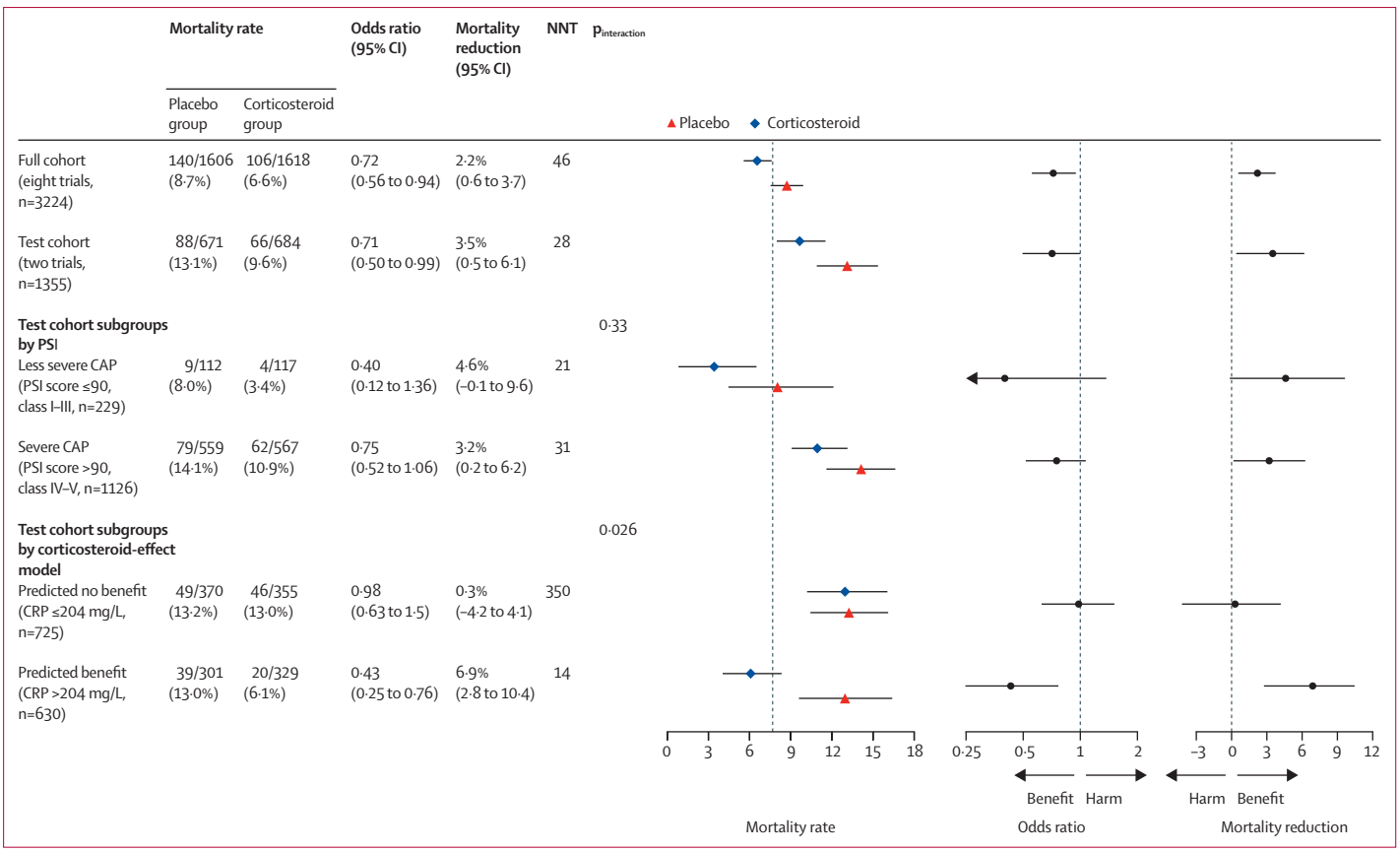


Figure 3: **Overall effect of adjuvant therapy with corticosteroids and results of external validation of the corticosteroid-effect model and PSI regarding 30-day mortality in the test cohort**
Heterogeneity of treatment effect on the relative, odds ratio scale and the absolute, mortality risk difference scale. For the relative scale, we added the $p_{interaction}$ value. CAP=community-acquired pneumonia. CRP=C-reactive protein. NNT=number needed to treat. PSI=pneumonia severity index.

