Coagulation phenotypes in sepsis and effects of recombinant human thrombomodulin: an analysis of three multicentre observational studies

Daisuke Kudo1*, Tadahiro Goto2, Ryo Uchimido3, Mineji Hayakawa4, Kazuma Yamakawa5, Toshikazu Abe6,7, Atsushi Shiraishi8 and Shigeki Kushimoto1

Abstract

Background: A recent randomised trial showed that recombinant thrombomodulin did not benefit patients who had sepsis with coagulopathy and organ dysfunction. Several recent studies suggested presence of clinical phenotypes in patients with sepsis and heterogenous treatment effects across different sepsis phenotypes. We examined the latent phenotypes of sepsis with coagulopathy and the associations between thrombomodulin treatment and the 28-day and in-hospital mortality for each phenotype.

Methods: This was a secondary analysis of multicentre registries containing data on patients (aged ≥ 16 years) who were admitted to intensive care units for severe sepsis or septic shock in Japan. Three multicentre registries were divided into derivation (two registries) and validation (one registry) cohorts. Phenotypes were derived using k-means with coagulation markers, platelet counts, prothrombin time/international normalised ratios, fibrinogen, fibrinogen/fibrin-degradation-products (FDP), D-dimer, and antithrombin activities. Associations between thrombomodulin treatment and survival outcomes (28-day and in-hospital mortality) were assessed in the derived clusters using a generalised estimating equation.

Results: Four sepsis phenotypes were derived from 3694 patients in the derivation cohort. Cluster dA (n = 323) had severe coagulopathy with high FDP and D-dimer levels, severe organ dysfunction, and high mortality. Cluster dB had severe disease with moderate coagulopathy. Clusters dC and dD had moderate and mild disease with and without coagulopathy, respectively. Thrombomodulin was associated with a lower 28-day (adjusted risk difference [RD]: −17.8% [95% CI −28.7 to −6.9%]) and in-hospital (adjusted RD: −17.7% [95% CI −27.6 to −7.8%]) mortality only in cluster dA. Sepsis phenotypes were similar in the validation cohort, and thrombomodulin treatment was also associated with lower 28-day (RD: −24.9% [95% CI −49.1 to −0.7%]) and in-hospital mortality (RD: −30.9% [95% CI −55.3 to −6.6%]).

*Correspondence: kudodaisuke@med.tohoku.ac.jp
1 Division of Emergency and Critical Care Medicine, Tohoku University Graduate School of Medicine, 1-1 Seiryo-machi, Aoba-ku, Sendai 980-8574, Japan

© The Author(s) 2021. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
Background

The global rate of sepsis-related mortality remains high and the annual age-standardised mortality owing to sepsis is 148.1 deaths per 100,000 of the global population [1]. Recombinant human thrombomodulin (rhTM) has anti-inflammatory and anticoagulation activities [2], and it has been suggested as an adjunct therapy for patients with sepsis, particularly those with sepsis-induced coagulopathy [3]. Nevertheless, a recent phase III randomised controlled trial revealed no beneficial effect of rhTM in patients with sepsis-induced coagulopathy [4]. This result can be explained by the heterogeneity of patients with sepsis and inappropriate criteria of coagulopathy [5] using the prothrombin time/international normalised ratio (PT-INR) and a platelet count based on subgroup analysis of an international phase II trial of rhTM [6].

Sepsis is a highly heterogeneous syndrome with variable aetiology and pathophysiology [7]. Thus, a specific therapy may benefit some, but not all, patients with sepsis. Several recent studies have classified sepsis into several phenotypes with distinct characteristics using cluster analysis [8–10], an unsupervised machine learning method that can identify relatively homogeneous groups in a heterogeneous population [11]. Furthermore, these studies indicated that specific therapies conferred benefits only in patients with specific phenotypes of sepsis [8–10].

Identifying the target patients receptive to rhTM treatment through grouping based on biomarker cut-offs is challenging. To address this issue, we examined latent sepsis phenotypes in terms of coagulopathy and identified which phenotypes would benefit from rhTM using machine learning approaches.

Methods

Study design and settings

Details of the methods and analytical processes in the present study are provided in the Supplemental Digital Content. This was a secondary analysis of the following multicentre registries: the Japan Septic Disseminated Intravascular Coagulation (JSEPTIC-DIC) study (UMIN-CTR ID: UMIN000012543) [12], Tohoku Sepsis Registry (UMIN000010297) [13], and Focused Outcomes Research in Emergency Care for Acute Respiratory Distress Syndrome, Sepsis, and Trauma (FORECAST) sepsis study (UMIN000019742) [14]. All three registries include information on consecutive patients admitted to ICUs for severe sepsis or septic shock [15, 16].

Briefly, the JSEPTIC-DIC study retrospectively reviewed data derived from 3195 consecutive patients with severe sepsis or septic shock, aged ≥16 years, admitted to 42 ICUs at 40 institutions in Japan, between January 2011 and December 2013 [12]. The Tohoku Sepsis Registry prospectively registered 616 consecutive patients who were admitted to ICUs with severe sepsis, or those who developed severe sepsis after admission to the ICUs or general wards at 10 institutions (three university hospitals and seven community hospitals) in the Tohoku District, Northern Japan, between January 2015 and December 2015 [13]. The multicentre prospective FORECAST sepsis study included 1184 consecutive patients aged ≥16 years, who were admitted to 59 ICUs in Japan with severe sepsis according to the sepsis-2 criteria [15], between January 2016 and March 2017 [14]. These studies were approved and the need for informed consent was waived by the institutional review boards at the participating hospitals.

