Clinical Infectious Diseases

MAJOR ARTICLE

Clinical sub-phenotypes of *Staphylococcus aureus* bacteraemia

Maaike C Swets[1,2], Zsuzsa Bakk[3], Annette C Westgeest[1], Karla Berry[4,5], George Cooper[4], Wynne Sim[6], Rui Shian Lee[6], Tze Yi Gan[6], William Donlon[6], Antonia Besu[6], Emily Heppenstall[4], Luke Tysall[7], Simon Dewar[5,7], Mark de Boer[1,8], Vance G Fowler Jr[9,10], David H Dockrell[4], Guy E Thwaites[11,12], Miquel Pujol[13,14,15], Natàlia Pallarès[16,17], Cristian Tebé[16], Jordi Carratalà[13,14,15,18], Alexander Szubert[19], Geert H Groeneveld[1,20], Clark D Russell[4,7*]

[1]Department of Infectious Diseases, Leiden University Medical Center, Leiden University, Leiden, The Netherlands; [2]Roslin Institute, University of Edinburgh, Edinburgh, United Kingdom; [3]Department of Methodology and Statistics, Leiden University, Leiden, The Netherlands; [4]Centre for Inflammation Research, Institute for Regeneration and Repair, The University of Edinburgh, Edinburgh, United Kingdom; [5]Clinical Infection Research Group, Western General Hospital, Edinburgh, United Kingdom; [6]Edinburgh Medical School, The University of Edinburgh, Edinburgh, United Kingdom; [7]Medical Microbiology, Royal Infirmary of Edinburgh, Edinburgh, United Kingdom; [8]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands; [9]Division of Infectious Diseases and International Health, Department of Medicine, Duke University School of Medicine, Durham, North Carolina, USA; [10]Duke Clinical Research Institute, Durham, North Carolina, USA; [11]Oxford University Clinical Research Unit, Ho Chi Minh city, Vietnam; [12]Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom; [13]Department of Infectious Diseases, Bellvitge University Hospital, L´Hospitalet de LLobregat, Barcelona, Spain; [14]Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain; [15]Centro de Investigación Biomédica en Red de Enfermedades Infecciosas (CIBERINFEC), Instituto de Salud Carlos III, Madrid, Spain; [16]Biostatistics Support and Research Unit, Germans

*Correspondence: Clark D. Russell MBChB PhD. Centre for Inflammation Research, Institute for Regeneration and Repair, South Building, Room G.10, 4-5 Little France Drive, Edinburgh, EH16 4UU, Scotland, United Kingdom. E-mail: clark.russell@ed.ac.uk

Alternative corresponding author: Maaike C Swets MD. Department of Infectious Diseases, Leiden University Medical Center, Leiden University, Leiden, The Netherlands. E-mail: m.c.swets@lumc.nl

DOI: 10.1093/cid/ciae338

Trias i Pujol Research Institute and Hospital (IGTP), Badalona, Spain; [17]Department of Basic Clinical Practice, School of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain; [18]Department of Clinical Sciences, School of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain; [19]MRC Clinical Trials Unit, University College London, United Kingdom; [20]Department of Internal Medicine- Acute Internal Medicine, Leiden University Medical Center, Leiden, The Netherlands.

Background: *Staphylococcus aureus* bacteraemia (SAB) is a clinically heterogeneous disease. The ability to identify sub-groups of patients with shared traits (sub-phenotypes) is an unmet need that could allow patient stratification for clinical management and research. We aimed to test the hypothesis that clinically-relevant sub-phenotypes can be reproducibly identified amongst patients with SAB.

Methods: We studied three cohorts of hospitalised adults with monomicrobial SAB: a UK retrospective observational study (Edinburgh cohort, n=458), the UK ARREST randomised trial (n=758), and the Spanish SAFO randomised trial (n=214). Latent class analysis was used to identify sub-phenotypes using routinely-collected clinical data, without considering outcomes. Mortality and microbiologic outcomes were then compared between sub-phenotypes.

Results: Included patients had predominantly methicillin-susceptible SAB (1366/1430,95.5%). We identified five distinct, reproducible clinical sub-phenotypes: (A) SAB associated with older age and comorbidity, (B) nosocomial intravenous catheter-associated SAB in younger people without comorbidity, (C) community-acquired metastatic SAB, (D) SAB associated with chronic kidney disease, and (E) SAB associated with injection drug use. Survival and microbiologic outcomes differed between the sub-phenotypes. 84-day mortality was highest in sub-phenotype A, and lowest in B and E. Microbiologic outcomes were worse in sub-phenotype C. In a secondary analysis of the ARREST trial, adjunctive rifampicin was associated with increased 84-day mortality in sub-phenotype B and improved microbiologic outcomes in sub-phenotype C.

