The Lancet Rheumatology

Immunoglobulin, Glucocorticoid, or combination therapy for Multisystem Inflammatory Syndrome in Children: A propensity weighted cohort study

--Manuscript Draft--

Manuscript Number:	TLRHEU-D-22-00664R1
Article Type:	FAST-TRACK Article (Original Research)
Keywords:	Multisystem Inflammatory Syndrome in Children; Immunomodulator therapy; Intravenous Immunoglobulin; Steroids; Propensity score analysis
Corresponding Author:	Michael Levin Imperial College London London, UNITED KINGDOM
First Author:	Samuel Channon-Wells, MMath
Order of Authors:	Samuel Channon-Wells, MMath
	Ortensia Vito, MSc
	Andrew J McArdle, MSc
	Eleanor G Seaby, BMBS
	Harsita Patel, MBBS
	Priyen Shah, MB BS
	Ekaterina Pazukhina, MSc
	Clare Wilson, MB BChir
	Claire Broderick, MBBS
	Giselle D'Souza, MSc
	Ilana Keren
	Ruud G Nijman, PhD
	Adriana Tremoulet, MAS
	Daniel Munblit, PhD
	Rolando Ulloa-Gutierrez, MD
	Michael J Carter, DPhil
	Padmanabhan Ramnarayan, MD
	Tisham De, PhD
	Clive Hoggart, PhD
	Elizabeth Whittaker, PhD
	Jethro A Herberg, PhD
	Myrsini Kaforou, PhD
	Aubrey J Cunnington, PhD
	Oleg Blyuss, PhD
	Michael Levin
Manuscript Region of Origin:	UNITED KINGDOM
Abstract:	Background SARS-CoV-2 associated Multisystem Inflammatory Syndrome in Children (MIS-C) has emerged as a serious illness in children world-wide. Immunoglobulin and/or glucocorticoids are currently recommended treatments.

Methods
The “Best Available Treatment Study” evaluated immunomodulatory treatments for MIS-C in an international observational cohort. Analysis of the first 614 patients was previously reported. Clinical and outcome data were collected onto a web-based database. Inverse probability weighting was used to compare primary treatments with intravenous immunoglobulin (IVIG), IVIG plus glucocorticoids (IVIG+G), or glucocorticoids alone, using IVIG as the reference treatment. Primary outcomes were: a composite of inotropic or ventilator support from the second day after treatment initiation, or death; and time-to-improvement on an ordinal clinical severity scale. Secondary outcomes included treatment escalation, clinical deterioration, fever, and coronary artery aneurysm occurrence and resolution.
Findings
After exclusions, 2009 children with clinically diagnosed MIS-C from 39 countries were enrolled between May 2020 and April 2022. 680 received primary treatment with IVIG; 698 IVIG+G; 487 glucocorticoids alone; 59 other combinations including biologics, and 85 no immunomodulator.
There were no significant differences between treatments for primary outcomes for the 1586 patients considered for primary analysis: adjusted odds ratios relative to IVIG for ventilation, inotropic support or death were 1·09 (95% confidence interval [CI] 0·75-1·58) and 0·93 (95% CI: 0·58-1·47) for IVIG+G and glucocorticoids alone respectively. Adjusted average hazard ratios for time-to-improvement were 1·04 (95% CI: 0·91-1·20) and 0·84 (95% CI: 0·70-1·00) for the same comparisons. Treatment escalation was less frequent for IVIG+G and glucocorticoids alone vs IVIG. Persistent fever was less common with IVIG+G compared with either IVIG or glucocorticoids alone. Coronary artery aneurysm occurrence and resolution did not differ significantly between treatment groups.
Interpretation
Recovery rates, including occurrence and resolution of coronary artery aneurysms, were similar for primary treatment with IVIG when compared to glucocorticoids or combination IVIG+G. Initial treatment with glucocorticoids appears to be a safe alternative to immunoglobulin or combined therapy, and may be advantageous in view of the cost and limited availability of IVIG in many countries.

Immunoglobulin, Glucocorticoid, or combination therapy for Multisystem Inflammatory Syndrome in Children: A propensity weighted cohort study

Samuel Channon-Wells MMath[1,5]*, Ortensia Vito MSc[1,5]*, Andrew J. McArdle MSc[1,5]*, Eleanor G. Seaby BMBS[1,2,3+], Harsita Patel MBBS[1,5+], Priyen Shah MB BS[1,5], Ekaterina Pazukhina MSc**, Clare Wilson MB BChir[1,5], Claire Broderick MBBS[1,5], Giselle D'Souza MSc[1,5], Ilana Keren[1], Ruud G. Nijman PhD[1,4,5], Adriana Tremoulet MAS[6], Daniel Munblit PhD[7,8], Rolando Ulloa-Gutierrez MD[9,10,11], Michael J Carter DPhil[12], Padmanabhan Ramnarayan MD[13], Tisham De PhD[1], Clive Hoggart PhD[14], Elizabeth Whittaker PhD[1,5,15], Jethro A. Herberg PhD[1,5,15], Myrsini Kaforou PhD[1,5], Aubrey J. Cunnington PhD[1,5,15], Oleg Blyuss PhD[16], Michael Levin FMedSci[1,5,15] and **The Best Available Treatment Study for MIS-C (BATS) consortium**

*,+ Contributed equally

[1] Department of Infectious Disease, Section of Paediatric Infectious Disease, Imperial College London, London, United Kingdom

[2] Genomic Informatics Group, University of Southampton, Southampton, United Kingdom

[3] Translational Genomics Group, Broad Institute of MIT and Harvard, Cambridge, MA, USA

[4] Department of Paediatric Emergency Medicine, Division of Medicine, St. Mary's hospital - Imperial College NHS Healthcare Trust, London, London, London, UK

[5] Centre for Paediatrics and Child Health, Imperial College, London, UK

[6] Department of Paediatrics, University of California San Diego/Rady Children's Hospital San Diego

[7] Department of Paediatrics and Paediatric Infectious Diseases, Institute of Child's Health, Sechenov First Moscow State Medical University (Sechenov University), Moscow, Russia

[8] Inflammation, Repair, and Development Section, National Heart and Lung Institute, Faculty of Medicine, Imperial College London, London, United Kingdom

[9] Servicio de Infectologia Pediatrica, Hospital Nacional de Niños "Dr. Carlos Sáenz Herrera", Centro de Ciencias Médicas, Caja Costarricense de Seguro Social (C.C.S.S.), San José, Costa Rica

[10] Instituto de Investigación en Ciencias Médicas UCIMED (IICIMED), San José, Costa Rica

[11] Cátedra de Pediatría, Facultad de Medicina, Universidad de Ciencias Médicas (UCIMED), San José, Costa Rica

[12] Department of Women and Children's Health, School of Life Course Sciences, King's College London, St Thomas' Hospital, SE1 7EH, London, UK

[13] Anaesthetics, Pain Medicine and Intensive Care (APMIC) Division, Department of Surgery and Cancer, Faculty of Medicine, Imperial College

[14] Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York City, NY USA

[15] Department of Paediatrics, Imperial College Healthcare NHS Trust, London, United Kingdom

[16] Wolfson Institute of Population Health, Queen Mary University of London

**No current institutional affiliation

Corresponding author: Professor Michael Levin, Department of Infectious Disease, Imperial College London. Email: m.levin@imperial.ac.uk

ABSTRACT

Background

SARS-CoV-2 associated Multisystem Inflammatory Syndrome in Children (MIS-C) has emerged as a serious illness in children world-wide. Immunoglobulin and/or glucocorticoids are currently recommended treatments.

Methods

The “Best Available Treatment Study” evaluated immunomodulatory treatments for MIS-C in an international observational cohort. Analysis of the first 614 patients was previously reported. Clinical and outcome data were collected onto a web-based database. Inverse probability weighting was used to compare primary treatments with intravenous immunoglobulin (IVIG), IVIG plus glucocorticoids (IVIG+G), or glucocorticoids alone, using IVIG as the reference treatment. Primary outcomes were: a composite of inotropic or ventilator support from the second day after treatment initiation, or death; and time-to-improvement on an ordinal clinical severity scale. Secondary outcomes included treatment escalation, clinical deterioration, fever, and coronary artery aneurysm occurrence and resolution.

Findings

After exclusions, 2009 children with clinically diagnosed MIS-C from 39 countries were enrolled between May 2020 and April 2022. 680 received primary treatment with IVIG; 698 IVIG+G; 487 glucocorticoids alone; 59 other combinations including biologics, and 85 no immunomodulator.

There were no significant differences between treatments for primary outcomes for the 1586 patients considered for primary analysis: adjusted odds ratios relative to IVIG for ventilation, inotropic support or death were 1·09 (95% confidence interval [CI] 0·75-1·58) and 0·93 (95% CI: 0·58-1·47) for IVIG+G and glucocorticoids alone respectively. Adjusted average hazard ratios for time-to-improvement were 1·04 (95% CI: 0·91-1·20) and 0·84 (95% CI: 0·70-1·00) for the same

comparisons. Treatment escalation was less frequent for IVIG+G and glucocorticoids alone vs IVIG. Persistent fever was less common with IVIG+G compared with either IVIG or glucocorticoids alone. Coronary artery aneurysm occurrence and resolution did not differ significantly between treatment groups.

Interpretation

Recovery rates, including occurrence and resolution of coronary artery aneurysms, were similar for primary treatment with IVIG when compared to glucocorticoids or combination IVIG+G. Initial treatment with glucocorticoids appears to be a safe alternative to immunoglobulin or combined therapy, and may be advantageous in view of the cost and limited availability of IVIG in many countries.

Funding and sponsors

The BATS writing group receive salary support from Imperial College, or are supported on individual grants, fellowships and project grants including the European Union's Horizon 2020 programme under GA No. 848196 DIAMONDS, Wellcome Trust (AJM, CB, Imperial 4i Wellcome Trust/NIHR Imperial BRC, CW; Wellcome Trust, MK 206508/Z/17/Z), the Medical Research Foundation (MRF-160-0008-ELP-KAFO-C0801), UK NIHR (RGN ACL-2018-021-007; MJC ACL-2018) and NIH (GA5R01AI128765). Enrolment of patients in London was supported by NIHR Imperial Biomedical Research Centre. Imperial College London was the main research sponsor for this study.

Role of the funding source

The funders and sponsor of the study had no role in study design, data collection, analysis, interpretation, or writing of the final manuscript.

Trial Registration Number

ISRCTN registry - ISRCTN69546370 https://doi.org/10.1186/ISRCTN69546370

Contributors

ML, AJM, OV, AC, EW, JH, MK, HP, PS, CW, RN, TD, and CH conceived the original study. OV, SCW, CB, ES, HP, GS and IK undertook data checks and quality control. OV, SCW and ML accessed and verified the underlying data reported in the manuscript. SCW, OB, AJM and ML conceived the current statistical analysis plan, which was reviewed by all co-authors prior to publication. SCW, ES, EP, and HP undertook the current analysis. SCW and ML prepared the manuscript. All authors had full access to all the data in the study. All authors read and approved the final manuscript and had final responsibility for the decision to submit for publication.

Conflict of Interest Statements/Declaration of interests

AT has provided unpaid consultancy work for Janssen Pharmaceuticals. DM has received grant support from the British Embassy in Moscow ('StopCOVID Cohort: Clinical Characterisation of

Russian Patients') and holds the following unpaid positions: Co-Chair of International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC) Global Paediatric Long COVID Working Group, Member of ISARIC working group on long-term follow-up in adults, Co-lead of the PC-COS project aiming to define the Core Outcome Set for Long-COVID, in collaboration with the WHO. MJC reports a personal fee from BioTest for speaking at the BioTest Immunology Forum 2022, Royal Society (www.biotest.com). EW holds the following unpaid positions: member of the paediatric steering committee for the RECOVERY trial; paediatric representative for NHS England working on the National paediatric virtual advisory network and expert advisory group for COVID treatment, Independent advisory group for COVID monoclonal antibodies, and co-lead for the pan-London Post-COVID service for children. All other authors declare no competing interests.

Ethics Committee approval

Please see "oversight" section in main text for relevant details.

RESEARCH IN CONTEXT

Evidence before this study

In the first wave of the COVID-19 pandemic paediatricians around the world rapidly identified and described a new inflammatory disorder, causing shock and multi-system failure in children approximately 4-6 weeks after SARS-CoV-2 infection. Faced with this new life-threatening disorder, termed multisystem inflammatory syndrome in children (MIS-C), with unknown pathophysiological mechanisms, paediatricians, national, and international paediatric bodies rapidly adopted treatments which are of benefit in other inflammatory disorders.

Based on the similarity in clinical features of MIS-C to Kawasaki Disease (KD), intravenous immunoglobulin (IVIG), the recognised treatment for KD, was adopted as the most widely used initial treatment, often combined with glucocorticoids and a range of biological agents. In the absence of data from randomised controlled trials (RCT), national and international organisations, including the World Health Organisation (WHO), American College of Rheumatology, and UK Royal College of Paediatrics and Child Health (RCPCH) produced treatment guidelines recommending IVIG as initial treatment, combined with glucocorticoids or biological agents for the most seriously ill or unresponsive patients.

We searched for publications on treatment of MIS-C (and the alternative name Paediatric Multisystem Inflammatory Syndrome Temporally associated with SARS-CoV-2 (PIMS-TS)) since April 2020 when the disorder was first recognised, until November 2022. In the extensive literature now published on MIS-C, there are many hundreds of observational studies, treatment recommendations and guidelines based on expert opinion, and reports of outcome after treatment. However, we found no RCTs, and only four propensity matched comparisons reporting outcomes after specific treatments, only two of which included comparison of glucocorticoids alone and IVIG, and all were based on relatively small patient cohorts.

Added value of this study

The Best Available Treatment Study (BATS) allowed us to compare treatment of MIS-C with IVIG alone, glucocorticoids alone, and combined glucocorticoids plus IVIG (combined therapy), in over 2000 patients from 39 different countries. This is the largest study to date of immunomodulator treatment options in MIS-C, including the largest cohort of patients treated initially with glucocorticoid monotherapy. After correcting for known confounders using propensity score weighting, initial treatment with glucocorticoid monotherapy or combined therapy demonstrated no significant difference to treatment with IVIG monotherapy in either time-to-improvement measured on an ordinal clinical severity scale, or in a composite outcome of inotropic support or ventilator support (invasive or non-invasive) from the second day after starting treatment or later, or death. Comparison of glucocorticoid monotherapy with combined therapy suggested a small benefit from combined therapy in time-to-improvement, but this appeared to be restricted to those who did not require inotropic and/or ventilatory support at baseline. Combined therapy was associated with faster fever resolution and less escalation of treatment, but with no other differences in secondary outcomes. Occurrence and resolution of coronary artery aneurysms was similar in all treatment groups, with the large majority of aneurysms resolving during follow up.

Implications of all the available evidence

Our study increases confidence that initial treatment of MIS-C with glucocorticoids is associated with similar outcomes to treatment with IVIG or combined therapy. In the context of all current observational data, there is, at best, only a small benefit in initial therapy combining IVIG and glucocorticoids compared to monotherapy with IVIG or glucocorticoids alone. Given the high cost and limited availability of IVIG in many countries this evidence supports initial glucocorticoid monotherapy as an acceptable alternative.

BACKGROUND

Since recognition in April 2020, Multi-system Inflammatory Syndrome in Children (MIS-C), temporally associated with SARS-CoV-2 infection,[1–4] has emerged as a rare but serious post-infectious illness.[5–8] In the absence of evidence from randomised controlled trials (RCT), treatment recommendations for the new disease were developed by clinical consensus in many countries. Based on similarity of MIS-C to Kawasaki disease (KD), for which Intravenous immunoglobulin (IVIG) is the established treatment,[9] national and international guidance has recommended IVIG as initial treatment, with addition of glucocorticoids and/or other immunomodulatory agents for patients with severe illness.[10,11]

While there have been no RCTs comparing treatments for MIS-C published to date, several observational studies using propensity score methods have suggested that combination treatment with IVIG and glucocorticoids was associated with improved cardiac outcomes.[12–14]

The Best Available Treatment Study (BATS) was initiated in the early months after first recognition of MIS-C and aimed to provide evidence for treatment recommendations by systematic data collection, and analysis of outcomes of treatments chosen by individual paediatricians responsible for patient care. In view of the urgent need for evidence to support treatment recommendations, analysis of the first 614 patients enrolled in BATS was reported in July 2021.[15] No significant differences in outcome were observed between patients treated with IVIG alone, glucocorticoids alone, or combination of IVIG and glucocorticoids (IVIG+G), although this may have been due to limited sample size. In this report, we compare the initial treatments for MIS-C in a much larger cohort of children, and also describe the outcomes of cardiac complications.

METHODS

Study Design

Details of the BATS propensity-weighted observational cohort study were described in the initial report.[15] Minor modifications of the data collection procedure and analysis plan were undertaken which are described below and in the published analysis plan and supplementary appendix. Briefly, paediatricians world-wide were invited to join BATS and upload data from patients with suspected MIS-C onto a web-based Research Electronic Data Capture database,[16] from June 2020 through to April 2022. As the spectrum of post-SARS-CoV-2 inflammatory disease was unknown when BATS was initiated,[3,5,17–19] and the reliability of the published criteria for MIS-C was unknown we invited recruitment of children with severe inflammatory illness after SARS-CoV-2 infection in addition to those meeting the USA Centre for Disease Control (CDC), WHO or UK case definitions.[20–22] De-identified longitudinal data were collected on presenting features, demography, laboratory findings, immunomodulatory (IVIG, glucocorticoids or biologicals) and supportive treatments. Treatments and daily data were collected by calendar day. Duration of admission, organ support required, and health status on discharge were recorded.

The original BATS case report form recorded no data on coronary artery aneurysms (CAA) after hospital discharge, and we therefore added an additional follow up questionnaire regarding CAA resolution (appendix p68).

Treatments and endpoints

The first calendar day of immunomodulatory treatment was defined as “day 0”, and subsequent treatment and outcomes defined relative to this. Primary treatment was defined as the immunomodulatory agent(s) initiated on day 0. Three primary treatment groups were large enough for weighted comparison according to our predefined sample-size estimations (appendix p69): IVIG alone, glucocorticoids alone, or IVIG+G. Two other groups were pre-

defined for additional analyses: those receiving other immunomodulator treatments (including in combination with IVIG and/or glucocorticoids), or no immunomodulator treatments.

Primary outcomes were modified from the previous analysis. The first primary outcome remained a composite of inotropic support or ventilator support (invasive or non-invasive) on day 2 or later, or death. However, the second primary outcome was altered from improvement on the ordinal severity scale by day 2, to time to improvement of at least one level on the ordinal clinical severity scale (ventilated and on inotropic support; ventilated; on inotropic support; receiving oxygen; no supportive therapy stratified by CRP level; and discharged – appendix p14). This modification was justified by the greater clinical relevance and additional statistical power of the time to event analysis.

Secondary outcomes included: immunomodulator escalation (any additional immunomodulator, a second dose of IVIG if primary treatment included IVIG, and if primary treatment included glucocorticoids, an increment of 5 mg/kg equivalent daily-dose of prednisolone)[23]; fever from day 2 onwards; individual components of the first primary outcome (death, or inotropic or ventilator support from day 2); CAA occurrence and resolution following treatment (coronary artery Z-score ≥2·5 or aneurysm documented)[24]; left ventricular (LV) dysfunction on echocardiography from day 2 onwards; no improvement in clinical severity scale at day 2; any increase in cardiorespiratory supportive therapy after day 0; therapeutic complications; and temporal dynamics of blood markers of inflammation and organ damage.

Analysis and Statistics

We applied inverse probability of treatment weighting (IPTW) using covariate-balancing propensity scores[25] to account for baseline differences between the three primary treatment groups. Confounding covariates were selected by expert consensus prior to analysis and were used in both covariate balancing and treatment effect estimation to produce doubly-robust estimates (appendix p18-20). As specified in the analysis plan, IVIG alone was the reference

treatment group. Weighted quasibinomial logistic-regression was used for dichotomous outcomes and weighted Cox-regression for time-to-event analyses. Outcomes were reported as adjusted odds ratios or average hazard ratios with 95% confidence intervals and p-values. P-value correction for multiple hypothesis testing was performed for the two primary outcomes and two treatment-group comparisons with the Bonferroni-Holm procedure (appendix p20).

All clinician-diagnosed MIS-C cases were included in analysis, with those meeting more restrictive definitions evaluated in subgroup or sensitivity analyses: restricting to patients meeting the WHO MIS-C criteria,[22] those meeting KD criteria; subgroups by age category and baseline inflammation; analysis by propensity score matching; and defining primary treatments as those received on days 0 and 1. Extensive additional subgroup and sensitivity analyses were performed as planned (appendix p21-22).

Inflammatory markers were plotted as percentages of each patient's peak value by admission day relative to treatment initiation. Smoothed curves with confidence intervals were weighted by the same approach and fitted using the generalized additive model method (appendix p18).

Oversight

BATS was designed by the study team at Imperial College London (members and roles in appendix p3). Patient data were collected by local investigators (consortium members in appendix p3-11). The updated statistical analysis plan was developed by the study management team and international advisory board, and analysis undertaken by the statistical group (appendix p3). The study was approved by the UK REC (20/HRA/2957) and registered with the international trial registry (ISRCTN69546370). Participating centers obtained ethical approval based on requirements in each country. The initial manuscript was drafted by the first and last authors and developed by all listed authors. The corresponding author, data management group, and analysis group had access to all data, vouching for the completeness and accuracy of data, and for fidelity to the protocol and analysis plan.

RESULTS

From 20th June 2020 to 25th April 2022, data from 2101 MIS-C patients from 39 countries and 121 sites were uploaded to BATS (appendix p48-51). 92 records were excluded, including four neonates and those with incomplete data, duplicate entries, or admission after the recruitment deadline (Fig1A). Of 2009 patients included for analysis, 680 received primary treatment with IVIG, 487 with glucocorticoids, 698 with combination IVIG+G, 59 received other immunomodulator combinations, and 85 received no immunomodulators (Fig1A). In the three main primary treatment groups, 579/1865 (31·0%) received additional immunomodulators by day 2, with 953/1865 (51.1%) receiving secondary agents in total. Treatment trajectories are described in detail (Fig1B, appendix p23).

Clinical and laboratory findings

Baseline clinical and laboratory findings showed some differences between primary treatment groups (Table 1, appendix p24,52). Patients in the no therapy group had significantly less derangement in laboratory markers of inflammation and organ dysfunction, while those in the combined IVIG+G and other treatments groups had the highest level of derangement overall. The combined IVIG+G and other immunomodulator groups had a higher proportion of patients receiving inotropes or ventilation on day 0 (appendix p55). Considering treatment received by day 2, a higher proportion of those on both IVIG+G or in whom biological agents were added were receiving inotropes or ventilated at baseline (appendix p55), but there were no major differences seen in blood markers between these groups (appendix p53).

1602/2009 (80·0%) patients met WHO MIS-C criteria (appendix p26). The most common missing criterion was evidence of SARS-CoV-2 exposure (appendix p56). SARS-CoV-2 antibody measurements were not tested in 406/2009 (20·4%), and negative in 259/2009 (13·0%). Bacteria were cultured in the blood of a small proportion of patients (appendix p26).

629/2009 (31·3%) overall, and 544/1602 (34·0%) of those meeting WHO MIS-C criteria also met the American Heart Association (AHA) definitions for complete KD (appendix p28,57).

Primary outcomes

Of 1865 patients in the three main treatment groups, 166 patients (9·0%) received immunomodulators prior to transfer to the reporting hospital and an additional 113 patients (6.1%) were missing baseline covariates, with a total of 1586 patients considered for our primary weighted analyses (Fig1A). Acceptable covariate balance was achieved for all IPTW outcome analyses (appendix p60-61,66). For the first primary outcome, receipt of inotropic support or ventilation on day 2 or later, or death, the adjusted odds ratios (OR) for patients receiving primary treatment with IVIG+G, or glucocorticoids alone as compared with IVIG were 1·09 (95% CI: 0·75-1·58, adjusted p-value 1·00) and 0·93 (95% CI: 0·58-1·47, adjusted p-value 1·00) respectively (Fig2A & 2C, appendix p29).

For the second primary outcome, time to improvement on the ordinal clinical severity scale, the adjusted average hazard ratio (AHR) for patients receiving IVIG+G vs IVIG was 1·04 (95% CI: 0·91-1·20, adjusted p-value 1·00) and for glucocorticoids alone vs IVIG was 0·84 (95% CI: 0·70-1·00, adjusted p-value 0·22, Fig2B, 2D & 3A, appendix p29), suggesting slower improvement in the glucocorticoid group. Subgroup analyses of time to improvement in severely ill children (requiring ventilatory or inotropic support at baseline), and those not requiring intensive support showed the suggested slower improvement in those receiving glucocorticoids vs combined treatment was confined to the less severely ill patients (AHR 1·06 (95% CI: 0·75-1·49) in the severe group vs 0·83 (95% CI: 0·62-1·11) in the milder group, Fig2B, 2D & 3B-C, appendix p34).

All sensitivity and subgroup analyses, including restricting to patients meeting WHO MIS-C criteria, showed no significant difference in the first primary outcome for the comparisons of IVIG+G or glucocorticoids alone with IVIG (Fig2A & 2C, appendix p32). For the second primary

outcome, in the subgroup of patients without significant comorbidities, the time-to-improvement was slower in the glucocorticoid group vs IVIG alone (AHR 0·82 (95% CI: 0·69-0·99), Fig2B & 2D, appendix p34) and the two-point time-to-improvement was slower in the IVIG+G group vs IVIG alone (AHR 0·87 (95% CI: 0·75-1·00)). All other planned sensitivity and subgroup analyses showed no significant difference in time-to-improvement for the comparisons of IVIG+G or glucocorticoids alone with IVIG alone.

Secondary outcomes

Escalation of immunomodulator treatment was less common in the IVIG+G and glucocorticoid groups compared to the IVIG group (OR 0·15 (95% CI: 0·11-0·20) & 0·68 (95% CI: 0·50-0·93) respectively, appendix p58). Persistent fever from day 2 was less common in patients receiving IVIG+G vs IVIG alone (OR 0·50 (95% CI: 0·38-0·67)), with no difference between the glucocorticoid or IVIG groups. In a post-hoc sensitivity analysis, there was no difference in persistent fever from day 3 between the IVIG+G and IVIG groups. Individual components of the composite outcome showed no differences between treatments (Fig2A & 2C, appendix p30). Of 1918 with reported echocardiograms, 236 (12·3%) had CAA at any time (13·6% in IVIG recipients, 8·9% glucocorticoid, and 12·9% IVIG+G (appendix p36)), with the largest disparity in aneurysm detection before starting immunomodulatory treatment (appendix p36). In the 705 patients with inpatient echocardiograms before and after treatment initiation (appendix p30) 50 (7.1%) had CAA present on the final echocardiogram before discharge, with no statistically significant difference apparent between groups after IPTW analysis, including for post-hoc analyses restricted to patients who did and did not meet complete KD criteria (appendix p58). Follow-up echocardiogram data were available in 196/236 (83·1%) patients with CAA during admission. Most CAA resolved during follow-up (92·9% total), with similar rates amongst primary treatment groups (appendix p36). Similar rates of resolution were seen when restricted to patients with follow-up by 6- and 12-weeks (appendix p37).

To establish if patients who did not receive IVIG were at greater risk of CAA, or had different rates of resolution, we explored CAA incidence in the glucocorticoid alone primary treatment group. 17/239 (7·1%) of those never receiving IVIG had CAA detected at any time during admission, compared with 24/221 (10·9%) who received IVIG later during admission. CAA were present at discharge in 5/239 (2·1%) of those without later IVIG and 9/221 (4·1%) of those receiving later IVIG treatment, with over 93% of CAA resolving in both groups on reported follow-up (appendix p38). No difference was seen between treatment groups in severity of CAA as judged by the distribution of z-scores (appendix p39). Larger z-scores were seen in younger patients (appendix p39).

Left ventricular dysfunction was reported in 202/1512 (13·4%) of patients with echocardiograms from day 2 onwards, with no difference between groups (appendix p30,58). There were no differences between IVIG+G or glucocorticoids vs IVIG for the secondary outcomes of no improvement by day 2 or increase in level of support after initiation of primary treatment. Death occurred in 8 (1·2% unadjusted), 10 (2·1%) and 5 (0·8%) patients in the complete IVIG+G, glucocorticoid and IVIG groups respectively.

Drug complications were reported in 59/1623 (3·6%) of patients receiving any glucocorticoids and 25/1658 (1·5%) patients receiving IVIG. Glucocorticoid complications were predominantly hypertension and hyperglycemia (appendix p40).

IVIG+G vs Glucocorticoids alone

A planned secondary analysis comparing glucocorticoids alone and combined IVIG+G demonstrated no difference in the first primary outcome, but a faster time-to-improvement for the IVIG+G group (AHR 1·25 (95% CI: 1·05-1·48), appendix p59). This was predominantly seen in the later days following treatment, and in those patients not requiring intensive support at baseline (Fig3A-C). Secondary outcomes for this comparison showed that escalation of primary

therapy and persistence of fever from day 2 were more common in the glucocorticoid alone group (appendix p30,59).

Effect of Immunomodulation on blood markers.

CRP declined more rapidly in patients receiving immunomodulators than in untreated patients (Fig4A). Comparison of primary treatment groups showed more rapid decline of CRP in the glucocorticoid and IVIG+G groups than in the IVIG treated patients (Fig4B). There was a suggestion of more rapid decline in troponin and ferritin in the glucocorticoid and combined treatment groups with a similar trend when restricting to those not receiving additional treatment between days 0 and 2 (Fig4C). Time course plots of other blood markers showed similar dynamics of blood markers between groups (appendix p62).

To investigate whether inadvertent inclusion of children with KD within BATS enrolment might have influenced treatment responses, we explored changes in blood markers separately in children most resembling KD. As KD is generally a disease in children aged 5-years and below, and MIS-C is often reported in older children, we compared those meeting AHA criteria for KD, and all children under 6 years ("KD-like"), with the remaining MIS-C patients.

The rate of decline in CRP was similar between the younger and older children and those fulfilling KD criteria treated with IVIG, with a suggestion of a more rapid decline in CRP in the non-KD-like patients receiving glucocorticoids alone (appendix p63).

DISCUSSION

Our comparison of treatment outcomes in an international cohort of 2009 children with MIS-C shows that treatment with glucocorticoids alone, or IVIG+G are not associated with significant differences in primary outcomes (requirements for inotropic support, ventilation on day two or beyond, or death; or rate of improvement on the ordinal severity scale) in comparison with IVIG alone. The findings are consistent with our preliminary report of 614 children.[15] However, the larger number of patients in each treatment group, increases the confidence in our findings. There was a non-significant trend towards a slower rate of improvement in patients treated with glucocorticoids alone in comparison with IVIG, but this comparison was confined to those with less severe illness at presentation. Reassuringly, we found no difference in CAA outcomes between primary treatment groups, with resolution seen in the vast majority of patients.

Our planned secondary analysis comparing glucocorticoids alone with combined IVIG+G demonstrated no difference in the first primary outcome, but a faster time-to-improvement for the IVIG+G group. This comparison was not adjusted for multiple hypothesis testing and the effect appears confined to those patients not requiring intensive support at baseline. Other Secondary endpoints, and thus also not corrected for multiple hypothesis testing, showed lower rates of treatment escalation and lower rates of fever on day 2 in the IVIG+G group.

A key question for clinicians is whether the potential incremental benefits of IVIG+G to reduce severity of illness and hasten resolution of fever are sufficient to justify the use of both agents. We note that the primary outcomes (progression or recovery from organ support) were chosen to select the most clinically important outcomes, whereas the secondary outcomes may detect less clinically important findings. Furthermore, we suggest that the finding of more common escalation of treatment for those on single agents, which was also observed in earlier studies,[12,13] may be biased by greater clinician readiness to add other treatments in seriously ill patients who do not

rapidly improve on monotherapy, whereas options to escalate treatment are fewer in patients treated with primary combination IVIG+G.

This question of whether combined IVIG+G is beneficial as compared to Glucocorticoids alone is relevant to both resource rich countries where IVIG is readily available and countries where IVIG has limited availability or cost imposes limitations in its use. For resource limited settings, our data suggests that primary treatment with glucocorticoids alone, is a safe alternative to IVIG or combined treatment, with IVIG being reserved for patients who fail to improve on glucocorticoids alone. For countries where IVIG cost is less prohibitive, the limited supply of IVIG and potential for combined treatments to have more side effects than single agents would argue for initial treatment with a single agent, and addition of second agents only in those who do not improve.

A higher proportion of patients receiving IVIG+G as primary treatment were receiving inotropes or ventilation at day 0, and had more deranged blood markers, suggesting more severely ill patients may have received IVIG+G. Importantly, key differences between treatment groups were adjusted for in the propensity score analysis. Children treated with IVIG+G had more rapid resolution of fever than children treated with IVIG or glucocorticoids alone. However no other clinically significant findings were more frequent in the IVIG+G group in comparison with either of the single agent treatment groups.

Patients who were initially treated with glucocorticoids or IVIG alone and then received additional treatment by day 2 were more likely to be receiving inotrope or ventilatory support at baseline. However, patients who received additional treatment did not differ substantially from patients who did not receive additional treatment across multiple biomarkers, suggesting that treatment with inotrope or ventilatory support influenced the clinical decision for administration of additional treatment. We have included adjustment for both baseline inotrope and ventilatory support in our IPTW analysis.

The use of IVIG as treatment for MIS-C has largely been driven by the similarity of MIS-C to KD, for which IVIG is the established treatment to reduce risk of CAA.[9] As coronary artery aneurysms

are observed in 10-20% of MIS-C cases,[13,15,26] there has been concern that failure to include IVIG in initial treatment would be associated with increased risk of CAA. We found that the incidence of CAA in patients receiving glucocorticoids as initial treatment was similar to the incidence of CAA in IVIG recipients (either IVIG or IVIG+G). Furthermore, the severity of CAA (as measured by z-score) and the proportion of patients undergoing complete resolution of CAA by time of discharge, or on follow up was similar in the glucocorticoid alone group to the IVIG and IVIG+G groups, including post-hoc analysis restricting to patients who never received IVIG. Our study thus provides reassurance that initial therapy with single agent glucocorticoids is not associated with increased risk of long-term coronary artery damage in MIS-C.

The American College of Rheumatologists currently recommends combined treatment with IVIG and glucocorticoids for MIS-C,[11] based on limited evidence of benefit from the USA and French propensity matched studies,[12,13] which showed lower rates of treatment escalation and improved cardiac function detected by echocardiogram with combined therapy. Neither of these studies included a glucocorticoid only group, and both were smaller than our current analysis.

We observed a more rapid decline in CRP in all three treatment groups as compared to patients not receiving immunomodulators. Although the curves for each treatment were overlapping, there was a non-significant trend to a more rapid decline in CRP, ferritin and troponin in the glucocorticoid containing groups.

Our study has several limitations. A key concern is the extent to which a retrospective comparison of outcomes following non-randomised choice of treatment can be used to guide clinical practice. We applied two different propensity score methods (weighting and matching), to remove bias caused by differences in severity, demography, or resource setting. We achieved good covariate balance between comparator groups using both approaches. However, other unmeasured differences might influence the results, and a large RCT would be the preferred approach to provide definitive answers. In addition, there is a risk of bias from the voluntary nature of data collection, as not all cases of MIS-C from each site were necessarily included in the study.

A second potential limitation is our use of the broad inclusion criteria of clinician diagnosed MIS-C. At the time BATS was initiated the accuracy of the published diagnostic criteria was unknown, and there were differences between the WHO, CDC and RCPCH criteria. Furthermore, availability of antibody testing for SARS-CoV-2 was limited in many countries. We therefore chose to include patients whose responsible clinicians considered them to have MIS-C, and in whom alternative diagnoses had been excluded. As we expected, our data confirms that the most commonly "missed" criteria to meet the WHO or CDC definitions of MIS-C was the presence of evidence of SARS-CoV-2 exposure. It is noteworthy that as the pandemic has evolved, and a high proportion of children have become SARS-CoV-2 antibody positive through natural infection or vaccination, the value of antibody against SARS-CoV-2 as evidence of recent infection has reduced. In view of the high rates of SARS-CoV-2 infection in schools, and the high proportion of asymptomatic childhood infection, a history of exposure to infection is of little value in diagnosis of MIS-C, and the WHO and CDC criteria may need to be re-evaluated. Despite these concerns, the large majority of patients in BATS did meet the WHO criteria, with only small differences in the proportions from each of the primary treatment groups. Our subgroup and sensitivity analyses did not find any difference in outcome when restricted to those meeting the WHO criteria, or the group with features overlapping KD.

An additional concern may be that the nature, severity, and epidemiology of MIS-C has changed over time, and with successive SARS-CoV-2 waves and introduction of childhood vaccinations against COVID-19. The disorder appears to have become less common in many countries as a high proportion of children have previous infection, and both natural infection and vaccination may reduce the incidence of MIS-C.[27] However, with SARS-CoV-2 now increasing in the previously unexposed population of China, there is likely to be a new wave of MIS-C and the findings reported here may be of considerable help to the clinicians experiencing this disease for the first time.

Other limitations include the wide variety of steroid dosing regimens used, and the large number of patients in whom additional treatments were added after the primary treatment. Although we have attempted to compare those remaining on a single agent, this group may have been less severely ill and therefore not representative of the treatment group overall. Additionally, after excluding patients with incomplete baseline covariates from the IPTW analysis, the final numbers of patients used for primary analyses were marginally below those stated in our sample-size calculations. However, the suggested effect sizes in these calculations are relatively arbitrary. More important is the final width of confidence intervals for treatment effects, which were generally small for our primary analyses. An additional limitation is the use of a composite primary outcome. This was necessitated by the relatively small numbers of patients with individual outcomes, and our aim to capture effects of treatment in patients across a wide spectrum of severity. As mitigation we evaluated the individual components of the composite score as secondary analyses. The time-to-improvement outcome also incurs the possibility of "built-in selection bias"[28] although we have attempted to isolate known factors that could incur such bias through extensive subgroup analyses. This limitation is relevant to all survival analysis, and would not be avoidable even for RCTs using the same outcome. Finally, we are not able to detect rare or longer-term effects of either IVIG or glucocorticoid administration.

The absence of significant differences between treatment groups poses several questions on the mechanisms underlying MIS-C. As IVIG and glucocorticoids have different possible modes of action in MIS-C,[29,30] the lack of difference between them, and the fact that dual therapy was not superior to single agent therapy is puzzling. One possible explanation might be different underlying disease processes in MIS-C, some of which respond to IVIG and some to glucocorticoids. If so, we would have expected that combination treatment would be superior to each treatment individually. Alternatively, glucocorticoids and IVIG may act at different points in the same causal pathway and with equal efficacy. This would explain the similar outcomes and lack of additive effect. A final possibility is that neither treatment has a significant effect on the

disease process. As the number of patients receiving no immunomodulator treatment was small and phenotypically distinct from those receiving immunomodulator treatment, we did not have an adequate “No treatment group” to evaluate this possibility. However, the more rapid decline in CRP in the treated vs untreated groups supports a beneficial effect of all three treatment regimes. In addition to IVIG and glucocorticoids, several other immunomodulatory agents were administered, including anti-IL1, anti-IL6 and anti-TNF agents. The numbers of patients receiving these agents were too low to enable IPTW comparison between them, or with IVIG, glucocorticoids and IVIG+G. Biologicals tended to be administered in combination with IVIG and glucocorticoids, and to more unwell patients.