Study population

We included all patients (aged ≥16 years), who were admitted to the ICUs with severe sepsis or septic shock as defined in the three registries, according to the International Sepsis Definitions Conference criteria [15, 16].

Phenotyping variables

We measured the following coagulation markers upon admission to the ICU for phenotyping: platelet counts, PT-INR, fibrinogen, fibrinogen/fibrin degradation products (FDP), D-dimer, and antithrombin activities.

Exposure

Patients were exposed to rhTM.

Outcomes

The main outcomes were 28-day and in-hospital mortality in the validation cohort. The secondary outcomes were ICU-free days, ventilator-free days, and the type of discharge in the validation cohort.

Conclusions: We identified four coagulation marker-based sepsis phenotypes. The treatment effects of thrombomodulin varied across sepsis phenotypes. This finding will facilitate future trials of thrombomodulin, in which a sepsis phenotype with high FDP and D-dimer can be targeted.

Keywords: Anticoagulants, Disseminated intravascular coagulation, Machine learning, Phenotype, Precision medicine, Thrombomodulin
Definitions
Ventilator-free days were defined as the number of days on which a patient did not require mechanical ventilation during the initial 28 days following enrolment. The number of ventilator-free days of patients, who died within day 28, was assigned as 0. ICU-free days were calculated similarly.

Statistical analysis
Analytical cohorts
We derived sepsis phenotypes from the JSEPTIC-DIC study (n = 3195) and Tohoku Sepsis Registry (n = 499) and validated the phenotypes using the FORECAST sepsis study (n = 1184).

Cluster derivation
We initially assessed the distribution and missingness in phenotyping variables. Non-normal data were log-transformed and scaled. Patients without 28-day mortality information were excluded. Missing data were imputed using the random forest method for each study cohort with the missForest package [17]. Random forest imputation is a nonparametric algorithm that accommodates nonlinearities and interactions and does not require the specification of a specific parametric model [18]. This approach generated single-point estimates by random draws from independent normal distributions centred on conditional means predicted by random forest. Random forest applies bootstrap aggregation of multiple regression trees to reduce the risk of overfitting and combines estimates from many trees [17]. Missingness was imputed using patient characteristics, laboratory data, outcomes, and other covariates, including in-hospital management (Additional file 1: Supplemental documents).

We applied k-means with Euclid distance, which is a basic and widely used machine learning-based clustering approach, to derive sepsis phenotypes [9, 11]. We then determined the optimal number of clusters using a consensus clustering approach that assessed both quantitatively and visually to estimate the number of unsupervised classes in a dataset by inducing sampling variability with sub-sampling [19]. In consensus clustering, we evaluated the separation of consensus matrix heatmaps using the elbow method, cumulative distribution function, and cluster-consensus plots. We visually evaluated the clustering using t-Distributed Stochastic Neighbour Embedding (t-SNE) for reducing dimensionality and visualising high-dimensional datasets [20]. We also derived phenotypes using a divisive hierarchical clustering approach as an alternative to k-means, for confirming the cluster consistency. The number of clusters was determined using the dendrogram and the elbow and gap statistic methods [21].

Evaluation of rhTM effects in derived phenotypes
We examined the associations between rhTM and the clinical outcomes in each derived cluster, using a generalised estimating equation to adjust for hospital-level variance. We analysed the associations after adjusting for the potential confounders of age, sex, comorbidities, and sequential organ failure assessment (SOFA) scores. In the derivation cohort, we did not adjust for the management before and after admission to the hospitals because the information on when each management was initiated was not available. We examined the cluster-level effects of rhTM by including the interaction term, rhTM use –x– cluster, in the model to examine different effects of rhTM across clusters. In addition, to confirm the robustness of the association of interest, we applied a Bayesian regression model to assess the associations between rhTM and the clinical outcomes for each derived cluster, based on k-means in the derivation cohort [21, 22]. Bayesian regression was performed using a Markov chain Monte Carlo procedure with four chains of 2000 iterations per chain. The results are shown as beta coefficients with 95% credible intervals and displayed as odds ratios with 95% credible intervals for simplicity.