Conclusions: We have identified reproducible and clinically-relevant sub-phenotypes within SAB, and provide proof-of-principle of differential treatment effects. Through clinical trial enrichment and patient stratification, these sub-phenotypes could contribute to a personalised medicine approach to SAB.

INTRODUCTION

Staphylococcus aureus bacteraemia (SAB) has long been recognised as a difficult-to-treat bacterial disease requiring prolonged antimicrobial treatment[1,2]. Major complications include the development of metastatic foci of infection (up to 37%[3]), recurrence of bacteraemia despite appropriate treatment (up to 10%[4]), and death (15-30% in-hospital[5,6]). Globally, *S. aureus* accounts

for the most overall deaths due to a bacterial pathogen, and specifically the most deaths associated with bacteraemia[7].

A defining clinical feature of SAB is heterogeneity. This encompasses patient characteristics (e.g. age and co-morbidity), pathogen characteristics, place of acquisition (community or hospital), source of bacteraemia, and extent of infection. Currently, there is no consensus on rationalising this clinical heterogeneity to achieve patient stratification, and clinical trials mainly consider SAB to be a single syndrome. Strategy trials in SAB, frequently investigating combination antimicrobial therapy, have so far not succeeded in identifying approaches that improve outcomes compared to standards of care[8-10]. However, because of the clinical heterogeneity intrinsic to SAB, it is possible we fail to identify sub-groups of patients who may differentially benefit (or suffer harm) from specific therapies[11]. In contrast, clinically-relevant sub-phenotypes have been identified in similarly heterogenous diseases including the acute respiratory distress syndrome[12], asthma[13], chronic obstructive pulmonary disease[14], and bronchiectasis[15]. In this context a sub-phenotype is considered to be a sub-group of patients with a specific disease who exhibit similar traits, such as clinical features, outcomes, or responses to treatment[16]. We aimed to test the hypothesis that clinically-relevant sub-phenotypes can be reproducibly identified amongst patients with SAB.

METHODS

Patient cohorts

Patients were included from three cohorts: a retrospective observational cohort study (Edinburgh cohort, n=458)[17], the ARREST multicentre, randomised, double-blind, placebo-controlled trial (n=758)[10], and the SAFO randomised clinical trial (n=214)[8]. The Edinburgh cohort included consecutive adults (≥18 years) with monomicrobial SAB diagnosed between 20/12/2019 and 23/08/2022 in three UK hospitals (**Supplementary Figure 1**). The ARREST trial recruited adults (≥18 years) with monomicrobial SAB in 29 UK hospitals between 10/12/2012 and 25/10/2016, and randomised participants to receive adjunctive rifampicin (600mg or 900mg/day) or placebo for up to 14 days, in addition to standard antibiotic treatment. Exclusion criteria included evidence of rifampicin non-susceptible *S. aureus*, contraindications to rifampicin, or if adjunctive rifampicin was considered mandatory. The SAFO trial recruited adults (≥18 years) with monomicrobial methicillin-susceptible *S. aureus* (MSSA) bacteraemia in 19 Spanish University hospitals between 31/05/2019 and 24/02/2022, and randomised participants to receive cloxacillin (2g 6x/day) plus fosfomycin (3g 4x/day), or cloxacillin alone, for the initial seven days of treatment. Exclusion criteria included Child-Pugh class C liver cirrhosis, moderate-severe heart failure, injection drug use (IDU), MRSA bacteraemia, penicillin allergy, and acute SARS-CoV-2 infection. Ethical approvals were obtained from the South East Scotland Research Ethics Committee 02 (23/SS/0025) for the Edinburgh cohort study, the London (Westminster) Research Ethics Committee (12/LO/0637) for the ARREST trial, and the Spanish Medicines and Healthcare

Products Regulatory Agency (AEMPS; 18-0905) and the Bellvitge University Hospital Ethics Committee (AC069/18) for the SAFO trial.