HTE was found across subgroups based on individual PSI classes, initial ICU admission, initial invasive mechanical ventilation, or any subgroup based on microbiological aetiology, although potential harm was noted in influenza (or any viral) infections, especially without a bacterial co-infection, with a trend towards significant HTE. After adjusting for corticosteroid-effect model subgroups, hydrocortisone showed significantly greater benefit than other corticosteroids, but we did not find treatment effect variation based on dose. We observed significantly better outcomes when corticosteroid treatment was initiated within 24 h, although this result was solely based on Dequin and colleagues' trial.[24] No significant HTE was observed for hospital-acquired infections across subgroups based on corticosteroid type, dose, or timing. Among patient subgroups by microbiological aetiology, we found consistent HTE between the subgroups identified by the corticosteroid-effect model, except for the viral infection group, in which we found point estimates that suggest harmful effect in both predicted benefit and predicted no benefit groups. Finally, the overall effect of corticosteroids on lengths of hospital stay and ICU stay, after excluding patients with 30-day mortality, and on hospital re-admissions, based on the three trials that tracked 30-day re-admissions, showed similarly significant results (appendix p 84).

Discussion

In the full cohort, totalling eight RCTs[15–17,20–24] and 3224 patients hospitalised with CAP, adjuvant corticosteroid therapy significantly reduced 30-day mortality (8·7% in the placebo group *vs* 6·6% in the corticosteroid group). Importantly, at external validation in two unseen RCTs (ie, the test cohort),[23,24] treatment effects varied significantly across patient subgroups classified by the proposed corticosteroid-effect model based on baseline CRP. No significant mortality reduction was observed in patients where the model predicted no benefit (ie, baseline CRP ≤204 mg/L), whereas adjuvant therapy with corticosteroids significantly reduced mortality for patients where the corticosteroid-effect model predicted benefit (ie, baseline CRP >204 mg/L; 13·0% in the placebo group *vs* 6·1% in the corticosteroid group). Our findings indicate that adjuvant corticosteroid therapy did not cause a significant increased risk of hospital-acquired infections or bleeding, but it significantly increased the risk of hospital re-admissions and hyperglycaemia.