Cluster validation and evaluation of rhTM effects
We predicted patient phenotypes in the FORECAST sepsis study as external data based on coagulation markers of clusters in the derivation cohort (derived from JSEPTIC-DIC and Tohoku Sepsis Registry). Predictions arose from the Euclidean distance from each patient to the centroid of each FORECAST phenotype. In each predicted cluster in the FORECAST sepsis study, we first described the frequency and clinical characteristics of the clusters. Thereafter, we used a generalised estimating equation to account for patient clustering within hospitals to assess associations between rhTM and clinical outcomes in each predicted cluster in the external data. The adjusted variables were age, sex, comorbidities, SOFA scores, and in-hospital management, including renal replacement therapy, and treatment with steroids, intravenous immunoglobulin, antithrombin, and vasopressors. As the FORECAST sepsis data included information on the time of management, we included management before or after admission as a covariate to estimate the effects of rhTM on clinical outcomes. For sensitivity analyses, we used a generalised estimating equation, applying the acute physiology and chronic health evaluation (APACHE II) score and source of infection as adjusted variables, instead of the SOFA score. The source of infection was categorised into respiratory, abdominal, skin and soft tissue, urinary tract, and others. As in the derivation cohort, we analysed the association by Bayesian regression with a Markov chain Monte Carlo procedure with four chains.
Values with $p < 0.05$ were considered statistically significant. All data were analysed using Stata version 14.1 (StataCorp, College Station, TX, USA) and R version 3.4.1 package for t-SNE (https://cran.r-project.org/web/packages/tsne/tsne.pdf) (R Foundation, Vienna, Austria).

Results

Patients in the derivation cohort

We excluded 117 patients without 28-day mortality information in the derivation cohort from the two multicentre registries, leaving 3694 patients who were eligible for analysis (3195 from JSEPTIC-DIC and 499 from the Tohoku Sepsis Registry). Table 1 summarises the patients’ characteristics. The median age was 72 years, and 40% of patients were female. Overall, rhTM was administered to 26.2% of patients (the distribution of the proportion of patients in the institutes is shown in Additional file 3: Figure S1). The in-hospital mortality and 28-day mortality rates were 32.1% and 20.4%, respectively.

Derivation of clinical sepsis phenotypes

We assessed the distributions and missingness among the phenotyping variables. Antithrombin activity was the most lacking variable, 51.3% and 45.8% in derivation and validation datasets, respectively (Table 2). According to clustering using k-means, a four-class model including the phenotype clusters derivation dA, dB, dC, and dD ("d" represents “derivation”) may be an optimal fit. The heat map matrix (Additional file 3: Figure S2), elbow method (Additional file 3: Figure S3), and cumulative distribution function curve (Additional file 3: Figure S4), indicated that the four-class model was optimal, whereas the cluster-consensus plot suggested that two, three, or four clusters were optimal (Additional file 3: Figure S5). The four-class model was supported by the t-SNE plot with clear separation (Fig. 1). Additional file 3: Figure S6 shows a cluster dendrogram obtained using a divisive hierarchical clustering approach. The elbow method showed that a two- or four-cluster model is optimal (Additional file 3: Figure S7), whereas the gap statistic method [21] showed that the four-cluster model was optimal (Additional file 3: Figure S8).

Patients in cluster dA were likely to have a severe physiological status and organ dysfunction (high APACHE II and SOFA scores), coagulopathy (low platelet counts, prolonged PT-INR, low fibrinogen, and extremely high FDP and D-dimer levels), high lactate levels, and high mortality (Table 1). Approximately 90% of patients in this cluster required vasopressors. The characteristics of patients in cluster dB were similar to that in cluster dA in terms of severity, but likely to have abdominal infection with normal white blood cell counts, moderate coagulopathy with moderate FDP and D-dimer levels, and low antithrombin activity. Patients in clusters dC and dD had moderate and mild disease, respectively. Although patients in cluster dC had coagulopathy with high FDP and D-dimer levels, those in cluster dD were likely to have respiratory infection without coagulopathy. The phenotypes were similar according to the four-cluster hierarchical clustering (Additional file 2: Table S1).

Evaluation of rhTM effects in the derivation cohort

Recombinant human thrombomodulin was administered to 128 (44.1%), 184 (31.5%), 334 (33.6%), and 210 (15.5%) patients in clusters dA, dB, dC, and dD, respectively. Clinical outcomes in cluster dA were better with than in those without rhTM (adjusted risk difference [RD], $-17.8\%$ [95% CI, $-28.7$ to $-6.9\%$] for 28-day mortality; RD, $-17.7\%$ [95% CI, $-27.6$ to $-7.8\%$] for in-hospital mortality (Table 3). Analysis of the rhTM effect modification across clusters using cluster dA as the reference showed that the effects of rhTM differed across clusters (all, $p < 0.05$), except for in-hospital mortality in cluster dB ($p = 0.31$). The associations were similar according to the four-cluster hierarchical clustering (Additional file 2: Table S2). Furthermore, rhTM treatment was associated with better clinical outcomes in cluster dA according to Bayesian regression (Additional file 2: Table S3).

Characteristics of phenotypes in the validation cohort

Additional file 2: Table S4 shows the patients’ characteristics in each cluster in the validation cohort. The median age was 73 years, 40% of the patients were women, and rhTM was administered to 21.2% of patients. In-hospital and 28-day mortality rates were 23.4% and 19.0%, respectively. These characteristics were similar to those in the derivation cohort but the rate of rhTM treatment and mortality were relatively lower.