The key question in interpreting clinical significance of this analysis is whether the findings are sufficiently robust to enable glucocorticoids to replace IVIG as primary treatment of MIS-C. The lack of significant difference in outcomes between patients treated with glucocorticoids as primary treatment, and those receiving IVIG or IVIG+G, and in particular the lack of difference in CAA severity, frequency, or resolution, suggests that initial treatment with glucocorticoids is a safe alternative to IVIG. A concern in adopting this approach is the difficulty in distinguishing MIS-C from KD, particularly in younger patients, and the possibility that IVIG will be withheld from children with KD because they are thought to have MIS-C. This concern highlights the need for a rapid diagnostic test to distinguish MIS-C from KD, as well as the need for urgent cardiology assessment in patients presenting with a suspected diagnosis of either disease. It also suggests that where clinical features closely resemble KD, particularly in younger children, retaining IVIG as a component of initial therapy is prudent.

MIS-C has emerged as an important childhood problem in low- and middle-income countries.[26,31] As IVIG is costly[32] and has limited availability in many countries, its use in preference to cheaper anti-inflammatory agents such as glucocorticoids should be supported by sound evidence. We did not find significant differences in outcome between treatment with glucocorticoids or IVIG as single agents or between the single and dual agent primary

treatments. Our findings suggest that glucocorticoids are not inferior to IVIG or combination IVIG+G as primary treatment of MIS-C, and their wide availability and lower cost would support their choice as initial treatment for MIS-C.

Figures and Tables

Figure 1A | Study flowchart

The study flow chart gives an overview of the total number of patients enrolled, excluded, and included for the analyses. Patients meeting the inclusion criteria are categorized by treatment groups (IVIG, Glucocorticoids, IVIG & Glucocorticoids, Other immunomodulator treatments [this includes: anti-tumor necrosis factor, anti-interleukin 1, anti-interleukin 6] and no immunomodulator treatments) and subdivided by our data-drive classification according to the WHO MIS-C criteria.

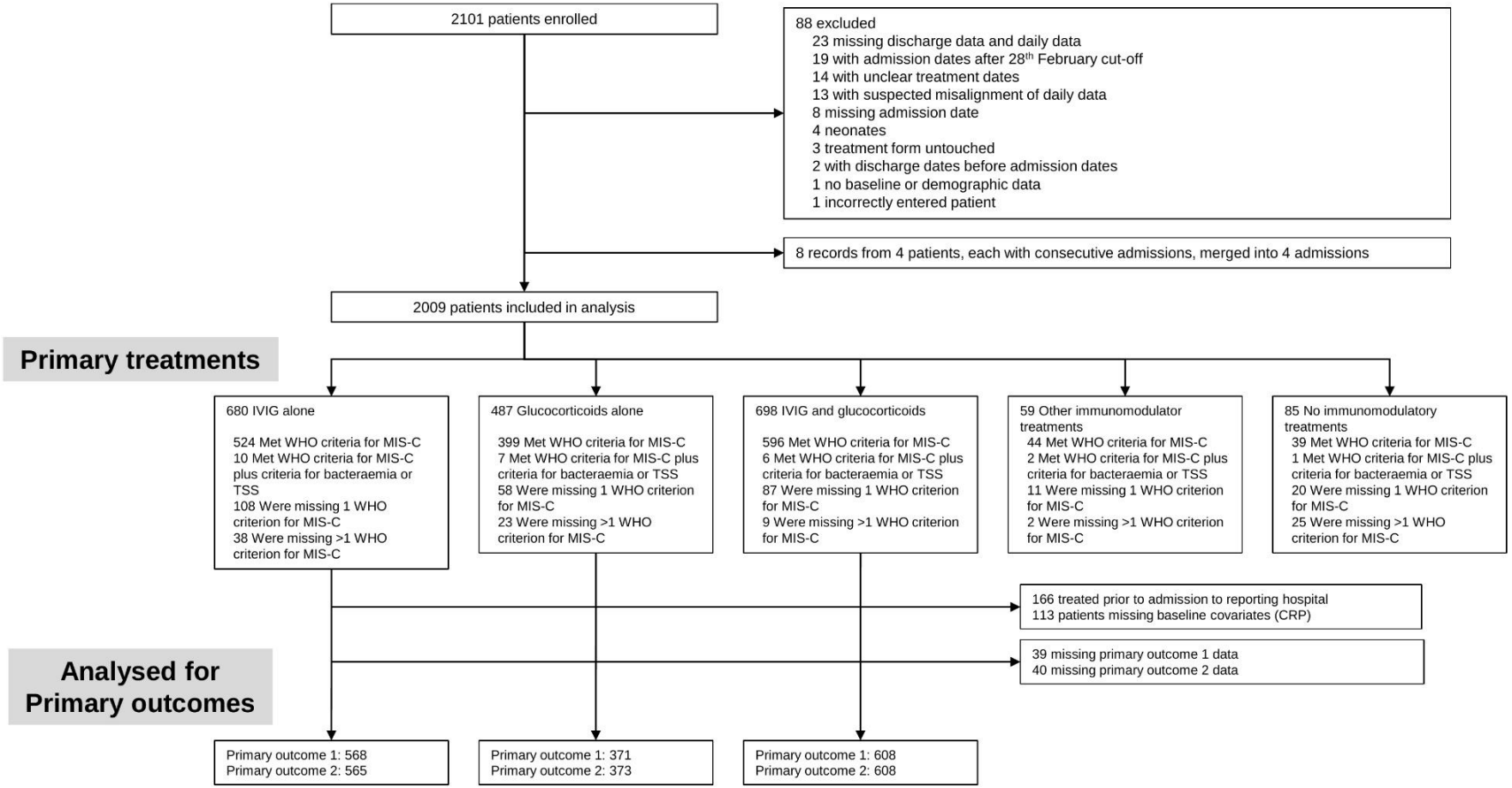


Figure 1B | Treatments received by patients over time following initiation of immunomodulator treatment

The Sankey diagram demonstrates the number of patients receiving cumulative therapies from days following initiation of immunomodulator treatment. Each vertical stack represents a different day in the patients' admission relative to starting immunomodulatory treatment (days 0 to 5), with day 0 representing the first day of immunomodulator treatment. The grey bands represent movement of patients between treatment groups from relative day 0 to 1, day 1 to 2, day 2 to 3, day 3 to 4 and day 4 to 5. The width of the grey bands is proportional to the number of patients (flow). The flow of patients is independent between time intervals; there is no continuous correspondence across days 1 to 5. The treatment groups are as stated. Of note, "Glucocorticoids" include intravenous and oral glucocorticoids (appendix p41). "Other" includes one or more other immunomodulatory treatment(s) given alone or in combination with Glucocorticoids and/or IVIG. Other immunomodulatory treatments include: anti-interleukin1, anti-interleukin 6, anti-tumour necrosis factor, cytokine adsorber (CytoSorb), granulocyte colony stimulating factor, colchicine, mesenchymal stem cells, convalescent plasma, cyclophosphamide, plasmapheresis and hydroxychloroquine

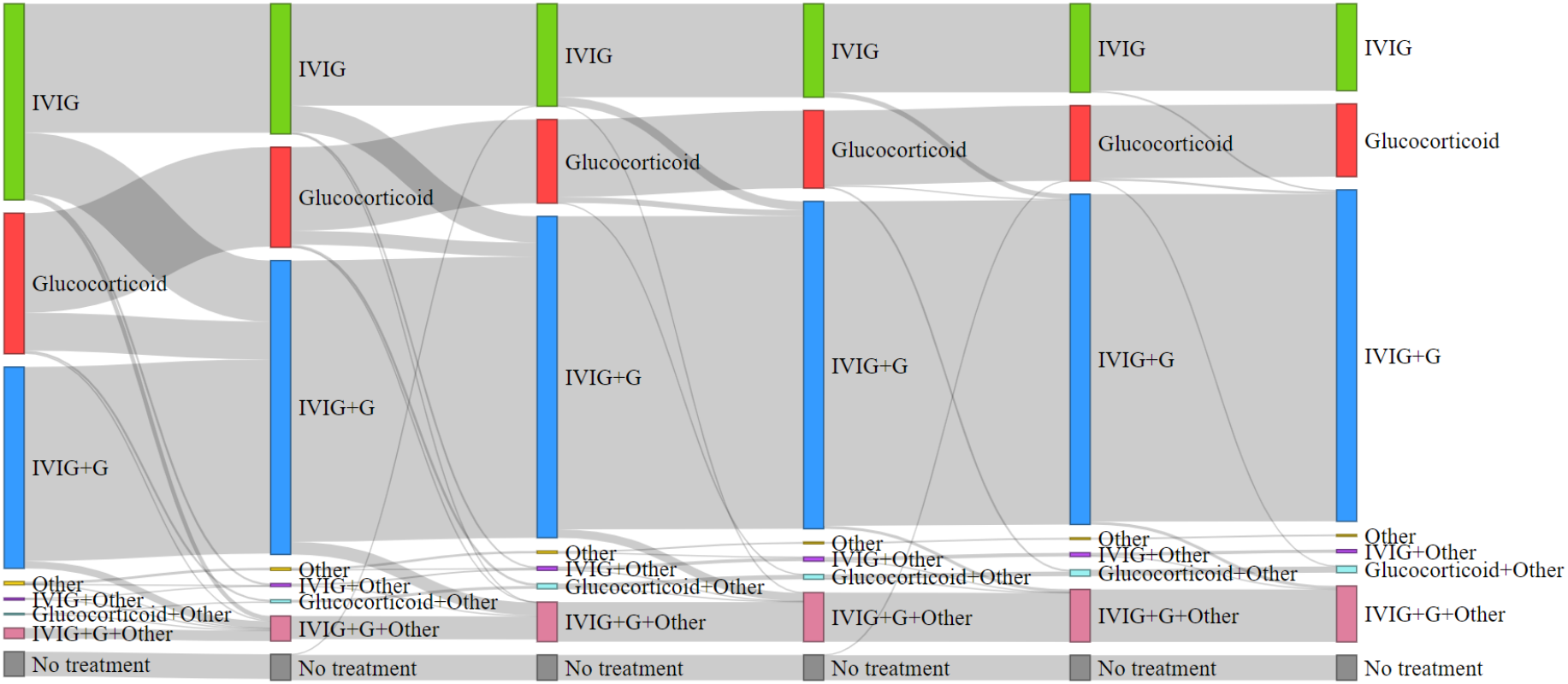


Figure 2 | Forest plots summarizing point estimates and 95% confidence intervals for primary analyses, including all subgroup and sensitivity analyses.

*Shown are outcomes for patients with suspected MIS-C who received IVIG plus glucocorticoids (Panels A & B) or glucocorticoids alone (Panels C & D) as compared with those who received IVIG alone (reference group, indicated by an odds ratio or average hazard ratio of 1·00). Displayed values are adjusted odds ratios or average hazard ratios (indicated on the x-axis). Panels A & C show the first primary outcome analyses, risk of inotropes, ventilation or death, and values to the right of the dotted line indicate superiority of IVIG alone. Panels B & D show the second primary outcome analyses, time to improvement in ordinal clinical severity score, with values to the left indicating superiority of IVIG alone. *indicates p-values corrected for multiple hypothesis testing using the Bonferroni-Holm procedure, observed p-value x4. Absolute numbers of patients included in each analysis can be found in appendix p29-32.*

Abbreviations: CRP: C-reactive protein; KD: Kawasaki Disease; WHO: World Health Organisation.

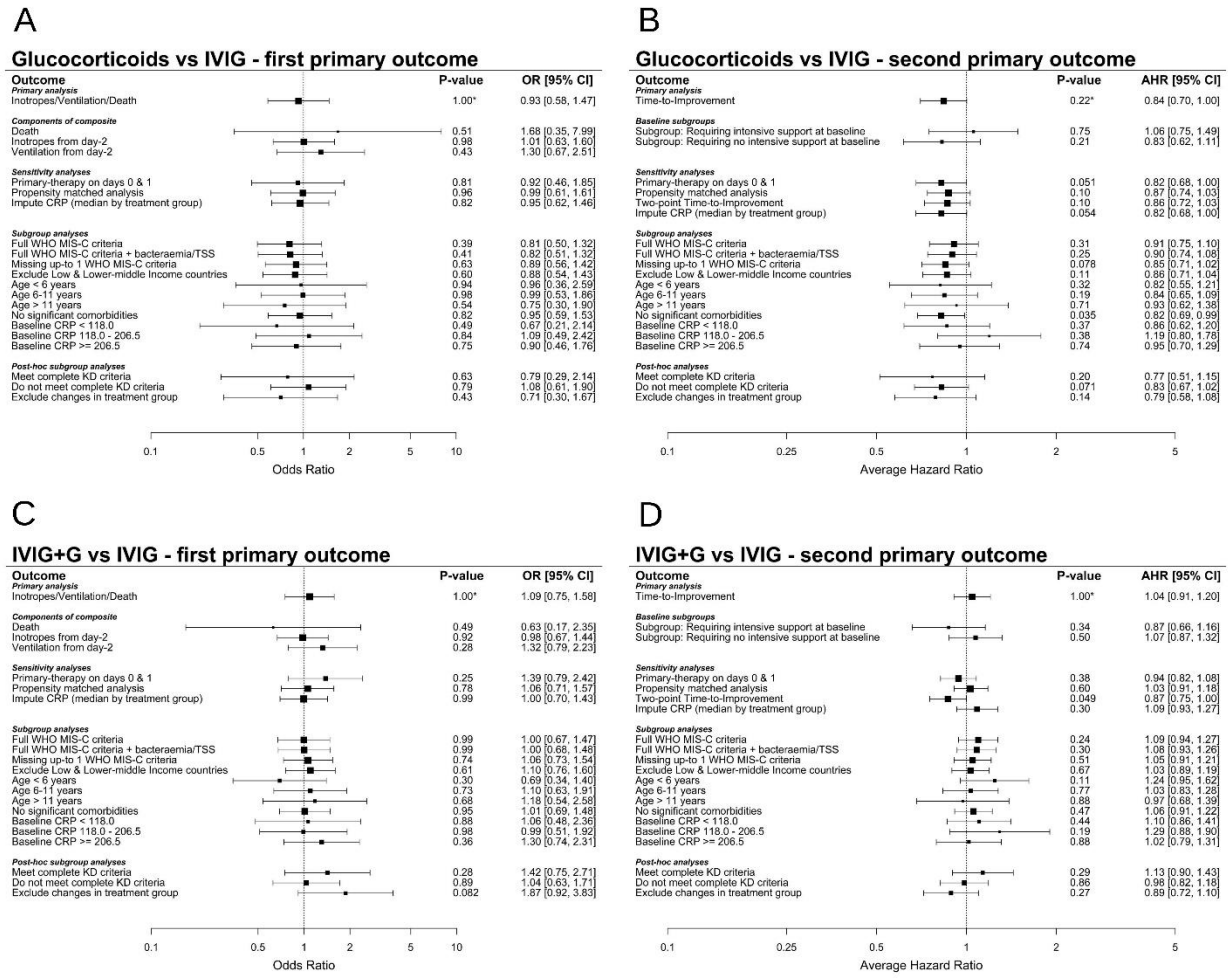


Figure 3 | Weighted clinical improvement over time

Panels A-C: Kaplan-Meier curves for the three main primary treatment groups showing time to one-point improvement in clinical severity on ordinal scale weighted by inverse probability of treatment, for (A) all patients, (B) subgroup of patients needing at least one of inotropes or ventilation at baseline, (C) subgroup of patients not requiring inotropes or ventilation at baseline. Tables below the Kaplan-Meier curves show the numbers at risk at the start of each day, and the number censored at this specific time point. Panel D: Clinical severity on ordinal scale, shown as proportional column charts from two days before treatment to 10 days after treatment, separated by primary treatment group, and weighted by inverse probability of treatment. Additional groups have been added for graphical purposes.

Abbreviations: CRP: C-reactive protein; ECMO: extracorporeal membrane oxygenation.

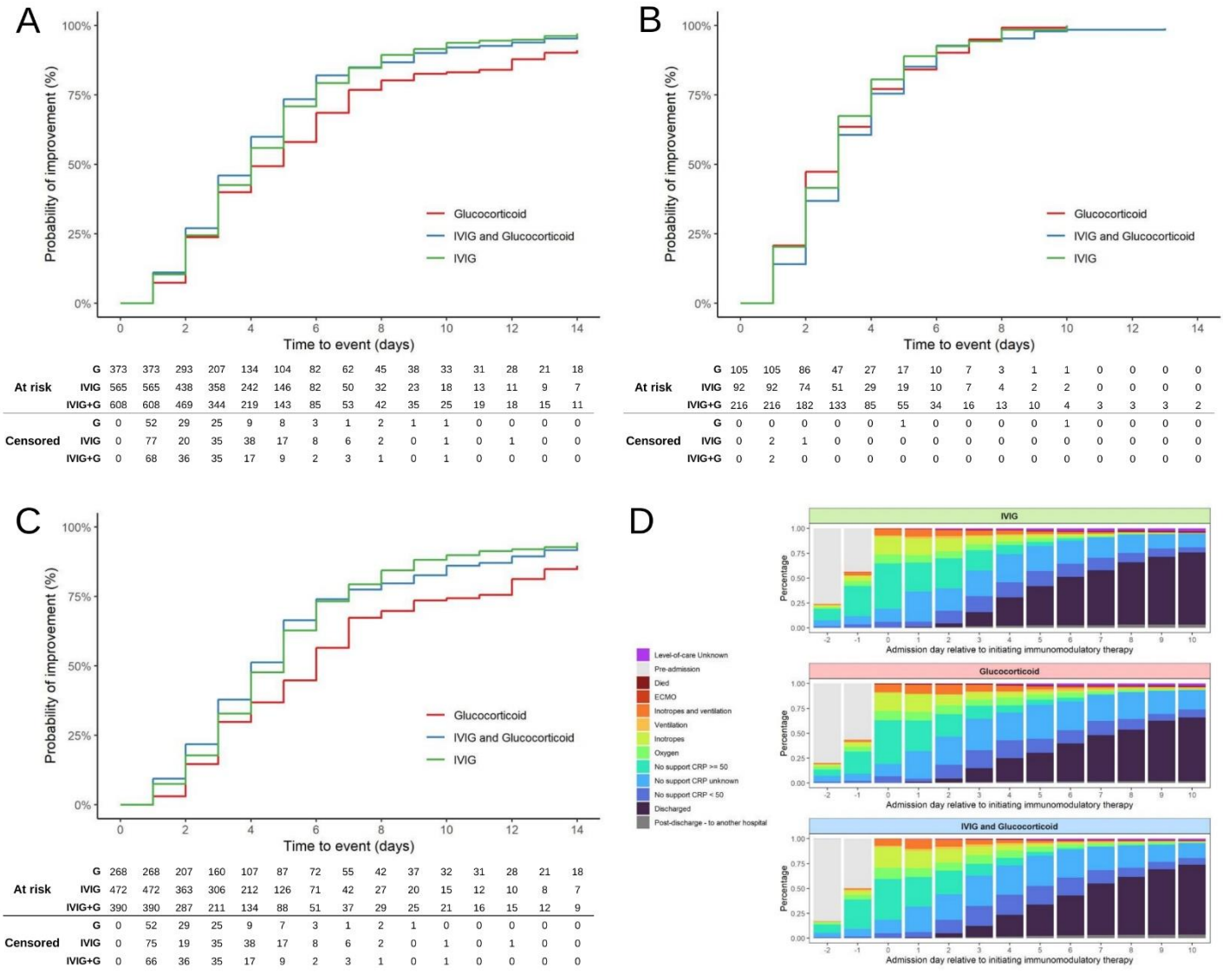


Figure 4 | Change in C-reactive protein (CRP), troponin and ferritin over time

Each of three key markers of inflammation (C-reactive protein, troponin, and ferritin) is plotted as a line and weighted by the covariate balancing propensity score. The levels are shown as a percentage of each patient's peak value, plotted by day relative to starting treatment. A generalized additive model was used to fit the curves. For each plot patients are only included if they had blood results available both before and after treatment initiation, and only if their last value up to treatment initiation was abnormal (CRP ≥ 8mg/L, troponin ≥ 14 ng/L, and ferritin ≥ 50 microgram/L). Panel A shows the fitted curves for the three measures in patients who received any immunomodulators, as compared with those who did not receive immunomodulators, using day of admission as relative admission day for patients not receiving immunomodulator treatment (NOTE: Curves for troponin in panel A were fitted using a loess model due to small sample numbers). Panel B shows the fitted curves for patients who received IVIG alone, IVIG plus glucocorticoids, and glucocorticoids alone as their primary treatment. Panel C shows the fitted curves for the three treatments combined in the patients whose primary treatment did not change between treatment initiation (day 0) and day 2.

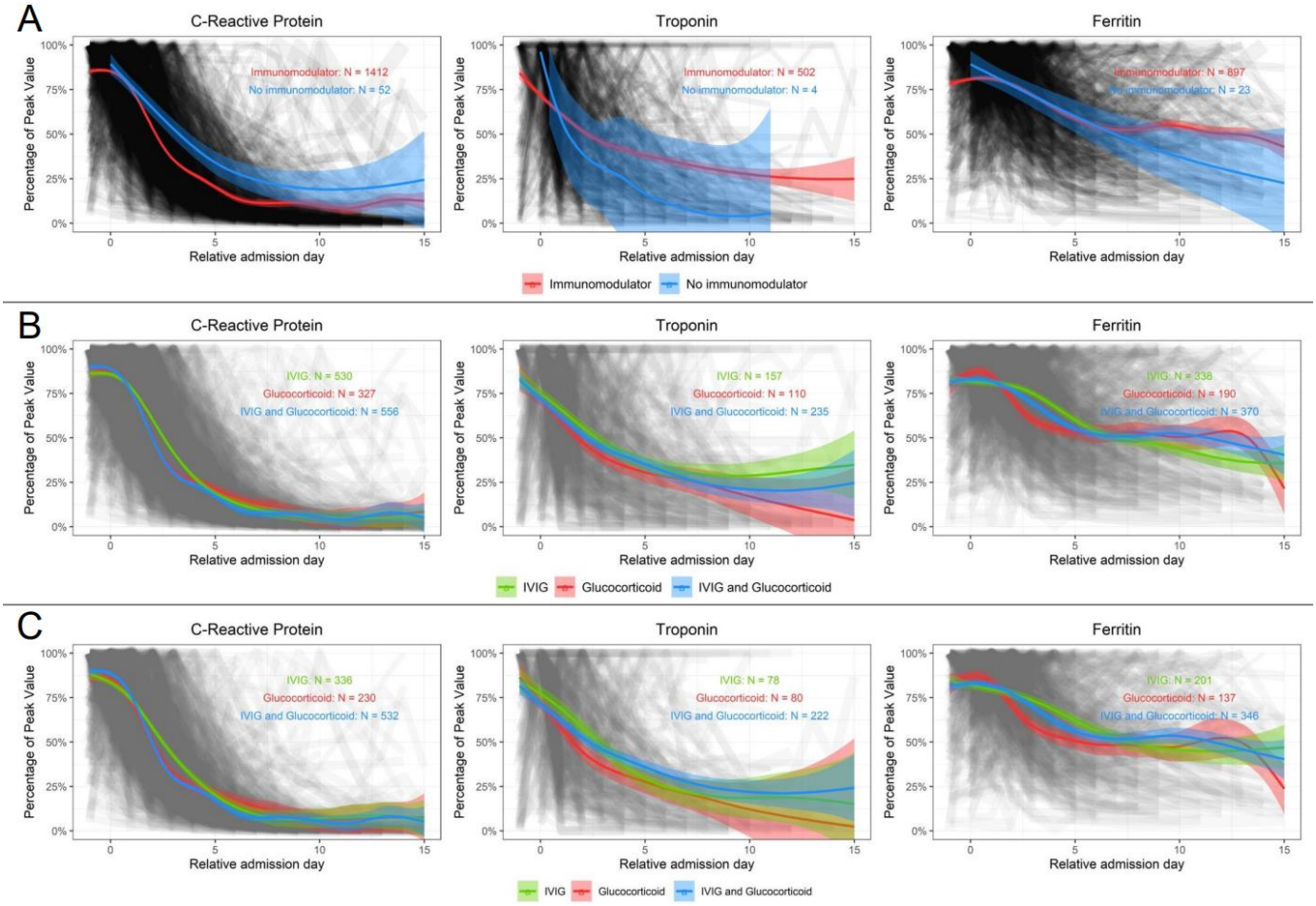


Table 1 | Clinical and demographic features in all treatment groups

Descriptive table of demographic features, clinical features and blood markers on admission, and proportion of patients meeting Kawasaki Disease criteria according to American Heart Association criteria. Patients with coronary artery aneurysms met the definition of Kawasaki Disease with less than 4 Kawasaki Disease clinical features. Patients were divided by treatment arm on day 0 (IVIG alone, glucocorticoid alone, IVIG+G, no treatment, and other (any other treatment combination including biologicals)). SARS-CoV-2 PCR data refer to test taken during admission. Organ support refers to receipt of ventilation, inotropes or ECMO on admission. Missing data (where applicable) are available in a full unabridged version in appendix p24.

Abbreviations: Ab: Antibody; CRP: C-reactive protein; ECMO: extracorporeal membrane oxygenation; PCR: polymerase chain reaction. ^Clinical and demographic features given as number and (%). +Numerical values given as median values and [interquartile ranges].

	Everyone (N=2009)	IVIG (N=680)	Glucocorticoids (N=487)	IVIG and Glucocorticoids (N=698)	Other (N=59)	No treatment (N=85)
+Age	8·0 [4·2 - 11]	6·8 [3·6 - 10]	8·8 [5·1 - 12]	8·4 [4·5 - 11]	11 [6·1 - 13]	7·3 [3·3 - 12]
^Proportion male	1191 (59·3%)	416 (61·2%)	288 (59·1%)	410 (58·7%)	44 (74·6%)	33 (38·8%)
^Proportion female	818 (40·7%)	264 (38·8%)	199 (40·9%)	288 (41·3%)	15 (25·4%)	52 (61·2%)
^Weight (age-adjusted z score ≥ 2)	299 (14·9%)	91 (13·4%)	70 (14·4%)	120 (17·2%)	10 (16·9%)	8 (9·41%)
^Ethnicity						
White	825 (41·1%)	290 (42·6%)	210 (43·1%)	272 (39·0%)	27 (45·8%)	26 (30·6%)
Latino	518 (25·8%)	161 (23·7%)	94 (19·3%)	222 (31·8%)	9 (15·3%)	32 (37·6%)
Black	212 (10·6%)	81 (11·9%)	34 (6·98%)	75 (10·7%)	13 (22·0%)	9 (10·6%)
Asian	131 (6·52%)	55 (8·09%)	36 (7·39%)	30 (4·30%)	4 (6·78%)	6 (7·06%)
Other or not known	323 (16·1%)	93 (13·7%)	113 (23·2%)	99 (14·2%)	6 (10·2%)	12 (14·1%)
^Significant comorbidity	108 (5·38%)	30 (4·41%)	32 (6·57%)	33 (4·73%)	4 (6·78%)	9 (10·6%)
^SARS-CoV-2 PCR positive	415 (20·8%)	131 (19·4%)	97 (20·0%)	148 (21·4%)	13 (22·0%)	26 (31·7%)
^SARS-CoV-2 Ab positive	1321 (66·5%)	412 (61·2%)	344 (71·4%)	492 (71·6%)	43 (72·9%)	30 (35·3%)
^Baseline requirement for ventilation/inotropes/ECMO	535 (26·6%)	117 (17·2%)	127 (26·1%)	252 (36·1%)	29 (49·2%)	10 (11·8%)
^Clinical features during admission						
Fever	1863 (92·7%)	653 (96·0%)	439 (90·1%)	649 (93·0%)	52 (88·1%)	70 (82·4%)
Sore throat	464 (25·5%)	159 (26·5%)	104 (22·9%)	175 (27·0%)	11 (21·6%)	15 (21·1%)
Cough	404 (21·1%)	125 (19·4%)	120 (25·3%)	131 (19·6%)	16 (30·8%)	12 (16·0%)
Respiratory distress	258 (13·3%)	70 (10·9%)	57 (11·9%)	112 (16·4%)	13 (23·6%)	6 (7·59%)
Abdominal pain	1211 (63·2%)	408 (63·9%)	289 (62·3%)	438 (64·8%)	37 (63·8%)	39 (48·1%)
Diarrhoea	882 (44·8%)	290 (43·9%)	195 (40·6%)	340 (49·4%)	23 (39·7%)	34 (41·5%)
Vomiting	1057 (54·0%)	330 (50·6%)	251 (52·3%)	408 (59·2%)	34 (60·7%)	34 (42·5%)
Headache	592 (32·8%)	199 (34·1%)	155 (35·0%)	203 (31·4%)	21 (38·9%)	14 (18·4%)
Irritability	355 (18·8%)	127 (20·2%)	69 (14·9%)	135 (20·2%)	10 (18·5%)	14 (18·4%)
Lethargy	655 (34·5%)	211 (33·3%)	186 (40·1%)	215 (32·1%)	23 (41·8%)	20 (26·7%)
^Proportion meeting Kawasaki Disease criteria	629 (31·3%)	265 (39·0%)	119 (24·4%)	225 (32·2%)	12 (20·3%)	8 (9·41%)
+Bloods on admission						
Lymphocytes (10^9/L)	1·2 [0·70 - 2·0]	1·3 [0·76 - 2·2]	1·2 [0·70 - 1·8]	1·1 [0·66 - 1·9]	0·86 [0·52 - 1·6]	1·8 [1·1 - 2·9]
Troponin (ng/L)	25 [6·1 - 80]	13 [5·0 - 43]	31 [9·8 - 100]	40 [10 - 110]	48 [10 - 270]	10 [2·0 - 38]
CRP (mg/L)	150 [85 - 220]	150 [85 - 210]	160 [75 - 220]	160 [90 - 230]	180 [97 - 280]	85 [23 - 180]
Ferritin (ug/L)	440 [230 - 860]	370 [210 - 650]	480 [260 - 970]	520 [260 - 960]	560 [340 - 1700]	280 [140 - 460]
Albumin (g/L)	32 [28 - 37]	34 [28 - 39]	32 [27 - 36]	32 [27 - 36]	32 [27 - 36]	35 [30 - 41]

REFERENCES

1 Riphagen S, Gomez X, Gonzalez-Martinez C, Wilkinson N, Theocharis P. Hyperinflammatory shock in children during COVID-19 pandemic. *Lancet.* 2020;**395:**1607-8.

2 Viner RM, Whittaker E. Kawasaki-like disease: emerging complication during the COVID-19 pandemic Comment. *Lancet.* 2020;**395:**1741-3.

3 Whittaker E, Bamford A, Kenny J, et al. Clinical Characteristics of 58 Children with a Pediatric Inflammatory Multisystem Syndrome Temporally Associated with SARS-CoV-2. *JAMA.* 2020;**324:**259–69.

4 Verdoni L, Mazza A, Gervasoni A, et al. An outbreak of severe Kawasaki-like disease at the Italian epicentre of the SARS-CoV-2 epidemic: an observational cohort study. *Lancet.* 2020;**395:**1771–8.

5 Feldstein LR, Tenforde MW, Friedman KG, et al. Characteristics and Outcomes of US Children and Adolescents with Multisystem Inflammatory Syndrome in Children (MIS-C) Compared with Severe Acute COVID-19. *JAMA.* 2021;**325:**1074–87.

6 Feldstein LR, Rose EB, Horwitz SM, et al. Multisystem Inflammatory Syndrome in U.S. Children and Adolescents. *N Engl J Med.* 2020;**383:**334–46.

7 Godfred-Cato S, Bryant B, Leung J, et al. COVID-19-Associated Multisystem Inflammatory Syndrome in Children - United States, March-July 2020. *MMWR Morb Mortal Wkly Rep.* 2020;**69:**1074–80.

8 Levin M. Childhood Multisystem Inflammatory Syndrome - A New Challenge in the Pandemic. *N Engl J Med.* 2020;**383:**393–5.

9 McCrindle BW, Rowley AH, Newburger JW, et al. Diagnosis, Treatment, and Long-Term Management of Kawasaki Disease: A Scientific Statement for Health Professionals From the American Heart Association. *Circulation.* 2017;**135(17):**e927–99.

10 Harwood R, Allin B, Jones CE, et al. A national consensus management pathway for paediatric inflammatory multisystem syndrome temporally associated with COVID-19 (PIMS-TS): results of a national Delphi process. *Lancet Child Adolesc Health.* 2021;**5:**133-141

11 Henderson LA, Canna SW, Friedman KG, et al. American College of Rheumatology Clinical Guidance for Multisystem Inflammatory Syndrome in Children Associated With SARS-CoV-2 and Hyperinflammation in Pediatric COVID-19: version 1. *Arthritis Rheumatol.* 2020;**72:**1791–805.

12 Ouldali N, Toubiana J, Antona D, et al. Association of Intravenous Immunoglobulins Plus Methylprednisolone vs Immunoglobulins Alone With Course of Fever in Multisystem Inflammatory Syndrome in Children. *JAMA.* 2021;**325:**855–64.

13 Son MBF, Murray N, Friedman K, et al. Multisystem Inflammatory Syndrome in Children - Initial Therapy and Outcomes. *N Engl J Med.* 2021;**385:**23-34.

14 Villacis-Nunez DS, Jones K, Jabbar A, et al. Short-term Outcomes of Corticosteroid Monotherapy in Multisystem Inflammatory Syndrome in Children. *JAMA Pediatr.* 2022;**176:**576-84.

15 McArdle AJ, Vito O, Patel H, et al. Treatment of Multisystem Inflammatory Syndrome in Children. *N Engl J Med.* 2021;**385:**11-22.

16 Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;**42:**377–81.

17 Toscano G, Palmerini F, Ravaglia S, et al. Guillain–Barré Syndrome Associated with SARS-CoV-2. *N Engl J Med.* 2020;**382:**2574–6.

18 LaRovere KL, Riggs BJ, Poussaint TY, et al. Neurologic Involvement in Children and Adolescents Hospitalized in the United States for COVID-19 or Multisystem Inflammatory Syndrome. *JAMA Neurol.* 2021;**78:** 536-47.

19 Tullie L, Ford K, Bisharat M, et al. Gastrointestinal features in children with COVID-19: an observation of varied presentation in eight children. *Lancet Child Adolesc Health.* 2020;**4(7):** e19-e20.

20 Centers for Disease Control and Prevention (CDC). Information for healthcare providers about multisystem inflammatory syndrome in children (MIS-C) https://www.cdc.gov/mis-c/hcp/. Date accessed: July 28, 2022

21 Royal College of Paediatrics and Child Health (RCPCH). Paediatric multisystem inflammatory syndrome temporally associated with COVID-19 (PIMS) - guidance for clinicians. 2020. https://www.rcpch.ac.uk/resources/paediatric-multisystem-inflammatory-syndrome-temporally-associated-covid-19-pims-guidance. Date accessed: December 10, 2022

22 World Health Organization. Multisystem inflammatory syndrome in children and adolescents temporally related to COVID-19. Scientific brief. May 2020. https://www.who.int/publications/i/item/multisystem-inflammatory-syndrome-in-children-and-adolescents-with-covid-19. Date accessed: December 10, 2022

23 National Institute for Health and Care Excellence (NICE). Equivalent anti-inflammatory doses of oral corticosteroids. June 2020. https://cks.nice.org.uk/topics/corticosteroids-oral/background-information/equivalent-anti-inflammatory-doses/. Date accessed: December 10, 2022

24 Lopez L, Colan SD, Stylianou Mario, et al. Relationship of Echocardiographic Z-Scores Adjusted for Body Surface Area to Age, Sex, Race, and Ethnicity: The Pediatric Heart Network Normal Echocardiogram Database. *Circ Cardiovasc Imaging.* 2017;**10(11):** e006979

25 Imai K, Ratkovic M. Covariate balancing propensity score. *J R Stat Soc Ser B Stat Methodol.* 2014;**76:**243–63.

26 Sood M, Sharma S, Sood I, Sharma K, Kaushik A. Emerging Evidence on Multisystem Inflammatory Syndrome in Children Associated with SARS-CoV-2 Infection: a systematic Review with Meta-analysis. *SN Compr Clin Med.* 2021;**3:**38–47.

27 Ouldali N, Bagheri H, Salvo F, et al. Hyper inflammatory syndrome following COVID-19 mRNA vaccine in children: A national post-authorization pharmacovigilance study. *Lancet Reg Health Eur.* 2022;**17**:100393.

28 Hernan MA. The hazards of hazard ratios. *Epidemiology*. 2010;**21(1):**13-5.

29 Zhu YP, Shamie I, Lee JC, et al. Immune response to intravenous immunoglobulin in patients with Kawasaki disease and MIS-C. *J Clin Invest.* 2021;**131(20):** e147076.

30 Patel H, McArdle A, Seaby E, Levin M, Whittaker E. The immunopathogenesis of SARS-CoV-2 infection in children: diagnostics, treatment and prevention. *Clin Transl Immunol.* 2022;**11:** e1405.

31 Irfan O, Muttalib F, Tang K, Jiang L, Lassi ZS, Bhutta Z. Clinical characteristics, treatment and outcomes of paediatric COVID-19: A systematic review and meta-analysis. *Arch Dis Child.* 2021;**106:**440-8.

32 Johnson SC, Williams DC, Brinton D, Chew M, Simpson A, Andrews AL. A Cost Comparison of Infliximab Versus Intravenous Immunoglobulin for Refractory Kawasaki Disease Treatment. *Hosp Pediatr.* 2021;**11:** 88-93.