Variables and definitions

Comorbidities were defined according to the Charlson Comorbidity Index. Acquisition of infection was categorised according to the definitions used by Friedman and colleagues[18]. The source of infection was the most likely portal of entry of *S. aureus* into the bloodstream. Metastatic infection was defined as the presence of foci of infection remote from the portal of entry thought to have arisen through haematogenous dissemination. All-cause 84-day mortality was recorded in all cohorts. In the Edinburgh cohort, persistent bacteraemia was defined as a further positive blood culture during treatment >96h after the index blood culture and recurrent bacteraemia was defined as a further positive blood culture with the same *S. aureus spa* type within 90 days of stopping treatment[3]. In the ARREST cohort, microbiologic failure was defined as ongoing signs and symptoms of infection and growth of *S. aureus* from blood or a sterile site for >14 days from randomisation. Recurrence was defined as growth of *S. aureus* from a sterile site after >7 days of apparent clinical improvement. These were combined into a composite microbiologic outcome referred to as composite microbiologic failure. In the SAFO trial, persistent bacteraemia was documented at days 3 and 7 after randomisation.

Statistical analyses

Latent class analysis (LCA) was used to look for homogenous sub-groups within the larger heterogeneous cohorts of SAB using indicator variables selected based on availability and potential clinical relevance (consensus opinion of CDR, MS, ACW and GHG)[19]. Baseline patient and microbiologic variables were considered as class-defining variables which were first determined using data from the Edinburgh cohort, then this model was applied to the ARREST and SAFO cohorts. The classes were formed without any consideration of clinical or microbiological outcomes. We excluded variables with >10% missing data, categorical variables with >50% co-linearity, any variable with <10% positivity unless considered of high clinical relevance, and any variable contributing <0·5% to the clustering[20]. Non-normally distributed values were log transformed for the LCA. Cases with missing values were handled with full information maximum likelihood (FIML), which is generally the preferred method for dealing with missing data in LCA[20]. With FIML, data is not imputed, but all available information is used for calculation of the likelihood contribution of each respondent to the estimation of the model parameters[20].

Model selection was based on a combination of statistical criteria and clinical knowledge. The statistical criteria used were the Bayesian Information Criteria (BIC), number of classes, and size of smallest class. The BIC is a statistical measure that provides information on the model fit, and is best at identifying the correct number of classes if a combination of continuous and categorical data is used[20]. A decrease in the BIC suggests that the addition of more classes is worth the added

model complexity[21]. To avoid a local maximum, in which case it would be difficult to replicate our findings, 16 random starting values were used, and 50 iterations for each start value. Those solutions were checked to make sure that the same maximum likelihood solution was found. When setting the seed, a fixed starting point for random number generation was established, which ensures reproducibility across different runs of the analysis. After identification of classes, we estimated the posterior probability of class membership for each of the identified classes for each individual, and assigned the individual to the class with the highest probability[20]. Given that LCA is a probabilistic method, there is a certain degree of uncertainty in class assignment, which can lead to classification errors. For example, an individual may have a 0·9 chance of belonging to class one, and a 0·1 chance of belonging to class two. This individual is then assigned to class one. We correct for misclassification error using the bias-adjusted three-step LCA[22]. LCA was done using the Latent GOLD 6.0 statistical software package[23].

Cohort characteristics were compared using contingency tables for categorical variables (Chi-square or Fisher's exact test), and Mann Whitney or Kruskal-Wallis tests for continuous variables (which Shapiro-Wilk tests demonstrated to be not normally distributed). To compare class-defining variables between sub-phenotypes, z-scores were calculated ($z = \frac{(value\ for\ subphenotype - mean\ for\ variable)}{standard\ deviation\ for\ variable}$). Additional meta-data not included as class-defining variables was compared between patients stratified by predicted sub-phenotype membership. Unadjusted one-year survival was compared using a Kaplan-Meier survival curve and log-rank test, performed using the *survminer*[24] and *ggplot2*[25] packages in R (RStudio version Version 2023·06·1+524). Unless otherwise stated, analyses and data visualisation were done using GraphPad Prism Version 10·0·3 for macOS.

RESULTS

Cohort characteristics

Characteristics of the Edinburgh, ARREST and SAFO cohorts are compared in **Table 1**. In comparison with the Edinburgh cohort, patients in ARREST were more likely to have SAB originating from skin or soft tissue infection (SSTI), and patients in SAFO were more likely to have an intravenous catheter as the source of bacteraemia. Consistent with previous comparisons of real-life patient cohorts with trial cohorts in SAB[9], 84-day mortality was lower in the ARREST and SAFO control arms compared to the Edinburgh cohort. Patients in the Edinburgh and ARREST cohorts predominantly had infection with MSSA (441/458 and 711/758 respectively), and the SAFO trial exclusively recruited people with MSSA bacteraemia.