We used only coagulation markers to predict clusters in the validation cohort, and the characteristics were similar to those in the derivation cohort (“v” represents “validation”). Similar to the patients in cluster dA, those in cluster vA were likely to have a severe physiological status and organ dysfunction (high APACHE II and SOFA scores), coagulopathy (low platelet counts, prolonged PT-INR, low fibrinogen, and extremely high FDP and D-dimer levels), high lactate levels, and moderately high mortality. Patients in cluster vB had a high mortality rate with moderate coagulopathy and moderate FDP and D-dimer levels. Patients in clusters vC and vD had moderate and mild disease, respectively. Patients in cluster vC had coagulopathy with high FDP and D-dimer levels, whereas those in cluster vD did not have coagulopathy.
Table 1  Characteristics of patients in the derivation cohort according to clusters

| Variables                  | Overall | Cluster dA | Cluster dB | Cluster dC | Cluster dD | p*  |
|----------------------------|---------|------------|------------|------------|------------|-----|
| n = 3694                   | n = 323 | n = 629    | n = 1147   | n = 1595   |            |     |
| Age*, median (IQR)         | 72.0 (62.0, 81.0) | 72.0 (58.0, 80.0) | 72.0 (63.0, 81.0) | 73.0 (63.0, 81.0) | 72.0 (62.0, 80.0) | 0.25 |
| Sex*, female               | 4168 (39.7%) | 164 (50.8%) | 268 (42.6%) | 483 (42.1%) | 553 (34.7%) | <0.001 |
| Body weight kg, median (IQR)| 54.7 (46.6, 64.2) | 55.0 (47.5, 64.0) | 52.0 (45.0, 61.1) | 55.0 (46.8, 64.3) | 55.0 (47.0, 65.0) | <0.001 |
| Comorbidity*               |         |            |            |            |            |     |
| Liver                      | 149 (4.0%) | 28 (8.7%)  | 73 (11.6%) | 30 (2.6%)  | 18 (1.1%)  | <0.001 |
| Respiratory                | 141 (3.8%) | 8 (2.5%)   | 23 (3.7%)  | 43 (3.7%)  | 67 (4.2%)  | 0.52  |
| Cardiovascular             | 316 (6.6%) | 20 (6.2%)  | 49 (7.8%)  | 97 (8.5%)  | 150 (9.4%) | 0.23  |
| Renal                      | 306 (8.3%) | 24 (7.4%)  | 50 (7.9%)  | 111 (9.7%) | 121 (7.6%) | 0.23  |
| Immunodeficiency           | 709 (19.2%) | 62 (19.2%) | 119 (18.9%) | 233 (20.3%) | 295 (18.5%) | 0.69  |
| Infection site             |         |            |            |            |            | <0.001 |
| ISTH DIC score             |         |            |            |            |            |     |
| ISTH DIC score ≥ 5         | 1430 (39.2%) | 291 (90.7%) | 384 (62.0%) | 653 (57.5%) | 102 (6.5%) | <0.001 |
| Management                |         |            |            |            |            |     |
| rhTM                      | 969 (29.3%) | 128 (44.1%) | 184 (31.5%) | 334 (33.6%) | 210 (15.8%) | <0.001 |
| Vasopressor use            | 2789 (75.5%) | 289 (89.5%) | 558 (88.7%) | 882 (76.9%) | 1060 (66.5%) | <0.001 |
| Renal replacement therapy  | 971 (26.3%) | 135 (41.8%) | 220 (35.0%) | 339 (29.6%) | 277 (17.4%) | <0.001 |
| Steroids                   | 894 (24.2%) | 112 (34.7%) | 214 (34.1%) | 285 (24.8%) | 283 (17.7%) | <0.001 |
| Intravenous immunoglobin   | 1088 (29.5%) | 116 (35.9%) | 239 (38.0%) | 362 (31.6%) | 371 (23.3%) | <0.001 |
Evaluation of the effects of rhTM in the validation cohort

All 1184 patients in the FORECAST sepsis study dataset were analysed for validation. Recombinant human thrombomodulin was administered to 44 (44.4%), 54 (31.2%), 98 (26.3%), and 46 (9.3%) patients in clusters vA, vB, vC, and vD, respectively. Clinical outcomes in cluster vA were better than in those without rhTM (adjusted RD, −24.9% [95% CI −49.1 to −0.7%] for 28-day mortality; RD −30.9% [95% CI −55.3 to −6.6%] for in-hospital mortality; Table 3). In contrast, rhTM was not associated with better outcomes in the other clusters. As secondary outcomes, rhTM was associated with increased number of ventilator-free days (6.7 days [95% CI 0.8–12.7 days]) in Cluster vA, but not with the number of ICU-free days or discharge location in any of the clusters (Additional file 2: Table S5). In the sensitivity analyses, rhTM was associated with better outcomes when APACHE II score and source of infection were applied as adjusted variables, instead of the SOFA score (adjusted RD, −0.24% [95% CI −0.45 to −0.02%] for 28-day mortality; RD −0.26% [95% CI −0.47 to −0.04%] for in-hospital mortality; Additional file 2: Table S6). The associations between rhTM, and 28-day and in-hospital mortalities were consistent with the findings of the Bayesian regression analysis (Additional file 2: Table S3 and Additional file 3: Figure S9).

Discussion

This secondary analysis of the sepsis registries identified four phenotypes with various coagulation features among patients with severe sepsis. Treatment with rhTM was associated with lower in-hospital mortality rates only in the phenotype with severe coagulopathy, characterised by low platelet counts, extremely high levels of FDP and D-dimer (phenotype clusters dA and vA), and severe organ dysfunction. These results were not identified in the other phenotypes.