DATA SHARING STATEMENT

Question	Response
Will individual participant data be available (including data dictionaries)?	Yes
Rationale for data sharing statement	BATS has collected de-identified data from multiple institutions in many countries. Each institution has signed an agreement with Imperial College on data security. We will need to assess requests for data on a case-by-case basis to ensure that the data that are provided fall within the existing agreements within the consortium.
What data in particular will be shared?	De-identified clinical and laboratory findings and response to treatment for the cohort included in this study. Any data provided will be de-identified and will conform to the agreements within the consortium for data sharing.
What other documents will be available?	The study handbook and statistical analysis plans are available at the ISRCTN registry at the following link: https://doi.org/10.1186/ISRCTN69546370
When will data availability start?	On publication of the manuscript. However, as approval for all data will have to be obtained from the consortium and partner institutions, approximately 3 months may be required before the data is provided.
When will data availability end?	Two years after publication
To whom will data be available?	Legitimate researchers and clinicians from medical and academic institutions.
For what types of analyses?	Only for academic and clinical research.
By what mechanism will data be made available?	On request to the corresponding author.
	Data requestors will need to sign a data access agreement

Immunoglobulin, Glucocorticoid, or combination therapy for Multisystem Inflammatory Syndrome in Children—: A propensity weighted cohort study

Samuel Channon-Wells ~~MBBS,~~ MMath[1,5]*, Ortensia Vito MSc[1,5]*, Andrew J. McArdle ~~MB BChir,~~ MSc[1,5]*, Eleanor G. Seaby ~~BMBS, MMedSc[1]~~BMBS[1,2,3+], Harsita Patel MBBS[1,5+], Priyen Shah MB BS[1,5], Ekaterina Pazukhina MSc**, Clare Wilson MB BChir[1,5], Claire Broderick MBBS[1,5], Giselle D'Souza MSc[1,5], Ilana Keren[1], Ruud G. Nijman ~~MD,~~ PhD[1,4,5], Adriana Tremoulet ~~MD,~~ MAS[6], Daniel Munblit ~~MD,~~ PhD[7,8], Rolando Ulloa-Gutierrez MD[9,10,11], Michael J Carter ~~BM BCh,~~ DPhil[12], Padmanabhan Ramnarayan ~~MBBS,~~ MD[13], Tisham De PhD[1], Clive Hoggart PhD[14], Elizabeth Whittaker ~~MB BAO BCh,~~ PhD[1,5,15], Jethro A. Herberg PhD[1,5,15], Myrsini Kaforou PhD[1,5], Aubrey J. Cunnington ~~BM BCh,~~ PhD[1,5,15], Oleg Blyuss PhD[16], Michael Levin FMedSci[1,5,15] and **The Best Available Treatment Study for MIS-C (BATS) consortium**

*,+ Contributed equally

[1] Department of Infectious Disease, Section of Paediatric Infectious Disease, Imperial College London, London, United Kingdom

[2] Genomic Informatics Group, University of Southampton, Southampton, United Kingdom

[3] Translational Genomics Group, Broad Institute of MIT and Harvard, Cambridge, MA, USA

[4] Department of Paediatric Emergency Medicine, Division of Medicine, St. Mary's hospital - Imperial College NHS Healthcare Trust, London, London, London, UK

[5] Centre for Paediatrics and Child Health, Imperial College, London, UK

[6] Department of Paediatrics, University of California San Diego/Rady Children's Hospital San Diego

[7] Department of Paediatrics and Paediatric Infectious Diseases, Institute of Child's Health, Sechenov First Moscow State Medical University (Sechenov University), Moscow, Russia

[8] Inflammation, Repair, and Development Section, National Heart and Lung Institute, Faculty of Medicine, Imperial College London, London, United Kingdom

[9] Servicio de Infectologia Pediatrica, Hospital Nacional de Niños "Dr. Carlos Sáenz Herrera", Centro de Ciencias Médicas, Caja Costarricense de Seguro Social (C.C.S.S.), San José, Costa Rica

[10] Instituto de Investigación en Ciencias Médicas UCIMED (IICIMED), San José, Costa Rica

[11] Cátedra de Pediatría, Facultad de Medicina, Universidad de Ciencias Médicas (UCIMED), San José, Costa Rica

[12] Department of Women and Children's Health, School of Life Course Sciences, King's College London, St Thomas' Hospital, SE1 7EH, London, UK

[13] Anaesthetics, Pain Medicine and Intensive Care (APMIC) Division, Department of Surgery and Cancer, Faculty of Medicine, Imperial College

[14] Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York City, NY USA

[15] Department of Paediatrics, Imperial College Healthcare NHS Trust, London, United Kingdom

[16] Wolfson Institute of Population Health, Queen Mary University of London

**No current institutional affiliation

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

Corresponding author: Professor Michael Levin, Department of Infectious Disease, Imperial College London. Email: m.levin@imperial.ac.uk

ABSTRACT

Background

SARS-CoV-2 associated Multisystem Inflammatory Syndrome in Children (MIS-C) has emerged as a serious illness in children world-wide. Immunoglobulin and/or glucocorticoids are currently recommended treatments.

Methods

The “Best Available Treatment Study” evaluated immunomodulatory treatments for MIS-C in an international observational cohort. Analysis of the first 614 patients was previously reported. Clinical and outcome data were collected onto a web-based database. Inverse probability weighting was used to compare primary treatments with intravenous immunoglobulin (IVIG), IVIG plus glucocorticoids (IVIG+G), or glucocorticoids alone, using IVIG as the reference treatment. Primary outcomes were: a composite of inotropic or ventilator support from the second day after treatment initiation, or death; and time-to-improvement on an ordinal clinical severity scale. Secondary outcomes included treatment escalation, clinical deterioration, fever, and coronary artery aneurysm occurrence and resolution.

Findings

After exclusions, 2009 children with clinically diagnosed MIS-C from 39 countries were enrolled between May 2020 and April 2022. 680 received primary treatment with IVIG; 698 IVIG+G; 487 glucocorticoids alone; 59 other combinations including biologics, and 85 no immunomodulator.

There were no significant differences between treatments for primary outcomes for the 1586 patients considered for primary analysis: adjusted odds ratios relative to IVIG for ventilation, inotropic support or death were 1·09 (95% confidence interval [CI] 0·75-1·58) and 0·93 (95% CI: 0·58-1·47) for IVIG+G and glucocorticoids alone respectively. Adjusted average hazard ratios for time-to-improvement were 1·04 (95% CI: 0·91-1·20) and 0·84 (95% CI: 0·70-1·00) for the same

comparisons. Treatment escalation was less frequent for IVIG+G and glucocorticoids alone vs IVIG. Persistent fever was less common with IVIG+G compared with either IVIG or glucocorticoids alone. Coronary artery aneurysm occurrence and resolution did not differ significantly between treatment groups.

Interpretation

Recovery rates, including occurrence and resolution of coronary artery aneurysms, were similar for primary treatment with IVIG when compared to glucocorticoids or combination IVIG+G. Initial treatment with glucocorticoids appears to be a safe alternative to immunoglobulin or combined therapy, and may be advantageous in view of the cost and limited availability of IVIG in many countries.

Funding and sponsors

The BATS writing group receive salary support from Imperial College, or are supported on individual grants, fellowships and project grants including the European Union's Horizon 2020 programme under GA No. 848196 DIAMONDS, Wellcome Trust (AJM, CB, Imperial 4i Wellcome Trust/NIHR Imperial BRC, CW; Wellcome Trust, MK 206508/Z/17/Z), the Medical Research Foundation (MRF-160-0008-ELP-KAFO-C0801), UK NIHR (RGN ACL-2018-021-007; MJC ACL-2018) and NIH (GA5R01AI128765). Enrolment of patients in London was supported by NIHR Imperial Biomedical Research Centre. Imperial College London was the main research sponsor for this study.

Role of the funding source

The funders and sponsor of the study had no role in study design, data collection, analysis, interpretation, or writing of the final manuscript.

Trial Registration Number

ISRCTN registry - ISRCTN69546370 https://doi.org/10.1186/ISRCTN69546370

Contributors

ML, AJM, OV, AC, EW, JH, MK, HP, PS, CW, RN, TD, and CH conceived the original study. OV, SCW, CB, ES, HP, GS and IK undertook data checks and quality control. OV, SCW and ML accessed and verified the underlying data reported in the manuscript. SCW, OB, AJM and ML conceived the current statistical analysis plan, which was reviewed by all co-authors prior to publication. SCW, ES, EP, and HP undertook the current analysis. SCW and ML prepared the manuscript. All authors had full access to all the data in the study. All authors read and approved the final manuscript~~.~~ and had final responsibility for the decision to submit for publication.

Conflict of Interest Statements/Declaration of interests

~~These will be uploaded when acquired from all authors.~~

AT has provided unpaid consultancy work for Janssen Pharmaceuticals. DM has received grant support from the British Embassy in Moscow ('StopCOVID Cohort: Clinical Characterisation of Russian Patients') and holds the following unpaid positions: Co-Chair of International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC) Global Paediatric Long COVID Working Group, Member of ISARIC working group on long-term follow-up in adults, Co-lead of the PC-COS project aiming to define the Core Outcome Set for Long-COVID, in collaboration with the WHO. MJC reports a personal fee from BioTest for speaking at the BioTest Immunology Forum 2022, Royal Society (www.biotest.com). EW holds the following unpaid positions: member of the paediatric steering committee for the RECOVERY trial; paediatric representative for NHS England working on the National paediatric virtual advisory network and expert advisory group for COVID treatment, Independent advisory group for COVID monoclonal antibodies, and co-lead for the pan-London Post-COVID service for children. All other authors declare no competing interests.

Ethics Committee approval

Please see "oversight" section in main text for relevant details.

RESEARCH IN CONTEXT

Evidence before this study

In the first wave of the COVID-19 pandemic paediatricians around the world rapidly identified and described a new inflammatory disorder, causing shock and multi-system failure in children approximately 4-6 weeks after SARS-CoV-2 infection. Faced with this new life-threatening disorder, termed multisystem inflammatory syndrome in children (MIS-C), with unknown pathophysiological mechanisms, paediatricians, national, and international paediatric bodies rapidly adopted treatments which are of benefit in other inflammatory disorders.

Based on the similarity in clinical features of MIS-C to Kawasaki Disease (KD), intravenous immunoglobulin (IVIG), the recognised treatment for KD, was adopted as the most widely used initial treatment, often combined with glucocorticoids and a range of biological agents. In the absence of data from randomised controlled trials (RCT), national and international organisations, including the World Health Organisation (WHO), ~~USA Centre for Disease Control (CDC),~~American College of Rheumatology, and UK Royal College of Paediatrics and Child Health (RCPCH) produced treatment guidelines recommending IVIG as initial treatment, combined with glucocorticoids or biological agents for the most seriously ill or unresponsive patients.

We searched for publications on treatment of MIS-C (and the alternative name Paediatric Multisystem Inflammatory Syndrome Temporally associated with SARS-CoV-2 (PIMS-TS)) since April 2020 when the disorder was first recognised, until November 2022. In the extensive literature now published on MIS-C, there are many hundreds of observational studies, treatment recommendations and guidelines based on expert opinion, and reports of outcome after treatment. However, we found no RCTs, and only four propensity matched comparisons reporting outcomes after specific treatments, only two of which included comparison of glucocorticoids alone and IVIG, and all were based on relatively small patient cohorts.

Added value of this study

The Best Available Treatment Study (BATS) allowed us to compare treatment of MIS-C with IVIG alone, glucocorticoids alone, and combined glucocorticoids plus IVIG (combined therapy), in over 2000 patients from 39 different countries. This is the largest study to date of immunomodulator treatment options in MIS-C, including the largest cohort of patients treated initially with glucocorticoid monotherapy. After correcting for known confounders using propensity score weighting, initial treatment with glucocorticoid monotherapy or combined therapy demonstrated no significant difference to treatment with IVIG monotherapy in either time-to-improvement measured on an ordinal clinical severity scale, or in a composite outcome of inotropic support or ventilator support (invasive or non-invasive) from the second day after starting treatment or later, or death. Comparison of glucocorticoid monotherapy with combined therapy suggested a small benefit from combined therapy in time-to-improvement, but this appeared to be restricted to those who did not require inotropic and/or ventilatory support at baseline. Combined therapy was associated with faster fever resolution and less escalation of treatment, but with no other differences in secondary outcomes. Occurrence and resolution of coronary artery aneurysms was similar in all treatment groups, with the large majority of aneurysms resolving during follow up.

Implications of all the available evidence

Our study increases confidence that initial treatment of MIS-C with glucocorticoids is associated with similar outcomes to treatment with IVIG or combined therapy. In the context of all current observational data, there is, at best, only a small benefit in initial therapy combining IVIG and glucocorticoids compared to monotherapy with IVIG or glucocorticoids alone. Given the high cost and limited availability of IVIG in many countries this evidence supports initial glucocorticoid monotherapy as an acceptable alternative.

BACKGROUND

Since recognition in April 2020, Multi-system Inflammatory Syndrome in Children (MIS-C), temporally associated with SARS-CoV-2 infection,[1–4] has emerged as a rare but serious post-infectious illness.[5–8] In the absence of evidence from randomised controlled trials (RCT), treatment recommendations for the new disease were developed by clinical consensus in many countries. Based on similarity of MIS-C to Kawasaki disease (KD), for which Intravenous immunoglobulin (IVIG) is the established treatment,[9] national and international guidance has recommended IVIG as initial treatment, with addition of glucocorticoids and/or other immunomodulatory agents for patients with severe illness.[10,11]

While there have been no RCTs comparing treatments for MIS-C published to date, several observational studies using propensity score methods have suggested that combination treatment with IVIG and glucocorticoids was associated with improved cardiac outcomes.[12–14]

The Best Available Treatment Study (BATS) was initiated in the early months after first recognition of MIS-C and aimed to provide evidence for treatment recommendations by systematic data collection, and analysis of outcomes of treatments chosen by individual paediatricians responsible for patient care. In view of the urgent need for evidence to support treatment recommendations, analysis of the first 614 patients enrolled in BATS was reported in July 2021.[15] No significant differences in outcome were observed between patients treated with IVIG alone, glucocorticoids alone, or combination of IVIG and glucocorticoids (IVIG+G), although this may have been due to limited sample size. In this report, we compare the initial treatments for MIS-C in a much larger cohort of children, and also describe the outcomes of cardiac complications.

METHODS

Study Design

Details of the BATS propensity-weighted observational cohort study were described in the initial report.[15] Minor modifications of the data collection procedure and analysis plan were undertaken which are described below and in the published analysis plan and supplementary ~~methods~~appendix.

Briefly, paediatricians world-wide were invited to join BATS and upload data from patients with suspected MIS-C onto a web-based Research Electronic Data Capture database.~~[16]~~[16] from June 2020 through to April 2022. As the spectrum of post-SARS-CoV-2 inflammatory disease was unknown when BATS was initiated,[3,~~65~~,17–19] and the reliability of the published criteria for MIS-~~C~~[20–22]C was unknown we invited recruitment of children with severe inflammatory illness after SARS-CoV-2 infection in addition to those meeting the ~~CDC,~~USA Centre for Disease Control (CDC), WHO or UK case definitions.[20–22] De-identified longitudinal data were collected on presenting features, demography, laboratory findings, immunomodulatory (IVIG, glucocorticoids or biologicals) and supportive treatments. Treatments and daily data were collected by calendar day. Duration of admission, organ support required, and health status on discharge were recorded ~~(see supplementary methods).~~.

The original BATS case report form recorded no data on coronary artery aneurysms (CAA) after hospital discharge, and we therefore added an additional follow up questionnaire regarding CAA resolution (~~supplementary~~ appendix ~~A~~p68).

Treatments and endpoints

The first calendar day of immunomodulatory treatment was defined as “day 0”, and subsequent treatment and outcomes defined relative to this. Primary treatment was defined as the immunomodulatory agent(s) initiated on day 0. Three primary treatment groups were large

enough for weighted comparison according to our predefined sample-size estimations (~~supplementary~~ appendix ~~B~~p69): IVIG alone, glucocorticoids alone, or IVIG+G. Two other groups were pre-defined for additional analyses: those receiving other immunomodulator treatments (including in combination with IVIG and/or glucocorticoids), or no immunomodulator treatments.

Primary outcomes were modified from the previous analysis. The first primary outcome remained a composite of inotropic support or ventilator support (invasive or non-invasive) on day 2 or later, or death. However, the second primary outcome was altered from improvement on the ordinal severity scale by day 2, to time to improvement of at least one level on the ordinal clinical severity scale (ventilated and on inotropic support; ventilated; on inotropic support; receiving oxygen; no supportive therapy stratified by CRP level; and discharged – ~~see supplementary methods).~~ appendix p14). This modification was justified by the greater clinical relevance and additional statistical power of the time to event analysis.

Secondary outcomes included: immunomodulator escalation (any additional immunomodulator, a second dose of IVIG if primary treatment included IVIG, and if primary treatment included glucocorticoids, an increment of 5 mg/kg equivalent daily-dose of prednisolone)[23]; fever from day 2 onwards; individual components of the first primary outcome (death, or inotropic or ventilator support from day 2); CAA occurrence and resolution following treatment (coronary artery Z-score ≥2·5 or aneurysm documented)[24]; left ventricular (LV) dysfunction on echocardiography from day 2 onwards; no improvement in clinical severity scale at day 2; any increase in cardiorespiratory supportive therapy after day 0; therapeutic complications; and temporal dynamics of blood markers of inflammation and organ damage.

Analysis and Statistics

We applied inverse probability of treatment weighting (IPTW) using covariate-balancing propensity scores[25] to account for baseline differences between the three primary treatment

groups. Confounding covariates were selected by expert consensus prior to analysis and were used in both covariate balancing and treatment effect estimation to produce doubly-robust estimates (appendix p18-20). As specified in the analysis plan, IVIG alone was the reference treatment group. Weighted ~~generalized linear models were~~quasibinomial logistic-regression was used for dichotomous outcomes and weighted Cox-regression for time-to-event analyses. Outcomes were reported as adjusted odds ratios or average hazard ratios with 95% confidence intervals and p-values. P-value correction for multiple hypothesis testing was performed for the two primary outcomes and two treatment-group comparisons with the Bonferroni-Holm procedure (~~supplementary methods~~appendix p20).

All clinician-diagnosed MIS-C cases were included in analysis, with those meeting more restrictive definitions evaluated in subgroup or sensitivity analyses: restricting to patients meeting the ~~World Health Organisation (WHO)~~WHO MIS-C criteria,[22] those meeting KD criteria; subgroups by age category and baseline inflammation; analysis by propensity score matching ~~(full details in supplementary methods);;~~ and defining primary treatments as those received on days 0 and 1. Extensive additional subgroup and sensitivity analyses were performed as planned (~~supplementary methods~~appendix p21-22).

Inflammatory markers were plotted as percentages of each patient's peak value by admission day relative to treatment initiation. Smoothed curves with confidence intervals were weighted by the same approach and fitted using the generalized additive model method (~~see supplementary methods~~appendix p18).

Oversight

BATS was designed by the study team at Imperial College London (members and roles in ~~supplement~~appendix p3). Patient data were collected by local investigators (~~Consortium~~consortium members ~~appendix~~ in ~~supplement~~appendix p3-11). The updated statistical analysis plan was developed by the study management team and international

advisory board, and analysis undertaken by the statistical group (~~memberships in supplement).~~ appendix p3). The study was approved by the UK REC (20/HRA/2957) and registered with the international trial registry (ISRCTN69546370). Participating centers obtained ethical approval based on requirements in each country. The initial manuscript was drafted by the first and last authors and developed by all listed authors. The corresponding author, data management group, and analysis group had access to all data, vouching for the completeness and accuracy of data, and for fidelity to the protocol and analysis plan.

RESULTS

From 20th June 2020 to 25th April 2022, data from 2101 MIS-C patients from 39 countries and 121 sites were uploaded to BATS (~~Figs S1, S2, S3~~appendix p48-51). 92 records were excluded, including four neonates and those with incomplete data, duplicate entries, or admission after the recruitment deadline (Fig1A). Of 2009 patients included for analysis, 680 received primary treatment with IVIG, 487 with glucocorticoids, 698 with combination IVIG+G, 59 received other immunomodulator combinations, and 85 received no immunomodulators (Fig1A). In the three main primary treatment groups, 579/1865 (31·0%) received additional immunomodulators by day 2, with ~~722~~953/1865 (~~38·7~~51.1%) receiving secondary agents in total. Treatment trajectories are described in detail (Fig1B ~~and Table S1~~, appendix p23).

Clinical and laboratory findings

Baseline clinical and laboratory findings showed some differences between primary treatment groups (~~Tables~~Table 1, ~~S2 and FigS4~~appendix p24,52). Patients in the no therapy group had significantly less derangement in laboratory markers of inflammation and organ dysfunction, while those in the combined IVIG+G and other treatments groups had the highest level of derangement overall. The combined IVIG+G and other immunomodulator groups had a higher proportion of patients receiving inotropes or ventilation on day 0 (~~FigS6A~~appendix p55). Considering treatment received by day 2, a higher proportion of those on both IVIG+G or in whom biological agents were added were receiving inotropes or ventilated at baseline (~~FigS6B~~appendix p55), but there were no major differences seen in blood markers between these groups (~~FigS5~~appendix p53).

1602/2009 (80·0%) patients met WHO MIS-C criteria (~~Table S3~~appendix p26). The most common missing criterion was evidence of SARS-CoV-2 exposure (~~FigS7~~appendix p56). SARS-CoV-2 antibody measurements were not tested in 406/2009 (20·4~~%,~~%), and negative in

259/2009 (13·0%.%). Bacteria were cultured in the blood of a small proportion of patients (Table S3). appendix p26). 629/2009 (31·3%%) overall, and 544/1602 (34·0%%) of those meeting WHO MIS-C criteria also met the American Heart Association (AHA) definitions for complete KD (Table S4, FigS8appendix p28,57).

Primary outcomes

Of 1865 patients considered for weighted analysis, 169in the three main treatment groups, 166 patients (9·10%) received immunomodulators prior to transfer to the reporting hospital and an additional 191113 patients (10·26.1%) were missing baseline covariates and were excluded from, with a total of 1586 patients considered for our primary weighted analyses. (Fig1A). Acceptable covariate balance was achieved for all IPTW outcome analyses (FigS11-12, S16appendix p60-61,66). For the first primary outcome, receipt of inotropic support or ventilation on day 2 or later, or death, the adjusted odds ratios (OR) for patients receiving primary treatment with IVIG+G, or glucocorticoids alone as compared with IVIG were 1·09 (95% CI: 0·75-1·58, adjusted p-value 1·00) and 0·93 (95% CI: 0·58-1·47, adjusted p-value 1·00) respectively (Fig2A & 2C, Table S5Aappendix p29).

For the second primary outcome, time to improvement on the ordinal clinical severity scale, the adjusted average hazard ratio (AHR) for patients receiving IVIG+G vs IVIG was 1·04 (95% CI: 0·91-1·20, adjusted p-value 1·00) and for glucocorticoids alone vs IVIG was 0·84 (95% CI: 0·70-1·00, adjusted p-value 0·22, Fig2B, 2D & 3A, Table S5Bappendix p29), suggesting slower improvement in the glucocorticoid group. Subgroup analyses of time to improvement in severely ill children (requiring ventilatory or inotropic support at baseline), and those not requiring intensive support showed the suggested slower improvement in those receiving glucocorticoids vs combined treatment was confined to the less severely ill patients (AHR 1·06 (95% CI: 0·75-1·49) in the severe group vs 0·83 (95% CI: 0·62-1·11) in the milder group, Fig2B, 2D & 3B-C, Table S5Eappendix p34).

All sensitivity and subgroup analyses, including restricting to patients meeting WHO MIS-C criteria, showed no significant difference in the first primary outcome for the comparisons of IVIG+G or glucocorticoids alone with IVIG (Fig2A & 2C, ~~Table S5D~~appendix p32). For the second primary outcome, in the subgroup of patients without significant comorbidities, the time-to-improvement was slower in the glucocorticoid group vs IVIG alone (AHR 0·82 (95% CI: 0·69-0·99), Fig2B & 2D, ~~Table S5E~~appendix p34) and the two-point time-to-improvement was slower in the IVIG+G group vs IVIG alone (AHR 0·87 (95% CI: 0·75-1·00)). All other planned sensitivity and subgroup analyses showed no significant difference in time-to-improvement for the comparisons of IVIG+G or glucocorticoids alone with IVIG alone.

Secondary outcomes

Escalation of immunomodulator treatment was less common in the IVIG+G and glucocorticoid groups compared to the IVIG group (OR 0·15 (95% CI: 0·11-0·20) & 0·68 (95% CI: 0·50-0·93) respectively, ~~FigS9A-B~~appendix p58). Persistent fever from day 2 was less common in patients receiving IVIG+G vs IVIG alone (OR 0·50 (95% CI: 0·38-0·67)), with no difference between the glucocorticoid or IVIG groups. In ~~an unplanned~~a post-hoc sensitivity analysis, there was no difference in persistent fever from day 3 between the IVIG+G and IVIG groups. Individual components of the composite outcome showed no differences between treatments (Fig2A & 2C, ~~Table S5C~~appendix p30).

Of 1918 with reported echocardiograms, 236 (12·3%) had CAA at any time (13·6% in IVIG recipients, 8·9% glucocorticoid, and 12·9% IVIG+G (~~Table S6B~~appendix p36)), with the largest disparity in aneurysm detection before starting immunomodulatory treatment (~~Table S6A~~appendix p36). In the 705 patients with inpatient echocardiograms before and after treatment initiation (~~Table S5C~~appendix p30) 50 (7.1%) had CAA present on the final echocardiogram before discharge, with no statistically significant difference apparent between

groups after IPTW analysis, including for ~~unplanned subgroup~~post-hoc analyses restricted to patients who did and did not meet complete KD criteria (~~Fig~~S9A ~~B~~appendix p58).

Follow-up echocardiogram data were available in 196/236 (83·1%) patients with CAA during admission. Most CAA resolved during follow-up (92·9% total), with similar rates amongst primary treatment groups (~~Table S6B~~appendix p36). Similar rates of resolution were seen when restricted to patients with follow-up by 6- and 12-weeks (~~Tables S6C-D~~appendix p37).

To establish if patients who did not receive IVIG were at greater risk of CAA, or had different rates of resolution, we explored CAA incidence in the glucocorticoid alone primary treatment group. 17/239 (7·1~~%~~%) of those never receiving IVIG had CAA detected at any time during admission, compared with 24/221 (10·9~~%~~%) who received IVIG later during admission. ~~Addition of IVIG as secondary treatment is more likely in severely ill patients, those with CAA detected, and those not rapidly improving on primary treatment. Therefore, those receiving later IVIG are likely to have had more severe disease.~~ CAA were present at discharge in 5/239 (2·1~~%~~%) of those without later IVIG and 9/221 (4·1~~%~~%) of those receiving later IVIG treatment, with over 93% of CAA resolving in both groups on reported follow-up (~~Table S6E~~appendix p38). No difference was seen between treatment groups in severity of CAA as judged by the distribution of z-scores (~~Table S7A~~appendix p39). Larger z-scores were seen in younger patients (~~Table S7B~~appendix p39).

Left ventricular dysfunction was reported in 202/1512 (13·4%) of patients with echocardiograms from day 2 onwards, with no difference between groups (~~Fig S9A-B, Table S5C).~~appendix p30,58). There were no differences between IVIG+G or glucocorticoids vs IVIG for the secondary outcomes of no improvement by day 2 or increase in level of support after initiation of primary treatment. Death occurred in 8~~,~~ (1·2% unadjusted), 10 (2·1%) and 5 (0·8%) patients in the complete IVIG+G, glucocorticoid and IVIG groups respectively.

Drug complications were reported in 59/1623 (3·6%) of patients receiving any glucocorticoids and 25/1658 (1·5%) patients receiving IVIG. Glucocorticoid complications were predominantly hypertension and hyperglycemia (~~Table S8~~appendix p40).

IVIG+G vs Glucocorticoids alone

A planned secondary analysis comparing glucocorticoids alone and combined IVIG+G demonstrated no difference in the first primary outcome, but a faster time-to-improvement for the IVIG+G group (AHR 1·25 (95% CI: 1·05-1·48), ~~FigS10~~appendix p59). This was predominantly seen in the later days following treatment, and in those patients not requiring intensive support at baseline (Fig3A-C). Secondary outcomes for this comparison showed that escalation of primary therapy and persistence of fever from day 2 were more common in the glucocorticoid alone group (~~FigS10, Table S5C~~appendix p30,59).

Effect of Immunomodulation on blood markers.

CRP declined more rapidly in patients receiving immunomodulators than in untreated patients (Fig4A). Comparison of primary treatment groups showed more rapid decline of CRP in the glucocorticoid and IVIG+G groups than in the IVIG treated patients (Fig4B). There was a suggestion of more rapid decline in troponin and ferritin in the glucocorticoid and combined treatment groups with a similar trend when restricting to those not receiving additional treatment between days 0 and 2 (Fig4C). Time course plots of other blood markers showed similar dynamics of blood markers between groups (~~FigS13~~appendix p62).

To investigate whether inadvertent inclusion of children with KD within BATS enrolment might have influenced treatment responses, we explored changes in blood markers separately in children most resembling KD. As KD is generally a disease in children aged 5-years and below, and MIS-C is often reported in older children, we compared those meeting AHA criteria for KD, and all children under 6 years (“KD-like”), with the remaining MIS-C patients.

The rate of decline in CRP was similar between the younger and older children and those fulfilling KD criteria treated with IVIG, with a suggestion of a more rapid decline in CRP in the non-KD-like patients receiving glucocorticoids alone (~~FigS14~~appendix p63).

DISCUSSION

Our comparison of treatment outcomes in an international cohort of 2009 children with MIS-C shows that treatment with glucocorticoids alone, or IVIG+G are not associated with significant differences in primary outcomes (requirements for inotropic support, ventilation on day two or beyond, or death; or rate of improvement on the ordinal severity scale) in comparison with IVIG alone. The findings are consistent with our preliminary report of 614 children.[15] However, the larger number of patients in each treatment group, increases the confidence in our findings. There was a non-significant trend towards a slower rate of improvement in patients treated with glucocorticoids alone in comparison with IVIG, but this comparison was confined to those with less severe illness at presentation. Reassuringly, we found no difference in CAA outcomes between primary treatment groups, with resolution seen in the vast majority of patients.

Our planned secondary analysis comparing glucocorticoids alone with combined IVIG+G demonstrated no difference in the first primary outcome, but a faster time-to-improvement for the IVIG+G group. This comparison was not adjusted for multiple hypothesis testing and the effect appears confined to those patients not requiring intensive support at baseline. Other Secondary endpoints, and thus also not corrected for multiple hypothesis testing, showed lower rates of treatment escalation and lower rates of fever on day 2 in the IVIG+G group.

A key question for clinicians is whether the potential incremental benefits of IVIG+G to reduce severity of illness and hasten resolution of fever are sufficient to justify the use of both agents. We note that the primary outcomes (progression or recovery from organ support) were chosen to select the most clinically important outcomes, whereas the secondary outcomes may detect less clinically ~~important findings~~important findings. Furthermore, we suggest that the finding of more common escalation of treatment for those on single agents, which was also observed in earlier studies,[12,13] may be biased by greater clinician readiness to add other treatments in seriously ill

patients who do not rapidly improve on monotherapy, whereas options to escalate treatment are fewer in patients treated with primary combination IVIG+G.

This question of whether combined IVIG+G is beneficial as compared to Glucocorticoids alone is relevant to both resource rich countries where IVIG is readily available and countries where IVIG has limited availability or cost imposes limitations in its use. For resource limited settings, our data suggests that primary treatment with glucocorticoids alone, is a safe alternative to IVIG or combined treatment, with IVIG being reserved for patients who fail to improve on glucocorticoids alone. For countries where IVIG cost is less prohibitive, the limited supply of IVIG and potential for combined treatments to have more side effects than single agents would argue for initial treatment with a single agent, and addition of second agents only in those who do not improve.

A higher proportion of patients receiving IVIG+G as primary treatment were receiving inotropes or ventilation at day 0, and had more deranged blood markers, suggesting more severely ill patients may have received IVIG+G. Importantly, key differences between treatment groups were adjusted for in the propensity score analysis. Children treated with IVIG+G had more rapid resolution of fever than children treated with IVIG or glucocorticoids alone. However no other clinically significant findings were more frequent in the IVIG+G group in comparison with either of the single agent treatment groups.

~~In keeping with earlier studies,[12,13] we observed a higher rate of treatment escalation (addition of other treatments) in the patients receiving single agents as primary treatment. This may be explained by a greater readiness to escalate when only one treatment was given, fewer options to escalate when initial treatment contained two agents, by failure of initial treatment, or by greater severity of illness in those receiving additional treatment.~~ Patients who were initially treated with glucocorticoids or IVIG alone and then received additional treatment by day 2 were more likely to be receiving inotrope or ventilatory support at baseline. However, patients who received additional treatment did not differ substantially from patients who did not receive additional treatment across multiple biomarkers, suggesting that treatment with inotrope or ventilatory support influenced the

clinical decision for administration of additional treatment. We have included adjustment for both baseline inotrope and ventilatory support in our IPTW analysis.

The use of IVIG as treatment for MIS-C has largely been driven by the similarity of MIS-C to KD, for which IVIG is the established treatment to reduce risk of CAA.[9] As coronary artery aneurysms ~~(CAA)~~ are observed in 10-20% of MIS-C cases,[13,15,26] there has been concern that failure to include IVIG in initial treatment would be associated with increased risk of CAA. We found that the incidence of CAA in patients receiving glucocorticoids as initial treatment was similar to the incidence of CAA in IVIG recipients (either IVIG or IVIG+G). Furthermore, the severity of CAA (as measured by z-score) and the proportion of patients undergoing complete resolution of CAA by time of discharge, or on follow up was similar in the glucocorticoid alone group to the IVIG and IVIG+G groups, including post-hoc analysis restricting to patients who never received IVIG. Our study thus provides reassurance that initial therapy with single agent glucocorticoids is not associated with increased risk of long-term coronary artery damage in MIS-C.

The American College of Rheumatologists currently recommends combined treatment with IVIG and glucocorticoids for MIS-C,[11] based on limited evidence of benefit from the USA and French propensity matched studies,[12,13] which showed lower rates of treatment escalation and improved cardiac function detected by echocardiogram with combined therapy. Neither of these studies included a glucocorticoid only group, and both were smaller than our current analysis. ~~Our planned primary endpoints of death or requirement for inotropes or ventilation on day 2, and time to improvement, did not show any significant difference when comparing either glucocorticoids alone or IVIG+G with IVIG alone. A planned secondary analysis comparing glucocorticoids alone and combined IVIG+G demonstrated no difference in the first primary outcome, but a faster time-to-improvement for the IVIG+G group. This comparison was not adjusted for multiple hypothesis testing and was confined to those patients not requiring intensive support at baseline. Secondary endpoints, not corrected for multiple hypothesis testing, showed higher rates of treatment escalation and lower rates of fever on day 2 in the IVIG+G group. While there is thus suggestive~~

~~evidence of benefit of combined therapy based on these secondary outcomes, a key question for clinicians managing patients with MIS-C is whether the lower rate of treatment escalation and potentially slower resolution of the illness as detected on the ordinal severity scale, and persistent fever, warrants continued use of both IVIG and glucocorticoids as primary treatment. This question~~

~~is relevant to both resource-rich countries where IVIG is readily available and countries where IVIG has limited availability or cost imposes limitations in its use. For resource-limited settings, our data suggests that primary treatment with glucocorticoids alone, is a safe alternative to IVIG or combined treatment, with IVIG being reserved for patients who fail to improve on glucocorticoids alone. For countries where IVIG cost is less prohibitive, the limited supply of IVIG and potential for combined treatments to have more side effects than single agents would argue for initial treatment with a single agent, and addition of second agents only in those who do not improve.~~

We observed a more rapid decline in CRP in all three treatment groups as compared to patients not receiving immunomodulators. Although the curves for each treatment were overlapping, there was a non-significant trend to a more rapid decline in CRP, ferritin and troponin in the glucocorticoid containing groups.

Our study has several limitations. A key concern is the extent to which a retrospective comparison of outcomes following non-randomised choice of treatment can be used to guide clinical practice. We applied two different propensity score methods (weighting and matching), to remove bias caused by differences in severity, demography, or resource setting. We achieved good covariate balance between comparator groups using both approaches. However, other unmeasured differences might influence the results, and a large RCT would be the preferred approach to provide definitive answers. In addition, there is a risk of bias from the voluntary nature of data collection, as not all cases of MIS-C from each site were necessarily included in the study.

A second potential limitation is our use of the broad inclusion criteria of clinician diagnosed MIS-C. At the time BATS was initiated the accuracy of the published diagnostic criteria was unknown, and there were differences between the WHO, CDC and RCPCH criteria. Furthermore, availability of antibody testing for SARS-CoV-2 was limited in many countries. We therefore chose to include patients whose responsible clinicians considered them to have MIS-C, and in whom alternative diagnoses had been excluded. As we expected, our data confirms that the most commonly "missed" criteria to meet the WHO or CDC definitions of MIS-C was the presence of evidence of SARS-CoV-2 exposure. It is noteworthy that as the pandemic has evolved, and a high proportion of children have become SARS-CoV-2 antibody positive through natural infection or vaccination, the value of antibody against SARS-CoV-2 as evidence of recent infection has reduced. In view of the high rates of SARS-CoV-2 infection in schools, and the high proportion of asymptomatic childhood infection, a history of exposure to infection is of little value in diagnosis of MIS-C, and the WHO and CDC criteria may need to be re-evaluated. Despite these concerns, the large majority of patients in BATS did meet the WHO criteria, with only small differences in the proportions from each of the primary treatment groups. Our subgroup and sensitivity analyses did not find any difference in outcome when restricted to those meeting the WHO criteria, or the group with features overlapping KD.

~~Another limitation is~~An additional concern may be that the nature, severity, and epidemiology of MIS-C has changed over time, and with successive SARS-CoV-2 waves and introduction of childhood vaccinations against COVID-19. The disorder appears to have become less common in many countries as a high proportion of children have previous infection, and both natural infection and vaccination may reduce the incidence of MIS-C.[27] However, with SARS-CoV-2 now increasing in the previously unexposed population of China, there is likely to be a new wave of MIS-C and the findings reported here may be of considerable help to the clinicians experiencing this disease for the first time.

Other limitations include the wide variety of steroid dosing regimens used, and the large number of patients in whom additional treatments were added after the primary treatment. Although we have attempted to compare those remaining on a single agent, this group may have been less severely ill and therefore not representative of the treatment group overall. Additionally, after excluding patients with incomplete baseline covariates from the IPTW analysis, the final numbers of patients used for primary analyses were marginally below those stated in our sample-size calculations. However, the suggested effect sizes in these calculations are relatively arbitrary. More important is the final width of confidence intervals for treatment effects, which were generally small for our primary analyses. ~~A third~~An additional limitation is the use of a composite primary outcome. This was necessitated by the relatively small numbers of patients with individual outcomes, and our aim to capture effects of treatment in patients across a wide spectrum of severity. As mitigation we evaluated the individual components of the composite score as secondary analyses. The time-to-improvement outcome also incurs the possibility of "built-in selection bias"[28] although we have attempted to isolate known factors that could incur such bias through extensive subgroup analyses. This limitation is relevant to all survival analysis, and would not be avoidable even for RCTs using the same outcome. Finally, we are not able to detect rare or longer-term effects of either IVIG or glucocorticoid administration.

The absence of significant differences between treatment groups poses several questions on the mechanisms underlying MIS-C. As IVIG and glucocorticoids have different possible modes of action in MIS-C,[~~27,28~~29,30] the lack of difference between them, and the fact that dual therapy was not superior to single agent therapy is puzzling. One possible explanation might be different underlying disease processes in MIS-C, some of which respond to IVIG and some to glucocorticoids. If so, we would have expected that combination treatment would be superior to each treatment individually. Alternatively, glucocorticoids and IVIG may act at different points in the same causal pathway and with equal efficacy. This would explain the similar outcomes and lack of additive effect. A final possibility is that neither treatment has a significant effect on the

disease process. As the number of patients receiving no immunomodulator treatment was small and phenotypically distinct from those receiving immunomodulator treatment, we did not have an adequate “No treatment group” to evaluate this possibility. However, the more rapid decline in CRP in the treated vs untreated groups supports a beneficial effect of all three treatment regimes. In addition to IVIG and glucocorticoids, several other immunomodulatory agents were administered, including anti-IL1, anti-IL6 and anti-TNF agents. The numbers of patients receiving these agents were too low to enable IPTW comparison between them, or with IVIG, glucocorticoids and IVIG+G. Biologicals tended to be administered in combination with IVIG and glucocorticoids, and to more unwell patients.