Identification of sub-phenotypes using latent class analysis

Eighteen class-defining variables were included in the final latent class analysis. Despite co-linearity, both creatinine and chronic kidney disease were included because creatinine provides

additional information on the presence of acute kidney injury (correlation coefficient 0·66). After determination of contributing variables (**Supplementary Figure 2**) using the Edinburgh cohort, latent class models with one to seven classes were fitted (**Table 2**). For the Edinburgh and ARREST cohorts, the BIC and clinical interpretability favoured the five-class model and the size of the smallest class was acceptable (>5% of total population).

The five classes identified by LCA in the Edinburgh and ARREST cohorts represented distinct clinical sub-phenotypes when considering their association with class-defining clinical variables, and were replicated in the two analyses (**Figure 1**). Sub-phenotype A was associated with older age, co-morbidity, and SAB from unknown or SSTI source. Sub-phenotype B was associated with nosocomial SAB, bacteraemia originating from an intravenous catheter, younger age, less co-morbidity, and lack of any metastatic foci. Sub-phenotype C was associated with community-acquired SAB from unknown source, with higher CRP, and with the presence of metastatic foci of infection. Sub-phenotype D was associated with chronic kidney disease, intravenous catheter source, and nosocomial or healthcare associated acquisition. In the Edinburgh cohort, 17/39 predicted members of this sub-phenotype received haemodialysis. Sub-phenotype E was associated with community-acquired SAB, younger age, IDU, liver disease, and with endocarditis. In the Edinburgh cohort, the source of SAB in 32/37 predicted members of this sub-phenotype was IDU (categorised as 'other' source in the LCA since this category did not exist in the classification used in ARREST). In the ARREST cohort, SSTI is the source enriched in sub-phenotype E consistent with acquisition through IDU. In the Edinburgh cohort, 14/37 predicted members of this sub-phenotype had infected deep vein thrombophlebitis and 3/37 had an infected pseudoaneurysm.

The Infectious Diseases Society of America (IDSA)[26] and Fowler *et al*[27] definitions of complicated SAB were applied to the Edinburgh cohort, identifying a lower proportion of patients meeting both definitions of complicated SAB in sub-phenotype B (nosocomial intravenous catheter SAB) and a higher proportion in C (community-acquired metastatic SAB; **Supplementary Figure 3A**). Two current definitions of 'low risk' SAB were also applied to the Edinburgh cohort (the SABATO trial eligibility criteria[28,29] and the definition used by Hendriks *et al*[30]), with patients meeting these definitions predominantly predicted to belong to sub-phenotype B (**Supplementary Figures 3B and C**). The distribution of *spa* type inferred clonal complexes (**Supplementary Figure 3D**) did not differ substantially between sub-phenotypes in the Edinburgh cohort. Predicted members of sub-phenotypes A (older co-morbid SAB) and D (CKD SAB) in the Edinburgh cohort had the highest Charlson Comorbidity Index, with the predicted members of sub-phenotypes B and E having the lowest (**Supplementary Figure 3E**).

The SAFO trial included a substantially smaller number of participants than the Edinburgh and ARREST cohorts, applied more stringent inclusion/exclusion criteria (excluding MRSA infection, moderate-severe heart failure, and IDU), and did not record baseline CRP. Combining model fit parameters and interpretability, a four-class model was favoured. As expected, sub-phenotype E (IDU SAB) was not identified but the other four classes identified were similar to sub-phenotypes A-D identified in the Edinburgh and ARREST cohorts (**Figure 1**). SAFO sub-phenotype A was

associated with older age, vascular disease, and SAB originating from SSTI. SAFO sub-phenotype B was associated with younger patients with nosocomial SAB from intravenous catheter source. SAFO sub-phenotype C was associated with community-acquired metastatic SAB from an unknown source. SAFO sub-phenotype D was associated with healthcare-associated SAB, chronic kidney disease, and intravenous catheter source. Predicted members of SAFO sub-phenotypes A and D had the highest Charlson Comorbidity Index, and B had the lowest, consistent with the associations seen in the Edinburgh cohort (**Supplementary Figure 4**).