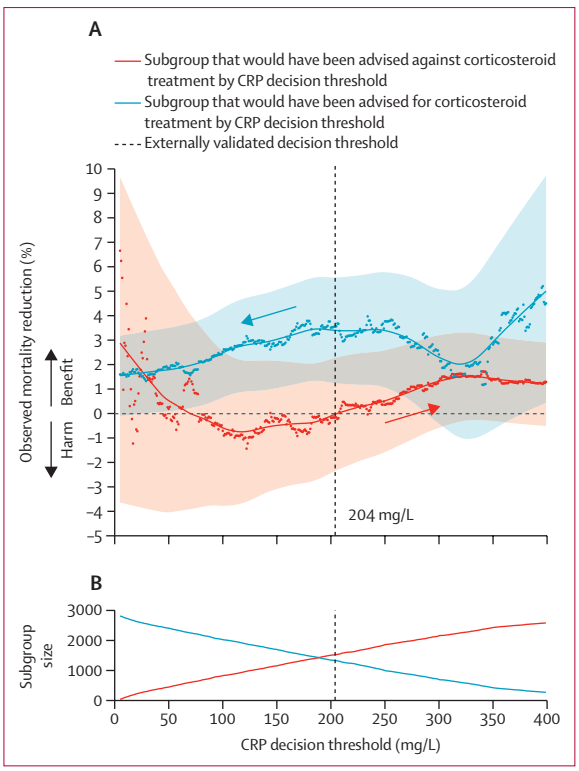


Figure 4: **The absolute treatment effects (A) and numbers of patients (B) in the patient subgroups resulting from different CRP decision thresholds**
Only patients from the full cohort with non-missing CRP values were included in the plot (n=2857). Following the framework by Dahabreh and colleagues,[40] individual patients experience either benefit, harm, or no effect from corticosteroids. If the CRP threshold was set too high, patients who benefit from corticosteroid treatment were classified into the untreated subgroup, as indicated by the upward movement of the orange line (see orange arrow). If set too low, patients who experience no effect, or even harm, from corticosteroid treatment were classified into the treated subgroup, as indicated by the downward movement of the blue line (see blue arrow). The CRP threshold based on our proposed corticosteroid-effect model (ie, 204 mg/L) is marked by the dashed line. Decision thresholds around 200 mg/L show similar, significant heterogeneity of treatment effect. The smooth curves including 95% CIs were estimated by a linear loess smoother. The 95% CIs were calculated through bootstrapping (1000 bootstrap samples). CRP=C-reactive protein.

In the full cohort, we observed slightly less relative benefit (in terms of ORs) but greater absolute benefit (mortality risk reduction) in patients with CAP in PSI class IV–V than in those in PSI class I–III. This seemingly counterintuitive finding might be explained by risk magnification,[41] in which relatively homogeneous relative treatment benefits lead to greater absolute benefits for patients with higher initial mortality risk. A model such as PSI that accurately predicts baseline mortality would suffice to identify patients with most (absolute) benefit from corticosteroids. However, when relative treatment effect is heterogeneous, patient characteristics act as effect modifiers. Our external validation results confirmed baseline CRP as a key effect modifier, showing non-linear effect modification (appendix pp 17–18). The conflicting results from the two recent trials[23,24] in our test cohort (both focused on severe CAP, but used different criteria for severe) might be explained by differing CRP distributions: patients included in Dequin and colleagues' trial,[24] with positive findings for the effect of hydrocortisone, had relatively high CRP, whereas those included in Meduri and colleagues' trial,[23] which reported no significant effect of methylprednisolone, had relatively low CRP (appendix p 11).

In this study, the effect modelling strategy was better suited for HTE than risk modelling, although effect modelling is more prone to overfitting and vulnerable to false discoveries of effect modification.[6,42] This is especially problematic in (pooled) randomised trials with small sample sizes and few events. This overfitting tendency is reflected in the poor performance of non-linear effect modelling methods, which outperformed the linear Tian method in the train cohort (apparent validation) but underperformed in unseen RCTs (appendix p 76). We recommend preregistering trained effect models, as in our study, to prevent test data from influencing any choices during model training and to reduce overfitting risks.[43]

In our sensitivity analysis for aggregation bias, Riley and colleagues'[31] adjustment method reduced the significant interaction between treatment effect and subgroups in the test cohort (appendix p 85). However, the test cohort includes only two trials, limiting accurate modelling of between-trial heterogeneity and the reliability of the adjustment. Notably, the same interaction direction was observed in seven of eight individual trials (appendix pp 107–109). Thus, although we cannot fully exclude aggregation bias, it is unlikely that the interaction is driven primarily by between-trial rather than within-trial effects.[31]

Although corticosteroid treatment is advised in current guidelines for acute respiratory distress syndrome[44] and septic shock,[45] two recent guidelines[3,4] on corticosteroid treatment for CAP present conflicting perspectives: the 2023 European Respiratory Society, European Society of Intensive Care Medicine, European Society of Clinical Microbiology and Infectious Diseases, and Latin American Thoracic Association (ERS/ESICM/ESCMID/ALAT) guideline[3] suggest corticosteroids only in the presence of shock, whereas the 2024 Society of Critical Care Medicine (SCCM) guideline[4] advocates their use for all severe CAP cases. However, the 2024 SCCM guideline[4] lacks a unified definition of severe CAP[5] and bases its recommendation on an aggregate data meta-analysis, classifying whole trial populations as severe or less severe, which is problematic if one wants to draw conclusions about heterogeneity at the individual patient level. Our individual patient data meta-analysis, which allows for subgroup analysis at the individual patient level, showed that on a relative scale (ie, ORs), patients with severe CAP do not benefit more than those with less severe CAP, whether based on PSI, CURB-65, or need for initial ICU admission or invasive mechanical ventilation.[26] Instead, our externally validated corticosteroid-effect model selected CRP as the most important predictor for treatment benefit.