The key question in interpreting clinical significance of this analysis is whether the findings are sufficiently robust to enable glucocorticoids to replace IVIG as primary treatment of MIS-C. The lack of significant difference in outcomes between patients treated with glucocorticoids as primary treatment, and those receiving IVIG or IVIG+G, and in particular the lack of difference in CAA severity, frequency, or resolution, suggests that initial treatment with glucocorticoids is a safe alternative to IVIG. A concern in adopting this approach is the difficulty in distinguishing MIS-C from KD, particularly in younger patients, and the possibility that IVIG will be withheld from children with KD because they are thought to have MIS-C. This concern highlights the need for a rapid diagnostic test to distinguish MIS-C from KD, as well as the need for urgent cardiology assessment in patients presenting with a suspected diagnosis of either disease. It also suggests that where clinical features closely resemble KD, particularly in younger children, retaining IVIG as a component of initial therapy is prudent.

MIS-C has emerged as an important childhood problem in low- and middle-income countries.[26,~~29~~31] As IVIG is ~~costly[30]~~costly[32] and has limited availability in many countries, its use in preference to cheaper anti-inflammatory agents such as glucocorticoids should be supported by sound evidence. We did not find significant differences in outcome between treatment with glucocorticoids or IVIG as single agents or between the single and dual agent primary

treatments. Our findings suggest that glucocorticoids are not inferior to IVIG or combination IVIG+G as primary treatment of MIS-C, and their wide availability and lower cost would support their choice as initial treatment for MIS-C.

Figures and Tables

Figure 1A | Study flowchart

The study flow chart gives an overview of the total number of patients enrolled, excluded, and included for the analyses. Patients meeting the inclusion criteria are categorized by treatment groups (IVIG, Glucocorticoids, IVIG & Glucocorticoids, Other immunomodulator treatments [this includes: anti-tumor necrosis factor, anti-interleukin 1, anti-interleukin 6] and no immunomodulator treatments) and subdivided by our data-drive classification according to the WHO MIS-C criteria.

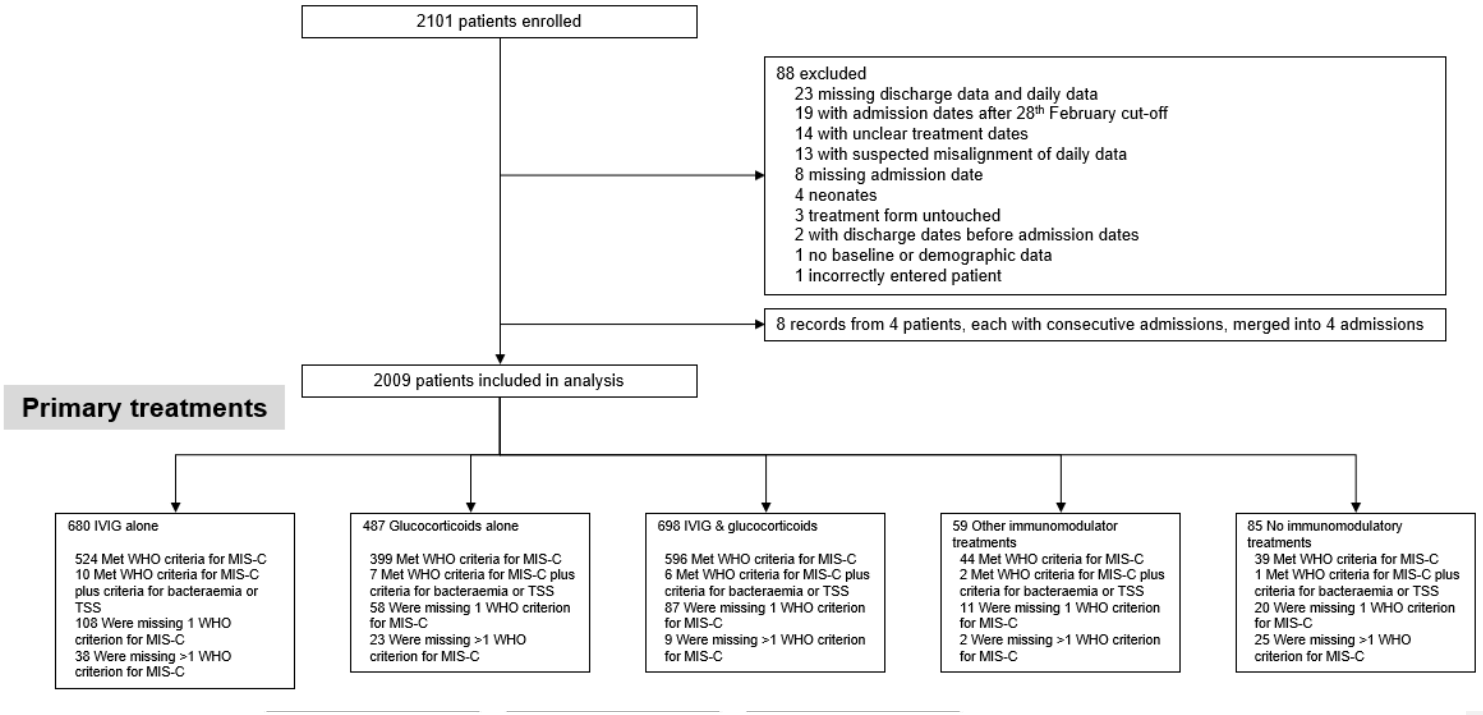


Figure 1B | Treatments received by patients over time following initiation of immunomodulator treatment

The Sankey diagram demonstrates the number of patients receiving cumulative therapies from days following initiation of immunomodulator treatment. Each vertical stack represents a different day in the patients' admission relative to starting immunomodulatory treatment (~~Days~~days 0 –to 5), with day 0 representing the first day of immunomodulator treatment. The grey bands represent movement of patients between treatment groups from relative day 0 to 1, day 1 to 2, day 2 to 3, day 3 to 4 and day 4 to 5. The width of the grey bands is proportional to the number of patients (flow). The flow of patients is independent between time intervals; there is no continuous correspondence across days 1 to 5. The treatment groups are as stated. Of note, "Glucocorticoids" include intravenous and oral glucocorticoids (~~Table S9~~appendix p41). "Other" includes one or more other immunomodulatory treatment(s) given alone or in combination with Glucocorticoids and/or IVIG. Other immunomodulatory treatments include: anti-interleukin1, anti-interleukin 6, anti-tumour necrosis factor, cytokine adsorber (CytoSorb), granulocyte colony stimulating factor, colchicine, mesenchymal stem cells, convalescent plasma, ~~Cyclophosphamide, Plasmapheresis and Hydroxychloroquine~~cyclophosphamide, plasmapheresis and hydroxychloroquine

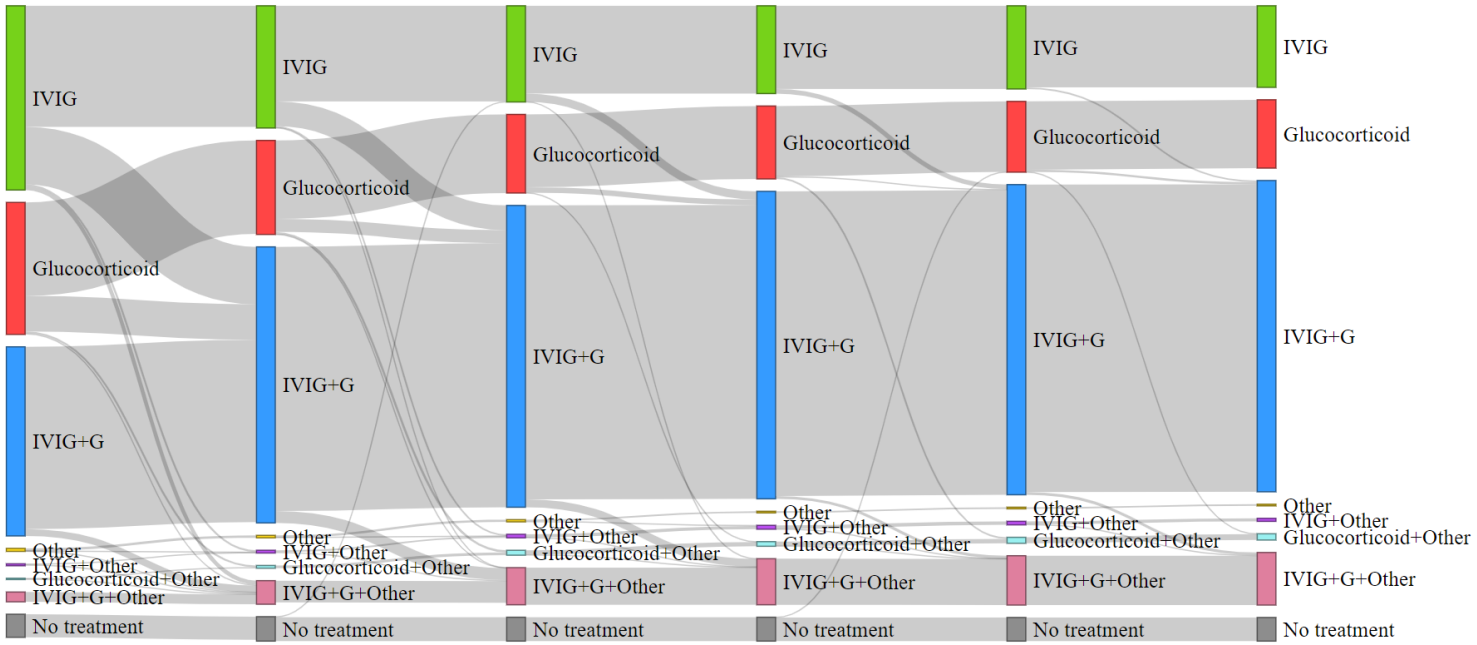
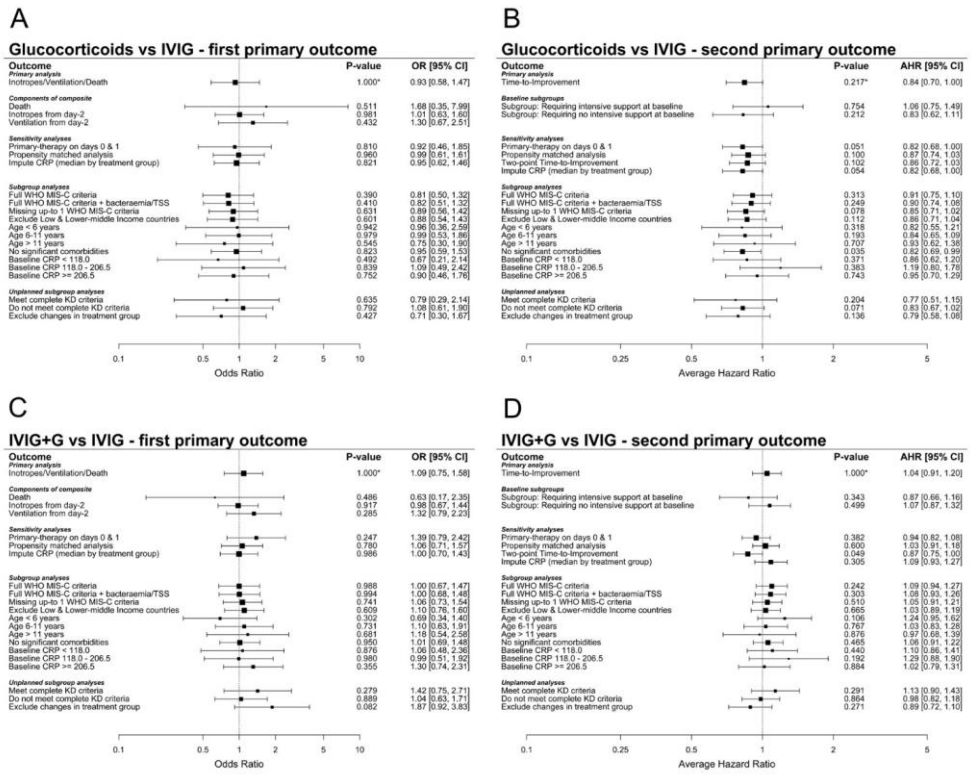


Figure 2 | Forest plots summarizing point estimates and 95% confidence intervals for primary analyses, including all subgroup and sensitivity analyses.

*Shown are outcomes for patients with suspected MIS-C who received IVIG plus glucocorticoids (Panels A & B) or glucocorticoids alone (Panels C & D) as compared with those who received IVIG alone (reference group, indicated by an odds ratio or average hazard ratio of 1·00). Displayed values are adjusted odds ratios or average hazard ratios (indicated on the x-axis). Panels A & C show the first primary outcome analyses, risk of inotropes, ventilation or death, and values to the right of the dotted line indicate superiority of IVIG alone. Panels B & D show the second primary outcome analyses, time to improvement in ordinal clinical severity score, with values to the left indicating superiority of IVIG alone. *indicates p-values corrected for multiple hypothesis testing using the Bonferroni-Holm procedure, observed p-value x4. Absolute numbers of patients included in each analysis can be found in appendix p29-32.*

Abbreviations: CRP: C-reactive protein; KD: Kawasaki Disease; WHO: World Health Organisation.



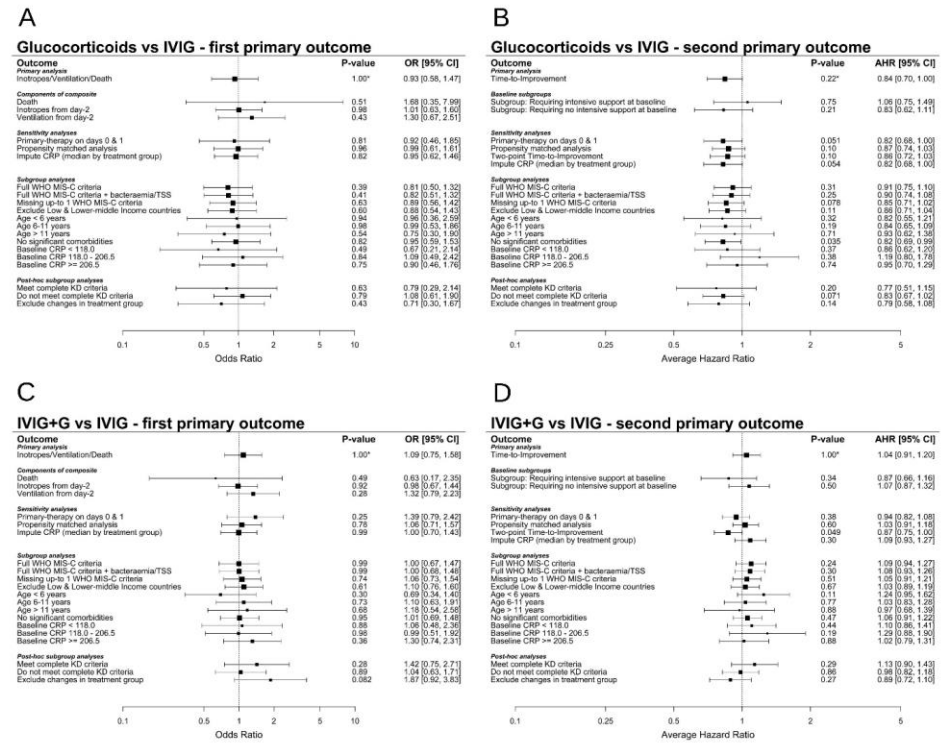
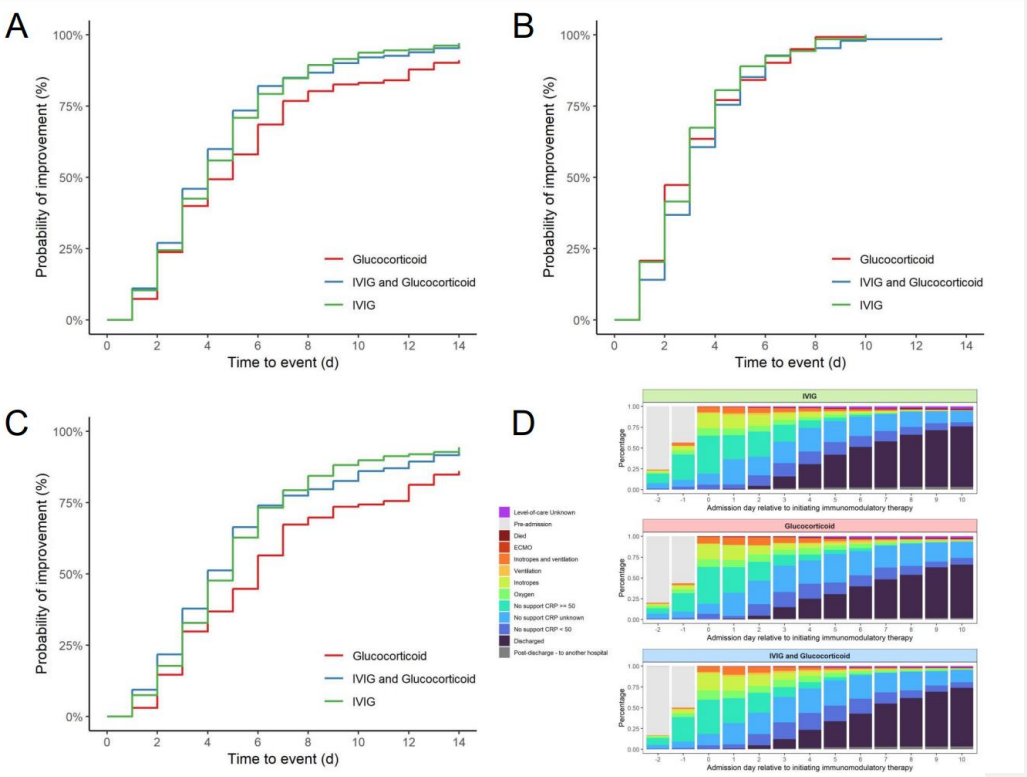
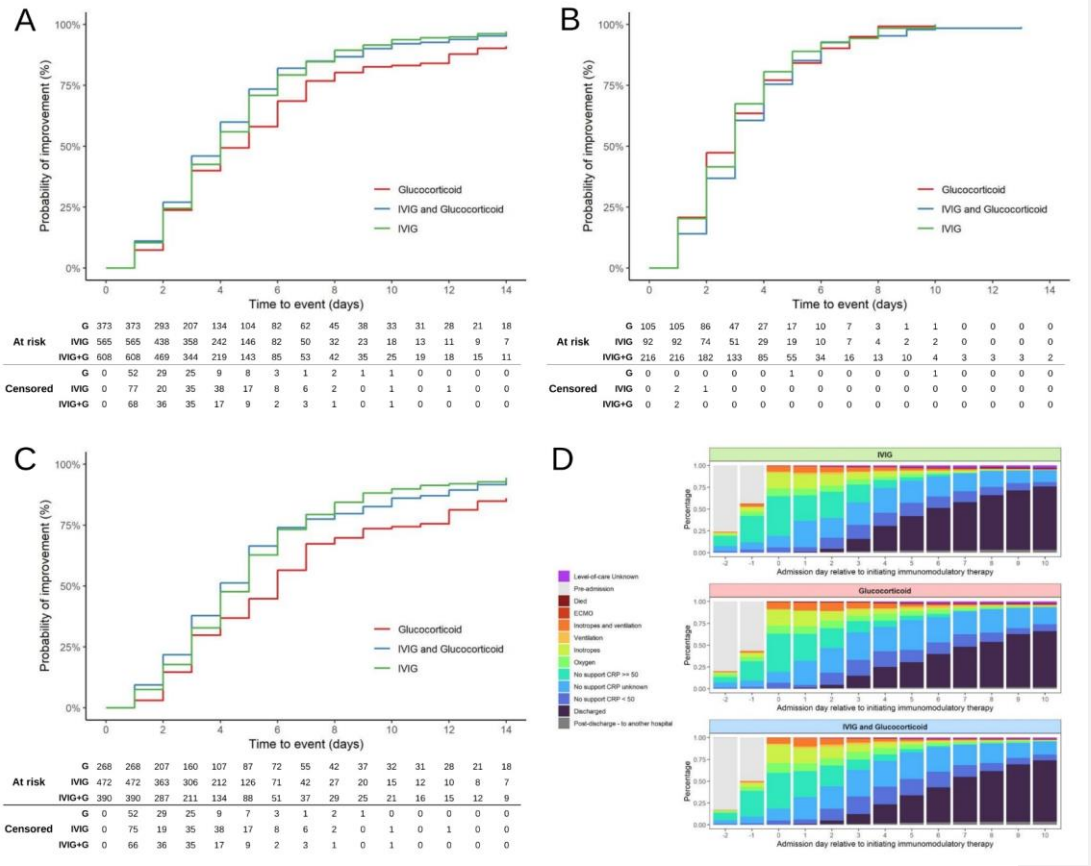

A
Glucocorticoids vs IVIG - first primary outcome
Outcome
P-value
OR [95% CI]
Primary analysis
Inotropes/Ventilation/Death
1.00*
0.93 [0.58, 1.47]
Components of composite
Death
0.51
1.68 [0.35, 7.99]
Inotropes from day-2
0.98
1.01 [0.63, 1.60]
Ventilation from day-2
0.43
1.30 [0.67, 2.51]
Sensitivity analyses
Primary-therapy on days 0 & 1
0.81
0.92 [0.46, 1.85]
Propensity matched analysis
0.96
0.99 [0.61, 1.61]
Impute CRP (median by treatment group)
0.82
0.95 [0.62, 1.46]
Subgroup analyses
Full WHO MIS-C criteria
0.39
0.81 [0.50, 1.32]
Full WHO MIS-C criteria + bacteraemia/TSS
0.41
0.82 [0.51, 1.32]
Missing up-to 1 WHO MIS-C criteria
0.63
0.89 [0.56, 1.42]
Exclude Low & Lower-middle Income countries
0.60
0.88 [0.54, 1.43]
Age < 6 years
0.94
0.96 [0.36, 2.59]
Age 6-11 years
0.98
0.99 [0.53, 1.86]
Age > 11 years
0.54
0.75 [0.30, 1.90]
No significant comorbidities
0.82
0.95 [0.59, 1.53]
Baseline CRP < 118.0
0.49
0.67 [0.21, 2.14]
Baseline CRP 118.0 - 206.5
0.84
1.09 [0.49, 2.42]
Baseline CRP >= 206.5
0.75
0.90 [0.46, 1.76]
Post-hoc subgroup analyses
Meet complete KD criteria
0.63
0.79 [0.29, 2.14]
Do not meet complete KD criteria
0.79
1.08 [0.61, 1.90]
Exclude changes in treatment group
0.43
0.71 [0.30, 1.67]
0.1
0.5
1
2
5
10
Odds Ratio
B
Glucocorticoids vs IVIG - second primary outcome
Outcome
P-value
AHR [95% CI]
Primary analysis
Time-to-Improvement
0.22*
0.84 [0.70, 1.00]
Baseline subgroups
Subgroup: Requiring intensive support at baseline
0.75
1.06 [0.75, 1.49]
Subgroup: Requiring no intensive support at baseline
0.21
0.83 [0.62, 1.11]
Sensitivity analyses
Primary-therapy on days 0 & 1
0.051
0.82 [0.68, 1.00]
Propensity matched analysis
0.10
0.87 [0.74, 1.03]
Two-point Time-to-Improvement
0.10
0.86 [0.72, 1.03]
Impute CRP (median by treatment group)
0.054
0.82 [0.68, 1.00]
Subgroup analyses
Full WHO MIS-C criteria
0.31
0.91 [0.75, 1.10]
Full WHO MIS-C criteria + bacteraemia/TSS
0.25
0.90 [0.74, 1.08]
Missing up-to 1 WHO MIS-C criteria
0.078
0.85 [0.71, 1.02]
Exclude Low & Lower-middle Income countries
0.11
0.86 [0.71, 1.04]
Age < 6 years
0.32
0.82 [0.55, 1.21]
Age 6-11 years
0.19
0.84 [0.65, 1.09]
Age > 11 years
0.71
0.93 [0.62, 1.38]
No significant comorbidities
0.035
0.82 [0.69, 0.99]
Baseline CRP < 118.0
0.37
0.86 [0.62, 1.20]
Baseline CRP 118.0 - 206.5
0.38
1.19 [0.80, 1.78]
Baseline CRP >= 206.5
0.74
0.95 [0.70, 1.29]
Post-hoc subgroup analyses
Meet complete KD criteria
0.20
0.77 [0.51, 1.15]
Do not meet complete KD criteria
0.071
0.83 [0.67, 1.02]
Exclude changes in treatment group
0.14
0.79 [0.58, 1.08]
0.1
0.25
0.5
1
2
5
Average Hazard Ratio
C
IVIG+G vs IVIG - first primary outcome
Outcome
P-value
OR [95% CI]
Primary analysis
Inotropes/Ventilation/Death
1.00*
1.09 [0.75, 1.58]
Components of composite
Death
0.49
0.63 [0.17, 2.35]
Inotropes from day-2
0.92
0.98 [0.67, 1.44]
Ventilation from day-2
0.28
1.32 [0.79, 2.23]
Sensitivity analyses
Primary-therapy on days 0 & 1
0.25
1.39 [0.79, 2.42]
Propensity matched analysis
0.78
1.06 [0.71, 1.57]
Impute CRP (median by treatment group)
0.99
1.00 [0.70, 1.43]
Subgroup analyses
Full WHO MIS-C criteria
0.99
1.00 [0.67, 1.47]
Full WHO MIS-C criteria + bacteraemia/TSS
0.99
1.00 [0.68, 1.48]
Missing up-to 1 WHO MIS-C criteria
0.74
1.06 [0.73, 1.54]
Exclude Low & Lower-middle Income countries
0.61
1.10 [0.76, 1.60]
Age < 6 years
0.30
0.69 [0.34, 1.40]
Age 6-11 years
0.73
1.10 [0.63, 1.91]
Age > 11 years
0.68
1.18 [0.54, 2.58]
No significant comorbidities
0.95
1.01 [0.69, 1.48]
Baseline CRP < 118.0
0.88
1.06 [0.48, 2.36]
Baseline CRP 118.0 - 206.5
0.98
0.99 [0.51, 1.92]
Baseline CRP >= 206.5
0.36
1.30 [0.74, 2.31]
Post-hoc subgroup analyses
Meet complete KD criteria
0.28
1.42 [0.75, 2.71]
Do not meet complete KD criteria
0.89
1.04 [0.63, 1.71]
Exclude changes in treatment group
0.082
1.87 [0.92, 3.83]
0.1
0.5
1
2
5
10
Odds Ratio
D
IVIG+G vs IVIG - second primary outcome
Outcome
P-value
AHR [95% CI]
Primary analysis
Time-to-Improvement
1.00*
1.04 [0.91, 1.20]
Baseline subgroups
Subgroup: Requiring intensive support at baseline
0.34
0.87 [0.66, 1.16]
Subgroup: Requiring no intensive support at baseline
0.50
1.07 [0.87, 1.32]
Sensitivity analyses
Primary-therapy on days 0 & 1
0.38
0.94 [0.82, 1.08]
Propensity matched analysis
0.60
1.03 [0.91, 1.16]
Two-point Time-to-Improvement
0.049
0.87 [0.75, 1.00]
Impute CRP (median by treatment group)
0.30
1.09 [0.93, 1.27]
Subgroup analyses
Full WHO MIS-C criteria
0.24
1.09 [0.94, 1.27]
Full WHO MIS-C criteria + bacteraemia/TSS
0.30
1.08 [0.93, 1.26]
Missing up-to 1 WHO MIS-C criteria
0.51
1.05 [0.91, 1.21]
Exclude Low & Lower-middle Income countries
0.67
1.03 [0.89, 1.19]
Age < 6 years
0.11
1.24 [0.95, 1.62]
Age 6-11 years
0.77
1.03 [0.83, 1.28]
Age > 11 years
0.88
0.97 [0.68, 1.39]
No significant comorbidities
0.47
1.06 [0.91, 1.22]
Baseline CRP < 118.0
0.44
1.10 [0.86, 1.41]
Baseline CRP 118.0 - 206.5
0.19
1.29 [0.88, 1.90]
Baseline CRP >= 206.5
0.88
1.02 [0.79, 1.31]
Post-hoc subgroup analyses
Meet complete KD criteria
0.29
1.13 [0.90, 1.43]
Do not meet complete KD criteria
0.86
0.98 [0.82, 1.16]
Exclude changes in treatment group
0.27
0.89 [0.72, 1.10]
0.1
0.25
0.5
1
2
5
Average Hazard Ratio

Figure 3 | Weighted clinical improvement over time

Panels A-C: Kaplan-Meier curves for the three main primary treatment groups showing time to one-point improvement in clinical severity on ordinal scale weighted by inverse probability of treatment, for (A) all patients, (B) subgroup of patients needing at least one of inotropes or ventilation at baseline, (C) subgroup of patients not requiring inotropes or ventilation at baseline. Tables below the Kaplan-Meier curves show the numbers at risk at the start of each day, and the number censored at this specific time point. Panel D: Clinical severity on ordinal scale, shown as proportional column charts from two days before treatment to 10 days after treatment, separated by primary treatment group, and weighted by inverse probability of treatment. Additional groups have been added for graphical purposes. ~~Abbreviations: CRP: C-reactive protein; ECMO: extracorporeal membrane oxygenation.~~



Abbreviations: CRP: C-reactive protein; ECMO: extracorporeal membrane oxygenation.



Formatted: Justified, Line spacing: Double

Figure 4 | Change in C-reactive protein (CRP), troponin and ferritin over time

Each of three key markers of inflammation (C-reactive protein, troponin, and ferritin) is plotted as a line and weighted by the covariate balancing propensity score. The levels are shown as a percentage of each patient's peak value, plotted by day relative to starting treatment. A generalized additive model was used to fit the curves. For each plot patients are only included if they had blood results available both before and after treatment initiation, and only if their last value up to treatment initiation was abnormal (CRP ≥ 8mg/L, troponin ≥ 14 ng/L, and ferritin ≥ 50 microgram/L). Panel A shows the fitted curves for the three measures in patients who received any immunomodulators, as compared with those who did not receive immunomodulators, using day of admission as relative admission day for patients not receiving immunomodulator treatment (NOTE: Curves for troponin in panel A were fitted using a loess model due to small sample numbers). Panel B shows the fitted curves for patients who received IVIG alone, IVIG plus glucocorticoids, and glucocorticoids alone as their primary treatment. Panel C shows the fitted curves for the three treatments combined in the patients whose primary treatment did not change between treatment initiation (day 0) and day 2.

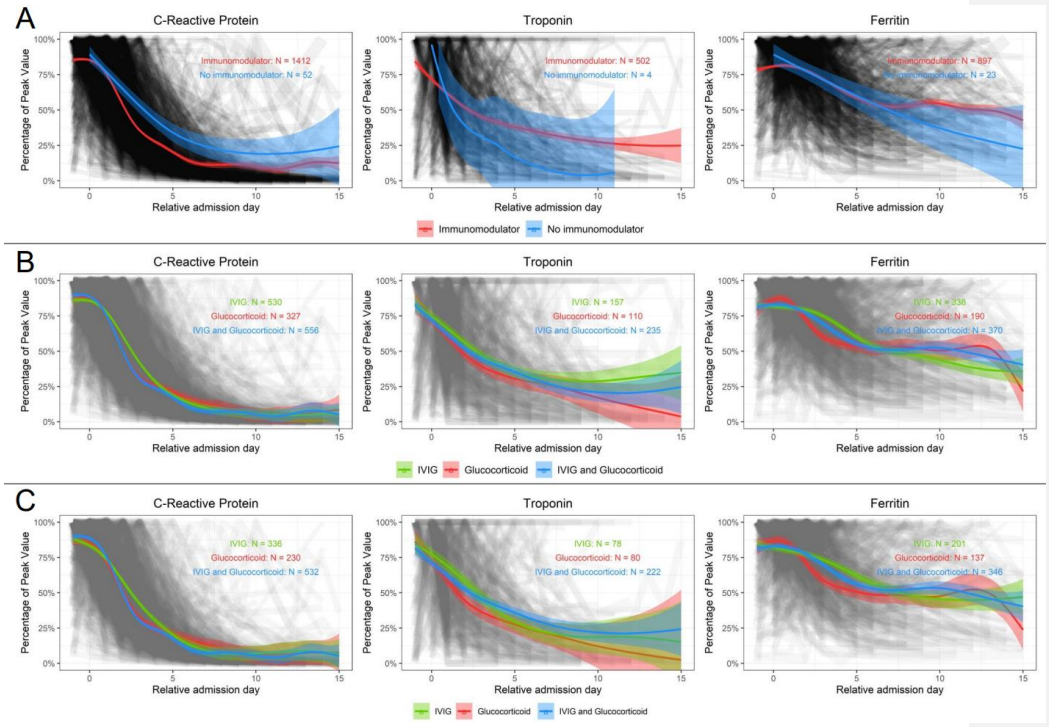


Table 1 | Clinical and demographic features in all treatment groups

Descriptive table of demographic features, clinical features and blood markers on admission, and proportion of patients meeting Kawasaki Disease criteria according to American Heart Association criteria. Patients with coronary artery aneurysms met the definition of Kawasaki Disease with less than 4 Kawasaki Disease clinical features. Patients were divided by treatment arm on day 0 (IVIG alone, glucocorticoid alone, IVIG+G, no treatment, and other (any other treatment combination including biologicals)). SARS-CoV-2 PCR data refer to test taken during admission. Organ support refers to receipt of ventilation, inotropes or ECMO on admission. Missing data (where applicable) are available in a full unabridged version in ~~Table S2~~appendix p24.

Abbreviations: Ab: Antibody; CRP: C-reactive protein; ECMO: extracorporeal membrane oxygenation; PCR: polymerase chain reaction. ^Clinical and demographic features given as ~~raw values~~number and (%). ⁺Numerical values given as median values and [interquartile ranges].

	Everyone (N=2009)	IVIG (N=680)	Glucocorticoids (N=487)	IVIG and Glucocorticoids (N=698)	Other (N=59)	No treatment (N=85)
⁺Age	8.0 [4.2 - 11]	6.8 [3.6 - 10]	8.8 [5.1 - 12]	8.4 [4.5 - 11]	11 [6.1 - 13]	7.3 [3.3 - 12]
^Proportion male	1191 (59.3%)	416 (61.2%)	288 (59.1%)	410 (58.7%)	44 (74.6%)	33 (38.8%)
^Weight (age-adjusted z score ≥ 2)	299 (14.9%)	91 (13.4%)	70 (14.4%)	120 (17.2%)	10 (16.9%)	8 (9.41%)
^Ethnicity						
White	825 (41.1%)	290 (42.6%)	210 (43.1%)	272 (39.0%)	27 (45.8%)	26 (30.6%)
Latino	518 (25.8%)	161 (23.7%)	94 (19.3%)	222 (31.8%)	9 (15.3%)	32 (37.6%)
Black	212 (10.6%)	81 (11.9%)	34 (6.98%)	75 (10.7%)	13 (22.0%)	9 (10.6%)
Asian	131 (6.52%)	55 (8.09%)	36 (7.39%)	30 (4.30%)	4 (6.78%)	6 (7.06%)
Other or not known	323 (16.1%)	93 (13.7%)	113 (23.2%)	99 (14.2%)	6 (10.2%)	12 (14.1%)
^Significant comorbidity	108 (5.38%)	30 (4.41%)	32 (6.57%)	33 (4.73%)	4 (6.78%)	9 (10.6%)
^SARS-CoV-2 PCR positive	415 (20.8%)	131 (19.4%)	97 (20.0%)	148 (21.4%)	13 (22.0%)	26 (31.7%)
^SARS-CoV-2 Ab positive	1321 (66.5%)	412 (61.2%)	344 (71.4%)	492 (71.6%)	43 (72.9%)	30 (35.3%)
^Baseline requirement for ventilation/inotropes/ECMO	535 (26.6%)	117 (17.2%)	127 (26.1%)	252 (36.1%)	29 (49.2%)	10 (11.8%)
^Clinical features during admission						
Fever	1863 (92.7%)	653 (96.0%)	439 (90.1%)	649 (93.0%)	52 (88.1%)	70 (82.4%)
Sore throat	464 (25.5%)	159 (26.5%)	104 (22.9%)	175 (27.0%)	11 (21.6%)	15 (21.1%)
Cough	404 (21.1%)	125 (19.4%)	120 (25.3%)	131 (19.6%)	16 (30.8%)	12 (16.0%)
Respiratory distress	258 (13.3%)	70 (10.9%)	57 (11.9%)	112 (16.4%)	13 (23.6%)	6 (7.59%)
Abdominal pain	1211 (63.2%)	408 (63.9%)	289 (62.3%)	438 (64.8%)	37 (63.8%)	39 (48.1%)
Diarrhea	882 (44.8%)	290 (43.9%)	195 (40.6%)	340 (49.4%)	23 (39.7%)	34 (41.5%)
Vomiting	1057 (54.0%)	330 (50.6%)	251 (52.3%)	408 (59.2%)	34 (60.7%)	34 (42.5%)
Headache	592 (32.8%)	199 (34.1%)	155 (35.0%)	203 (31.4%)	21 (38.9%)	14 (18.4%)
Irritability	355 (18.8%)	127 (20.2%)	69 (14.9%)	135 (20.2%)	10 (18.5%)	14 (18.4%)
Lethargy	655 (34.5%)	211 (33.3%)	186 (40.1%)	215 (32.1%)	23 (41.8%)	20 (26.7%)
^Proportion meeting Kawasaki Disease criteria	629 (31.3%)	265 (39.0%)	119 (24.4%)	225 (32.2%)	12 (20.3%)	8 (9.41%)
⁺Bloods on admission						
Lymphocytes (10^9/L)	1.2 [0.70 - 2.0]	1.3 [0.76 - 2.2]	1.2 [0.70 - 1.8]	1.1 [0.66 - 1.9]	0.86 [0.52 - 1.6]	1.8 [1.1 - 2.9]
Troponin (ng/L)	25 [6.1 - 80]	13 [5.0 - 43]	31 [9.8 - 100]	40 [10 - 110]	48 [10 - 270]	10 [2.0 - 38]
CRP (mg/L)	150 [85 - 220]	150 [85 - 210]	160 [75 - 220]	160 [90 - 230]	180 [97 - 280]	85 [23 - 180]
Ferritin (ug/L)	440 [230 - 860]	370 [210 - 650]	480 [260 - 970]	520 [260 - 960]	560 [340 - 1700]	280 [140 - 460]
Albumin (g/L)	32 [28 - 37]	34 [28 - 39]	32 [27 - 36]	32 [27 - 36]	32 [27 - 36]	35 [30 - 41]

	Everyone (N=2009)	IVIG (N=680)	Glucocorticoids (N=487)	IVIG and Glucocorticoids (N=698)	Other (N=59)	No treatment (N=85)
⁺Age	8·0 [4·2 - 11]	6·8 [3·6 - 10]	8·8 [5·1 - 12]	8·4 [4·5 - 11]	11 [6·1 - 13]	7·3 [3·3 - 12]
^Proportion male	1191 (59·3%)	416 (61·2%)	288 (59·1%)	410 (58·7%)	44 (74·6%)	33 (38·8%)
^Proportion female	818 (40·7%)	264 (38·8%)	199 (40·9%)	288 (41·3%)	15 (25·4%)	52 (61·2%)

^Weight (age-adjusted z score ≥ 2)	299 (14·9%)	91 (13·4%)	70 (14·4%)	120 (17·2%)	10 (16·9%)	8 (9·41%)
^Ethnicity						
White	825 (41·1%)	290 (42·6%)	210 (43·1%)	272 (39·0%)	27 (45·8%)	26 (30·6%)
Latino	518 (25·8%)	161 (23·7%)	94 (19·3%)	222 (31·8%)	9 (15·3%)	32 (37·6%)
Black	212 (10·6%)	81 (11·9%)	34 (6·98%)	75 (10·7%)	13 (22·0%)	9 (10·6%)
Asian	131 (6·52%)	55 (8·09%)	36 (7·39%)	30 (4·30%)	4 (6·78%)	6 (7·06%)
Other or not known	323 (16·1%)	93 (13·7%)	113 (23·2%)	99 (14·2%)	6 (10·2%)	12 (14·1%)
^Significant comorbidity	108 (5·38%)	30 (4·41%)	32 (6·57%)	33 (4·73%)	4 (6·78%)	9 (10·6%)
^SARS-CoV-2 PCR positive	415 (20·8%)	131 (19·4%)	97 (20·0%)	148 (21·4%)	13 (22·0%)	26 (31·7%)
^SARS-CoV-2 Ab positive	1321 (66·5%)	412 (61·2%)	344 (71·4%)	492 (71·6%)	43 (72·9%)	30 (35·3%)
^Baseline requirement for ventilation/inotropes/ECMO	535 (26·6%)	117 (17·2%)	127 (26·1%)	252 (36·1%)	29 (49·2%)	10 (11·8%)
^Clinical features during admission						
Fever	1863 (92·7%)	653 (96·0%)	439 (90·1%)	649 (93·0%)	52 (88·1%)	70 (82·4%)
Sore throat	464 (25·5%)	159 (26·5%)	104 (22·9%)	175 (27·0%)	11 (21·6%)	15 (21·1%)
Cough	404 (21·1%)	125 (19·4%)	120 (25·3%)	131 (19·6%)	16 (30·8%)	12 (16·0%)
Respiratory distress	258 (13·3%)	70 (10·9%)	57 (11·9%)	112 (16·4%)	13 (23·6%)	6 (7·59%)
Abdominal pain	1211 (63·2%)	408 (63·9%)	289 (62·3%)	438 (64·8%)	37 (63·8%)	39 (48·1%)
Diarrhoea	882 (44·8%)	290 (43·9%)	195 (40·6%)	340 (49·4%)	23 (39·7%)	34 (41·5%)
Vomiting	1057 (54·0%)	330 (50·6%)	251 (52·3%)	408 (59·2%)	34 (60·7%)	34 (42·5%)
Headache	592 (32·8%)	199 (34·1%)	155 (35·0%)	203 (31·4%)	21 (38·9%)	14 (18·4%)
Irritability	355 (18·8%)	127 (20·2%)	69 (14·9%)	135 (20·2%)	10 (18·5%)	14 (18·4%)
Lethargy	655 (34·5%)	211 (33·3%)	186 (40·1%)	215 (32·1%)	23 (41·8%)	20 (26·7%)
^Proportion meeting Kawasaki Disease criteria	629 (31·3%)	265 (39·0%)	119 (24·4%)	225 (32·2%)	12 (20·3%)	8 (9·41%)
+Bloods on admission						
Lymphocytes (10^9/L)	1·2 [0·70 - 2·0]	1·3 [0·76 - 2·2]	1·2 [0·70 - 1·8]	1·1 [0·66 - 1·9]	0·86 [0·52 - 1·6]	1·8 [1·1 - 2·9]
Troponin (ng/L)	25 [6·1 - 80]	13 [5·0 - 43]	31 [9·8 - 100]	40 [10 - 110]	48 [10 - 270]	10 [2·0 - 38]
CRP (mg/L)	150 [85 - 220]	150 [85 - 210]	160 [75 - 220]	160 [90 - 230]	180 [97 - 280]	85 [23 - 180]
Ferritin (ug/L)	440 [230 - 860]	370 [210 - 650]	480 [260 - 970]	520 [260 - 960]	560 [340 - 1700]	280 [140 - 460]
Albumin (g/L)	32 [28 - 37]	34 [28 - 39]	32 [27 - 36]	32 [27 - 36]	32 [27 - 36]	35 [30 - 41]

REFERENCES

1 Riphagen S, Gomez X, Gonzalez-Martinez C, Wilkinson N, Theocharis P. Hyperinflammatory shock in children during COVID-19 pandemic. *Lancet.* 2020;**395:**1607-8.