Clinical outcomes of SAB sub-phenotypes

Differences in 84-day mortality and microbiologic outcomes were observed between the sub-phenotypes (**Figure 2A; Supplementary Table 1**). In both the Edinburgh cohort and ARREST placebo arm (n=388), 84-day mortality was highest in sub-phenotype A and lowest in sub-phenotypes E and B. In the Edinburgh cohort, people assigned to sub-phenotypes A and D had the lowest one-year survival, whereas those assigned to E had the highest (**Supplementary Figure 5**). In the Edinburgh cohort, sub-phenotype C was associated with increased rates of persistent or recurrent bacteraemia (**Figure 2B**). In the ARREST placebo arm cohort, sub-phenotype C was also associated with a higher rate of composite microbiologic failure (**Figure 2C**). In both cohorts, sub-phenotype B was associated with lower rates of microbiologic failure. The smaller number of patients in the SAFO control arm (n=110) limited our ability to compare outcomes between the sub-phenotypes, but similar patterns were observed (**Supplementary Figure 6**). People assigned to sub-phenotypes A and D had higher 84-day mortality. Patients assigned to sub-phenotype B had the lowest mortality and lowest rate of persistent bacteraemia. Persistent bacteraemia at day 7 was uncommon but mostly occurred in sub-phenotype C.

Secondary analysis of the effect of adjunctive rifampicin treatment stratified by SAB sub-phenotype

An application of stratification of patients with SAB into sub-phenotypes is to enrich clinical trial design. Within the ARREST cohort, we considered each sub-phenotype separately and within each compared the effect of adjunctive rifampicin on 84-day mortality and composite microbiologic failure (**Figure 3; Supplementary Table 2**). Patients assigned to sub-phenotype B and randomised to adjunctive rifampicin had a higher 84-day mortality rate compared to patients randomised to placebo (odds ratio (OR) 18·8, 95% confidence interval (CI) 1·1–334·4, p=0·006). In sub-phenotype C, randomisation to adjunctive rifampicin was associated with reduced composite microbiologic failure (OR 0·17, 95% CI 0·04–0·8, p=0·02).

DISCUSSION

In hospitalised patients with predominantly MSSA bacteraemia, five sub-phenotypes can be identified using routinely-available clinical data. These sub-phenotypes differ in survival and

microbiologic outcomes. In a hypothesis generating secondary analysis of the ARREST trial, differential treatment effects were observed. Adjunctive rifampicin was associated with increased 84-day mortality in one sub-phenotype (nosocomial intravenous catheter SAB) and an improved microbiologic outcome in another (community-acquired metastatic SAB).

Our findings permit several observations about SAB from an unbiased standpoint. Sub-phenotype B (nosocomial intravenous catheter SAB) represents patients at low risk of adverse outcomes. One hundred and thirty two (28·8%) patients were predicted to belong to this sub-phenotype in the Edinburgh cohort, whereas 71 (15·5%) met the inclusion criteria for the SABATO trial[28,29] and 83 (18·1%) met the Hendriks et al definition of low-risk SAB[30], with the majority predicted to belong to sub-phenotype B. Sub-phenotype B could therefore represent a rational target for expanded investigation of earlier oral switch in SAB, providing a data-driven definition of 'low-risk' SAB. Furthermore, the risks of adjunctive agents might outweigh the limited potential to improve on already good outcomes, as exemplified by our finding that adjunctive rifampicin potentially caused increased mortality in this sub-phenotype. Inclusion of this sub-phenotype in trials of combination therapy should be done cautiously. Sub-phenotype E (IDU SAB) was associated with complicated disease but despite this, low mortality. Sub-phenotype C (community-acquired metastatic SAB) had complicated disease and worse microbiologic outcomes, but without clear predisposing factors. Patients in this sub-phenotype had a lower Charlson Comorbidity Index and generally lacked an obvious source for bacteraemia. The possible benefit of adjunctive rifampicin warrants further investigation in this sub-phenotype, in addition to alternative adjunctive agents including antimicrobials (e.g. clindamycin, currently being evaluated in the adjunctive treatment domain of the SNAP trial[31]) and anti-staphylococcal lysins (Exebacase)[32]. These sub-phenotypes also provide a framework for investigation of immunobiology in SAB, and could facilitate identification of treatable traits, for example defective phagocyte responses that could be therapeutically re-calibrated[11].

Our study has several strengths. The sub-phenotypes were replicated in analysis of an observational cohort and a large trial cohort with permissive inclusion criteria. Four of the sub-phenotypes could also be identified in a smaller trial with more restrictive inclusion criteria. Trial populations of SAB differ from real-life cohorts, including patient characteristics and mortality rates[9,29]. It is therefore re-assuring that despite the differences between the cohorts (**Table 1**), the core features of the identified sub-phenotypes were reproducible, suggesting the findings are generalisable. Outcomes differed between sub-phenotypes but were not included as class-defining variables, and overall the association between sub-phenotype and outcome was consistent across the cohorts. To allow prospective sub-phenotype prediction of individual patients, future work will aim to identify a sub-set of variables that can be used as predictive markers of sub-phenotype membership.