Hence, although we concur with the 2024 SCCM guideline[4] that adjuvant corticosteroid therapy should be considered in a CAP subgroup, we emphasise the importance of clearly defining this subgroup, and suggest considering corticosteroids for those with increased baseline CRP concentrations. Although the threshold based on our corticosteroid-effect model was 204 mg/L, the optimal threshold might be in a wider range of thresholds around this value, as we (post hoc) observed significant mortality reduction for treated subgroups resulting from CRP thresholds around 200 mg/L, splitting patients into subgroups of approximately equal numbers of patients (figure 4). Our results showed point estimates suggesting harmful effects of corticosteroids in patients with influenza or other viral infections (regardless of baseline CRP ≤204 mg/L or >204 mg/L), aligning with existing evidence from primarily observational studies.[46,47] Although this post-hoc and non-significant finding calls for further investigation, we recommend caution when treating patients with CAP who have viral infections using corticosteroids, particularly in the absence of bacterial co-infection.

The trial by Dequin and colleagues[24] was the only one among the included studies that implemented a response-dependent treatment regimen, administering corticosteroids for either 8 days or 14 days based on patient improvement by day 4, as assessed using predefined criteria. The strong positive effect found in this trial suggests the potential benefit from such regimens, although further research is needed to confirm this. Moreover, we observed a significantly greater benefit from hydrocortisone, based on the trials by Dequin and colleagues[24] and Confalonieri and colleagues,[15] suggesting that it might be more effective than other corticosteroid types. This post-hoc finding should also be interpreted with caution, and further research is needed to confirm it. Finally, we observed a significantly greater benefit from corticosteroid treatment started within 24 h, and even treatment harm for patients treated after 48 h, suggesting that patients should be treated within 48 h and ideally within 24 h after hospital admission. This finding, however, was also found post hoc, based on only one trial,[24] and therefore should be interpreted with caution.

The trial by Torres and colleagues[21] included only patients with CRP higher than 150 mg/L, motivated by its suggested prognostic value. In addition to its diagnostic[48] and prognostic value[49] (which could offer risk stratification, but not treatment guidance), CRP has also been hypothesised as a predictor of benefit from corticosteroids.[50,51] The most extensive evidence to date regarding CRP as a predictor for corticosteroid benefit in patients hospitalised with CAP comes from Briel and colleagues,[7] who compared CRP concentrations below 188 mg/L with those at or above this threshold in their earlier individual patient data meta-analysis, but found no significant HTE. Thus, our study provides the first evidence of CRP's use in guiding corticosteroid treatment. Instead of hypothesising a specific effect modifier a priori and assessing its credibility,[52] we performed a fully data-driven method, allowing the corticosteroid-effect model to select interactions independent of subjective human judgements. Although other metrics have been proposed to evaluate discrimination for benefit,[30,53] we introduced the AUC-benefit metric, which avoids one-to-one matching and fixed decision thresholds. Despite not being directly clinically interpretable, it is useful for comparing discriminative peformance across HTE models.

Our study has limitations. First, the included trials show some differences, including treatment timing, dose, and duration, which makes the pooling of all these trials an approach that is debatable. Second, previous antibiotic treatment might have affected baseline CRP values. In Meduri and colleagues' trial,[23] patients were enrolled up to 4 days after hospital presentation, with most receiving antibiotics within the first 6 h, although it is unclear how many had CRP measured after antibiotic treatment. By contrast, in the majority of other RCTs, CRP was measured within the first day of hospital presentation, as reported in the original publications or confirmed by corresponding authors (appendix p 38). Third, in three of the four trials from which we collected hyperglycaemia data, no standardised definition was used (appendix p 39), complicating interpretation. Fourth, additional factors such as radiological abnormalities, which could improve individualised treatment predictions, were unavailable. Fifth, cytokine and chemokine data (available for some trials) were not comparable across trials due to differences in measurement techniques. Therefore, inflammatory profiles could not be assessed in more detail. We acknowledge that CRP values were measured using different laboratory techniques across trials, with patients from 95 hospitals. However, our external validation procedure showed the effectiveness of the proposed corticosteroid-effect model, even when trained and tested on patients with CRP measured using various methods, suggesting its usefulness despite this limitation—which will persist as long as variation in laboratory techniques across hospitals remains. Adjuvant therapy with corticosteroids in patients with CAP significantly reduced 30-day mortality. Risk reduction of 30-day mortality by adjuvant therapy with corticosteroids varies significantly across subgroups classified by our newly developed and externally validated corticosteroid-effect model, driven by CRP. Application of the current model has the potential to guide a more personalised decision-making process in this patient population.