2 Viner RM, Whittaker E. Kawasaki-like disease: emerging complication during the COVID-19 pandemic Comment. *Lancet.* 2020;**395:**1741-3.

3 Whittaker E, Bamford A, Kenny J, et al. Clinical Characteristics of 58 Children with a Pediatric Inflammatory Multisystem Syndrome Temporally Associated with SARS-CoV-2. *JAMA.* 2020;**324:**259–69.

4 Verdoni L, Mazza A, Gervasoni A, et al. An outbreak of severe Kawasaki-like disease at the Italian epicentre of the SARS-CoV-2 epidemic: an observational cohort study. *Lancet.* 2020;**395:**1771–8.

~~55 Feldstein LR, Rose EB, Horwitz SM, et al. Multisystem Inflammatory Syndrome in U.S. Children and Adolescents. *N Engl J Med* 2020;**383:**334–46.~~

~~6~~ Feldstein LR, Tenforde MW, Friedman KG, et al. Characteristics and Outcomes of US Children and Adolescents with Multisystem Inflammatory Syndrome in Children (MIS-C) Compared with Severe Acute COVID-19. *JAMA.* 2021;**325:**1074–87.

6 Feldstein LR, Rose EB, Horwitz SM, et al. Multisystem Inflammatory Syndrome in U.S. Children and Adolescents. ~~7 Levin M. Childhood Multisystem Inflammatory Syndrome - A New Challenge in the Pandemic. *N Engl J Med* 2020;**383:**393–5~~334–46.

~~8~~7 Godfred-Cato S, Bryant B, Leung J, et al. COVID-19-Associated Multisystem Inflammatory Syndrome in Children - United States, March-July 2020. *MMWR Morb Mortal Wkly Rep.* 2020;**69:**1074–80.

8 Levin M. Childhood Multisystem Inflammatory Syndrome - A New Challenge in the Pandemic. *N Engl J Med.* 2020;**383:**393–5.

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

Formatted: Dutch (Netherlands)

9 McCrindle BW, Rowley AH, Newburger JW, et al. Diagnosis, Treatment, and Long-Term Management of Kawasaki Disease: A Scientific Statement for Health Professionals From the American Heart Association. *Circulation*. 2017;**135(17):**e927–99.

10 Harwood R, Allin B, Jones CE, et al. A national consensus management pathway for paediatric inflammatory multisystem syndrome temporally associated with COVID-19 (PIMS-TS): results of a national Delphi process. *Lancet Child Adolesc Health*. 2021;**5:**133-141

11 Henderson LA, Canna SW, Friedman KG, et al. American College of Rheumatology Clinical Guidance for Multisystem Inflammatory Syndrome in Children Associated With SARS-CoV-2 and Hyperinflammation in Pediatric COVID-19: version 1. *Arthritis Rheumatol*. 2020;**72:**1791–805.

12 Ouldali N, Toubiana J, Antona D, et al. Association of Intravenous Immunoglobulins Plus Methylprednisolone vs Immunoglobulins Alone With Course of Fever in Multisystem Inflammatory Syndrome in Children. *JAMA*. 2021;**325:**855–64.

13 Son MBF, Murray N, Friedman K, et al. Multisystem Inflammatory Syndrome in Children - Initial Therapy and Outcomes. *N Engl J Med*. 2021;**385:**23-34.

14 Villacis-Nunez DS, Jones K, Jabbar A, et al. Short-term Outcomes of Corticosteroid Monotherapy in Multisystem Inflammatory Syndrome in Children. *JAMA Pediatr*. 2022;**176:**576-84.

15 McArdle AJ, Vito O, Patel H, et al. Treatment of Multisystem Inflammatory Syndrome in Children. *N Engl J Med*. 2021;**385:**11-22.

16 Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;**42:**377–81.

17 Toscano G, Palmerini F, Ravaglia S, et al. Guillain–Barré Syndrome Associated with SARS-CoV-2. *N Engl J Med*. 2020;**382:**2574–6.

18 LaRovere KL, Riggs BJ, Poussaint TY, et al. Neurologic Involvement in Children and Adolescents Hospitalized in the United States for COVID-19 or Multisystem Inflammatory Syndrome. *JAMA Neurol*. 2021;**78:** 536-47.

19 Tullie L, Ford K, Bisharat M, et al. Gastrointestinal features in children with COVID-19: an observation of varied presentation in eight children. *Lancet Child Adolesc Health*. 2020;**4(7):** e19-e20.

~~20~~20 Centers for Disease Control and Prevention (CDC). Information for healthcare providers about multisystem inflammatory syndrome in children (MIS-C) https://www.cdc.gov/mis-c/hcp/. Date accessed: July 28, 2022

21 Royal College of Paediatrics and Child Health~~.~~ (RCPCH). Paediatric multisystem inflammatory syndrome temporally associated with COVID-19 (PIMS) - guidance for clinicians. 2020 ~~(~~. https://www.rcpch.ac.uk/resources/paediatric-multisystem-inflammatory-syndrome-temporally-associated-covid-19-pims-guidance~~)~~. Date accessed: December 10, 2022

~~21 Centers for Disease Control and Prevention. Information for healthcare providers about multisystem inflammatory syndrome in children (MIS-C) (https://www.cdc.gov/mis-c/hcp/).~~

22 World Health Organization. Multisystem inflammatory syndrome in children and adolescents temporally related to COVID-19. Scientific brief. May 2020 ~~(~~. https://www.who.int/publications/i/item/multisystem-inflammatory-syndrome-in-children-and-adolescents-with-covid-19~~)~~. Date accessed: December 10, 2022

23 National Institute for Health and Care Excellence~~.~~ (NICE). Equivalent anti-inflammatory doses of oral corticosteroids. ~~(~~June 2020. https://cks.nice.org.uk/topics/corticosteroids-oral/background-information/equivalent-anti-inflammatory-doses~~)~~./. Date accessed: December 10, 2022

24 Lopez L, Colan SD, Stylianou Mario, et al. Relationship of Echocardiographic Z-Scores Adjusted for Body Surface Area to Age, Sex, Race, and Ethnicity: The Pediatric Heart Network Normal Echocardiogram Database. *Circ Cardiovasc Imaging*. 2017;**10(11):** e006979

25 Imai K, Ratkovic M. Covariate balancing propensity score. *J R Stat Soc Ser B Stat Methodol*. 2014;**76:**243–63.

26 Sood M, Sharma S, Sood I, Sharma K, Kaushik A. Emerging Evidence on Multisystem Inflammatory Syndrome in Children Associated with SARS-CoV-2 Infection: a systematic Review with Meta-analysis. *SN Compr Clin Med*. 2021;**3:**38–47.

~~27~~27 Ouldali N, Bagheri H, Salvo F, et al. Hyper inflammatory syndrome following COVID-19 mRNA vaccine in children: A national post-authorization pharmacovigilance study. *Lancet Reg Health Eur.* 2022;**17**:100393.

28 Hernan MA. The hazards of hazard ratios. *Epidemiology*. 2010;**21(1):**13-5.

29 Zhu YP, Shamie I, Lee JC, et al. ~~(2021).~~ Immune response to intravenous immunoglobulin in patients with Kawasaki disease and MIS-C. *J Clin Invest*. 2021;**131(20):** e147076.

~~28~~30 Patel H, McArdle A, Seaby E, Levin M, Whittaker E. The immunopathogenesis of SARS-CoV-2 infection in children: diagnostics, treatment and prevention. *Clin Transl Immunol*. 2022;**11:** e1405.

~~29~~31 Irfan O, Muttalib F, Tang K, Jiang L, Lassi ZS, Bhutta Z. Clinical characteristics, treatment and outcomes of paediatric COVID-19: A systematic review and meta-analysis. *Arch Dis Child*. 2021;**106:**440-8.

~~30~~32 Johnson SC, Williams DC, Brinton D, Chew M, Simpson A, Andrews AL. A Cost Comparison of Infliximab Versus Intravenous Immunoglobulin for Refractory Kawasaki Disease Treatment. *Hosp Pediatr*. 2021;**11:** 88-93.

DATA SHARING STATEMENT

Question	Response
Will individual participant data be available (including data dictionaries)?	Yes
Rationale for data sharing statement	BATS has collected de-identified data from multiple institutions in many countries. Each institution has signed an agreement with Imperial College on data security. We will need to assess requests for data on a case-by-case basis to ensure that the data that are provided fall within the existing agreements within the consortium.
What data in particular will be shared?	De-identified clinical and laboratory findings and response to treatment for the cohort included in this study. Any data provided will be de-identified and will conform to the agreements within the consortium for data sharing.
What other documents will be available?	The study handbook and statistical analysis plans are available at the ISRCTN registry at the following link: https://doi.org/10.1186/ISRCTN69546370
When will data availability start?	On publication of the manuscript. However, as approval for all data will have to be obtained from the consortium and partner institutions, approximately 3 months may be required before the data is provided.
When will data availability end?	Two years after publication
To whom will data be available?	Legitimate researchers and clinicians from medical and academic institutions.
For what types of analyses?	Only for academic and clinical research.
By what mechanism will data be made available?	On request to the corresponding author. Data requestors will need to sign a data access agreement

Formatted: Line spacing: Multiple 1.15 li

Editors' specific comments to the Author:

1. Please indicate whether any of the authors are full professors and also carefully check the spelling of all names and accuracy of affiliations

All authors have checked the spelling of their names and their affiliations. The following authors are full professors: ML and AC.

2. Tables should be supplied in a separate Word file (not Excel). Each row of data should be in a separate line. Please ensure that rows and columns are not tabbed; data should be entered in cell form.

This will be supplied with our revisions. Apologies for this error in our previous submission.

3. Please include a row for female sex in table 1.

Table 1 has been amended as requested (page 30)

4. Please ensure that all p values are provided to two significant figures, unless p<0.0001 (note number of decimal places).

We have amended p-values to the requested format.

5. If available, please include number censored (as well as number at risk) in each group for each time point on the Kaplan-Meier curves.

We have included these in the KM curves for each time point for figure 3 (page 28)

6. Please ensure the reporting adheres to STROBE reporting guidelines and upload a completed checklist with your revision.

We have cross-referenced our work against the STROBE reporting guidelines, and have added additional details to both the main text and supplement where details were missing. Our report now adheres to these guidelines, and a completed checklist will be uploaded with our revision.

7. Figures: Please supply fully editable files for all figures (eg, EPS files, AI files, PDF files, depending on software used to produce them). Please see our artwork guidelines for full details of acceptable file formats: https://www.thelancet.com/pb/assets/raw/Lancet/authors/artwork-guidelines-1641398207410.pdf

These have all been generated in PDF form and will be supplied with the submission of our revisions.

8. Authors contributions. Please update your authors contributions statement to confirm the following:

- all authors had full access to all the data in the study;
- all authors had final responsibility for the decision to submit for publication;
- and at least two identified authors directly accessed and verified the underlying data reported in the manuscript.

These changes have been made as requested (pages 5-6).

9. All authors are required to provide a signed author contribution statement form, available to be downloaded from https://els-jbs-prod-

cdn.literatumonline.com/pb/assets/raw/Lancet/authors/tlrheum-author-signatures-1555082865647.pdf (completed forms should be uploaded with your revised manuscript).

We have completed all these forms as requested and will upload them with the rest of our revisions.

10. All authors must complete and return an ICMJE conflict of interest form, available from http://www.icmje.org/conflicts-of-interest/ The declarations on the forms should exactly match those stated in the declarations section of the paper.

We have completed all these forms as requested and will upload them with the rest of our revisions. A Declaration of Interests statement has been added to the main text as requested.

11. Please supply the appendix as a single pdf file, with numbered pages. Please do not include a cover page with details of the paper, as we add a cover page to the appendix file before publication with this information. The protocol can be included in your appendix if you wish. When citing the appendix in the text, please do not refer to specific table/figure numbers. Rather simply state the page number. Eg, "(appendix p XX)".

We will supply the appendix in the required format, and have amended references to the appendix in our main manuscript as requested.

12. As your author line includes a study group (eg, 'on behalf of the XXXX trial study group'), collaborators' names and affiliations may be listed in the appendix. Additionally, if you wish the names of collaborators within a study group to appear on PubMed, please upload with your revision *a separate Word document* with a list of names of the study group members presented as a two-column table. First and middle names or initials should be placed in the first column, and surnames in the second column. Names should be ordered as you wish them to appear on PubMed. The table will not be included in the paper itself – it's simply used to make sure that PubMed adds the names correctly.

We have supplied a table with our revisions, titled: "BATS consortium table"

Reviewers' Comments:

Note that reviewer numbers are allocated by the system at invitation and not at completion of reviews, so some numbers might be missing.

- In your point-by-point reply to the reviewers', please indicate the text changes which have been made (if any) and the line number on the tracked changes manuscript at which your change can be found. [Line numbers can be added to your word document using the 'page layout' tab. Please select continuous numbers.]
- Please do not use the ''Comments' function in your word document.
- When interpreting editorial points made by reviewers, please remember that we will edit the final manuscript if accepted.

Reviewer #1:

The authors present data from BATS - a retrospective cohort from across the world of MIS-C treatment. The manuscript addresses an important issue - although thankfully MIS-C is less of an issue now. Overall the manuscript is well written.

This is essentially a large report of the previous publication NEJM in 2021 by the same group. They do acknowledge this and state this is a much larger cohort which is true.
The retrospective nature of this and self reporting does present some problems as opposed to prospective and more robustly collected data.

I would like to see a split of the countries which contributed the data. Is use of GC alone a feature of inability to afford IVIG (which is very expensive). In several Asian, African and other LMIC settings, patients/parents pay out of pocket for medications. Whilst this does not undermine the data, important to acknowledge this.

We thank the reviewer for raising this point. We have provided a detailed overview of the number of sites and patients recruited from each country in supplementary figures 1 and 2A-C (supplement pages 48-50). We have discussed the limited availability of IVIG due to cost in our discussion (page 19), and how this relates to use of GC. In addition, unadjusted distributions of treatment choices by World Bank Resource Group can be found in supplementary figure S11C (supplement page 60), which shows that use of GC is proportionally more common in LM/L-ICs, supporting the hypothesis that inability to afford IVIG is a driver for GC use. As we have adjusted for Resource Group in our IPTW analysis, we feel further discussion of this point is not required. Please also see response to comment below, which is relevant to this point.

It is likely the IVIG + GC is more likely in Western Europe and North America. if anything, this is more likely to show the strength of giving GC alone which is more cost-effective than IVIG or IVIG +GC.

We thank the reviewer for this interesting comment. We have supplied an additional supplementary figure (SF17, supplement page 67) which demonstrates the proportion of patients from each treatment group in different geographical regions. This shows a relatively even spread of treatments across regions. The exceptions are North America, where use of any regime without IVIG was uncommon, as well as Africa and Asia – both regions contributing very few patients to BATS. We hope the addition of this supplementary figure provides adequate explanation of these points.

Figure 1B whilst make look pretty hardly adds anything of value and does not merit presenting in my personal opinion.

We thank the reviewer for their opinion on this graphic. We believe this figure succinctly demonstrates the movement of patients from specific treatment groups during their treatment course, and merits inclusion on this basis. In particular, figure 1B is important for understanding the escalation of patients on monotherapy into combined therapy, as highlighted by a separate reviewer (reviewer 5, comment number 3). If there is room for this figure to be retained in the paper this would be our preference.

I would also like to know if any side effects of IVIG or GC were asked for or reported.

We requested any potential drug complications for all patients, asking sites to report the suspected drug/s, and side effects using both free text and a pre-defined checklist. Complications are reported in detail in supplementary table 8 (supplement page 40), and a brief overview is provided in the main results section (page 16), with reference to this table for further information to the reader.

Please can the dose of IVIG be clarified and documented.

We thank the reviewer for their comment. IVIG doses were reported in g/kg. The median dose of IVIG was 2.0 g/kg, with IQR of (2.0 – 2.0). We have supplied an additional supplement table with this information (table S9B, supplement page 41)

GC - was it IVMP or oral steroids - if so for how long. It is likely in this format of retrospective data collection multiple dosing regimens are likely to have been used - need clarity on mean dose of steroids and differing dosing regimens.

We agree that this is an important limitation of the study. GC dosing details are provided in supplementary table 9A (supplement page 41). We included both oral and IV GC, excluding low-dose oral GC, or low-dose IV hydrocortisone (as it is commonly used as an adjunct to inotropic therapy in sick children). These criteria are described in detail in the supplementary methods (supplement page 13, under “treatment definitions”). The lengths of GC courses are difficult to ascertain in our cohort, as data was only collected during hospital stay, and many patients were given weaning GC on discharge. We therefore do not feel it is meaningful to report the hospital length of GC courses as this ignores this crucial follow up period. We have added a statement to the discussion to highlight this limitation (page 22, first paragraph – “Other limitations include the wide variety of steroid dosing regimens used,…”).

Reviewer #2:

This study assesses immunoglobulin, glucocorticoid, or combination therapy for Multisystem Inflammatory Syndrome in children using a cohort study. There are several major and minor concerns about the statistics and methodology of the paper.

Major
Methods
1) Page 11, last paragraph, lines 1-3: Please clarify how you identified confounders. A minimally sufficient set of confounders should be selected using a causal directed acyclic graph.

We thank the reviewer for raising this issue, and we apologise for the lack of clarity in the main methods section. The list of confounders selected was reached by consensus within the study team and international advisory board, with substantial expertise in Paediatric Infectious Disease, Clinical Trials, and Medical Statistics. Whilst we did not use a formal causal DAG, we considered an extensive list of confounders (including all those defined in original protocol and SAP 1 from the previous NEJM paper, with additional confounders) and selected the final list those through consensus of our expert study team. These were reported in our pre-analysis SAP (https://www.isrctn.com/ISRCTN69546370), and were selected based on the criteria of variables relating to treatment assignment and/or the disease outcomes, but excluding those that are believed to only relate to the treatment assignment, as these are known to inflate the variance of treatment effect estimations (see Brookhart et al. 2006, Am J Epidemiol. *Variable selection for propensity score models*). Full details are described in the supplement on pages 18-19, and an additional statement briefly explaining this is included in the main methods section (page 11, final paragraph – "Confounding covariates were selected by expert consensus prior to analysis…").

2) Page 11, last paragraph, lines 1-3: Please clarify if you have adjusted for confounders by including them in the outcome model as well.

We did adjust for the confounders in the outcome model also. This is an important point for the analysis, and thank you for highlighting this omission from the main methods. We have added a clarifying statement from the supplementary methods into the main methods section (page 11, last paragraph) as follows: *"Confounding covariates… were used in both covariate balancing and treatment effect estimation to produce doubly-robust estimates".* Here we have referred to the outcome model as "treatment effect estimation" as we feel this is more intuitive to the lay audience (Leite W. Practical propensity score methods using R. 1st ed. SAGE Publications Lt; 2017. Chapter 1; page 7).

3) Page 11, last paragraph, line 4: Please clarify the generalized linear model you used: binary logistic regression?

We used weighted quasibinomial logistic regression. We have added a corresponding statement to the main methods section (page 12) from the supplement.

4) Page 11, last paragraph, lines 4-5: The linearity assumption underlying generalized linear and Cox regression models for quantitative predictors should be assessed.

We thank the reviewer for this comment and are happy to provide these details. The linearity assumption for both the generalised linear models and Cox-regression models was assessed for quantitative predictors (age and baseline CRP) by visualisation each predictor against the linear predictors for each model (for example, the logit of the outcome for logistic regression). These demonstrated no non-linear relationships. We have therefore not applied transformations to these predictors. In fact, for all primary analyses there was no relationship with age, reflecting the fact that, after covariate balancing, age was not statistically significant in either the glm or Cox-regression models. However, the inclusion of age in the final multivariable treat effect estimation models is justified based on clinical grounds due to its clinical importance in the outcomes from MIS-C. An accompanying statement has been added to the supplementary methods (supplement page 20, first section).

5) Page 11, last paragraph, lines 4-5: Please clarify the origin, start, and end times for survival analysis.

Thank you for highlighting this ambiguity in our explanation. The start time is the first day the patient receives treatment. The origin (time when patients are first deemed to be at risk) is the first day after treatment initiation. The end time is defined as either the time to improvement (as outlined in the supplement on page 16) or alternatively the patient is right censored (explanations of censoring also found in supplement page 16). This has been clarified in the supplement on page 16.

6) Page 11, last paragraph, line 5: Risk ratios (RRs) are generally preferred to odds ratios (ORs) due to ease of interpretation, collapsibility, and less susceptibility to sparse-data bias, and so should be estimated in cohort studies. I suggest that the authors estimate adjusted RRs with 95% confidence intervals (CIs) using modified Poisson regression model e.g., see the following paper:
Zou G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. Am J Epidemiol. 2004;159(7):702-706.

We thank the reviewer for this useful discussion point. We have two reasons why we prefer the reporting of odds-ratios in this instance. Most importantly, to enable the reader to compare our previously published results with this new paper we feel strongly that we should report the same effect measure as used in the NEJM BATS paper (McArdle et al. NEJM. 2021). Although we acknowledge and agree completely that in a traditional prospective cohort study the risk ratio is preferable, the BATS study does not neatly fit within this category, for two main reasons: 1) Data was not necessarily collected prospectively for all patient episodes, as we allowed retrospective case reporting, and 2) not all patients with MIS-C at each site were entered onto the BATS database. For these reasons, we feel reporting the odds ratio is actually preferred. It is also well accepted and understood by clinicians and readers of medical literature, and as stated above will ensure consistency with our previous report.

7) Supplementary appendix - BATS second analysis, page 13, last two paragraphs: Naïve methods of handling missing data such as interpolation are subject to selection bias and should be avoided. Standard methods such as inverse probability weighting and multiple imputation should be used under missing at random assumption.

We understand this valid concern from the reviewer, which we paid substantial attention to during our analysis plan. However, our imputation/interpolation methods (described in the supplement

pages 13-14) are motivated by clinical experience, and we strongly believe they are preferred in this context. This is best demonstrated by an example. Suppose a patient was reported to have been receiving ventilatory support on day 1, with missing respiratory data on day 2, and then reported to be receiving ventilatory support on day 3, before having this support removed by the end of day 3. Clinical experience, both with MIS-C patients and general patients on ICU, tells us that by far the most likely value for this missing variable is for the patient to be receiving ventilatory support on day 2. In addition, experience from our previous analysis informs us that complete datasets within BATS demonstrate a very low (for example, for those receiving ventilatory support, less than 2% of complete records have discontinuous periods of ventilatory support).

Furthermore, we believe this data is most likely missing-not-at-random, rather than missing-at-random. To see why, it is important to consider the very time-consuming data entry task required for a single BATS patient, with significant data entry required for each day of a patient admission. Given this, if we look again at respiratory support as an example, we believe it is much less likely that missing values will be present when there is a change in ventilatory support, or supplemental oxygen. These clinical changes are important markers of a patient's clinical status, and are likely to be better recorded in clinical records and more accurately by those entering data. Because of this, we feel the proposed statistical methods for imputation are less appropriate. We are therefore confident that our imputation/interpolation methods, which are very similar to those published in the previous NEJM BATS paper, are sensible, pragmatic, and less biased than purely statistical methods.

8) Supplementary appendix - BATS second analysis, page 13, last two paragraphs: Related to the previous comment, please report the range for proportions of missing data.

We are happy to provide these data, and have added an additional table to the supplement (table S15, supplement page 47). This table shows the total numbers and percentages of missing values for each variable where imputation/interpolation were carried out for the final analysis, both before and after these approaches were applied, and separated by total number of missing days and patients with any missing data.

9) Supplementary appendix - BATS second analysis, page 18, fifth paragraph: The correct specification of the exposure model should be assessed.

We thank the reviewer for their comment on this important matter. The model specification has been assessed in multiple different ways during our analysis. Firstly, the covariates selected for the model have been selected after lengthy consideration. Importantly, these were all assessed by multiple experts within our study team and international advisory board to ensure these represented the most clinically important confounders, whilst assuring we only included covariates believed to be relevant to the outcome (see response to major comment #1 from same reviewer). These covariates were used for CBPS estimation, and the covariate balance was rigorously assessed at this stage. This is arguably the most important diagnostic for model specification in CBPS methods (Wyss et al. 2014. Am J Epidemiol. "The Role of Prediction Modeling in Propensity Score Estimation: An Evaluation of Logistic Regression, bCART, and the Covariate-Balancing Propensity Score").

The same covariate list was taken forward to treatment effect estimation to produce doubly robust estimates (see response to major comment #2 from same reviewer). We assessed the model specification at this stage through multiple standard approaches, including assessment of collinearity, as well as core model metrics (see response to major comment #4 from same reviewer). Whilst it may have been statistically possible to marginally improve model specification by reviewing covariates that could be removed at this stage, this is a relatively arbitrary process. We took a different approach in including all covariates for this second stage. We believe this is vitally important from a clinical perspective, to include all variables believed to be of significance to the outcome. This also ensures we produce doubly robust estimates, allowing for a second-chance to capture relationships between the covariates and outcomes.

We are confident this approach is supported by the literature on this topic. We note that of primary importance in PS analysis is "using study design and subject matter expertise to gain an understanding of the underlying causal structure" (Wyss et al. 2014. Am J Epidemiology). Indeed, there does not appear to be a consensus in the literature regarding the most appropriate evaluation of model specification for treatment effect estimation following PS calculation, but instead more recent focus in PS methods has been motivated by the goal of minimizing covariate imbalance, which is a key metric of PS model performance. In our analysis we have placed significant focus on initial covariate selection, through expert consensus. We note the covariate list is similar to other studies on this topic using PS methods (Son et al. 2021 NEJM; Ouldali et al. 2021 JAMA). We also note results from the same study above (Wyss et al.) which demonstrate that the CBPS methodology is relatively robust to minor model mis-specification, both in terms of covariate balance and estimation bias.

We therefore conclude that our approach, motivated by inclusion of the most clinically relevant variables, could potentially include some minor model mis-specification. However, our model assessments demonstrate highly effective covariate balancing between groups, and subsequent models used for treatment effect estimation pass all reasonable diagnostics. Any model mis-specification is therefore likely to be minor, and crucially from a clinical perspective we have ensured we include what we believe to be all relevant confounders.

10) Supplementary appendix - BATS second analysis, page 18, last paragraph, last two lines: It is not clear if you used IPTW or doubly-robust (DR) estimators. The latter requires both exposure and outcome modeling.

We thank the reviewer for highlighting this ambiguity. We used doubly robust estimators. We have altered the text in the supplement to clarify this point (supplement pages 18 final paragraph, and page 20, title of first subsection).

11) Supplementary appendix - BATS second analysis, page 18, second paragraph, lines 3-5: Please clarify the spline function you used in the generalized additive model.

We used the geom_smooth function from the ggplot2 R package with the "gam" method, as stated with references in the supplement (supplement page 18, second paragraph). This method automatically estimates the degree of smoothing using penalized regression splines with smoothing parameters selected by a restricted maximum likelihood approach. We are using this package to aid

smoothed visualisations of the data, and we feel these details are not necessary for the general reader, and the interested reader can find these details in the relevant package documentation. We have clarified the exact implementation of the geom_smooth function in the supplement to aid such interested readers.

12) Supplementary appendix - BATS second analysis, page 20, first paragraph: Please clarify the target for which you estimated the treatment effect using propensity score matching.

We thank the reviewer for this comment and apologise for this accidental omission. The estimand used for the matched analysis was the Average Treatment Effect in the Treated (ATT). This decision was necessitated by limitations in the software used for this analysis. A sensitivity analysis using the other available estimand, Average Treatment Effect in the Controls (ATC) demonstrated no significant difference in the findings. We attach a table to our submission to demonstrate these results. We have added an explanatory statement to the relevant section of the supplement (supplement page 20) to clarify the methods used and explain reasons for this. Given the similarity in results with both the ATT and ATC targets we are confident this slight discrepancy in methods does not invalidate our findings.

13) Supplementary appendix - BATS second analysis, page 20, first paragraph, line 6: Please clarify the type of regression analysis you used.

We thank the reviewer for spotting this accidental omission of detail and have clarified the type of regression analysis as suggested (supplement page 20).

14) Supplementary appendix - BATS second analysis, page 20, second paragraph, lines 2-3 from bottom: Proportional hazards assumption is still required as you have included the confounders in the Cox model.

We thank the reviewer for this comment and are happy to describe our analyses in more detail. For the cox-regression models we assessed the proportional hazards assumption by 1) Schoenfeld tests for non-proportionality both for individual covariates and a global Schoenfeld test; and 2) assessing plots of the scaled Schoenfeld residuals versus time (Grambsch PM and Therneau TM, 1994, Biometrika. *Proportional hazards tests and diagnostics based on weighted residuals).* For the main time-to-event analyses, the Schoenfeld tests demonstrated a possible violation of proportional hazards assumption in two covariates (baseline inotropic status and meeting KD criteria, p = 0.0174 and 0.0068 respectively) and for the global Schoenfeld test (p = 0.025). However, assessment of the residual plots demonstrated no significant time dependence for any of the coefficients. We believe the statistically significant non-proportionality is of no practical significance to the model assumptions, given the flat profiles of the Schoenfeld residual plots. This is likely a statistical artifact from the large sample size – an effect that has been previously noted (Therneau and Grambsch, Modeling Survival Data: Extending the Cox Model, Published London: Springer, 2000 – Section 6.5). We have therefore not incorporated any time-dependent interaction terms. We have added an explanatory statement to the supplement (supplement page 20) in line with the reviewer's suggestion.

Note: The implementation of the Schoenfeld tests for non-proportionality were implemented through the cox.zph function of the R-package "survival", written by Therneau et al. This package was accidentally omitted from our initial supplementary reference list, and this has now been corrected.

Results

15) The Results section suffers from overreliance on significance testing or point estimation which should be avoided, and results should be interpreted in light of appropriate association measures such as RR and hazard ratio (HR) estimates with 95% CIs along with P-values e.g., see the following paper:

Greenland S, Mansournia MA, Joffe M. To curb research misreporting, replace significance and confidence by compatibility: A Preventive Medicine golden jubilee article. Preventive Medicine. 2022 Jul 3:107127.

We thank the reviewer for raising this issue. We agree that we may have been overly negative about the possible benefit of combined therapy with IVIG and GC based on the results of significance testing. We note that the comparison of GC alone with Combined IVIG+GC was a secondary analysis, and therefore not adjusted for multiple hypothesis testing. The confidence intervals reported are thus likely to be an overly precise. However, we agree that there is a suggestion that combined therapy with IVIG+GC does result in more rapid recovery (albeit in the less sick patients) more rapid fever resolution and lower rate of escalation to additional therapy. We have therefore modified the discussion to reflect this potential benefit.

16) Please report median (IQR) of the follow-up time.

Although we agree that the follow-up time is typically important for time-to-event analyses, we feel this is less important in our context. The follow-up time was the time in days from initiation of treatment to either discharge or death. Due to variation in practices and healthcare resources across the international community recruiting to BATS we feel this measure is not a clinically useful measure for the range of severity seen in MIS-C. This was part of the motivation behind the ordinal scale used, which attempts to capture more clinically meaningful outcomes for each patient. The outcomes described in the ordinal scale are typically captured during inpatient hospital care, except for the censored patients. If the reviewer still feels this is important then we are happy to provide the median and IQR for follow-up time for both the time-to-event analyses and the coronary artery aneurysm data.

17) Please report censoring proportion, and reasons for censoring.

We thank the reviewer for this comment and are happy to provide these. The numbers censored at each timepoint have been included in the main figure 3 (page28). The total numbers censored can be found in supplementary table S5B for the primary time-to-event analyses (supplement page 29), and in supplementary table S5E for additional analyses (supplement pages 34-35). Reasons for censoring the 2nd primary outcome for all patients included in the main analysis are reported in a new supplementary table 14 (supplementary page 46).

18) Table 1 and Table S3, etc: Please report mean (SD) for description of quantitative variables with reasonably symmetrical distribution.

We thank the reviewer for this careful observation. However, we feel this change is not necessary as it will significantly reduce the consistency for readers between this paper and the previous report of BATS results in the NEJM in the majority of cases. For tables reporting CAA z-scores or drug doses, these are typically not symmetrically distributed, so although these tables are new to this analysis the appropriate metric to report is the median and interquartile range.

19) Tables S10-11, etc: Please report 95% CI for the beta coefficient instead of standard error.

We have replaced the beta coefficient with SE as suggested and removed the z-scores to improve readability (Tables S10-13, supplement pages 42-45).

20) Tables S10-11, etc: Please omit redundant information such as Z statistics values, log-likelihood, etc.

The suggested change has been made (supplement pages 42-45)

21) Figures S4-5: Please clarify the definition of outliers in your Box-and-whisker plot.

We have added this detail to the figure captions (supplement pages 52-54). The outliers are defined as any point more than 1.5*IQR way from the upper and lower quartiles.

22) Figures S4-5: Please do not dichotomize P-values at 1% or 5% levels. Please report corrected P-values rather than their range; any P-value less than 0.001 should be reported as"<0.001".

We have adjusted the figures as suggested (supplement pages 52-54). Please note, this may reduce readability as there are multiple comparisons on each individual panel.

Discussion

23) An important limitation of the study is built-in selection bias in HRs which should be acknowledged e.g., see the following paper:
Hernán MA. The hazards of hazard ratios. Epidemiology 2010; 21:13-15.

We thank the reviewer for their thought-provoking comment and reference. In our context, the potential for this built-in bias (in relationship to treatment assignment) could occur when studying the time-to-improvement outcome if there is a specific subpopulation of patients who respond very differently to one of the studied treatments. To incur this selection bias this subgroup of patients would have to improve more rapidly than others in the same treatment arm. From the clinical experience of our expert study group, the known factors that facilitate this bias have largely been corrected for in our analysis (age, baseline level of inflammation, clinical criteria for Kawasaki disease, etc), and in particular, we have conducted extensive subgroup analyses in an attempt to isolate such effects and remove their influence from our analysis (figure 2, page 27). This analysis did not provide any results to suggest these factors have played an important part in influencing the time to improvement in our main analysis.

This potential for this in-built selection bias still remains, considering unknown factors that could influence treatment response. Studying these factors in such a non-randomised observational study is not possible, and we have commented on this limitation already in our discussion (page 20). We note that the proposed mechanism for selection bias here is not unique to our study, and is an inherent issue with survival analysis, which cannot be adjusted for even in an RCT setting unless all confounders are adjusted for. However, the presence of unknown confounders remains in any such

study, and hence this remains a limitation of all similar analyses. We note that in our study we chose to investigate two distinct primary outcomes, both adjusted across multiple comparisons, as well as a range of secondary outcomes. These extra analyses provide us with multiple avenues to identify treatment effect, and help increase the robustness of our analysis. Since our final conclusions are not wholly reliant on the presented time-to-event analysis we believe this reduces the potential impact the proposed built-in bias could have.

In summary, we believe that our analysis has attempted to address known factors that could generate such built-in bias through multiple subgroup analyses. The possibility of built-in selection bias due to unknown susceptibility factors remains for the time-to-event analyses, but this is a general limitation of the methodology, not inherent to our study. As such, we have added a comment to this effect to our discussion highlighting this as a methodological limitation (page 22 – "The time-to-improvement outcome also incurs the possibility of "built-in selection bias"28 although we have attempted to isolate known factors that could incur such bias through extensive subgroup analyses. This limitation is relevant to all survival analysis, and would not be avoidable even for RCTs using the same outcome.").