Our study has important limitations. First, despite using model parameters such as BIC there is a degree of subjectivity with the class selection based on clinical interpretability. Second, the class-defining variables included were restricted to routinely available clinical data. Inclusion of

inflammation biomarkers could provide biological insights. Third, the included cohorts were from countries with a low prevalence of MRSA. The USA300 MRSA clone is prevalent in the USA and independently associated with metastatic disease[3]. Replication in a cohort with higher MRSA prevalence will be required. Fourth, the cohorts differed in inclusion criteria and the variables available for analysis. Fifth, the definitions of microbiologic outcomes used in the cohorts were different, preventing direct comparison. Sixth, the Edinburgh cohort was a retrospective observational study, without structured prospective monitoring of microbiologic outcomes. Detection of persistent or recurrent SAB was opportunistic, relying on healthcare attendance and blood cultures being taken, so is likely to be subject to ascertainment bias and under-ascertainment of these outcomes. Finally, although receipt of adjunctive rifampicin was randomised in the ARREST trial, reducing the risk of confounding, the numbers within each sub-phenotype were relatively small so these results must be interpreted as strictly hypothesis-generating. Overall, it remains possible the sub-phenotypes will not be replicable in other patient cohorts, or that additional sub-phenotypes may exist (e.g. associated with MRSA infection), or that outcomes/treatment responses could differ. We are conducting further replication studies to address these questions.

In summary, our findings support the hypothesis that clinically-relevant sub-phenotypes do exist within SAB, and suggest that patient stratification within SAB clinical trials is required to identify strategies to improve outcomes for patients. This could inform a personalised medicine approach to SAB.

NOTES

Funding sources

ARREST was funded by the National Institutes of Health Research's Health Technology Assessment (Project number 10/104/2). The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research, or the Department of Health. The MRC Clinical Trials Unit at UCL is supported by funding from the MRC (MC_UU_00004/05). MP received a research grant from Laboratorios ERN. VGF was supported by Grant R01-AI165671 from the National Institutes of Health. This work was supported by the AMR cross-council funding from the MRC to the SHIELD consortium "Optimising Innate Host Defence to Combat Antimicrobial Resistance" (PI: DHD). CDR was supported by an Edinburgh Clinical Academic Track (ECAT)/Wellcome Trust PhD Training Fellowship for Clinicians award (214178/Z/18/Z). This research was funded in part by Wellcome 214178/Z/18/Z. For the purpose of open access, the author has applied a creative commons attribution (CC BY) licence to any author accepted manuscript version arising.

Potential conflicts of interest

CT received speaker fees from Gedeon Richter. VGF reports Grants/research support: Astra Zeneca; MedImmune; Merck; ContraFect, Karius, Genentech, Regeneron, Basilea. Paid Consultant: Astra Zeneca; GSK; Armata; Debiopharm; Genentech; Basilea Affinergy, Janssen, ContraFect, Destiny. Royalties: UpToDate.

Author contributions

CDR conceived the study. CDR, GET, MP, JNP, CT, JC, AS, KB, GC, WS, RSL, TYG, WD, AB, EH, LT curated the data. MCS, ZB, ACW, GHG, CDR did the investigation. MCS, ZB, CDR did the formal analysis. CDR, MCS and GHG supervised the study. MCS and CDR wrote the original draft of the manuscript. All authors reviewed and edited the manuscript. CDR and MCS validated the study data. MCS, CDR, AS, GET, MP, JNP, CT, JC had access to the raw data. The corresponding author had full access to all the data and final responsibility for the decision to submit for publication.

Data sharing

Please contact the corresponding author to discuss access to the dataset. We would welcome opportunities to share data and contribute to collaborative analyses.

Acknowledgements

This work used data provided by patients and collected by the NHS as part of their care. We gratefully acknowledge the work of the ARREST and SAFO trial teams (listed in the Supplement).