Contributors

JMS, PAvdZ, MJTR, JHK, and HE developed the concept and designed the study. DS, MC, ALG, HK, GS-C, WJWB, CAB, and AT collected the data. JMS and PAvdZ conducted systematic searches. JMS did the

statistical analysis. JHK and MJTR were responsible for statistical supervision. JMS and JHK accessed and verified the data. JMS, JHK, MJTR, PAvdZ, and HE interpreted the results. JMS, PAvdZ, JHK, MJTR, and HE wrote the first draft of the manuscript, and all authors critically revised the manuscript for important intellectual content. JHK, MJTR, PAvdZ, and HE supervised the study. All authors were responsible for the decision to submit for publication.

Declaration of interests

P-FD declares funding from the French Ministry of Health for the CAPE COD trial, whose data are used in this Article; loan of equipment and supply of devices for a therapeutic trial from Aerogen and Fisher & Paykel; consulting fees from Aridis Pharmaceuticals; payment or honoraria from Vidal; and support for attending meetings or travel from AOP Health. GS-C declares an institutional research grant from Grifols; lecture honoraria from TEVA, Zambon, Boehringer Ingelheim, and Insmed; travel or conference grants from Chiesi, Boehringer Ingelheim, TEVA, Pari, and Insmed; participation on advisory boards for Zambon and Insmed; and is an HERMES examination committee member (European Respiratory Society) and Director of the bronchiectasis integrated research projects board (Spanish respiratory society). EWS declares royalties for a book with Springer. CAB declares participation on an advisory board for Novartis Ocular regarding brolucizumab in age-related macular degeneration or diabetic macular oedema (August, 2022). AT declares consulting fees and payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing, and educational events from Pfizer, Poliphor, MSD, Janssen, and OM Pharma. DvK declares a Local Erasmus MC University Medical Center Grant. CMS is funded by the Swiss National Science Foundation (grant number 323530_221860). All other authors declare no competing interests.

Data sharing

Data from individual trials were provided by trial groups for the specific purpose of conducting this individual patient data meta-analysis. Any requests by other researchers for those data should be directed to the responsible party for individual trials.

References

1 GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 2020; **396:** 1204–22.

2 Kellum JA, Kong L, Fink MP, et al. Understanding the inflammatory cytokine response in pneumonia and sepsis: results of the Genetic and Inflammatory Markers of Sepsis (GenIMS) Study. *Arch Intern Med* 2007; **167:** 1655–63.

3 Martin-Loeches I, Torres A, Nagavci B, et al. ERS/ESICM/ESCMID/ALAT guidelines for the management of severe community-acquired pneumonia. *Intensive Care Med* 2023; **49:** 615–32.

4 Chaudhuri D, Nei AM, Rochwerg B, et al. 2024 focused update: guidelines on use of corticosteroids in sepsis, acute respiratory distress syndrome, and community-acquired pneumonia. *Crit Care Med* 2024; **52:** e219–33.

5 Pirracchio R, Venkatesh B, Legrand M. Low-dose corticosteroids for critically ill adults with severe pulmonary infections: a review. *JAMA* 2024; **332:** 318–28.

6 Kent DM, Paulus JK, van Klaveren D, et al. The Predictive Approaches to Treatment effect Heterogeneity (PATH) statement. *Ann Intern Med* 2020; **172:** 35–45.

7 Briel M, Spoorenberg SMC, Snijders D, et al. Corticosteroids in patients hospitalized with community-acquired pneumonia: systematic review and individual patient data metaanalysis. *Clin Infect Dis* 2018; **66:** 346–54.