Minor

24) Table 1, footnote, line 2 from bottom: Please change "raw values" to "number" or "frequency".

The suggested change has been made (page 30)

25) Table 1 and Table S3: The term "Proportion" before "male" and "meeting Kawasaki Disease criteria" is redundant and should be omitted.

We thank the reviewer for this suggestion. However, we feel this language is less ambiguous to the lay reader, and so have not made the suggested change.

26) Tables S10-11, etc: Please avoid spurious precision in the presentation of numbers e.g., report beta estimates with two decimals.

Many thanks for raising this point. We have reviewed these and other numeric results and addressed this issue.

27) Tables S10-11, Figure S10 etc: P-values of "0.000000" and "0.000" do not make sense. Any P-value less than 0.001 should be reported as"<0.001".

We thank the reviewer for spotting this inconsistency in the supplementary material. We have amended p-values throughout to the requested format.

Reviewer #3:

Thank you for the opportunity to review this interesting manuscript.
The reported work is unique due to the large collection of data of 2009 children with MIS-C from different areas around the globe conferring it greater generalisability than any previous study.
A further strength consists in the inclusion of coronary artery findings. The findings indicate that treatment with glucocorticoids is at least not inferior to IVIG or IVIG plus glucocorticoids, which is relevant given the limited availability (or lack of availability) of IVIG in different healthcare settings.

Major comments
1. this study of over 2000 patients includes the 614 contained in the original report (McArdle et al NEJM 2021). Were analyses performed excluding the original 614? For example, it could be that multiple treatments were less common in more recent times as clinicians became more confident in handling the disease, and patients presenting earlier with lower severity. Did the proportion of treatment escalation/additional treatment decrease?

We thank the reviewer for raising the important issue of how MIS-C has changed over time and with successive waves of SARS-CoV-2. We included the initial 614 patients in this analysis, as our analysis plan was to undertake early analysis (which we appreciated would lack power) in view of the urgent clinical need for data on management of a new disease. This approach was supported by WHO as initial guidance on management from WHO was based on the preliminary report. Our published analysis plan was to undertake further analyses as the numbers of cases in the study increased .

We agree with the reviewer that MIS-C has changed over time, both as clinicians became more familiar with the disorder, and probably as population levels of immunity have changed due to previous infection and vaccination. We are in the process of preparing a report on how clinical features, severity and treatment has change over time. We plan to submit this report as a separate letter or short report, following publication of the current manuscript. We do not feel it would be possible to adequately present the data on changes in MIS-C over time within the current manuscript which is already long and complex. In light of this comment, we have also added a paragraph in the discussion on the changing spectrum and epidemiology of MIS-C.

2. The data collection for BATS was, as it appears, highly pragmatic and voluntary. Risks of bias with data collection (despite the large sample size) are not well described. Do the authors have data on how many patients with MIS-C per center were NOT included in BATS? Please add such information or state the lack of it as a limitation.

We thank the reviewer for this comment. Unfortunately, we do not have this information and agree this is a potential source of bias, although most likely towards the inclusion of more unwell patients, who are the most relevant population to study for immunomodulator treatment decisions. We have added a sentence in the discussion to highlight this limitation (page 20, final paragraph)

3. case definition: as per the methods page 10 it seems that "children with severe inflammatory illness after SARS-CoV-2 infection" were included even if they did not meet CDC RPCH or WHO criteria for MIS-C? Were sensitivity analyses conducted restricted to children meeting criteria for MIS-C?

We thank the reviewer for highlighting this potential limitation with our analysis. We used a broad case definition for inclusion, as at the time the study was initiated none of the published criteria for MIS-C were validated. We did categorise each patient using the WHO MIS-C criteria using their clinical and laboratory data entered into the BATS database (see supplement page 15 for this data-driven approach). For each of the primary outcomes we report a sensitivity analysis restricting to this cohort of patients (figure 2, page 27), with none of the primary outcomes showing any difference between the primary comparison groups.

We have attempted to discuss this issue at length (page 21, first paragraph), noting that the majority of patients who did not meet the WHO criteria missed just one criterion (table S3, supplement page 26), and the reason for this was in most cases not having a confirmed COVID-19 exposure (figure S7, supplement page 56). We have argued in our discussion that for many patients this information could have been missing due to either lack of availability of SARS-COV-2 antibody measurements (especially early in the pandemic), or due to under reporting of exposure to SARS-COV-2 due to high frequency of asymptomatic infections in children.

Furthermore, because our data-driven approach could only use data entered onto the BATS database, it is possible that the proportion meeting the WHO criteria was in fact even higher, but we could not characterise them as such due to possible missing data – for example, lack of SARS-COV-2 exposure as described above.

4. How many BATS patients had CAA follow-up information? Differences across the reasons? Could there be follow-up bias for example in sites in South America, where IVIG availability is more restricted?

This is an important consideration. The majority of BATS data collection was during in-hospital care. We collected additional data on Coronary Artery Aneurysm follow-up only for patients who had CAAs on the final reported echocardiogram from the recruiting hospital (main manuscript page 10, supplement pages 17 and 68). The numbers with missing follow-up information are found in supplementary tables 6B-D (supplement pages 36-37, also includes how we define “not having follow-up”), and represent a relatively low proportion of those with CAA at discharge (under 17%). In addition, the follow-up rates were similar in the three treatment groups. A breakdown by region demonstrated the lowest follow-up rates are from South America (53%), Southern Europe (71%) and Central America (77%). Although this may allow for the potential for follow-up bias, a plot of the primary treatment groups by region (new supplementary figure SF17, supplement page 67) shows that in fact North and Eastern Europe have a much higher rate of primary therapy without IVIG than these countries. This therefore makes any potential follow-up bias highly unlikely to be significantly dependent on IVIG therapy. We therefore have not amended the text.

5. admission to ICU, ICU length of stay, and hospital length of stay are objective outcomes with relevance to patients. Post-hoc analyses on these outcomes would increase the relevance of the findings.

We thank the reviewer for this suggestion. During the initial study design, we specifically avoided these outcomes due to different admission/discharge criteria – for both hospital and ICU – across different sites, countries and continents. These outcomes are likely to be biased due to these discrepancies, which are often resource-driven rather than purely clinical. The ordinal scale was

chosen specifically to avoid this issue, as it captures more clinically relevant outcomes that are less biased by resource. In addition, the ordinal scale captures a measure of improvement across a wider spectrum of MIS-C patients. As experience managing cases of MIS-C increased it became apparent that a larger proportion of patients than in the initial reports of this new disorder did not require intensive care, and would therefore not be captured or provide information on the effect of treatment using ordinal scales solely based on levels of intensive care support. This was a further motivating factor in our carefully considered ordinal scale.

6. overall the result section is quite lengthy and in parts could be condensed with some materials shifted to the Appendix to facilitate readability.

We acknowledge the concerns raised from the reviewer. However, we believe that shortening the results session significantly would remove critical information that the general reader would require to interpret our broad findings and discussion. If the editors agree that this section should be reduced, then we will attempt to do so without removing critical information.

Minor comments

Introduction: ok

Methods:
- the authors refer to "pre-planned analyses". Was the SAP publicly available, or scripts uploaded on GitHub for example before completion of recruitment?

The SAPs and original protocols were all uploaded to the LANCET submission system, and are all publicly available on the ISRCTN clinical trials registry. The link is included on supplement page 3 at: https://www.isrctn.com/ISRCTN69546370

- clarity on case definition, see above.

Please see our response to Major comment 3 from the same reviewer, which we believe addresses this issue.

- handling missingness of data - what were the proportions of missing data (appendix page 13)? statistical review recommended

We thank the reviewer for this comment, with similar comments being raised by another reviewer. We are happy to provide these data, but to avoid another very large supplementary table we have restricted the full presentation of missing data to those variables where we used imputation/interpolation to adjust for missing data (as described in supplement pages 13-14). We have added an additional table to the supplement (table S15, page 47), which shows the total numbers and percentages of missing values for each variable, both before and after these approaches were applied, and separated by total number of missing days and patients with any missing data.

With regards to the validity of the methods used for handling missing data, we repeat our response to the other reviewer below:

We understand this valid concern from the reviewer, which we paid substantial attention to during our analysis plan. However, our imputation/interpolation methods (described in the supplement pages 13-14) are motivated by clinical experience, and we strongly believe they are preferred in this context. This is best demonstrated by an example. Suppose a patient was reported to have been receiving ventilatory support on day 1, with missing respiratory data on day 2, and then reported to be receiving ventilatory support on day 3, before having this support removed by the end of day 3. Clinical experience, both with MIS-C patients and general patients on ICU, tells us that by far the most likely value for this missing variable is for the patient to be receiving ventilatory support on day 2. In addition, experience from our previous analysis informs us that complete datasets within BATS demonstrate a very low (for example, for those receiving ventilatory support, less than 2% of complete records have discontinuous periods of ventilatory support).

Furthermore, we believe this data is most likely missing-not-at-random, rather than missing-at-random. To see why, it is important to consider the very time-consuming data entry task required for a single BATS patient, with significant data entry required for each day of a patient admission. Given this, if we look again at respiratory support as an example, we believe it is much less likely that missing values will be present when there is a change in ventilatory support, or supplemental oxygen. These clinical changes are important markers of a patient's clinical status, and are likely to be better recorded in clinical records and more accurately by those entering data. Because of this, we feel the proposed statistical methods for imputation are less appropriate. We are therefore confident that our imputation/interpolation methods, which are very similar to those published in the previous NEJM BATS paper, are sensible, pragmatic, and less biased than purely statistical methods.

- the primary and secondary endpoints were constructed with modification of the original study. Naturally, for a new disease like MIS-C such outcomes could be suitable. Please consider adding rationale/limitations related to the outcomes in the discussion section.

We thank the reviewer for this suggestion. We chose outcomes that we believe are most important for clinicians. Other significant work in the field has focussed on other outcomes, such as echocardiographic changes (Son et al. NEJM 2021), which we feel are less clinically relevant than our choice of primary outcomes. In addition, the use of cardiac function, detected by echocardiography, as an outcome would exclude those countries where immediate and serial echocardiography is not easily available. We have discussed the rationale for the outcomes where we believe necessary in the supplementary methods (for example, the outcomes based on our ordinal severity scale, supplement page 14). Most of these outcomes were reported in the first BATS report in the NEJM (McArdle et al. 2021), and their rationale and limitations have been previously described. We have therefore not altered the manuscript based on this suggestion.

- the SAP and analysis section seem fine, but statistical review is recommended.

We have separately addressed the comments from all reviewers on statistical aspects of the paper.

- more detail on the follow-up schedule should be provided in the appendix. it is likely that follow-up practices were highly variable across sites, potentially reducing the validity of the data.

For this study almost all data was collected during inpatient care, and the follow-up period extends only up to discharge/death. However, excluding the CAA outcomes, all outcomes are typically captured well during this time. These outcomes are clinically relevant, and less affected by variations in practice across institutions than metrics such as hospital/ICU length of stay. Indeed, this formed part of the rationale for the choice of outcomes (see response to reviewer comment #5 from the same reviewer). For the CAA outcomes we have reported in detail how we collected this follow-up data (page 10, supplement pages 17 and 68) and the follow-up rates in each group (supplementary pages 36-37, tables S6B-D). We therefore feel that additional description of the follow-up is not required.

Results:

- what was the distribution of the 3 treatment groups (IVIG, IVIG+G, G) and the additional groups (other immunomodulators) by country/region? there is a risk of confounding by site - it seems that analyses adjusted for HIC/U/LMIC based on World Bank criteria which may not fully account for site effects.

We thank the reviewer for this interesting comment, which is very similar to that raised by another reviewer. We have supplied an additional supplementary figure (SF17, page 67) which demonstrates the proportion of patients from each treatment group in different geographical regions. This demonstrates a relatively even spread of treatments across regions. The exceptions are North America, where use of any regime without IVIG was uncommon, as well as Africa and Asia – both regions contributing very few patients to BATS. Whilst there is a risk of confounding by site, we feel that the main driver for IVIG availability is likely to be resource status, which was the motivated factor behind using the World Bank resource group as a confounder.

We believe part of the strength of our analyses comes from the broad international representation of our cohort, which is unique in size and diversity. Given the very wide spread of recruiting centres across many different countries, adjusting by either site or country would likely introduce many unnecessary variables to the model. Since many of these would not share any relationship with treatment assignment or outcome, this would have the unwanted effect of artificially increasing the variance of our estimates (see Brookhart et al. 2006, Am J Epidemiol. *Variable selection for propensity score models*), leading to significant inaccuracy in significance testing. We therefore feel our approach is the most practical way of handling the diversity of recruitment sites, and have not altered our analyses.

- as expdcted the no treatment group was less sick. How was day 0 defined in the not treatment group?

Those not receiving any treatment were only compared with other treatment groups in analyses of the rate of decline in blood variables. In this case the admission day was used for comparison, as described in the caption to figure 4 (page 29)

- diagnosis, treatments, and severity outcomes for MISC seemed to change after the first wave. Can the authors describe changes over time in this large cohort?

Please see response on this issue to another reviewer repeated here:

We thank the reviewer for raising the important issue of how MIS-C has changed over time and with successive waves of SARS-CoV-2. We included the initial 614 patients in this analysis, as our analysis plan was to undertake early analysis (which we appreciated would lack power) in view of the urgent clinical need for data on management of a new disease. This approach was supported by WHO as initial guidance on management from WHO was based on the preliminary report. Our published analysis plan was to undertake further analyses as the numbers of cases in the study increased.

We agree with the reviewer that MIS-C has changed over time, both as clinicians became more familiar with the disorder, and probably as population levels of immunity have changed due to previous infection and vaccination. We are in the process of preparing a report on how clinical features, severity and treatment has change over time. We plan to submit this report as a separate letter or short report, following publication of the current manuscript. We do not feel it would be possible to adequately present the data on changes in MIS-C over time within the current manuscript which is already long and complex. In light of this comment, we have also added a paragraph in the discussion on the changing spectrum and epidemiology of MIS-C.

- timing of aneurysm diagnosis?

This is an interesting question, but unfortunately difficult to answer meaningfully from our data. We do present some data on this, looking at whether or not CAAs were first diagnosed before or after treatment initiation (supplementary table 6A, supplement page 36). This shows that, as expected, the % with CAA diagnosed before treatment were higher in the primary treatment regimens containing IVIG. This reflects inbuilt bias in the relationship between primary treatment and timing of CAA diagnosis, as the presence of CAAs is a strong motivator to initiate IVIG therapy.

Furthermore, the timing of diagnosis of coronary artery dilatation or aneurysms was likely to be influenced by the availability of echocardiographic and cardiology expertise, and thus affected by local resources and practice. These factors are also tied up with both IVIG availability and usage. For example, in North America usage of IVIG was very high for primary therapy, and the availability and frequency of echocardiograms is also higher in this population. Adjusting for these features whilst looking at the exact timing of CAA diagnosis is not practical for our dataset, and so we feel reporting these values for our analysis would not lead to interpretable and meaningful results.

Due to these factors, we felt that it would be difficult to establish if any treatment prevented CAA, as the number of patients with echocardiography results before any treatment is relatively low. However, from the clinical point of view the issue of whether treatment affected persistence or resolution of CAA is of more importance, and we focused analysis on whether CAA were present on discharge and if they resolved differently in each treatment group.

- page 16. "Addition of IVIG as secondary treatment is more likely in severely ill patients, those with CAA detected, and those not rapidly improving on primary treatment." it is not clear if this is an interpretation or finding (tense change)? Post-hoc analyses? Coudl the authors provide data and 95%-confidence intervals for these exploratory statements?

We thank the reviewer for this comment, and agree this is ambiguous. We have not analysed our data for these findings, and this statement was from our direct experience and clinical intuition. We

agree this should not be present in the results section, and so we have removed this sentence and an additional relevant sentence from the results section.

- please provide % in addition to absolute numbers where appropriate.

We have added percentages to the reported unadjusted death numbers for the three primary treatment groups (page 17). We felt addition of percentages was not necessary for other reported absolute numbers where it is not already reported.

Discussion:
The discussion could be shortened and condensed.

As described in our response to comment #6 from the same reviewer, we acknowledge the concerns raised regarding the length of the current manuscript. However, we again feel that additional significant reduction of the discussion would reduce readability of this substantial set of analyses, with many complexities meriting discussion. If required by the editors we will review and shorten if required.

Figures:
Excellent

Tables:
Ok

Reviewer #4:

Channon-Wells et al report follow-up data with a larger cohort of patients with presumed MIS-C in the Best Available Treatment Study (BATS). The original publication from BATS in the NEJM included the first 614 patients enrolled in the study. The current manuscript includes over 2,000 children from 39 countries, making it an important addition to the field.
The authors compare initial immunomodulatory therapy (started on day 0/first calendar day of immunomodulatory therapy). The primary outcomes were a composite of inotropic support/ventilator support on day 2 or later or death (same as prior primary endpoint in NEJM) and time to improvement in at least one level of the ordinal clinical scale. There was no statistically significant difference amongst the treatment groups (combination therapy compared to IVIG alone and glucocorticoid monotherapy compared to IVIG alone) in the primary outcomes; however, there was a trend towards slower time to improvement in the glucocorticoid alone group vs. IVIG alone. For secondary analyses, there was less treatment escalation in the combination vs IVIG alone groups and glucocorticoid vs. IVIG alone groups. Fever was less common at day 2 in the combination group compared to the IVIG alone group. When comparing combination therapy with glucocorticoids alone in secondary analyses, there was faster time to improvement in the IVIG + glucocorticoid group. Similarly, there was less treatment escalation and persistence of fever at day 2 in the combination group vs. glucocorticoids alone.
For CAAs, the incidence, severity, and rate of resolution were similar across groups during the follow up period. Compared to the initial study, this version provides information about CAAs after hospital discharge, which will be of great interest. However, only patients with known CAAs had results reported for echocardiograms after hospital discharge.

Major concerns:
1)One major unresolved question in the field is whether glucocorticoid monotherapy is equivalent to combination therapy with IVIG and glucocorticoids. In this study, a greater proportion of the patients in the IVIG + glucocorticoid group met the WHO case definition for MIS-C, were ventilated and/or treated with inotropes at day 0, had elevated inflammatory markers, and had CAAs (even before initiation of immunomodulators) compared to those in the glucocorticoid group alone. Thus, patients were sicker in the combination therapy group. This discrepancy in the treatment groups coupled with the finding that combination therapy was associated with faster improvement (a primary endpoint), less treatment escalation, and less fever than glucocorticoid alone all indicate that combination therapy is superior to glucocorticoids alone. Yet, the interpretation drawn by the authors is that these two treatment approaches are equivalent.

We thank the reviewer for raising this important concern. We note that as comparison of combined IVIG+GC with GC alone was a secondary analysis, the results were not corrected for multiple hypothesis testing which would decrease the significance of the reported findings. We also note that reduction of fever, and escalation of therapy are “soft endpoints” and that that the faster rate of improvement with combination therapy affected predominantly the less ill patients. However, we agree with the reviewer that we may have been excessively negative about potential benefits of dual therapy. We have therefore reworded the discussion to make the conclusion more balanced.

2) As the COVID19 pandemic has evolved, the incidence and severity of MIS-C have decreased, which may be in part due to vaccination and/or change in SARS-CoV-2 variants. The authors do not provide any information about vaccination or time of illness. Further, if treatment groups had an uneven distribution of patients over time, this would greatly confound results and this point is left totally unaddressed by the authors.

We thank the reviewer for this comment which has also been made by two other reviewers . We repeat the response here:

We agree with the reviewer that MIS-C has changed over time, both as clinicians became more familiar with the disorder, and probably as population levels of immunity have changed due to previous infection and vaccination. We are in the process of preparing a report on how clinical features, severity and treatment has change over time. We plan to submit this report as a separate letter or short report, following publication of the current manuscript. We do not feel it would be possible to adequately present the data on changes in MIS-C over time within the current manuscript which is already long and complex. In light of this comment, we have also added a paragraph in the discussion on the changing spectrum and epidemiology of MIS-C.

With regard to an effect of vaccination on MIS-C presentation we note that the publication of the French experience (Ouldali et al. The Lancet Regional Health- Europe 2022;17: 100393 – reference also added to discussion) showed MIS-C can occur after vaccination, but is very rare and as vaccines were not widely available to the paediatric population during BATS data collection this has probably had little effect on the disease to date.

Minor points

1) In the research in context section, the COVID-19 treatment guidelines panel from NIH provides treatment guidelines in the USA for COVID19/MIS-C, not CDC

Thank you for highlighting this fact. We have changed the reference (page 7) to the American College of Rheumatology, who provided the first clinical guidance for MIS-C in the USA

2) A greater proportion of patients in the combination group (~85%) met the WHO case definition for MISC compared to IVIG alone (~77%) and glucocorticoidds alone (~81%). This should be stated in the results section and addressed in the discussion.

We agree with the reviewer that there are some differences in the % of patients meeting the WHO criteria between groups, although the differences were fairly small. We have amended a sentence in the discussion to this effect, but have not added directly to the results section in light of limited space and the fact that this effect is relatively small, and the fact that our results do not change when restricting to the patients meeting the WHO MIS-C case definition.

3) The inclusion criteria for the BATS study is broad. Patients do not need to meet existing case definitions and can be referred by pediatricians and not the specialists that typically diagnose and treat patients with MIS-C (ICU, rheumatology, cardiology, ID). The concern is that children with other febrile illness of childhood were erroneously induced in the study. The fact that13% of patients tested negative for SARS-CoV-2 antibodies further supports this concern. This should be addressed in greater detail in the discussion.

We thank the reviewer for highlighting this potential limitation with our analysis. We used a broad case definition for inclusion, as at the time the study was initiated none of the international criteria for MIS-C were validated. We categorised each patient using the WHO MIS-C criteria in a data driven way, using only their clinical and laboratory data entered into the BATS database (see supplement page 15 for this data-driven approach). Importantly, despite a proportion of patients not meeting this data-driven criteria, our sensitivity analysis restricting to the cohort of patients who did (figure 2, page 27) showed no difference between the primary comparison groups for both primary outcomes.

We have attempted to discuss this issue at length (page 21, first paragraph), noting that the majority of patients who did not meet the WHO criteria missed just one criterion (table S3, supplement page 26), and the reason for this was in most cases not having a confirmed COVID-19 exposure (figure S7, supplement page 56). We have argued in our discussion that for many patients this information could have been missing due to either lack of availability of SARS-COV-2 antibody measurements (especially early in the pandemic), or due to under reporting of exposure to SARS-COV-2 due to high frequency of asymptomatic infections in children.

Furthermore, because our data-driven approach could only use data entered onto the BATS database, it is possible, and we feel quite likely, that the proportion meeting the WHO criteria was in fact even higher, but we could not characterise them as such due to possible missing data – for example, lack of SARS-COV-2 exposure as described above. In the discussion we note that as the pandemic has evolved, the presence of SARS-COV-2 antibody has become less useful as a diagnostic for MIS-C, due to widespread asymptomatic infection.

With regards to the diagnostic accuracy of cases recruited to BATS, our experience of running the study is that the majority of centres and clinicians entering patients are from the specialties more experienced with MIS-C (ICU, Infectious Diseases, Rheumatology, Cardiology). Since MIS-C has a broad phenotype that can be mimicked by other diseases there is no feasible way to design a trial with 100% diagnostic accuracy for recruited MIS-C patients. We feel our cohort represents the spectrum of patients treated for MIS-C, and therefore our results are clinically meaningful for those managing this disorder. We have currently not amended our discussion considering our above viewpoint, and we would appreciate any further comments from the reviewer.

4) There seems to be some discrepancies on the numbers patients in the glucocorticoid alone group who later received IVIG. In table S1, the number is 191. In table S6E, it is 230. Either way, a significant proportion of the 487 children in the glucocorticoid along group received additional treatment with IVIG. This should be addressed in the results section and not in a supplemental table.

We thank the reviewer for this comment. We apologise for the error, and are grateful for this being highlighted. We have updated table S1 in the supplement page 23, and the corresponding results section of the main manuscript (page 12). We agree that many of the patients initially commenced on single agent therapy had other agents added. We show this in the Sankey plot which graphically illustrates the crossing over of treatments and also the timing of the changes. We hope this figure provides the reviewer with a clear indication in the main results of the addition of treatments and timing.

5) Follow up echocardiograms were only reported for the 236 patients with known CAAs. Can the authors report on follow up echocardiograms in all patients (not just those with CAAs).

Unfortunately, we do not have this data, as this was not part of the original data collection protocol. We pragmatically chose to request follow-up data from patients with known CAAs at discharge to limit the burden on centres entering data, in an effort to maximize the quality and impact of the returned additional data. It is worth noting that sites were not funded to enter patients, and so we relied entirely on good will for this venture. Based on both the clinical experience of the study team and the available follow-up data we have in patients with CAAs at discharge, we believe that thankfully the rate of newly diagnosed CAAs after discharge is low.

6) In the discussion, the authors discuss that combination therapies can have more side effects. Yet, there was no difference in adverse events reported in the combination vs. monotherapy groups in the results section of the study.

We agree with the reviewer that theoretically combined therapy should result in more side effects as there will be negative effects of both the steroid and IVIG component. However, many of the side effects of IVIG relate to the administration of large fluid volumes to patients with already compromised cardiac function, and less frequently to allergy, haemolytic anaemia or immune mediated effects. Our data capture tool may not easily have enabled volume overload to be detected as a side effect. Late effects of corticosteroids such as secondary infection or avascular bone necrosis may also not have been captured by our data collection process. Despite these limitations in data capture, we feel it is reasonable to expect increased side effects for patients receiving polypharmacy, based primarily on clinical experience and intuition.

Reviewer #5: Manuscript ID: TLRHEU-D-22-00664

Thank you for giving me the opportunity to review this manuscript.
This is a very well-designed study comparing outcomes depending on initial immunomodulatory treatments for MIS-C. This is the largest study on this topic.
The use of the propensity score approach is a very important strength, allowing to adequately adjust for differences in baseline characteristics between groups. Given the lack of randomized study, this is likely to provide the highest level of evidence. For all these reasons, I strongly support publication of this study.
I have the following remarks/questions that should be addressed before acceptance.

1) The main previously published study showing more favorable outcomes for IVIG plus steroids compared to IVIG alone for MIS-C is the Overcoming study (Son et al, NEJM 2021, 10.1056/NEJMoa2102605). The main outcome of this study was a composite outcome of hemodynamic support or cardiac dysfunction on day 2 or after. Because the present study found discrepant results compared to this study, it would be helpful to analyze the same outcome to allow comparison. I suggest adding this analysis among the secondary outcomes, and discussing these different results.

We thank the reviewer for this comment. However, we feel this additional composite outcome is not especially clinically meaningful, and its addition is likely to complicate an already extensive analysis. As the rapid availability of echocardiography on a sequential basis is largely unavailable in much of the world outside US centres, the coverage of echocardiographic data is less complete in our cohort compared to that of Son et al. In addition, since echocardiographic findings are heavily influence by current volume status and haemodynamic support, we feel that it is a less clinically informative outcome compared to the provision of haemodynamic support. We hope to assure the reviewer that in any case, it is very likely that this additional composite outcome would also not show any significant difference between treatment groups, since the individual outcomes are reported and show no difference (haemodynamic support – see figure 2, page 27 – and LV dysfunction – see figure S9, supplement page 58). We have therefore not added this analysis to our report.

2) We have seen that MIS-C frequency and severity decreased over time, which may be due to the different circulating variants. Did this study capture different waves of SARS-CoV-2? If so, would it be possible to analyze separately the different waves? One explanation of the lack of difference between groups might be related to less severe cases compared to the historic variant analyzed by Son et al.

We thank the reviewer for this important point which has also been raised by other reviewers we repeat here our response to the other reviewers :

We agree with the reviewer that MIS-C has changed over time, both as clinicians became more familiar with the disorder, and probably as population levels of immunity have changed due to previous infection and vaccination. We are in the process of preparing a report on how clinical features, severity and treatment has change over time. We plan to submit this report as a separate letter or short report, following publication of the current manuscript. We do not feel it would be possible to adequately present the data on changes in MIS-C over time within the current

manuscript which is already long and complex. In light of this comment, we have also added a paragraph in the discussion on the changing spectrum and epidemiology of MIS-C.

3) Among the reported limits of this study, the fact that an important number of MIS-C cases in the IVIG or steroids alone groups received an escalation before analyzing the outcomes on day 2 is of concern. I really appreciated the discussion provided by the authors to discuss this point, which is fairly reported in the limit section. I was just wondering if analyzing the main outcome considering the escalation as a failure would provide different findings? If so, it would be very interesting to add this is the results section as an exploratory analysis and to add this in the discussion.

We thank the reviewer for this important question. We discussed at length including escalation of therapy as a primary outcome. However, we felt that it was likely to be biased as patients treated with a single agent, who were not rapidly improving would be more likely to have a second agent added than those already on dual therapy. We agree that we might have found a statistical benefit of combined therapy over single agent if we had included escalation as a failure of first treatment but felt this might be a biased conclusion for the reasons mentioned.

We have however reported escalation of therapy as a secondary outcome, and included restricting analysis to those who did not escalate as sensitivity analyses, so we feel we have addressed the issue in a manner that can let readers consider the issue.

4) Regarding the comparison between steroids alone and IVIG plus steroids: one of the two main outcomes (time-to-improvement) found differences between the 2 groups, favoring the combination therapy. Again, I appreciated the elegant discussion on this finding. However, in this context, I am not sure that it is completely appropriate to state in the first sentence of the discussion that no differences were found between groups for the main outcomes. I fully understand and agree that for resource limited settings, steroids alone may be considered as the first-line therapy, but I think the conclusions of the paper should be driven by the main outcome results, and may highlight this point. Especially in the abstract, it is stated that no difference were found for the main outcomes, which seems incorrect.

We agree with the reviewer that we may have been excessively cautious in reporting the lack of significant difference in primary outcomes between Glucocorticoids and combination therapy as showing no difference between the two. However, we note that the comparison of Glucocorticoids and IVIG +GC was not a planned primary analysis, and was therefore not subject to multiple hypothesis correction, and the precision of the differences detected would be greater, and may not reach significance, with correction. Furthermore, as we discussed neither the lower rate of fever, or more rapid recovery in the less ill patients may be overly significant clinically in comparison to the other non-significant results. That said, we agree that a more open discussion of these findings would be preferable. We have therefore altered the wording in the discussion to be more open to a possible benefit of combination therapy

5) This study included 2009 MIS-C cases, but it seems that only 1505 were included in the main analysis, based on the number of excluded cases described in the results section. This number is not

reported anywhere in the manuscript (sorry if I am wrong). To meet reporting guidelines, it is required to clearly report the number of cases included in the main analysis in the abstract, the result section, and in the flow chart (Figure 1).

We thank the reviewer for drawing our attention to this. We have updated the flow diagram (figure 1A page 25) to include this information, and have edited the text in the results section on page 13. We have added a brief summary of these numbers to the abstract (page 3), with further detail precluded due to space limitations.

6) In the same way, the number of included cases is lacking in the Figure 2, 3 and 4. I can not find easily the number of included cases for each analysis in these figures, neither in the main manuscript. I suggest adding, for each analysis, the number of included cases, at least in the legend of the figures.

We thank the reviewer for this comment. Given the number of analyses presented, it is not possible to add these numbers in figure 2 whilst retaining readability. We have instead added a sentence to the caption (page 27) directing the reader to this detailed information in the supplementary material. We remark that the reader will obtain a sense of the respective size of datasets from the width of the confidence intervals and can then refer to the supplement if the exact numbers are of particular interest. The numbers for each analysis in figure 3 can be found in the tables below each plot as the numbers at risk at time zero (page 28). These tables were accidentally omitted on the first copy of this figure in our first submission, for which we apologise. This omission has been corrected. For figure 4 the numbers are already included within the plot panels (page 29).

Editorial points - IMPORTANT:

- The following points list items that **must be included before the manuscript can be considered further**. Addressing them at this stage reduces the risk of errors and delays later.
- Please read the requirements below carefully and consult me or https://www.thelancet.com/preparing-your-manuscript, for further details or clarification if needed.
- Please note that not every point below will be relevant to your manuscript.

We thank the editor for providing this checklist. We have written specific responses only to points we feel are relevant to our manuscript, or have not been addressed during the editor's specific comments or reviewer comments above.

Authorship and reporting guidelines:

1. Please check that all author name spellings and affiliations are correct.
2. Please indicate any authors who are full professors.
3. Please list the highest degree for each author (one degree only, please).

We have amended these to display just the highest degree for each individual (page 1).

4. Please follow the appropriate EQUATOR network reporting guidelines and include the corresponding checklist(s). These include: CONSORT reporting guidelines for randomised trials (http://www.consort-statement.org), STROBE for observational studies, PRISMA for systematic reviews, STARD for diagnostic studies, CHEERS for economic evaluations and RECORD for routinely collected health data. *Lancet* specific guidelines for reporting RCT and systematic reviews and meta analyses are available here:
   http://www.thelancet.com/pb/assets/raw/Lancet/authors/Rctguidelines.pdf
   https://thelancet.com/pb/assets/raw/Lancet/authors/metaguidelines.pdf
5. *The Lancet Rheumatology* endorses the SAGER guidelines for reporting of sex and gender information in study design, data analyses, results and interpretation of findings: https://www.equator-network.org/reporting-guidelines/sager-guidelines/. For all study types, we encourage correct use of the terms sex (when reporting biological factors) and gender (when reporting identity, psychosocial, or cultural factors). Where possible, report the sex and/or gender of study participants, and describe the methods used to determine sex and gender. Separate reporting of data by demographic variables, such as age and sex, facilitates pooling of data for subgroups across studies and should be routine, unless inappropriate. Discuss the influence or association of variables, such as sex and/or gender, on your findings, where appropriate, and the limitations of the data.

Title/summary:

6. Please ensure that the title of the paper is non-declamatory (i.e, it describes the aim of the study rather than the findings) and that it includes a description of the study type (e.g,. a randomised controlled trial).

We have added a description of the study methodology in the title of our paper (page 1).

7. For trials, please limit the summary to pre-defined primary endpoints and safety endpoints.
8. For trials, please state the trial registration number.
9. Please report the sex/gender, age, and ethnicity of the study population (n/N [%]) in the summary if applicable.

We have not added these data to our summary, as we feel this is not necessary for our particularly study. Sex, age and ethnicity have been described within the paper at relevant points.

Methods:

10. At the end of the methods section please state the role of the funder in: data collection, analysis, interpretation, writing of the manuscript and the decision to submit.
11. Please explain any deviations from the protocol.
12. Please ensure that all outcomes specified in the protocol (including all secondary outcomes) are reported in the manuscript. If there are any secondary endpoints that cannot be included, please mention these explicitly and explain why and where they will be made available.
10. If any exploratory outcomes are reported that were not pre-specified, please make it clear that these analyses were post-hoc.

We have adjusted the language where necessary to ensure clarity regarding planned and post-hoc analyses.

11. Please use rINNs for drug names. For genes and proteins, authors can use their preferred terminology so long as it is in current use by the community, but should provide the preferred name from Uniprot (http://www.uniprot.org/uniprot/) for proteins and HUGO (http://www.genenames.org) for genes at first use to assist non-specialists.

We have reviewed all drug names and amended these where necessary.

12. For drug studies, please ensure that details of doses, route of delivery, and schedule are included.

We have described drug doses, routes of delivery and schedule where possible. However, given the observational nature of this study we have not predefined any of these features, and in addition dosing schedules given prior to admission, or after discharge/transfer are not recorded.

Results:

16. For the main outcome measures, please include a result for each group, plus a point estimate (eg, RR, HR) with a measure of precision (e.g, 95% CI) for the absolute difference between groups, in both the Summary and the main Results section of the paper.

We have reviewed our manuscript and no changes were required.

17. p-values should be given to two significant figures, but no longer than 4 decimal places (e.g. $p<0.0001$).

18. Please provide absolute numbers to accompany all percentages. Percentages should be rounded to whole numbers unless the study population is very large (>1000 individuals).

We have reviewed our manuscript and appendix and made amendments where necessary to present the data in the requested format.

19. ~~Please provide absolute numbers to accompany all percentages. Percentages should be rounded to whole numbers unless the study population is very large (>10 000 individuals).~~ DUPLICATE from above

20. ~~Please give 95% confidence intervals for hazard ratios/odds ratios.~~ DUPLICATE from above

21. For means, please provide standard deviation (or error, as appropriate).

We have reviewed our manuscript and no changes were required.

22. Please provide interquartile ranges for medians.

We have reviewed our manuscript and no changes were required.

23. Please provide numbers at risk for Kaplan-Meier plots and ensure that plots include a measure of effect (e.g, log-rank p); estimates should be reported with 95% CIs

24. Where possible, we ask that you present data (primary/secondary outcomes, adverse events, patient-reported outcomes) disaggregated by biological sex (or gender, if collected). These data can be included in the main tables or presented in the appendix. For RCTs, please report the study as pre-specified in your protocol, but post-hoc assessment of outcomes disaggregated by sex/gender is welcomed.

We do not feel this is relevant to our study. We have adjusted for sex during our primary analyses, and these covariates are therefore balanced between these groups.

Discussion:

25. Please ensure that the Discussion contains a section on limitations of the study.

Our discussion also has a heavy discussion of the limitations of the study, which has been further enhanced on consideration of the reviewer comments. If a completely new section is requested, we are happy to separate out these comments, although we feel this may negatively impact the narrative and readability of the discussion

Additional requirements:

25. Please provide the figures in an editable format (eg, EPS files, PowerPoint files, depending on software used to produce them). If figures are composed of photographs or other images, high resolution files (300 dpi or greater) should be provided. More information can be found here: https://www.thelancet.com/for-authors/forms?section=artwork.

26. References should be in Vancouver style. For references with six authors or fewer, all authors should be listed. For those with seven or more authors, only the first three authors and 'et al' should be listed. Please ensure that reference numbering throughout the manuscript is not inserted with electronic referencing software, such as Endnote, as this is incompatible with our production system (if used, please convert to normal text before

resubmission). If the references “move” from the body text into tables or figures, please maintain the sequence of citation. Please ensure tables and figures are cited correctly in the body text to prevent the need for renumbering of references should the table and figure citations subsequently move. All web references should have the exact date they were last accessed. With your revised submission please enclose copies of any papers cited as being 'in-press', along with a copy of the acceptance letter from the journal. References that are "submitted" should be removed and citations in the text replaced with "(unpublished data; authors)".

We have reviewed all references, including those in the supplement, and ensured they all meet Vancouver styling as specified above.

27. If accepted, only 5-6 non-text items (figures, tables, or panels) can be accommodated in the main paper; additional material can be provided in a web appendix. Please indicate which items can go in a web appendix.

28. Please provide a research in context panel with 3 parts: Evidence before this study (which includes a description of how you searched for evidence and how you assessed the quality of that evidence); Added value of the study (this section should not simply repeat the results but indicate how it adds to the field); and Implications of all the available evidence.

29. At the end of the manuscript, please provide a Contributors statement that summarises the contribution of each author to the work. *The Lancet's* journals require that more than one author has directly accessed and verified the underlying data in all research articles. For research articles that are the result of an academic and commercial partnership, at least one of the authors named as having accessed and verified data must be from the academic team. Please state which author(s) have accessed and verified the data, and which author(s) were responsible for the decision to submit the manuscript.