The severity of coagulopathy is defined by the DIC scoring systems, such as the International Society on Thrombosis and Haemostasis (ISTH) scoring system for diagnosing overt DIC [23] and Japanese Association for Acute Medicine DIC scoring system [24], both of which have been applied in many studies. The difference between these systems and machine learning-based clustering is the use of a trivial cut-off. Table 1, Additional file 2: Tables S1, and S4 show that each phenotype cluster included patients with various ISTH DIC scores without a clear cut-off that overlapped with the other clusters. This suggests that clustering based on machine learning can detect novel phenotypes that cannot be identified using the conventional scoring systems.

Recombinant human thrombomodulin has anticoagulation effects and was shown to be beneficial for patients...
with sepsis and coagulopathy in observational studies and in a subgroup analysis of a phase II trial [3, 6]. The latest phase III trial focused on patients with sepsis with cardiovascular or respiratory dysfunction as well as coagulopathy according to subgroup analysis of the phase II trial [4]. However, the phase III trial did not identify a positive effect of rhTM on survival, suggesting that differentiating a subgroup that may benefit from rhTM is difficult using conventional methods with clear cut-offs. In our study, despite overlapping characteristics, various DIC scores, and differences in severity among clusters, cluster vA (dA) was the only phenotype in which rhTM was associated with better survival outcomes. This suggests that machine learning clustering can identify optimal clinical phenotypes for rhTM treatment. Additionally, the machine learning clustering described herein used only six variables, all of which are general markers that can be measured in most hospitals. Additionally, the results can be available soon after admission before deciding to administer rhTM in an emergency room or ICU.

Other studies using machine learning-based clustering for patients with sepsis also suggest that several specific therapies have beneficial effects only in patients with specific phenotypes. A Toll-like receptor 4 antagonist, protocol-based resuscitation, activated protein C, and fluid input affected each phenotype differently [9, 10]. The effectiveness of rhTM also varied across phenotypes in our study. Therefore, selecting an optimal clinical phenotype may be key to the success of a specific therapy for patients with sepsis. Including entire populations with sepsis may explain why previous randomised trials found no beneficial effects of adjunctive therapies [25–28]. The goal of precision/tailored medicine is to select the optimal therapy for patients, for which machine learning-based clustering can be effective. Although our study does not fully address the definite endotypes of coagulation in sepsis biologically or pathophysiologically, our findings improve the understanding of the true endotypes of sepsis with coagulation.

**Fig. 1** t-SNE plot. This t-Distributed Stochastic Neighbour Embedding (t-SNE) plot is a dimensionality reduction technique for graphically simplifying extensive datasets. Four clusters are plotted, and some patients are on the borderlines between clusters. Circles represent individual patients (green, cluster dA; purple, cluster dB; blue, cluster dC; red, cluster dD).
Table 3  Unadjusted and adjusted risk difference between recombinant thrombomodulin use and outcomes

| Outcomes | Cluster dA | p value | Cluster dB | p value | Cluster dC | p value | Cluster dD | p value |
|----------|------------|---------|------------|---------|------------|---------|------------|---------|
| Unadjusted association (vs. non rhTM use) |           |         |            |         |            |         |            |         |
| 28-Day death | $-10.8$ (−$21.5$ to $-0.1$) | 0.047 | 3.5 (−4.6 to 11.5) | 0.40 | $-1.9$ (−6.9 to 3.2) | 0.47 | 2.2 (−2.8 to 7.1) | 0.39 |
| In-hospital death | $-10.9$ (−$20.8$ to $-1.1$) | 0.03 | 1.6 (−7.2 to 10.3) | 0.73 | $-8.0$ (−13.8 to $-2.3$) | 0.01 | 0.8 (−$5.2$ to 6.8) | 0.78 |
| Adjusted association (vs. non rhTM use) |           |         |            |         |            |         |            |         |
| 28-Day death | $-17.8$ (−$28.7$ to $-6.9$) | 0.001 | 0.7 (−7.1 to 8.6) | 0.85 | $-3.1$ (−8.3 to 2.1) | 0.24 | $-0.7$ (−4.5 to 6.0) | 0.79 |
| In-hospital death | $-17.7$ (−$27.6$ to $-7.8$) | $<0.001$ | 0.2 (−7.9 to 8.3) | 0.97 | $-10.2$ (−15.9 to $-4.6$) | $<0.001$ | $-1.3$ (−7.6 to 4.9) | 0.67 |

In the validation cohort, the adjusted variables were age, sex, comorbidities, and sequential organ failure assessment (SOFA) scores.

In the derivation cohort, the adjusted variables were age, sex, comorbidities, SOFA scores, and in-hospital management, including renal replacement therapy, and treatment with steroids, intravenous immunoglobulin, antithrombin, and vasopressors.