REFERENCES

1. Wilson R, Hamburger M. Fifteen years' experience with staphylococcus septicemia in a large city hospital; analysis of fifty-five cases in the Cincinnati General Hospital 1940 to 1954. *Am J Med*. 1957;22(3):437-457.
2. Verway M, Brown KA, Marchand-Austin A, et al. Prevalence and Mortality Associated with Bloodstream Organisms: a Population-Wide Retrospective Cohort Study. *J Clin Microbiol.* 2022;60(4):e0242921.
3. Souli M, Ruffin F, Choi S-H, et al. Changing Characteristics of *Staphylococcus aureus* Bacteremia: Results From a 21-Year, Prospective, Longitudinal Study. *Clinical Infectious Diseases.* 2019;69(11):1868-1877.
4. Choi SH, Dagher M, Ruffin F, et al. Risk Factors for Recurrent *Staphylococcus aureus* Bacteremia. *Clin Infect Dis.* 2021;72(11):1891-1899.
5. Bai AD, Lo CKL, Komorowski AS, et al. *Staphylococcus aureus* bacteraemia mortality: a systematic review and meta-analysis. *Clinical Microbiology and Infection.* 2022;28(8):1076-1084.
6. Westgeest AC, Ruffin F, Kair JL, et al. The association of female sex with management and mortality in patients with *Staphylococcus aureus* bacteraemia. *Clin Microbiol Infect.* 2023;29(9):1182-1187.
7. Global mortality associated with 33 bacterial pathogens in 2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet.* 2022;400(10369):2221-2248.

8. Grillo S, Pujol M, Miró JM, et al. Cloxacillin plus fosfomycin versus cloxacillin alone for methicillin-susceptible *Staphylococcus aureus* bacteremia: a randomized trial. *Nat Med.* 2023;29(10):2518-2525.
9. Dolby HW, Clifford SA, Laurenson IF, Fowler VG, Russell CD. Heterogeneity in *Staphylococcus aureus* Bacteraemia Clinical Trials Complicates Interpretation of Findings. *J Infect Dis.* 2022;226(4):723-728.
10. Thwaites GE, Scarborough M, Szubert A, et al. Adjunctive rifampicin for *Staphylococcus aureus* bacteraemia (ARREST): a multicentre, randomised, double-blind, placebo-controlled trial. *Lancet.* 2018;391(10121):668-678.
11. Russell CD, Baillie JK. Treatable traits and therapeutic targets: Goals for systems biology in infectious disease. *Curr Opin Syst Biol.* 2017;2:140-146.
12. Calfee CS, Delucchi K, Parsons PE, Thompson BT, Ware LB, Matthay MA. Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. *The Lancet Respiratory Medicine.* 2014;2(8):611-620.
13. Green RH, Brightling CE, Bradding P. The reclassification of asthma based on subphenotypes. *Curr Opin Allergy Clin Immunol.* 2007;7(1):43-50.
14. Agusti A, Bel E, Thomas M, et al. Treatable traits: toward precision medicine of chronic airway diseases. *European Respiratory Journal.* 2016;47(2):410-419.
15. Aliberti S, Lonni S, Dore S, et al. Clinical phenotypes in adult patients with bronchiectasis. *European Respiratory Journal.* 2016;47(4):1113-1122.
16. Wildi K, Livingstone S, Palmieri C, LiBassi G, Suen J, Fraser J. The discovery of biological subphenotypes in ARDS: a novel approach to targeted medicine? *J Intensive Care.* 2021;9(1):14.
17. Russell CD, Berry K, Cooper G, et al. Distinct clinical endpoints of *Staphylococcus aureus* bacteraemia complicate assessment of outcome. *Clin Infect Dis*. 2024.
18. Friedman ND, Kaye KS, Stout JE, et al. Health care--associated bloodstream infections in adults: a reason to change the accepted definition of community-acquired infections. *Ann Intern Med.* 2002;137(10):791-797.
19. Vermunt JK, Magidson J. Latent class analysis. *The sage encyclopedia of social sciences research methods.* 2004;2:549--553.
20. Sinha P, Calfee CS, Delucchi KL. Practitioner's Guide to Latent Class Analysis: Methodological Considerations and Common Pitfalls. *Crit Care Med.* 2021;49(1):e63-e79.
21. Weller BE, Bowen NK, Faubert SJ. Latent Class Analysis: A Guide to Best Practice. *Journal of Black Psychology.* 2020;46(4):287-311.
22. Vermunt JK. Latent Class Modeling with Covariates: Two Improved Three-Step Approaches. *Political Analysis.* 2010;18(4):450-469.
23. Vermunt JK, Magidson J. LG-syntax user's guide: Manual for latent gold syntax module version 6.0. 2021. Arlington, MA: Statistical Innovations Inc
24. Kassambara A, Kosinski M, Biecek P. survminer: Drawing Survival Curves using 'ggplot2'. R package version 0.4.9. 2021.
25. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.
26. Liu C, Bayer A, Cosgrove SE, et al. Clinical practice guidelines by the infectious diseases society of america for the treatment of methicillin-resistant *Staphylococcus aureus* infections in adults and children: executive summary. *Clin Infect Dis.* 2011;52(3):285-292.