30. At the end of the manuscript please summarise the declaration of interests for each author.

31. As corresponding author, please confirm that all authors have seen and approved of the final text

32. If your author line has more than 20 authors, we very strongly encourage the use of a study group name. Collaborators' names and affiliations may be listed in the appendix. Additionally, if you wish the names of collaborators within a study group to appear on PubMed, please upload with your revision a list of names of all study group members presented as a two-column table in Word. First and middle names or initials should be placed in the first column, and surnames in the second column. Names should be ordered as you wish them to appear on PubMed. The table will not be included in the paper itself - it is simply used to make sure that PubMed adds the names correctly.

33. Please note our guideline length for research articles is 3500 words and 30 references. For RCTs, the text can be expanded to 4500 words.

34. All research articles must contain a data sharing statement, to be included at the end of the manuscript. For more information on these required statements see the Data sharing section of the Information for Authors (https://thelancet.com/pb-

assets/Lancet/authors/tlrheum-info-for-authors.pdf) and (https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(17)31282-5/fulltext)

We have included a data sharing statement at the end of our manuscript (page 35), adjusted to meet the needs of our study group from the example statements provided by the journal.

35. We require written and signed consent from any individuals who are cited in the acknowledgments section or as personal communications. The following format can be used and a signed statement uploaded on resubmission:

    - "I permit <corresponding author> et al to list my name in the acknowledgment section of their manuscript and I have seen a copy of the paper <full article title>."
    - "I permit <corresponding author> et al to cite a personal communication from me in their manuscript <full article title>."

Given the consortium's size, and the long-running nature of data collection it is not practical to obtain written consent from all of these past and present collaborators, some of whom may have moved from the roles they were in during their involvement with BATS. All participants who recruited patients into BATS we aware of the intention of acknowledging all consortium members. We have contacted all recruiting sites to request that they inform us of any of their members who explicitly do not want to be listed in the consortium's membership.

TECHNICAL INFORMATION:

When you submit the revised paper, please provide the following:

1. One "clean" copy of your manuscript
2. One copy where your changes are highlighted (tracked changes).
3. A separate, point by point response to the editorial and referee comments typed immediately following each specific point above. Please do not use
4. Any images and/or tables (even if no revisions have been made).

Please do NOT include a copy of your original manuscript. All text files should be supplied as MS Word files.

Please also supply the word count for the body of your paper and your abstract (word count for the body of your paper should not include abstract, references, figures or tables).

To enable readers to better appreciate research findings and to encourage full and transparent reporting of outcomes, *The Lancet* family journals offer to publish a webaddress in accepted paper that links to the study's protocol on the author's institutional website (see Lancet 2009; 373: 992). This is particularly encouraged for randomised controlled trials, but is welcome for all types of research.

To submit your revised manuscript, please visit *The Lancet Rheumatology's* Online Submission and Peer Review Website at: https://www.editorialmanager.com/tlrheu/ and enter your username and password.

Your username is: Your username is: m.levin@imperial.ac.uk
If you need to retrieve password details, please go to: click here to reset your password

After you have entered your account details, remember to click the 'Author' button. You will see a menu item call 'Submission Needing Revision'. You will find your submission record there.

We ask all authors of, and all contributors (including medical writers and editors) to specify their conflicts of interest (if any) and individual contributions to a manuscript under consideration at *The Lancet Rheumatology*. *The Lancet Rheumatology* will not publish any articles unless we have a completed author statement form, conflict of interest form, and the signatures of all authors. Please sign and complete the author statement form and the ICMJE conflicts of interest statement form, and either upload the signed copies in to EM with your manuscript, scan and email to rheumatology@lancet.com. In addition, please also include written consent of any cited individual(s) noted in acknowledgments or personal communications.

The Lancet Rheumatology's requirements are described in more detail at the Information for Authors page at: https://www.thelancet.com/pb-assets/Lancet/authors/tlrheum-info-for-authors.pdf

The editors may use such information as a basis for editorial decisions and will publish such disclosures if they are believed to be important to readers in judging the manuscript.

In summary, the signed forms we require are:

- Authors' contribution and signatures (hand-written and electronic signatures both accepted)
- Signed Conflict of interest statement for ALL authors (1 form per author)

Please also check whether you need to provide the following:

- Signed copyright permissions for previously published material
- Signed consent from individuals cited (by name) in the Acknowledgements
- Signed consent for use of cited personal communications

STROBE REPORTING CHECKLIST

Figure 1A

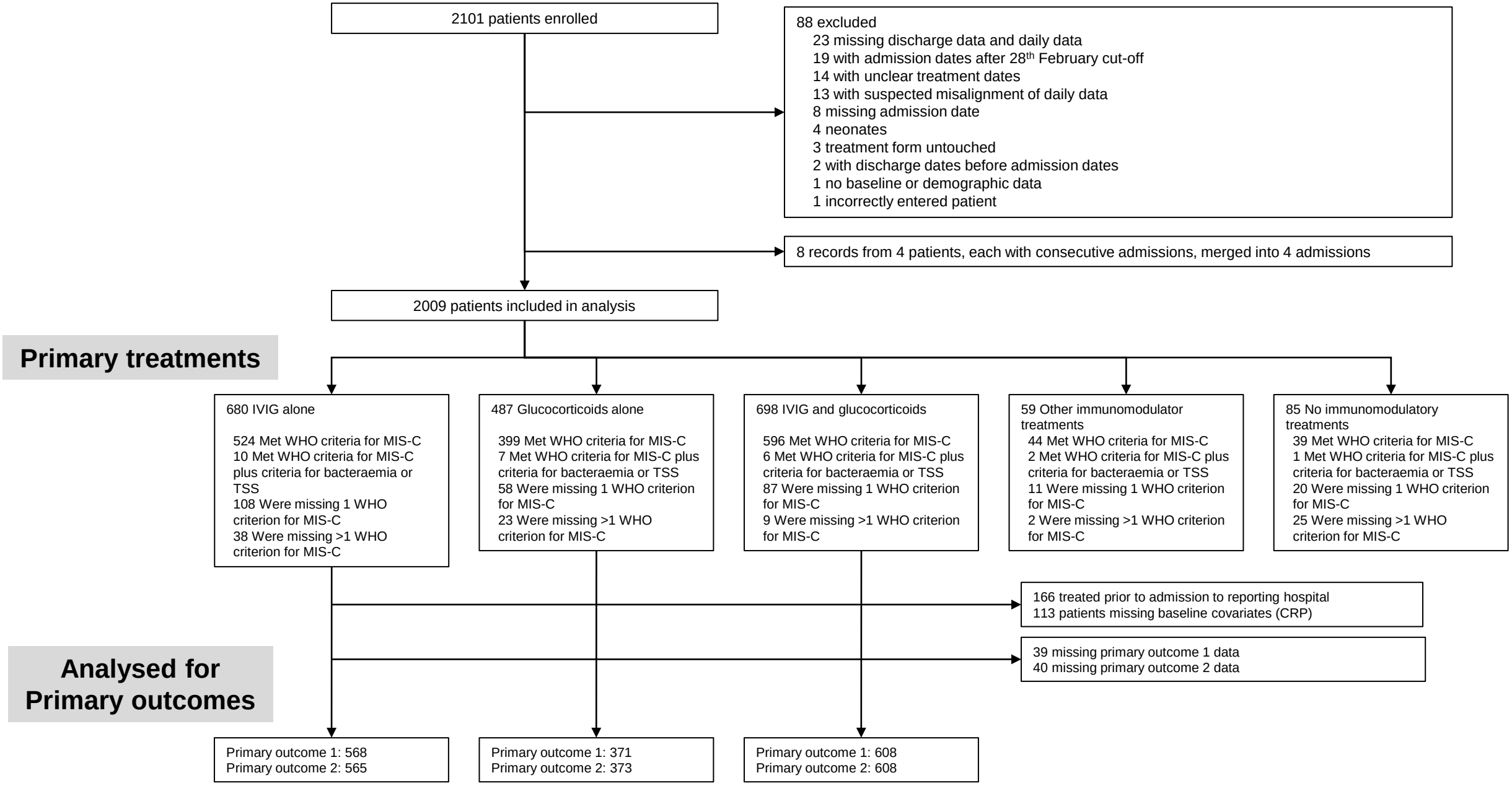

2101 patients enrolled
88 excluded
23 missing discharge data and daily data
19 with admission dates after 28th February cut-off
14 with unclear treatment dates
13 with suspected misalignment of daily data
8 missing admission date
4 neonates
3 treatment form untouched
2 with discharge dates before admission dates
1 no baseline or demographic data
1 incorrectly entered patient
8 records from 4 patients, each with consecutive admissions, merged into 4 admissions
2009 patients included in analysis
Primary treatments
680 IVIG alone
524 Met WHO criteria for MIS-C
10 Met WHO criteria for MIS-C plus criteria for bacteraemia or TSS
108 Were missing 1 WHO criterion for MIS-C
38 Were missing >1 WHO criterion for MIS-C
487 Glucocorticoids alone
399 Met WHO criteria for MIS-C
7 Met WHO criteria for MIS-C plus criteria for bacteraemia or TSS
58 Were missing 1 WHO criterion for MIS-C
23 Were missing >1 WHO criterion for MIS-C
698 IVIG and glucocorticoids
596 Met WHO criteria for MIS-C
6 Met WHO criteria for MIS-C plus criteria for bacteraemia or TSS
87 Were missing 1 WHO criterion for MIS-C
9 Were missing >1 WHO criterion for MIS-C
59 Other immunomodulator treatments
44 Met WHO criteria for MIS-C
2 Met WHO criteria for MIS-C plus criteria for bacteraemia or TSS
11 Were missing 1 WHO criterion for MIS-C
2 Were missing >1 WHO criterion for MIS-C
85 No immunomodulatory treatments
39 Met WHO criteria for MIS-C
1 Met WHO criteria for MIS-C plus criteria for bacteraemia or TSS
20 Were missing 1 WHO criterion for MIS-C
25 Were missing >1 WHO criterion for MIS-C
166 treated prior to admission to reporting hospital
113 patients missing baseline covariates (CRP)
39 missing primary outcome 1 data
40 missing primary outcome 2 data
Analysed for Primary outcomes
Primary outcome 1: 568
Primary outcome 2: 565
Primary outcome 1: 371
Primary outcome 2: 373
Primary outcome 1: 608
Primary outcome 2: 608

Figure 1B

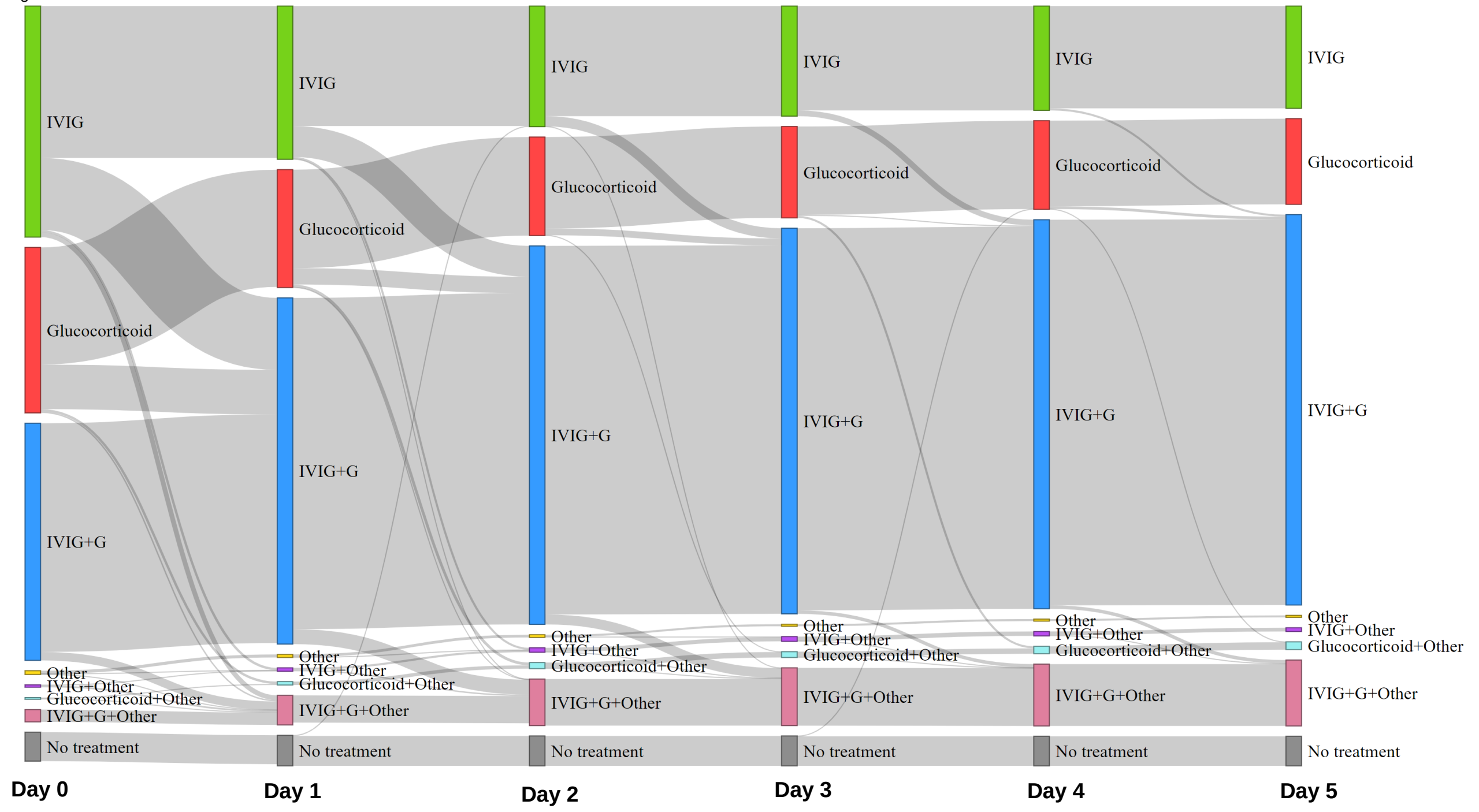
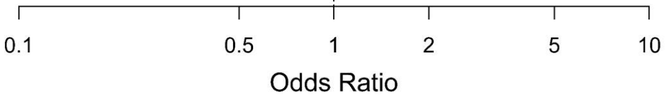

IVIG
Glucocorticoid
IVIG+G
Other
IVIG+Other
Glucocorticoid+Other
IVIG+G+Other
No treatment
Day 0
Day 1
Day 2
Day 3
Day 4
Day 5

Figure 2A-D

A

Glucocorticoids vs IVIG - first primary outcome

Outcome	P-value	OR [95% CI]
Primary analysis		
Inotropes/Ventilation/Death	1.00*	0.93 [0.58, 1.47]
Components of composite		
Death	0.51	1.68 [0.35, 7.99]
Inotropes from day-2	0.98	1.01 [0.63, 1.60]
Ventilation from day-2	0.43	1.30 [0.67, 2.51]
Sensitivity analyses		
Primary-therapy on days 0 & 1	0.81	0.92 [0.46, 1.85]
Propensity matched analysis	0.96	0.99 [0.61, 1.61]
Impute CRP (median by treatment group)	0.82	0.95 [0.62, 1.46]
Subgroup analyses		
Full WHO MIS-C criteria	0.39	0.81 [0.50, 1.32]
Full WHO MIS-C criteria + bacteraemia/TSS	0.41	0.82 [0.51, 1.32]
Missing up-to 1 WHO MIS-C criteria	0.63	0.89 [0.56, 1.42]
Exclude Low & Lower-middle Income countries	0.60	0.88 [0.54, 1.43]
Age < 6 years	0.94	0.96 [0.36, 2.59]
Age 6-11 years	0.98	0.99 [0.53, 1.86]
Age > 11 years	0.54	0.75 [0.30, 1.90]
No significant comorbidities	0.82	0.95 [0.59, 1.53]
Baseline CRP < 118.0	0.49	0.67 [0.21, 2.14]
Baseline CRP 118.0 - 206.5	0.84	1.09 [0.49, 2.42]
Baseline CRP >= 206.5	0.75	0.90 [0.46, 1.76]
Post-hoc subgroup analyses		
Meet complete KD criteria	0.63	0.79 [0.29, 2.14]
Do not meet complete KD criteria	0.79	1.08 [0.61, 1.90]
Exclude changes in treatment group	0.43	0.71 [0.30, 1.67]



B

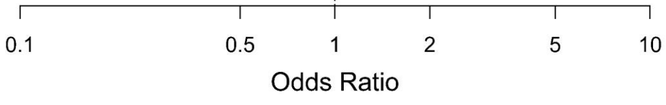
Glucocorticoids vs IVIG - second primary outcome

Outcome	P-value	AHR [95% CI]
Primary analysis		
Time-to-Improvement	0.22*	0.84 [0.70, 1.00]
Baseline subgroups		
Subgroup: Requiring intensive support at baseline	0.75	1.06 [0.75, 1.49]
Subgroup: Requiring no intensive support at baseline	0.21	0.83 [0.62, 1.11]
Sensitivity analyses		
Primary-therapy on days 0 & 1	0.051	0.82 [0.68, 1.00]
Propensity matched analysis	0.10	0.87 [0.74, 1.03]
Two-point Time-to-Improvement	0.10	0.86 [0.72, 1.03]
Impute CRP (median by treatment group)	0.054	0.82 [0.68, 1.00]
Subgroup analyses		
Full WHO MIS-C criteria	0.31	0.91 [0.75, 1.10]
Full WHO MIS-C criteria + bacteraemia/TSS	0.25	0.90 [0.74, 1.08]
Missing up-to 1 WHO MIS-C criteria	0.078	0.85 [0.71, 1.02]
Exclude Low & Lower-middle Income countries	0.11	0.86 [0.71, 1.04]
Age < 6 years	0.32	0.82 [0.55, 1.21]
Age 6-11 years	0.19	0.84 [0.65, 1.09]
Age > 11 years	0.71	0.93 [0.62, 1.38]
No significant comorbidities	0.035	0.82 [0.69, 0.99]
Baseline CRP < 118.0	0.37	0.86 [0.62, 1.20]
Baseline CRP 118.0 - 206.5	0.38	1.19 [0.80, 1.78]
Baseline CRP >= 206.5	0.74	0.95 [0.70, 1.29]
Post-hoc analyses		
Meet complete KD criteria	0.20	0.77 [0.51, 1.15]
Do not meet complete KD criteria	0.071	0.83 [0.67, 1.02]
Exclude changes in treatment group	0.14	0.79 [0.58, 1.08]

C

IVIG+G vs IVIG - first primary outcome

Outcome	P-value	OR [95% CI]
Primary analysis		
Inotropes/Ventilation/Death	1.00*	1.09 [0.75, 1.58]
Components of composite		
Death	0.49	0.63 [0.17, 2.35]
Inotropes from day-2	0.92	0.98 [0.67, 1.44]
Ventilation from day-2	0.28	1.32 [0.79, 2.23]
Sensitivity analyses		
Primary-therapy on days 0 & 1	0.25	1.39 [0.79, 2.42]
Propensity matched analysis	0.78	1.06 [0.71, 1.57]
Impute CRP (median by treatment group)	0.99	1.00 [0.70, 1.43]
Subgroup analyses		
Full WHO MIS-C criteria	0.99	1.00 [0.67, 1.47]
Full WHO MIS-C criteria + bacteraemia/TSS	0.99	1.00 [0.68, 1.48]
Missing up-to 1 WHO MIS-C criteria	0.74	1.06 [0.73, 1.54]
Exclude Low & Lower-middle Income countries	0.61	1.10 [0.76, 1.60]
Age < 6 years	0.30	0.69 [0.34, 1.40]
Age 6-11 years	0.73	1.10 [0.63, 1.91]
Age > 11 years	0.68	1.18 [0.54, 2.58]
No significant comorbidities	0.95	1.01 [0.69, 1.48]
Baseline CRP < 118.0	0.88	1.06 [0.48, 2.36]
Baseline CRP 118.0 - 206.5	0.98	0.99 [0.51, 1.92]
Baseline CRP >= 206.5	0.36	1.30 [0.74, 2.31]
Post-hoc subgroup analyses		
Meet complete KD criteria	0.28	1.42 [0.75, 2.71]
Do not meet complete KD criteria	0.89	1.04 [0.63, 1.71]
Exclude changes in treatment group	0.082	1.87 [0.92, 3.83]



D

IVIG+G vs IVIG - second primary outcome

Outcome	P-value	AHR [95% CI]
Primary analysis		
Time-to-Improvement	1.00*	1.04 [0.91, 1.20]
Baseline subgroups		
Subgroup: Requiring intensive support at baseline	0.34	0.87 [0.66, 1.16]
Subgroup: Requiring no intensive support at baseline	0.50	1.07 [0.87, 1.32]
Sensitivity analyses		
Primary-therapy on days 0 & 1	0.38	0.94 [0.82, 1.08]
Propensity matched analysis	0.60	1.03 [0.91, 1.18]
Two-point Time-to-Improvement	0.049	0.87 [0.75, 1.00]
Impute CRP (median by treatment group)	0.30	1.09 [0.93, 1.27]
Subgroup analyses		
Full WHO MIS-C criteria	0.24	1.09 [0.94, 1.27]
Full WHO MIS-C criteria + bacteraemia/TSS	0.30	1.08 [0.93, 1.26]
Missing up-to 1 WHO MIS-C criteria	0.51	1.05 [0.91, 1.21]
Exclude Low & Lower-middle Income countries	0.67	1.03 [0.89, 1.19]
Age < 6 years	0.11	1.24 [0.95, 1.62]
Age 6-11 years	0.77	1.03 [0.83, 1.28]
Age > 11 years	0.88	0.97 [0.68, 1.39]
No significant comorbidities	0.47	1.06 [0.91, 1.22]
Baseline CRP < 118.0	0.44	1.10 [0.86, 1.41]
Baseline CRP 118.0 - 206.5	0.19	1.29 [0.88, 1.90]
Baseline CRP >= 206.5	0.88	1.02 [0.79, 1.31]
Post-hoc analyses		
Meet complete KD criteria	0.29	1.13 [0.90, 1.43]
Do not meet complete KD criteria	0.86	0.98 [0.82, 1.18]
Exclude changes in treatment group	0.27	0.89 [0.72, 1.10]

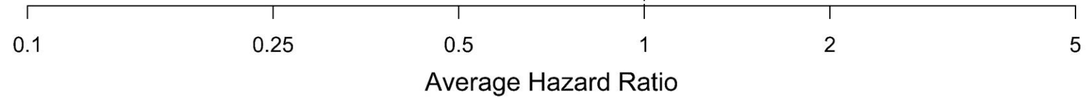


Figure 2A

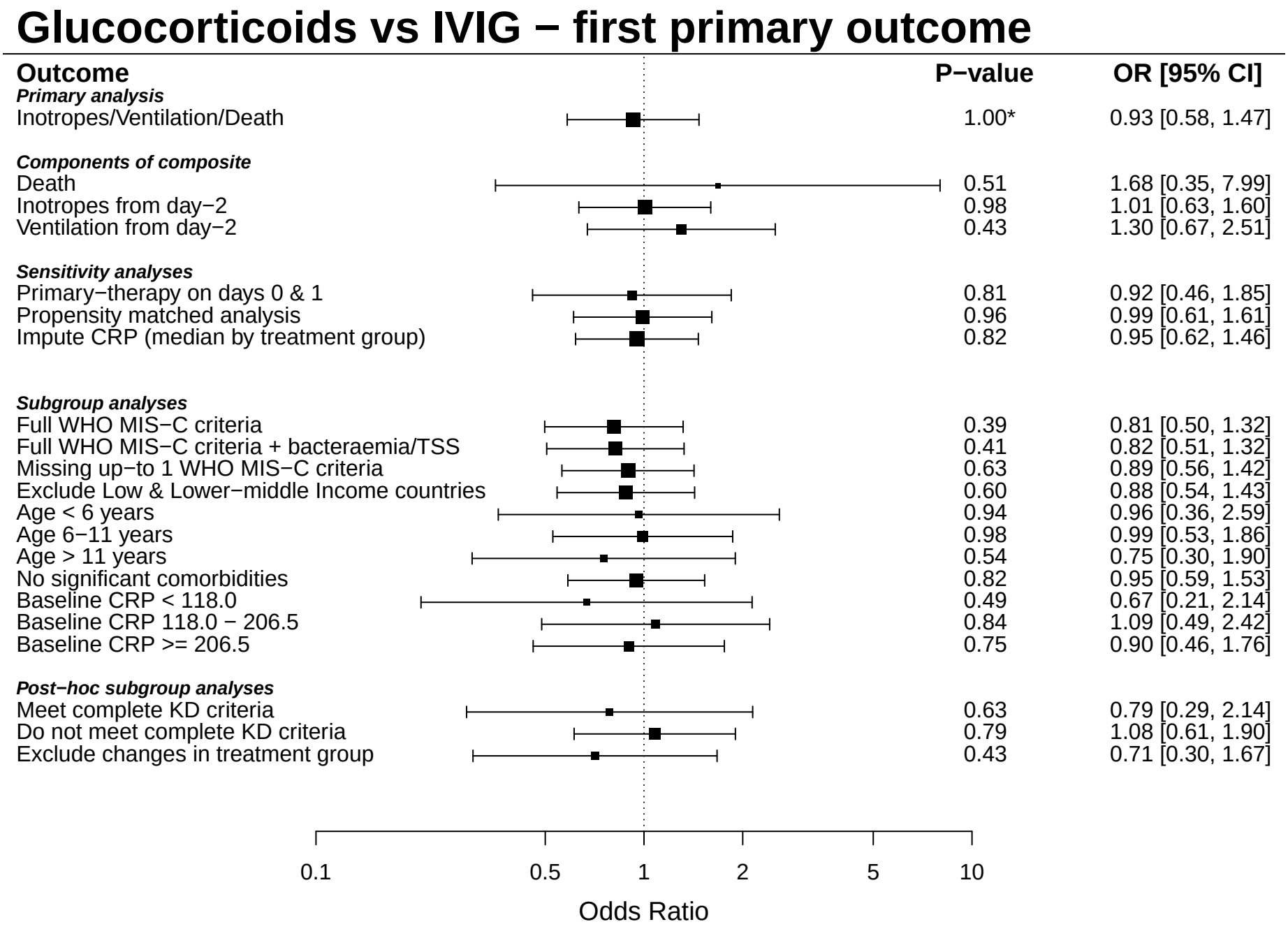

Glucocorticoids vs IVIG − first primary outcome
Outcome
P−value
OR [95% CI]
Primary analysis
Inotropes/Ventilation/Death
1.00*
0.93 [0.58, 1.47]
Components of composite
Death
0.51
1.68 [0.35, 7.99]
Inotropes from day−2
0.98
1.01 [0.63, 1.60]
Ventilation from day−2
0.43
1.30 [0.67, 2.51]
Sensitivity analyses
Primary−therapy on days 0 & 1
0.81
0.92 [0.46, 1.85]
Propensity matched analysis
0.96
0.99 [0.61, 1.61]
Impute CRP (median by treatment group)
0.82
0.95 [0.62, 1.46]
Subgroup analyses
Full WHO MIS−C criteria
0.39
0.81 [0.50, 1.32]
Full WHO MIS−C criteria + bacteraemia/TSS
0.41
0.82 [0.51, 1.32]
Missing up−to 1 WHO MIS−C criteria
0.63
0.89 [0.56, 1.42]
Exclude Low & Lower−middle Income countries
0.60
0.88 [0.54, 1.43]
Age < 6 years
0.94
0.96 [0.36, 2.59]
Age 6−11 years
0.98
0.99 [0.53, 1.86]
Age > 11 years
0.54
0.75 [0.30, 1.90]
No significant comorbidities
0.82
0.95 [0.59, 1.53]
Baseline CRP < 118.0
0.49
0.67 [0.21, 2.14]
Baseline CRP 118.0 − 206.5
0.84
1.09 [0.49, 2.42]
Baseline CRP >= 206.5
0.75
0.90 [0.46, 1.76]
Post−hoc subgroup analyses
Meet complete KD criteria
0.63
0.79 [0.29, 2.14]
Do not meet complete KD criteria
0.79
1.08 [0.61, 1.90]
Exclude changes in treatment group
0.43
0.71 [0.30, 1.67]
0.1
0.5
1
2
5
10
Odds Ratio

Figure 2B

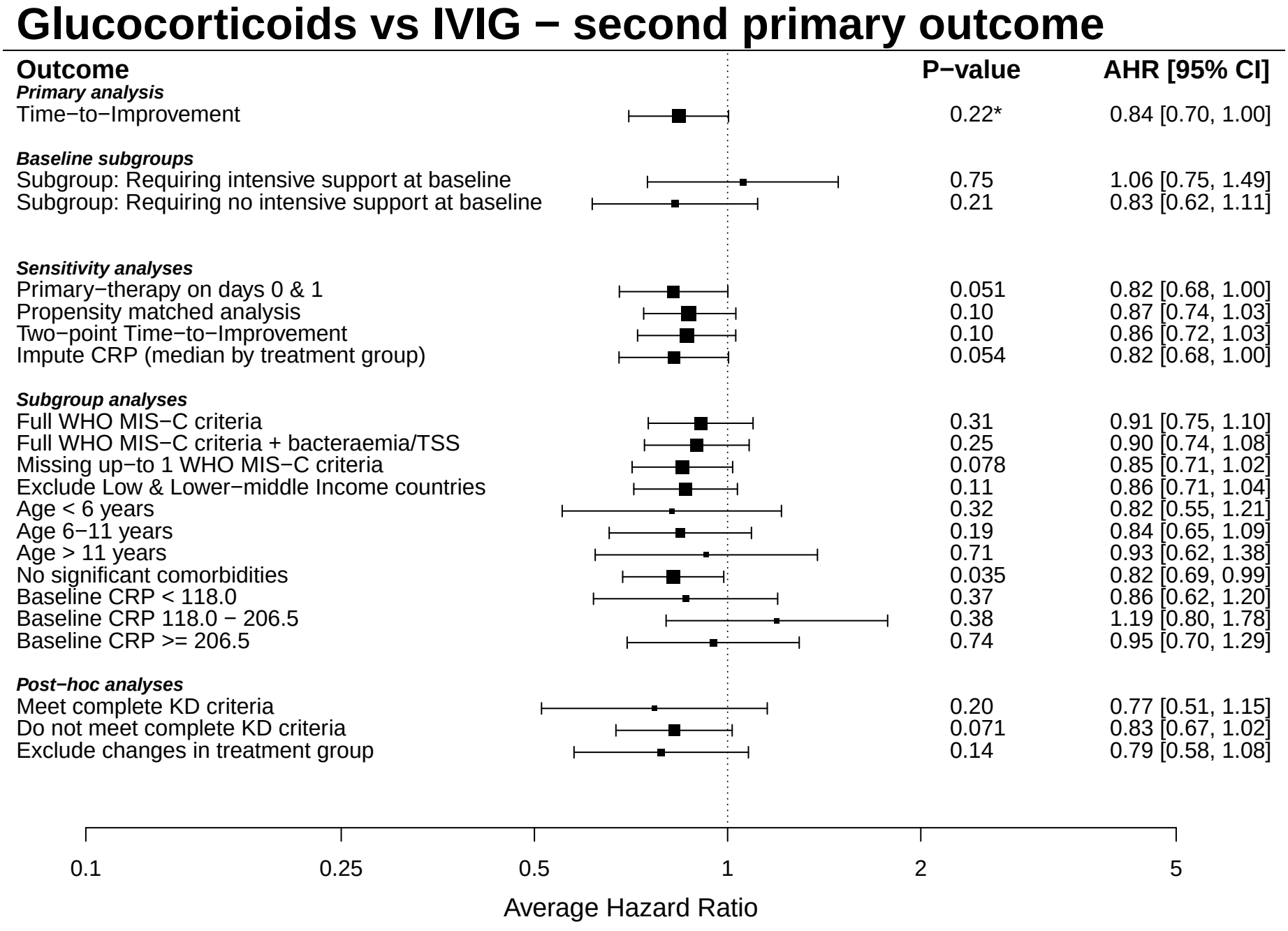

Glucocorticoids vs IVIG − second primary outcome
Outcome
P−value
AHR [95% CI]
Primary analysis
Time−to−Improvement
0.22*
0.84 [0.70, 1.00]
Baseline subgroups
Subgroup: Requiring intensive support at baseline
0.75
1.06 [0.75, 1.49]
Subgroup: Requiring no intensive support at baseline
0.21
0.83 [0.62, 1.11]
Sensitivity analyses
Primary−therapy on days 0 & 1
0.051
0.82 [0.68, 1.00]
Propensity matched analysis
0.10
0.87 [0.74, 1.03]
Two−point Time−to−Improvement
0.10
0.86 [0.72, 1.03]
Impute CRP (median by treatment group)
0.054
0.82 [0.68, 1.00]
Subgroup analyses
Full WHO MIS−C criteria
0.31
0.91 [0.75, 1.10]
Full WHO MIS−C criteria + bacteraemia/TSS
0.25
0.90 [0.74, 1.08]
Missing up−to 1 WHO MIS−C criteria
0.078
0.85 [0.71, 1.02]
Exclude Low & Lower−middle Income countries
0.11
0.86 [0.71, 1.04]
Age < 6 years
0.32
0.82 [0.55, 1.21]
Age 6−11 years
0.19
0.84 [0.65, 1.09]
Age > 11 years
0.71
0.93 [0.62, 1.38]
No significant comorbidities
0.035
0.82 [0.69, 0.99]
Baseline CRP < 118.0
0.37
0.86 [0.62, 1.20]
Baseline CRP 118.0 − 206.5
0.38
1.19 [0.80, 1.78]
Baseline CRP >= 206.5
0.74
0.95 [0.70, 1.29]
Post−hoc analyses
Meet complete KD criteria
0.20
0.77 [0.51, 1.15]
Do not meet complete KD criteria
0.071
0.83 [0.67, 1.02]
Exclude changes in treatment group
0.14
0.79 [0.58, 1.08]
0.1
0.25
0.5
1
2
5
Average Hazard Ratio

Figure 2C

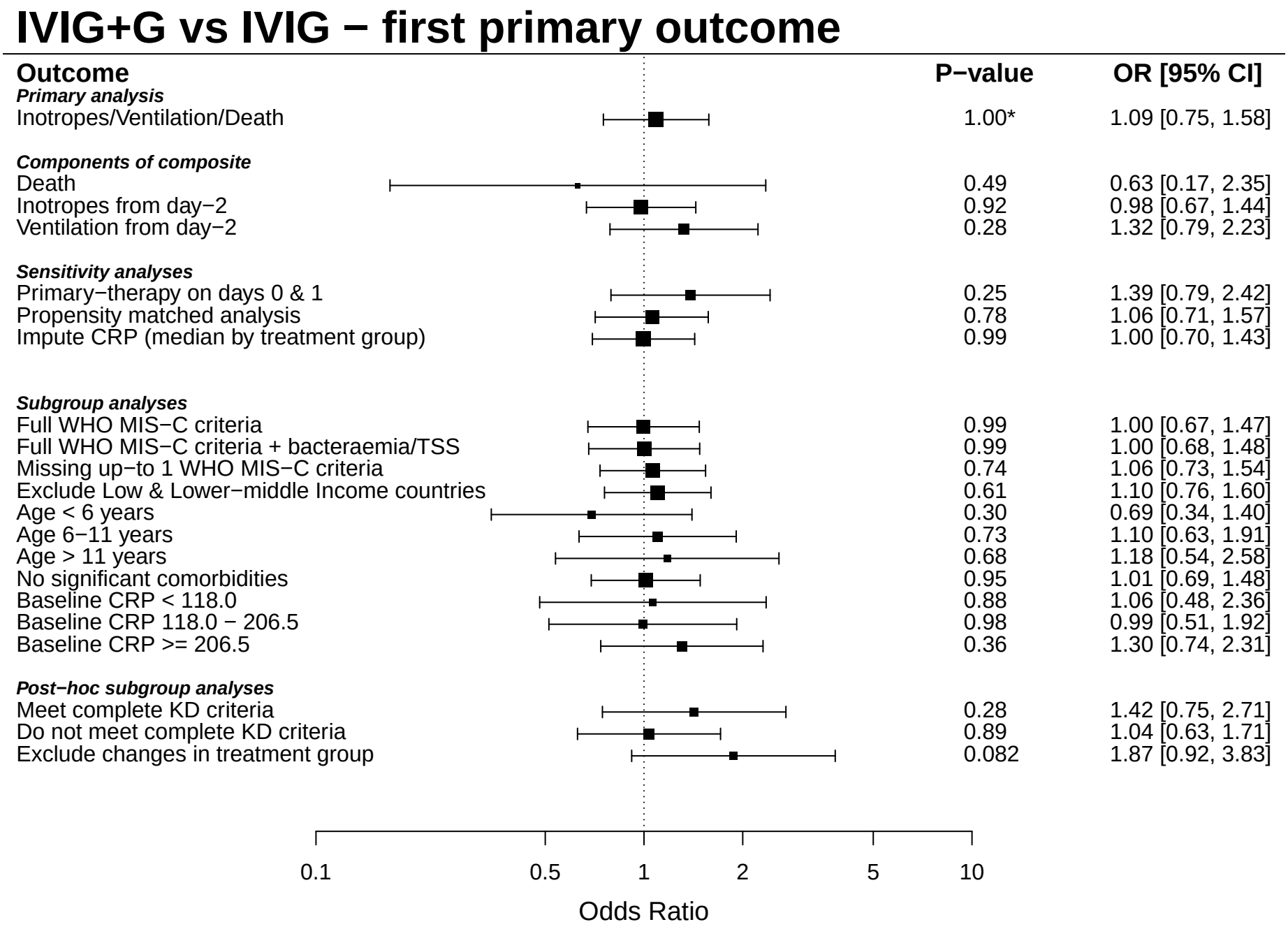

IVIG+G vs IVIG − first primary outcome
Outcome
P−value
OR [95% CI]
Primary analysis
Inotropes/Ventilation/Death
1.00*
1.09 [0.75, 1.58]
Components of composite
Death
0.49
0.63 [0.17, 2.35]
Inotropes from day−2
0.92
0.98 [0.67, 1.44]
Ventilation from day−2
0.28
1.32 [0.79, 2.23]
Sensitivity analyses
Primary−therapy on days 0 & 1
0.25
1.39 [0.79, 2.42]
Propensity matched analysis
0.78
1.06 [0.71, 1.57]
Impute CRP (median by treatment group)
0.99
1.00 [0.70, 1.43]
Subgroup analyses
Full WHO MIS−C criteria
0.99
1.00 [0.67, 1.47]
Full WHO MIS−C criteria + bacteraemia/TSS
0.99
1.00 [0.68, 1.48]
Missing up−to 1 WHO MIS−C criteria
0.74
1.06 [0.73, 1.54]
Exclude Low & Lower−middle Income countries
0.61
1.10 [0.76, 1.60]
Age < 6 years
0.30
0.69 [0.34, 1.40]
Age 6−11 years
0.73
1.10 [0.63, 1.91]
Age > 11 years
0.68
1.18 [0.54, 2.58]
No significant comorbidities
0.95
1.01 [0.69, 1.48]
Baseline CRP < 118.0
0.88
1.06 [0.48, 2.36]
Baseline CRP 118.0 − 206.5
0.98
0.99 [0.51, 1.92]
Baseline CRP >= 206.5
0.36
1.30 [0.74, 2.31]
Post−hoc subgroup analyses
Meet complete KD criteria
0.28
1.42 [0.75, 2.71]
Do not meet complete KD criteria
0.89
1.04 [0.63, 1.71]
Exclude changes in treatment group
0.082
1.87 [0.92, 3.83]
0.1
0.5
1
2
5
10
Odds Ratio