**rhTM** recombinant human thrombomodulin

**Limitations**

This study had several limitations. We used three registries that included different variables. Therefore, unmeasured confounders, and a lack of information such as the timing of rhTM administration may have biased our findings. Nevertheless, the data included detailed clinical information that is generally used for adjustment, and the results in the validation cohort accounted for management before or after admission. The number of missing variables for clustering was not small; therefore, missingness may have limited our findings. However, missing data imputation using the random forest approach is considered valid and valid imputation reduces bias, even when the proportion of missingness is high [29]. Our data did not include information on the long-term outcomes (i.e., 6-/12-months mortality) and SOFA score at several weeks after admission, although long-term outcomes are also important. Our data did not include the duration of rhTM administration, which was presumably 6 days, according to the generally prescribed dose and duration in Japan. We could not evaluate whether the phenotypes and efficacy of rhTM are applicable for patients with sepsis defined by the Sepsis-3 criteria [30], as three observational studies enrolled patients with sepsis using the Sepsis-2 definition [16], and the datasets did not include SOFA scores before admission. Heparin is commonly used for anticoagulant therapy, worldwide. However, we could not include heparin treatment for sepsis-induced coagulopathy in the model, because two of three registries did not collect the data. Only 5% (167/3195) of the patients were treated with heparin for coagulopathy with sepsis (excluding use for venous thromboembolism) in JSEPTIC-DIC study [12]. We need to develop a model to determine the phenotypes of individual patients to be able to perform a clinical trial in the future. Finally, our data were derived from Japanese patients; therefore, the generalisability of the results may be limited.

**Conclusions**

The findings derived using machine learning clustering indicated that rhTM can benefit only patients with a severe coagulopathy phenotype. Identifying patients for whom a therapy will have a beneficial effect can lead to precision/tailored medicine in critical care. To achieve this goal, the accuracy of phenotyping should be increased by analysing more patients, and through further validation. A randomised trial focusing on suitable phenotypes determined by effective phenotyping is warranted.

**Appendix: The core investigators FORECAST Study**

Satoshi Gando 1, 2), Yasuhiro Otomo 3), Shigeki Kushima 4), Hiroshi Ogura 5), Seitaro Fujishima 6), Atsushi Shiraishi 7), Daizoh Saitoh 8), Toshihiko Mayumi 9), Kiyotsugu Takuma 10), Taka-aki Nakada 11), Yasukazu Shiraishi 12), Takehiko Tarui 13), Toru Hifumi 14), Kohji
Okamoto 15), Yutaka Umemura 16), Joji Kotani 17), Yuichiro Sakamoto 18), Junichi Sasaki 19), Shin-ichiro Shiraishi 20), Ryosuke Tsuruta 21), Akiyoshi Hagiwara 22), Toshikazu Abe 23, 24, 25), Kazuma Yamakawa 26), Tomohiko Masuno 27), Naoshi Takeyama 28), Norio Yamashita 29), Hiroto Ikeda 30), Masashi Ueyama 31)

1) Division of Acute and Critical Care Medicine, Hokkaido University, Sapporo, Japan
2) Critical Care Medicine, Sapporo Higashi Tokushukai Hospital, Sapporo, Japan
3) Trauma and Acute Critical Care Center, Medical Hospital, Tokyo Medical and Dental University, Tokyo, Japan
4) Division of Emergency and Critical Care Medicine, Tohoku University Graduate School of Medicine, Sendai, Japan
5) Department of Traumatology and Acute Critical Medicine, Osaka University Graduate School of Medicine, Osaka, Japan
6) Center for General Medicine Education, Keio University School of Medicine, Tokyo, Japan
7) Emergency and Trauma Center, Kameda Medical Center, Kamogawa, Japan
8) Division of Traumatology, Research Institute, National Defense Medical College, Tokorozawa, Japan
9) Department of Emergency Medicine, School of Medicine, University of Occupational and Environmental Health, Kitakyushu, Japan
10) Emergency & Critical Care Center, Kawasaki Municipal Kawasaki Hospital, Kawasaki, Japan
11) Department of Emergency and Critical Care Medicine, Chiba University Graduate School of Medicine, Chiba, Japan
12) Department of Acute Medicine, Kawasaki Medical School, Kurashiki, Japan
13) Department of Trauma and Critical Care Medicine, Kyorin University School of Medicine, Tokyo, Japan
14) Department of Emergency and Critical Care Medicine, St. Luke's International Hospital, Tokyo, Japan
15) Department of Surgery, Center for Gastroenterology and Liver Disease, Kitakyushu City Yahata Hospital, Kitakyushu, Japan
16) Department of Disaster and Emergency Medicine, Kobe University Graduate School of Medicine, Kobe, Japan
17) Department of Traumatology and Acute Critical Medicine, Osaka University Graduate School of Medicine, Osaka, Japan
18) Emergency and Critical Care Medicine, Saga, University Hospital, Saga, Japan
19) Department of Emergency and Critical Care Medicine, Keio University School of Medicine, Tokyo, Japan
20) Department of Emergency and Critical Care Medicine, Aizu Chuo Hospital, Aizuwakamatsu, Japan
21) Advanced Medical Emergency & Critical Care Center, Yamaguchi, Ube, Japan
22) Department of Emergency Medicine, Niizashiki Chuo General Hospital, Niiza, Japan
23) Department of General Medicine, Juntendo University, Tokyo, Japan
24) Health Services Research and Development Center, University of Tsukuba, Tsukuba, Japan
25) Department of Health Services Research, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan
26) Division of Trauma and Surgical Critical Care, Osaka, General Medical Center, Osaka, Japan
27) Department of Emergency and Critical Care Medicine, Nippon Medical School, Tokyo, Japan
28) Advanced Critical Care Center, Aichi Medical University Hospital, Nagakute, Japan
29) Advanced Emergency medical Service Center, Kurume University Hospital, Kurume, Japan
30) Department of Emergency Medicine, Teikyo University School of Medicine, Tokyo, Japan
31) Department of Trauma, Critical Care Medicine, and Burn Center, Japan Community Healthcare Organization, Chukyo Hospital