8 Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted? *Stat Med* 2002; **21:** 1559–73.

9 Riley RD, Tierney JF, Stewart LA, eds. Individual participant data meta-analysis: a handbook for healthcare research. John Wiley & Sons, 2021.

10 Bica I, Alaa AM, Lambert C, van der Schaar M. From real-world patient data to individualized treatment effects using machine learning: current and future methods to address underlying challenges. *Clin Pharmacol Ther* 2021; **109:** 87–100.

11 Stewart LA, Clarke M, Rovers M, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Individual Participant Data: the PRISMA-IPD statement. *JAMA* 2015; **313:** 1657–65.

12 Smit J, Endeman H, van der Zee P. Heterogeneity of Treatment Effect of Corticosteroids in Community Acquired Pneumonia patients (HITEC-CAP): a individual patient data meta-analysis. PROSPERO CRD42022380746. 2022. https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=380746 (accessed Jan 7, 2025).

13 Smit JM, Van Der Zee PA, Stoof SCM, et al. Predicting individualized treatment effects of corticosteroids in community-acquired-pneumonia: a data-driven analysis of randomized controlled trials. *medRxiv* 2023; published online Oct 3. https://doi.org/10.1101/2023.10.03.23296132 (preprint).

14 Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019; **366:** l4898.

15 Confalonieri M, Urbino R, Potena A, et al. Hydrocortisone infusion for severe community-acquired pneumonia: a preliminary randomized study. *Am J Respir Crit Care Med* 2005; **171:** 242–48.

16 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:** 975–82.

17 Meijvis SCA, Hardeman H, Remmelts HHF, et al. Dexamethasone and length of hospital stay in patients with community-acquired pneumonia: a randomised, double-blind, placebo-controlled trial. *Lancet* 2011; **377:** 2023–30.

18 Sabry NA, Omar EE-D. Corticosteroids and ICU course of community acquired pneumonia in Egyptian settings. *Pharmacol Pharm* 2011; **2:** 73–81.

19 Nafae RM, Ragab MI, Amany FM, Rashed SB. Adjuvant role of corticosteroids in the treatment of community-acquired pneumonia. *Egypt J Chest Dis Tuberc* 2013; **62:** 439–45.

20 Blum CA, Nigro N, Briel M, et al. Adjunct prednisone therapy for patients with community-acquired pneumonia: a multicentre, double-blind, randomised, placebo-controlled trial. *Lancet* 2015; **385:** 1511–18.

21 Torres A, Sibila O, Ferrer M, 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; **313:** 677–86.

22 Wittermans E, Vestjens SMT, Spoorenberg SMC, 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:** 1–10.

23 Meduri GU, Shih M-C, Bridges L, et al. Low-dose methylprednisolone treatment in critically ill patients with severe community-acquired pneumonia. *Intensive Care Med* 2022; **48:** 1009–23.

24 Dequin P-F, Meziani F, Quenot J-P, et al. Hydrocortisone in severe community-acquired pneumonia. *N Engl J Med* 2023; **388:** 1931–41.

25 Fine MJ, Auble TE, Yealy DM, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med* 1997; **336:** 243–50.

26 Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax* 2003; **58:** 377–82.

27 Endeman H, Meijvis SCA, Rijkers GT, et al. Systemic cytokine response in patients with community-acquired pneumonia. *Eur Respir J* 2011; **37:** 1431–38.

28 van Klaveren D, Balan TA, Steyerberg EW, Kent DM. Models with interactions overestimated heterogeneity of treatment effects and were prone to treatment mistargeting. *J Clin Epidemiol* 2019; **114:** 72–83.

29 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; **109:** 1517–32.

30 Efthimiou O, Hoogland J, Debray TPA, et al. Measuring the performance of prediction models to personalize treatment choice. *Stat Med* 2023; **42:** 1188–206.

31 Riley RD, Debray TPA, Fisher D, 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:** 2115–37.

32 Morris TP, Walker AS, Williamson EJ, White IR. Planning a method for covariate adjustment in individually randomised trials: a practical guide. *Trials* 2022; **23:** 328.