Figure 2D

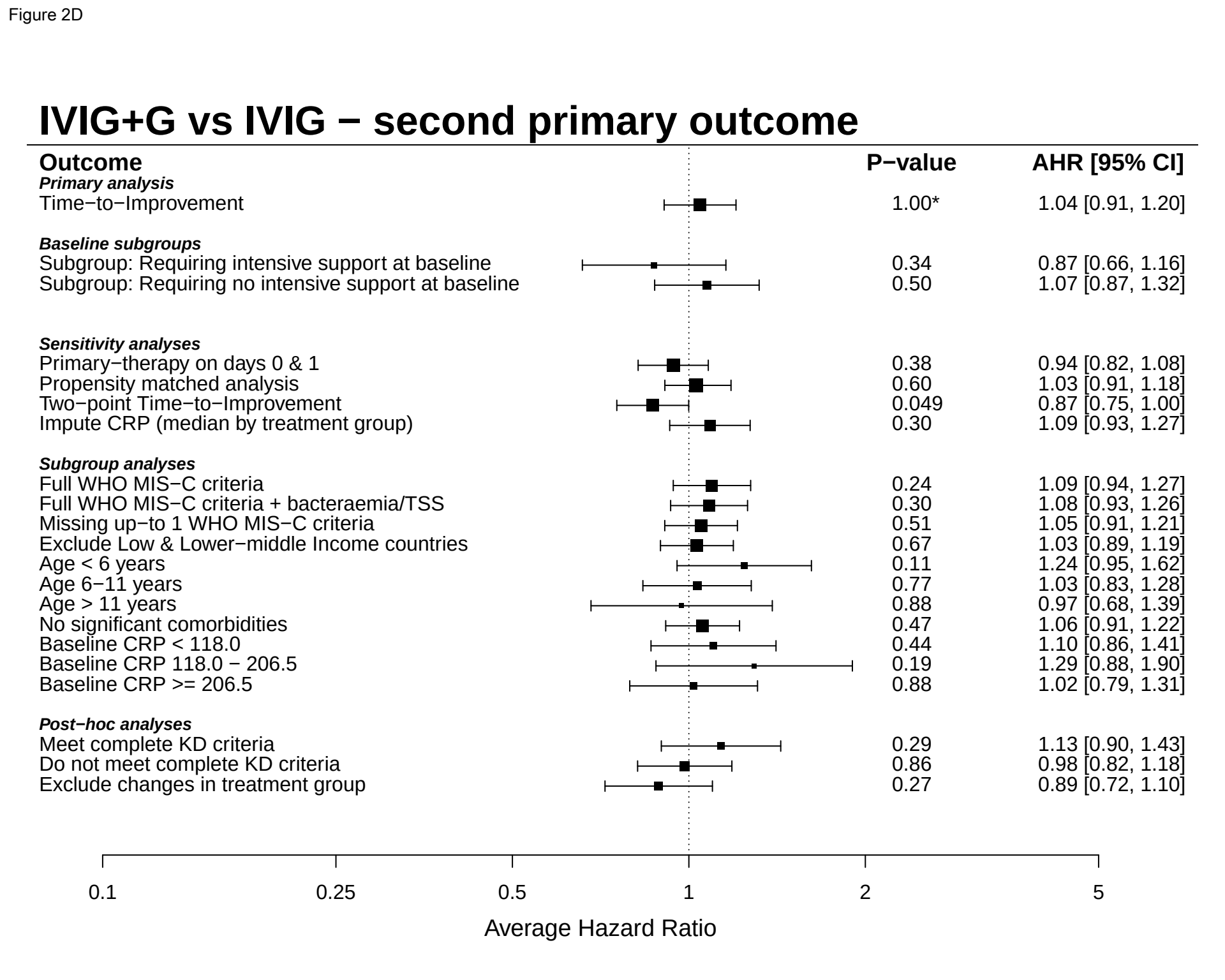
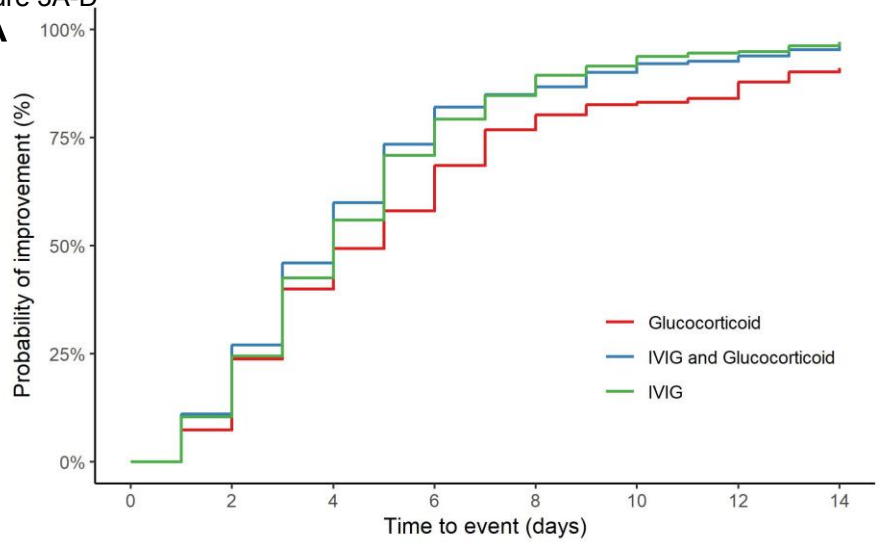

IVIG+G vs IVIG − second primary outcome
Outcome
P−value
AHR [95% CI]
Primary analysis
Time−to−Improvement
1.00*
1.04 [0.91, 1.20]
Baseline subgroups
Subgroup: Requiring intensive support at baseline
0.34
0.87 [0.66, 1.16]
Subgroup: Requiring no intensive support at baseline
0.50
1.07 [0.87, 1.32]
Sensitivity analyses
Primary−therapy on days 0 & 1
0.38
0.94 [0.82, 1.08]
Propensity matched analysis
0.60
1.03 [0.91, 1.18]
Two−point Time−to−Improvement
0.049
0.87 [0.75, 1.00]
Impute CRP (median by treatment group)
0.30
1.09 [0.93, 1.27]
Subgroup analyses
Full WHO MIS−C criteria
0.24
1.09 [0.94, 1.27]
Full WHO MIS−C criteria + bacteraemia/TSS
0.30
1.08 [0.93, 1.26]
Missing up−to 1 WHO MIS−C criteria
0.51
1.05 [0.91, 1.21]
Exclude Low & Lower−middle Income countries
0.67
1.03 [0.89, 1.19]
Age < 6 years
0.11
1.24 [0.95, 1.62]
Age 6−11 years
0.77
1.03 [0.83, 1.28]
Age > 11 years
0.88
0.97 [0.68, 1.39]
No significant comorbidities
0.47
1.06 [0.91, 1.22]
Baseline CRP < 118.0
0.44
1.10 [0.86, 1.41]
Baseline CRP 118.0 − 206.5
0.19
1.29 [0.88, 1.90]
Baseline CRP >= 206.5
0.88
1.02 [0.79, 1.31]
Post−hoc analyses
Meet complete KD criteria
0.29
1.13 [0.90, 1.43]
Do not meet complete KD criteria
0.86
0.98 [0.82, 1.18]
Exclude changes in treatment group
0.27
0.89 [0.72, 1.10]
0.1
0.25
0.5
1
2
5
Average Hazard Ratio

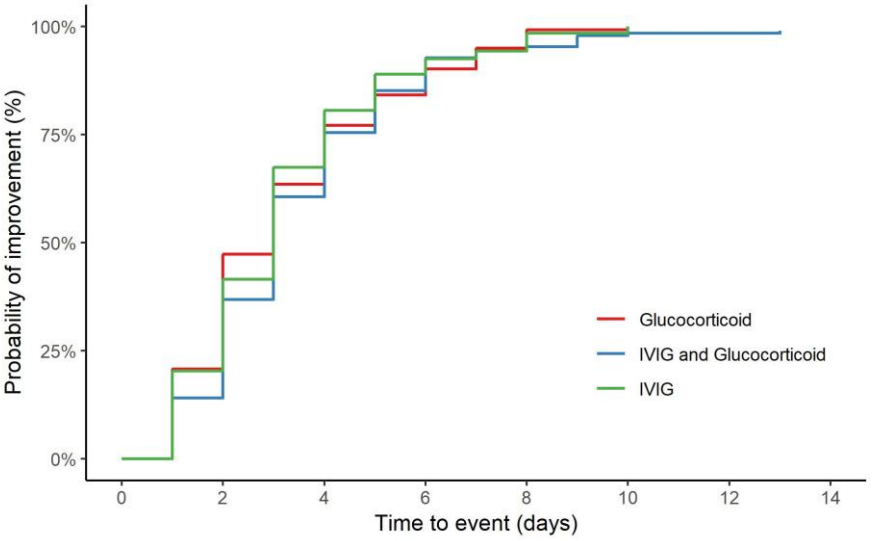
Figure 3A-D

A



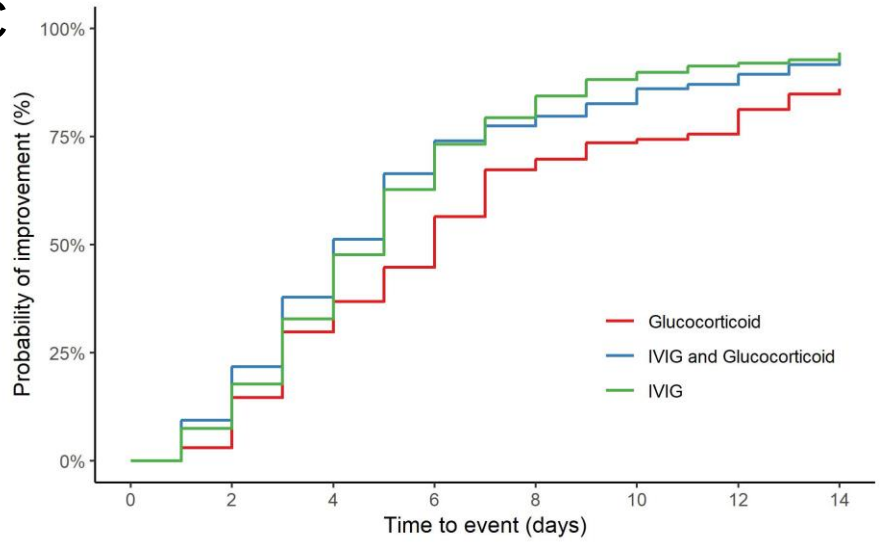
At risk	G	373	373	293	207	134	104	82	62	45	38	33	31	28	21	18
	IVIG	565	565	438	358	242	146	82	50	32	23	18	13	11	9	7
	IVIG+G	608	608	469	344	219	143	85	53	42	35	25	19	18	15	11
Censored	G	0	52	29	25	9	8	3	1	2	1	1	0	0	0	0
	IVIG	0	77	20	35	38	17	8	6	2	0	1	0	1	0	0
	IVIG+G	0	68	36	35	17	9	2	3	1	0	1	0	0	0	0

B



At risk	G	105	105	86	47	27	17	10	7	3	1	1	0	0	0	0
	IVIG	92	92	74	51	29	19	10	7	4	2	2	0	0	0	0
	IVIG+G	216	216	182	133	85	55	34	16	13	10	4	3	3	3	2
Censored	G	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0
	IVIG	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0
	IVIG+G	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0

C



At risk	G	268	268	207	160	107	87	72	55	42	37	32	31	28	21	18
	IVIG	472	472	363	306	212	126	71	42	27	20	15	12	10	8	7
	IVIG+G	390	390	287	211	134	88	51	37	29	25	21	16	15	12	9
Censored	G	0	52	29	25	9	7	3	1	2	1	0	0	0	0	0
	IVIG	0	75	19	35	38	17	8	6	2	0	1	0	1	0	0
	IVIG+G	0	66	36	35	17	9	2	3	1	0	1	0	0	0	0

D

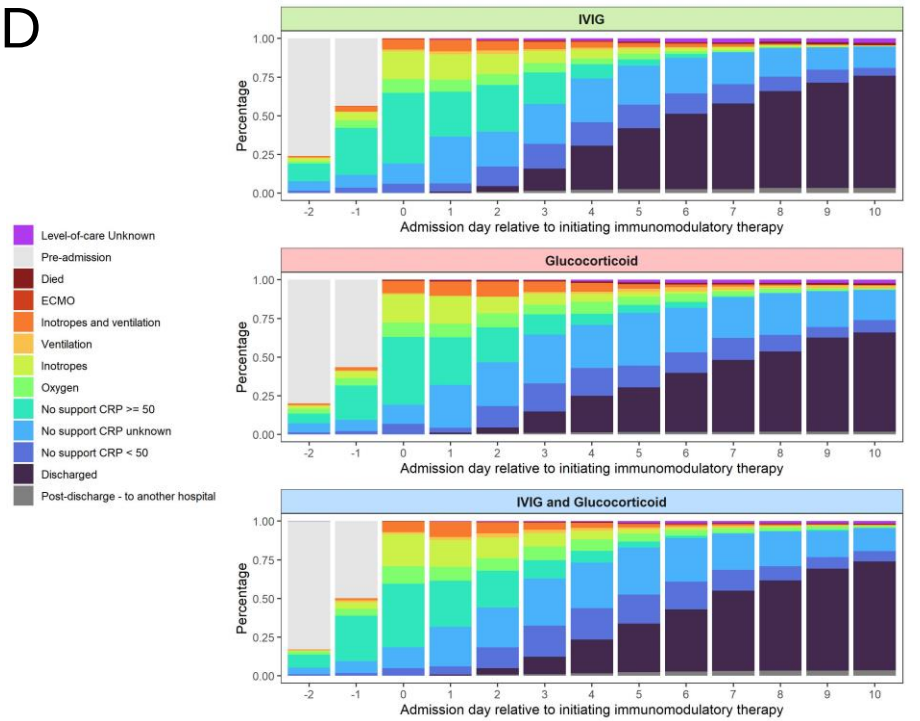


Figure 3A

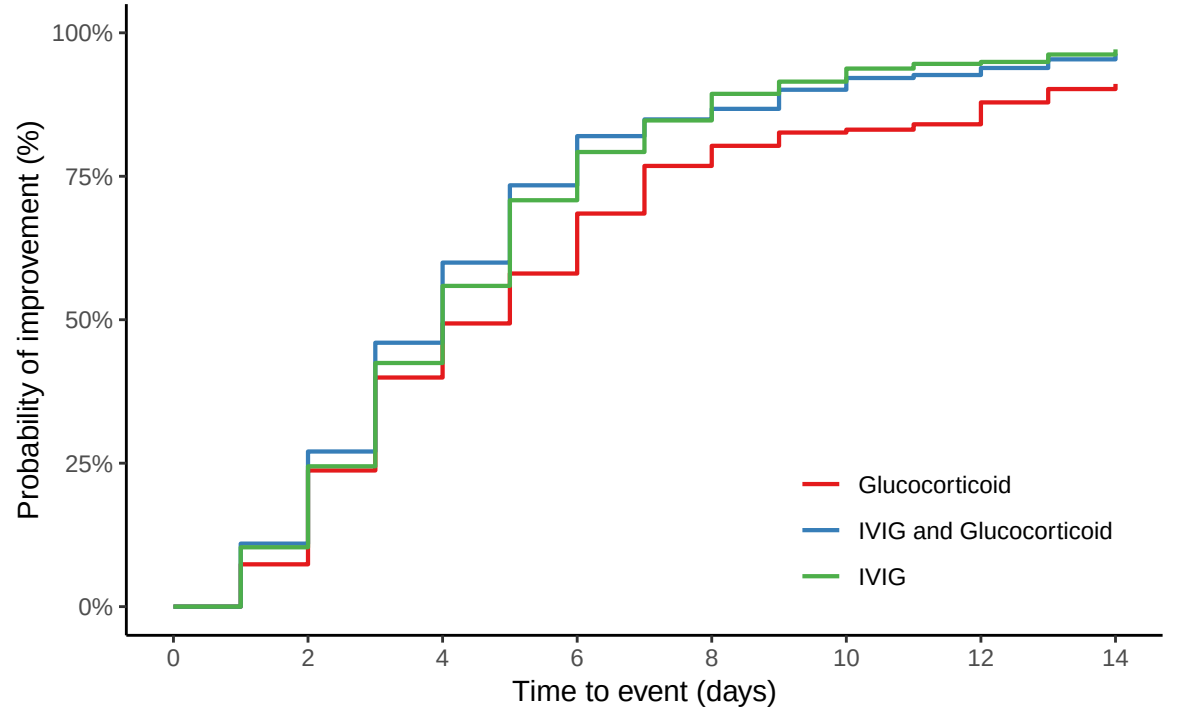

100%
75%
50%
25%
0%
Probability of improvement (%)
0
2
4
6
8
10
12
14
Time to event (days)
Glucocorticoid
IVIG and Glucocorticoid
IVIG

Figure 3B

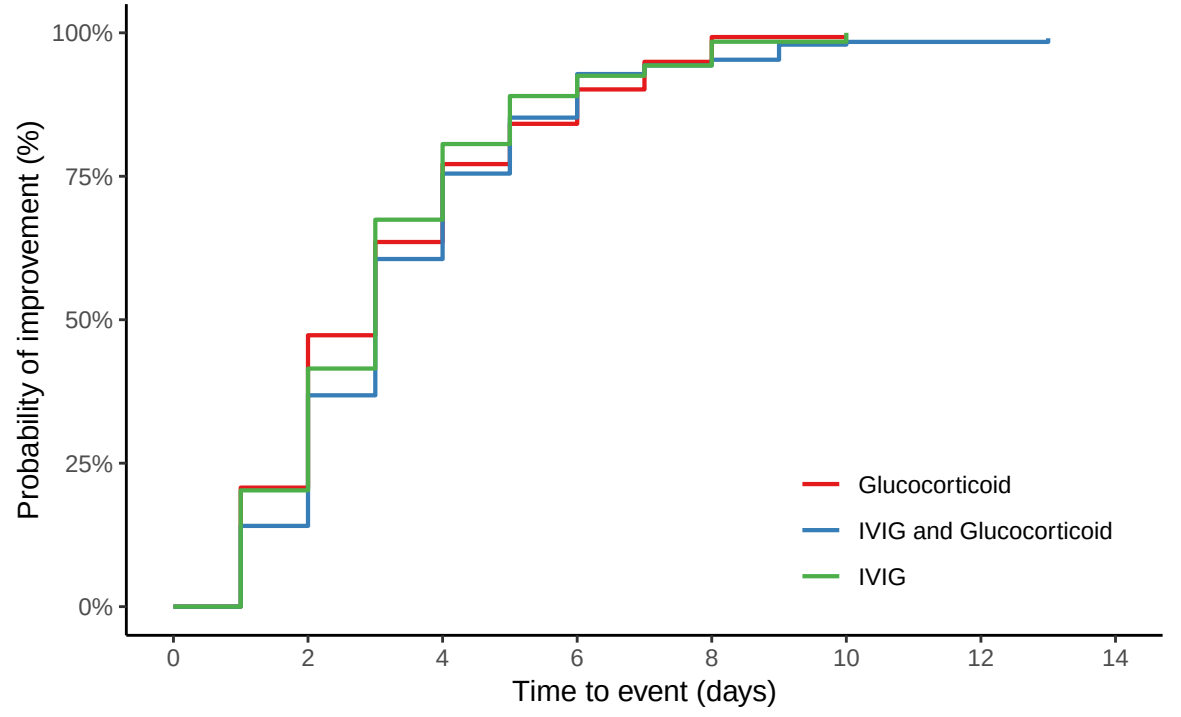

100%
75%
50%
25%
0%
Probability of improvement (%)
0
2
4
6
8
10
12
14
Time to event (days)
Glucocorticoid
IVIG and Glucocorticoid
IVIG

Figure 3C

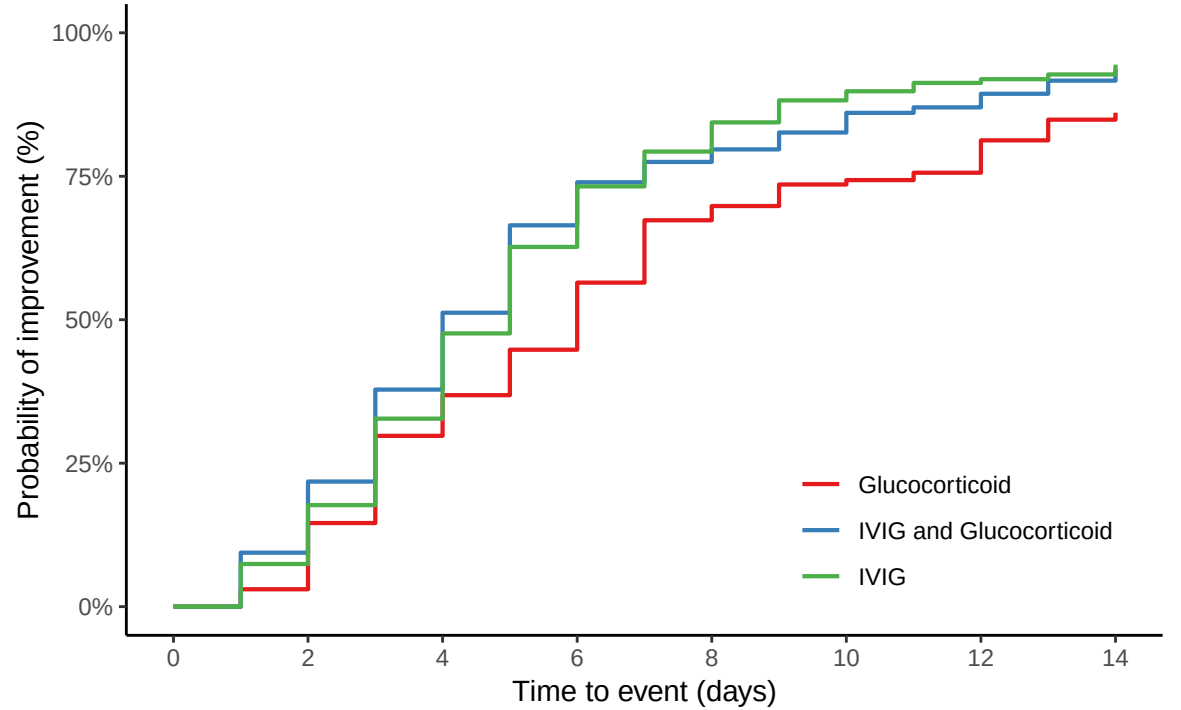

100%
75%
50%
25%
0%
Probability of improvement (%)
0
2
4
6
8
10
12
14
Time to event (days)
Glucocorticoid
IVIG and Glucocorticoid
IVIG

Figure 3D

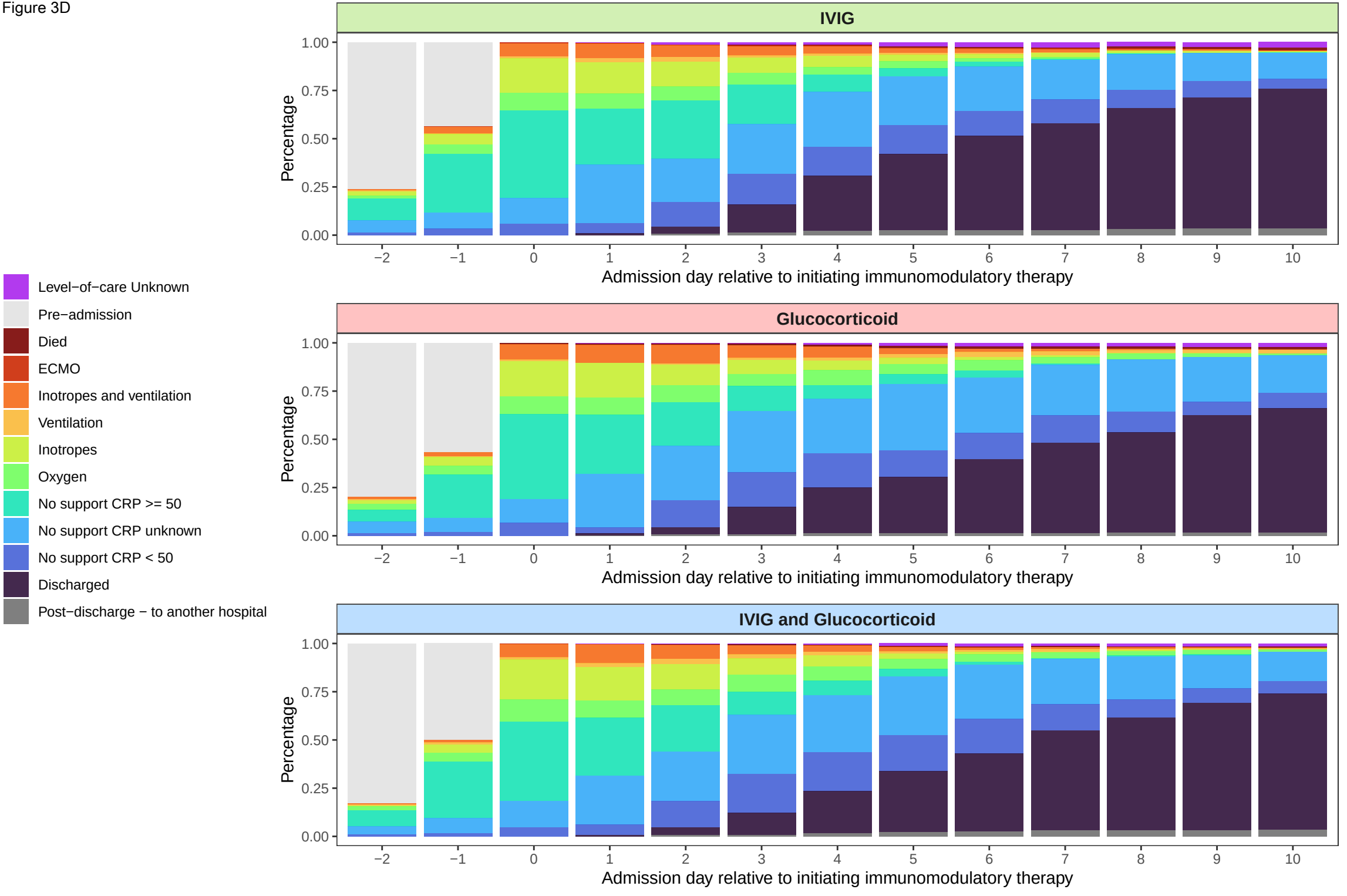

IVIG
Glucocorticoid
IVIG and Glucocorticoid
Percentage
1.00
0.75
0.50
0.25
0.00
−2
−1
0
1
2
3
4
5
6
7
8
9
10
Admission day relative to initiating immunomodulatory therapy
Level−of−care Unknown
Pre−admission
Died
ECMO
Inotropes and ventilation
Ventilation
Inotropes
Oxygen
No support CRP >= 50
No support CRP unknown
No support CRP < 50
Discharged
Post−discharge − to another hospital

Figure 4A

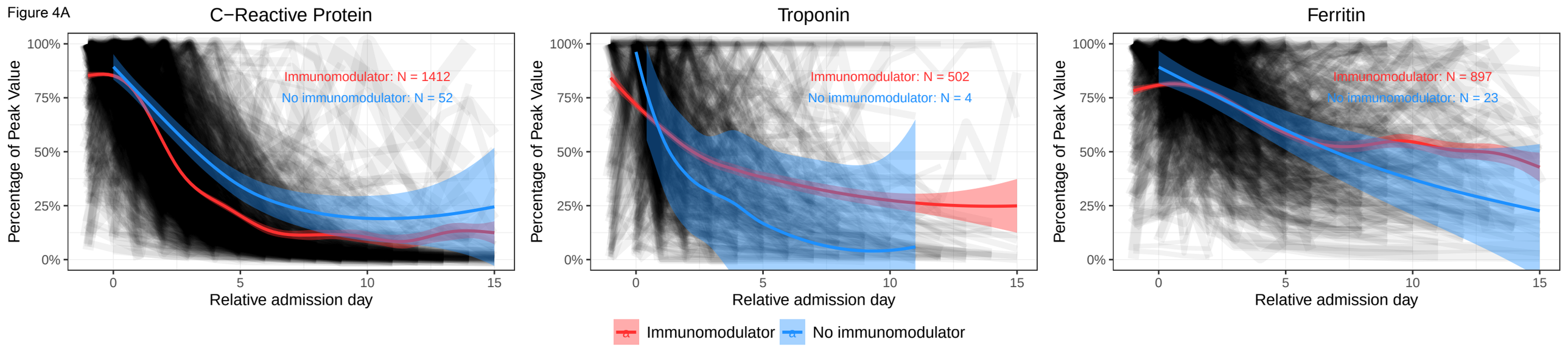

C−Reactive Protein
Troponin
Ferritin
Percentage of Peak Value
100%
75%
50%
25%
0%
0
5
10
15
Relative admission day
Immunomodulator: N = 1412
No immunomodulator: N = 52
Immunomodulator: N = 502
No immunomodulator: N = 4
Immunomodulator: N = 897
No immunomodulator: N = 23
Immunomodulator
No immunomodulator

Figure 4B

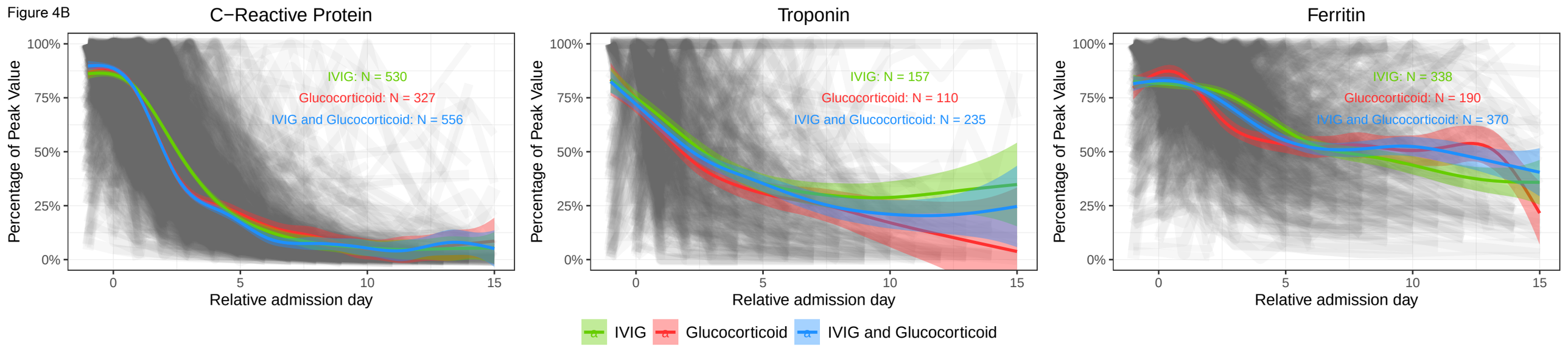

C−Reactive Protein
Troponin
Ferritin
Percentage of Peak Value
100%
75%
50%
25%
0%
0
5
10
15
Relative admission day
IVIG: N = 530
Glucocorticoid: N = 327
IVIG and Glucocorticoid: N = 556
IVIG: N = 157
Glucocorticoid: N = 110
IVIG and Glucocorticoid: N = 235
IVIG: N = 338
Glucocorticoid: N = 190
IVIG and Glucocorticoid: N = 370
IVIG
Glucocorticoid
IVIG and Glucocorticoid

Figure 4C

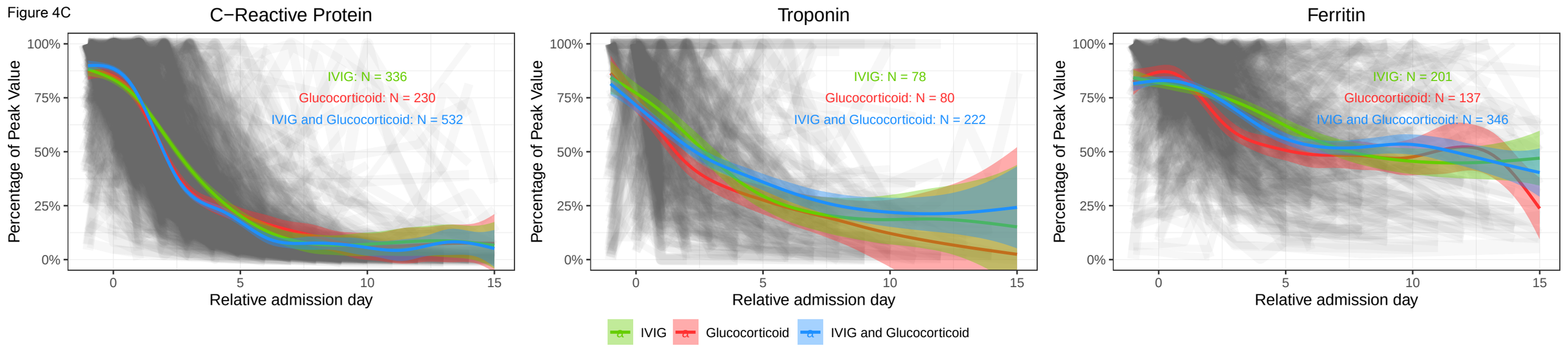

C−Reactive Protein
Troponin
Ferritin
Percentage of Peak Value
100%
75%
50%
25%
0%
0
5
10
15
Relative admission day
IVIG: N = 336
Glucocorticoid: N = 230
IVIG and Glucocorticoid: N = 532
IVIG: N = 78
Glucocorticoid: N = 80
IVIG and Glucocorticoid: N = 222
IVIG: N = 201
Glucocorticoid: N = 137
IVIG and Glucocorticoid: N = 346
IVIG
Glucocorticoid
IVIG and Glucocorticoid

Table 1 - word document format

	Everyone (N=2009)	IVIG (N=680)	Glucocorticoids (N=487)	IVIG and Glucocorticoids (N=698)	Other (N=59)	No treatment (N=85)
+Age	8·0 [4·2 - 11]	6·8 [3·6 - 10]	8·8 [5·1 - 12]	8·4 [4·5 - 11]	11 [6·1 - 13]	7·3 [3·3 - 12]
^Proportion male	1191 (59·3%)	416 (61·2%)	288 (59·1%)	410 (58·7%)	44 (74·6%)	33 (38·8%)
^Proportion female	818 (40·7%)	264 (38·8%)	199 (40·9%)	288 (41·3%)	15 (25·4%)	52 (61·2%)
^Weight (age-adjusted z score ≥ 2)	299 (14·9%)	91 (13·4%)	70 (14·4%)	120 (17·2%)	10 (16·9%)	8 (9·41%)
^Ethnicity						
White	825 (41·1%)	290 (42·6%)	210 (43·1%)	272 (39·0%)	27 (45·8%)	26 (30·6%)
Latino	518 (25·8%)	161 (23·7%)	94 (19·3%)	222 (31·8%)	9 (15·3%)	32 (37·6%)
Black	212 (10·6%)	81 (11·9%)	34 (6·98%)	75 (10·7%)	13 (22·0%)	9 (10·6%)
Asian	131 (6·52%)	55 (8·09%)	36 (7·39%)	30 (4·30%)	4 (6·78%)	6 (7·06%)
Other or not known	323 (16·1%)	93 (13·7%)	113 (23·2%)	99 (14·2%)	6 (10·2%)	12 (14·1%)
^Significant comorbidity	108 (5·38%)	30 (4·41%)	32 (6·57%)	33 (4·73%)	4 (6·78%)	9 (10·6%)
^SARS-CoV-2 PCR positive	415 (20·8%)	131 (19·4%)	97 (20·0%)	148 (21·4%)	13 (22·0%)	26 (31·7%)
^SARS-CoV-2 Ab positive	1321 (66·5%)	412 (61·2%)	344 (71·4%)	492 (71·6%)	43 (72·9%)	30 (35·3%)
^Baseline requirement for ventilation/inotropes/ECMO	535 (26·6%)	117 (17·2%)	127 (26·1%)	252 (36·1%)	29 (49·2%)	10 (11·8%)
^Clinical features during admission						
Fever	1863 (92·7%)	653 (96·0%)	439 (90·1%)	649 (93·0%)	52 (88·1%)	70 (82·4%)
Sore throat	464 (25·5%)	159 (26·5%)	104 (22·9%)	175 (27·0%)	11 (21·6%)	15 (21·1%)
Cough	404 (21·1%)	125 (19·4%)	120 (25·3%)	131 (19·6%)	16 (30·8%)	12 (16·0%)
Respiratory distress	258 (13·3%)	70 (10·9%)	57 (11·9%)	112 (16·4%)	13 (23·6%)	6 (7·59%)
Abdominal pain	1211 (63·2%)	408 (63·9%)	289 (62·3%)	438 (64·8%)	37 (63·8%)	39 (48·1%)
Diarrhea	882 (44·8%)	290 (43·9%)	195 (40·6%)	340 (49·4%)	23 (39·7%)	34 (41·5%)
Vomiting	1057 (54·0%)	330 (50·6%)	251 (52·3%)	408 (59·2%)	34 (60·7%)	34 (42·5%)
Headache	592 (32·8%)	199 (34·1%)	155 (35·0%)	203 (31·4%)	21 (38·9%)	14 (18·4%)
Irritability	355 (18·8%)	127 (20·2%)	69 (14·9%)	135 (20·2%)	10 (18·5%)	14 (18·4%)
Lethargy	655 (34·5%)	211 (33·3%)	186 (40·1%)	215 (32·1%)	23 (41·8%)	20 (26·7%)
^Proportion meeting Kawasaki Disease criteria	629 (31·3%)	265 (39·0%)	119 (24·4%)	225 (32·2%)	12 (20·3%)	8 (9·41%)
+Bloods on admission						
Lymphocytes (10^9/L)	1·2 [0·70 - 2·0]	1·3 [0·76 - 2·2]	1·2 [0·70 - 1·8]	1·1 [0·66 - 1·9]	0·86 [0·52 - 1·6]	1·8 [1·1 - 2·9]
Troponin (ng/L)	25 [6·1 - 80]	13 [5·0 - 43]	31 [9·8 - 100]	40 [10 - 110]	48 [10 - 270]	10 [2·0 - 38]
CRP (mg/L)	150 [85 - 220]	150 [85 - 210]	160 [75 - 220]	160 [90 - 230]	180 [97 - 280]	85 [23 - 180]
Ferritin (ug/L)	440 [230 - 860]	370 [210 - 650]	480 [260 - 970]	520 [260 - 960]	560 [340 - 1700]	280 [140 - 460]
Albumin (g/L)	32 [28 - 37]	34 [28 - 39]	32 [27 - 36]	32 [27 - 36]	32 [27 - 36]	35 [30 - 41]

Members list for Full Consortium, as requested by editors

Table of extra results - for reviewer 2 comment 12