Abbreviations
APACHE: Acute physiology and chronic health evaluation; FDP: Fibrinogen/fibrin degradation product; FORECAST: Focused Outcomes Research in Emergency Care for Acute Respiratory Distress Syndrome, Sepsis, and Trauma; ICU: Intensive care unit; ISTH: International Society on Thrombosis and Haemostasis; JSEPTIC-DIC: Japan Septic Disseminated Intravascular Coagulation; PT-INR: Prothrombin time/international normalised ratio; RD: Risk difference; rhTM: Recombinant human thrombomodulin; SOFA: Sequential organ failure assessment.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s13054-021-03541-5.

Additional file 1. Supplemental documents.
Additional file 2. Supplemental Tables.
Additional file 3. Supplemental Figures.

Acknowledgements
We thank the core investigators of the FORECAST sepsis study (Appendix) for providing the dataset. We are grateful to all investigators involved in the JSEPTIC-DIC study, Tohoku Sepsis Registry, and the FORECAST sepsis study for contributing to data collection and assessment.

Authors’ contributions
All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by DK, TG, RU, MH, TA, and AS. Statistical analysis was reviewed by RU, TA, and AS. The first draft of the manuscript was written by DK and TG. The manuscript was reviewed and edited by KY, TA, AS, and SK, and all authors commented on the previous versions of the manuscript. All authors read and approved the final manuscript. Funding acquisition: DK; supervision: SK.
Funding
This work (writing manuscript) was supported by JSPS KAKENHI Grant Number JP19H03755.

Availability of data and materials
The datasets generated and/or analysed during the original studies are available in the Scientific data, https://www.nature.com/articles/sdata2018243 (J-SEPTIC DIC study), and Mendeley Data, https://data.mendeley.com/datasets/vvv89kw3k571 (Tohoku Sepsis Registry). The dataset of the FORECAST sepsis study is not publicly available, based on the decision by the Japanese Association for Acute Medicine.

Declarations

Ethics approval and consent to participate
Three original studies were approved, and the need for informed consent was waived by the institutional review boards at the participating hospitals.

Consent for publication
Not applicable.

Competing interests
D.K., M.H., and S.K. received personal fees from Asahi Kasei Pharma Corporation. The other authors have no conflicts of interest to declare.

Author details
1 Division of Emergency and Critical Care Medicine, Tohoku University Graduate School of Medicine, 1-1 Seiryo-machi, Aoba-ku, Sendai 980-8574, Japan. 2 Department of Clinical Epidemiology and Health Economics, School of Public Health, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. 3 Intensive Care Unit, Tokyo Medical and Dental University School of Medicine, 1-1 Seiryo-machi, Aoba-ku, Sendai 980-8574, Japan. 4 Department of Emergency Medicine, Hokkaido University Hospital, Kita 14 Nishi-5, Kita-ku, Sapporo 061-8648, Japan. 5 Division of Emergency Medicine, Osaka Medical College, 2-7 Daigakumach, Takatsuki 569-8686, Japan. 6 Department of Emergency and Critical Care Medicine, Tsukuba Memorial Hospital, 1187-298 Kaname, Tsukuba 300-2622, Japan. 7 Health Services Research and Development Center, University of Tsukuba, 1-1-1 Tennodai, Tsukuba 305-8577, Japan. 8 Emergency and Trauma Center, Kameda Medical Center, 929 Higashimachi, Kamogawa 296-8602, Japan.