33 SciPy. kruskal. https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.mstats.kruskal.html#scipy.stats.mstats.kruskal (accessed Jan 7, 2025).

34 Athey S, Tibshirani J, Wager S. Generalized random forests. *Ann Stat* 2019; **47:** 1179–203.

35 Künzel SR, Sekhon JS, Bickel PJ, Yu B. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proc Natl Acad Sci USA* 2019; **116:** 4156–65.

36 Nie X, Wager S. Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika* 2021; **108:** 299–319.

37 Fernández-Serrano S, Dorca J, Garcia-Vidal C, et al. Effect of corticosteroids on the clinical course of community-acquired pneumonia: a randomized controlled trial. *Crit Care* 2011; **15:** R96.

38 Lloyd M, Karahalios A, Janus E, 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:** 1052–60.

39 Groenwold RHH, Moons KGM, Vandenbroucke JP. Randomized trials with missing outcome data: how to analyze and what to report. *CMAJ* 2014; **186:** 1153–57.

40 Dahabreh IJ, Hayward R, Kent DM. Using group data to treat individuals: understanding heterogeneous treatment effects in the age of precision medicine and patient-centred evidence. *Int J Epidemiol* 2016; **45:** 2184–93.

41 Harrell F. Viewpoints on heterogeneity of treatment effect and precision medicine. June 4, 2018. https://www.fharrell.com/post/hteview/ (accessed Jan 7, 2025).

42 van Klaveren D, Vergouwe Y, Farooq V, Serruys PW, Steyerberg EW. Estimates of absolute treatment benefit for individual patients required careful modeling of statistical interactions. *J Clin Epidemiol* 2015; **68:** 1366–74.

43 Steyerberg EW, Harrell FE Jr. Prediction models need appropriate internal, internal–external, and external validation. *J Clin Epidemiol* 2016; **69:** 245–47.

44 Qadir N, Sahetya S, Munshi L, et al. An update on management of adult patients with acute respiratory distress syndrome: an official American Thoracic Society Clinical Practice guideline. *Am J Respir Crit Care Med* 2024; **209:** 24–36.

45 Evans L, Rhodes A, Alhazzani W, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Crit Care Med* 2021; **49:** 1974–82.

46 Lansbury LE, Rodrigo C, Leonardi-Bee J, Nguyen-Van-Tam J, Lim WS. Corticosteroids as adjunctive therapy in the treatment of influenza: an updated Cochrane systematic review and meta-analysis. *Crit Care Med* 2020; **48:** e98–106.

47 Moreno G, Rodríguez A, Reyes LF, et al. Corticosteroid treatment in critically ill patients with severe influenza pneumonia: a propensity score matching study. *Intensive Care Med* 2018; **44:** 1470–82.

48 Ebell MH, Bentivegna M, Cai X, Hulme C, Kearney M. Accuracy of biomarkers for the diagnosis of adult community-acquired pneumonia: a meta-analysis. *Acad Emerg Med* 2020; **27:** 195–206.

49 Chalmers JD, Singanayagam A, Hill AT. C-reactive protein is an independent predictor of severity in community-acquired pneumonia. *Am J Med* 2008; **121:** 219–25.

50 Bouras M, Rello J, Roquilly A. Steroids in severe community-acquired pneumonia: dangerous, worthless, or miracle cure? The roller coaster of clinical trials. *Anaesth Crit Care Pain Med* 2023; **42:** 101253.

51 Odeyemi YE, Herasevich S, Chalmers SJ, et al. Biomarker-concordant steroid use in critically ill patients with pneumonia. *Mayo Clin Proc Innov Qual Outcomes* 2020; **4:** 649–56.

52 Schandelmaier S, Briel M, Varadhan R, et al. Development of the Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in randomized controlled trials and meta-analyses. *CMAJ* 2020; **192:** E901–06.

53 van Klaveren D, Steyerberg EW, Serruys PW, Kent DM. The proposed 'concordance-statistic for benefit' provided a useful metric when modeling heterogeneous treatment effects. *J Clin Epidemiol* 2018; **94:** 59–68.