Received: 4 January 2021 Accepted: 10 March 2021

Published online: 19 March 2021

References
1. Rudd KE, Johnson SC, Agesa KM, Shackelford KA, Tsoi D, Kiefe CL, et al. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. Lancet. 2020;395:200–11.
2. Abeayama K, Stern DM, Ito Y, Kawaihara K, Yoshimoto Y, Tanaka M, et al. The N-terminal domain of thrombomodulin sequesters high-mobility group-B1 protein, a novel anti-inflammatory mechanism. J Clin Invest. 2005;115:2567–74.
3. Yamakawa K, Ogura H, Fujimi S, Morikawa M, Ogawa Y, Mohri T, et al. Recombinant human soluble thrombomodulin in sepsis-induced disseminated intravascular coagulation: a multicenter propensity score analysis. Intensive Care Med. 2013;39:644–52.
4. Vincent JL, Francois B, Zabolotskikh I, Daga MK, Lascarrou JB, Kirov MY, et al. Effect of a recombinant human soluble thrombomodulin on mortality in patients with sepsis-associated coagulopathy: the SCARLET randomized clinical trial. JAMA. 2019;321:1993–2002.
5. van der Poll T. Recombinant human soluble thrombomodulin in patients with sepsis-associated coagulopathy: another negative sepsis trial? JAMA. 2019;321:1978–80.
6. Vincent JL, Ramesh MK, Ernest D, LaRosa SP, Pachl J, Akaiwa N, et al. A randomized, double-blind, placebo-controlled, Phase 2b study to evaluate the safety and efficacy of recombinant human soluble thrombomodulin, ART-123, in patients with sepsis and suspected disseminated intravascular coagulation. Crit Care Med. 2013;41:2069–79.
7. Cecconi M, Evans L, Levy M, Rhodes A. Sepsis and septic shock. Lancet. 2018;392:75–87.
8. Guilamet MCV, Bernauer M, Micek ST, Krolle MH. Cluster analysis to define distinct clinical phenotypes among septic patients with bloodstream infections. Medicine (Baltimore). 2019;98:e15276.
9. Seymour CW, Kennedy JN, Wang S, Chang CH, Elliott CF, Xu Z, et al. Deri-vation, validation, and potential treatment implications of novel clinical phenotypes for sepsis. JAMA. 2019;321:2003–17.
10. Zhang Z, Zhang G, Goyal H, Mo L, Hong Y. Identification of subclasses of sepsis that showed different clinical outcomes and responses to amount of fluid resuscitation: a latent profile analysis. Crit Care. 2018;22:347.
11. Steinel D. K-means clustering: a half-century synthesis. Br J Math Stat Psychol. 2006;59:1–34.
12. Hayakawa M, Yamakawa K, Saito S, Uchino S, Kudo D, Izuka Y, et al. Nationwide registry of sepsis patients in Japan focused on disseminated intravascular coagulation 2011–2013. Sci Data. 2018;5:180243.
13. Kudo D, Kushimoto S, Miyagawa N, Saito T, Hasegawa M, Ito F, et al. The impact of organ dysfunctions on mortality in patients with severe sepsis: a multicenter prospective observational study. J Crit Care. 2018;45:178–83.
14. Abe T, Ogura H, Shiroma A, Kushimoto S, Saito D, Fujishima S, et al. Characteristics, management, and in-hospital mortality among patients with severe sepsis in intensive care units in Japan: the FORECAST study. Crit Care. 2018;22:322.
15. Dellinger RP, Levy MM, Rhodes A, Annane D, Gerlach H, Opal SM, et al. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock. 2012. Crit Care Med. 2013;41:580–637.
16. Levy MM, Fink MP, Marshall JC, Abraham E, Angus D, Cook D, et al. 2001 SCCM/ESICM/ACCP/ATS/SIS international sepsis definitions conference. Intensive Care Med. 2003;29:530–8.
17. Stekhoven DJ, Buhlmann P. MissForest—non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012;28:112–8.
18. Shah AD, Bartlett JW, Carpenter J, Nicholas O, Hemingway H. Comparison of random forest and parametric imputation models for imputing missing data using MICE. A CALIBER study. Ann J Epidemiol. 2014;179:764–74.
19. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010;26:1572–8.
20. Hinton G. Visualizing data using t-SNE. J Mach Learn Res. 2008;9:2579–605.
21. Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. J R Stat Soc B. 2001;63:411–23.
22. Zampieri FG, Costa EL, Iwashyna TJ, Carvalho CRR, Damiani LP, Taniguchi L, et al. Heterogeneous effects of alveolar recruitment in acute respiratory distress syndrome: a machine learning reanalysis of the Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial. Br J Anaesth. 2019;123:88–95.
23. Taylor FB Jr, Toh CH, Hoots WK, Wada H, Levi M, Scientific Subcommittee on Disseminated Intravascular Coagulation (DIC) of the International Society on Thrombosis and Haemostasis (ISTH). Towards definition, clinical and laboratory criteria, and a scoring system for disseminated intravascular coagulation. Thromb Haemost. 2001;86:1327–30.
24. Gando S, Iba T, Eguchi Y, Otomoto Y, Okamoto K, Koseki K, et al. A multi-center, prospective validation of disseminated intravascular coagulation diagnostic criteria for critically ill patients: comparing current criteria. Crit Care Med. 2006;34:625–31.
25. Abraham E, Wunderink R, Silverman H, Perl TM, Naraswaty S, Levy H, et al. Efficacy and safety of monoclonal antibody to human tumor necrosis factor alpha in patients with sepsis syndrome. A randomized, controlled, double-blind, multicenter clinical trial. TNF-alpha MAb sepsis study group. JAMA. 1995;273:934–41.
26. Opal SM, Laterre PF, Francois B, LaRosa SP, Angus DC, Mira JP, et al. Effect of eritoran, an antagonist of MD2-TLR4, on mortality in patients with sepsis that showed different clinical outcomes and responses to amount of fluid resuscitation: a latent profile analysis. Crit Care Med. 2019;47:1154–62.
27. Ranieri VM, Thompson BT, Barie PS, Dhainaut JF, Douglas IS, Finfer S, et al. The heterogeneous effect of alveolar recruitment in acute respiratory distress syndrome: a machine learning reanalysis of the Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial. Br J Anaesth. 2019;123:88–95.
29. Madley-Dowd P, Hughes R, Tilling K, Heron J. The proportion of missing data should not be used to guide decisions on multiple imputation. J Clin Epidemiol. 2019;110:63–73.
30. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). JAMA. 2016;315:801–10.

**Publisher’s Note**
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.