27. Fowler VG, Jr., Olsen MK, Corey GR, et al. Clinical identifiers of complicated *Staphylococcus aureus* bacteremia. *Arch Intern Med.* 2003;163(17):2066-2072.
28. Kaasch AJ, López-Cortés LE, Rodríguez-Baño J, et al. Efficacy and safety of an early oral switch in low-risk *Staphylococcus aureus* bloodstream infection (SABATO): an international, open-label, parallel-group, randomised, controlled, non-inferiority trial. *Lancet Infect Dis.* 2024.
29. Cooper G, Dolby HW, Berry K, Russell CD. Eligibility of patients with *Staphylococcus aureus* bacteraemia for early oral switch. *Lancet Infect Dis.* 2024.
30. Hendriks MMC, Schweren KSA, Kleij A, et al. Low-risk *Staphylococcus aureus* bacteremia patients do not require routine diagnostic imaging: A multicenter retrospective cohort study. *Clin Infect Dis.* 2024.
31. Anpalagan K, Dotel R, MacFadden DR, et al. Does adjunctive clindamycin have a role in *Staphylococcus aureus* bacteremia? A protocol for the adjunctive treatment domain of the *S. aureus* Network Adaptive Platform (SNAP) randomized controlled trial. *Clin Infect Dis*. 2024.
32. Fowler VG, Jr., Das AF, Lipka-Diamond J, et al. Exebacase in Addition to Standard-of-Care Antibiotics for *Staphylococcus aureus* Bloodstream Infections and Right-Sided Infective Endocarditis: A Phase 3, Superiority-Design, Placebo-Controlled, Randomized Clinical Trial (DISRUPT). *Clin Infect Dis.* 2024.

TABLES AND FIGURES

Table 1: Characteristics of included patients

	Edinburgh cohort (n=458)	ARREST cohort (n=758)	SAFO cohort (n=214)	P-value
Age, years	68 (52-79)	65 (50-76)	65 (54-75)	0.08
Sex, male	292 (53·2)	NA[1]	150 (70·1)	0.11
Acquisition				<0.0001
Community acquired	181 (39·5)	485 (64·1)	78 (36·4)	
Healthcare associated	110 (24·0)	140 (18·5)	52 (24·3)	
Nosocomial	167 (36·5)	132 (17·4)	84 (39·3)	
Comorbidities				
Dementia	44 (9·6)	31 (4·1)	7 (3·3)	0.0002
Chronic kidney disease	35 (7·6)	138 (18·3)	18 (8·4)	<0.0001
Liver disease[2]	57 (12·4)	56 (7·4)	13 (6·1)	0.004
Vascular disease[3]	121 (26·4)	NA	59 (27·6)	0.77
Prosthetic cardiac material[4]	46 (10·0)	NA	16 (7·5)	0.31
Injection drug use	41 (9·0)	83 (11·1)	0	<0.0001
Vital signs				
Heart rate, beats per minute	98 (85-110)	94 (82-107)	NA	0.005
Temperature, °C	38.2 (37.6-38.8)	37.0 (37.0-38.0)	37.3 (36.5-38.3)	<0.0001
Laboratory measurements				
Haemoglobin, g/L	114 (100-129)	107 (93-122)	NA	<0.0001
Creatinine, µmol/L	90 (65-138)	80 (61-131)	80 (62-123)	0.05
C-reactive protein, mg/L	156 (69-273)	150 (87-218)	NA	0.15

	Edinburgh cohort (n=458)	ARREST cohort (n=758)	SAFO cohort (n=214)	P-value
SAB characteristics				
MRSA	17 (3·7)	47 (6·2)	0	<0.0001
Infective endocarditis	35 (7·6)	40 (5·3)	15 (7·0)	0.21
Other metastatic foci[5]	99 (21·6)	203 (26·8)	49 (22·9)	0.11
Source of bacteraemia				<0.0001
Unknown	163 (35·6)	221 (29·2)	70 (32·7)	
Intravenous catheter	93 (20·3)	141 (18·6)	68 (31·8)	
Skin or soft tissue infection	88 (19·2)	293 (38·7)	39 (18·2)	
Other	61 (13·3)[6]	55 (7·3)[7]	20 (9·3)	
Respiratory	28 (6·1)	29 (3·8)	4 (1·9)	
Urine	25 (5·5)	19 (2·5)	13 (6·1)	
All-cause 84-day mortality	121 (26·4)	56/388 (14·4)[8]	17/110 (15·5)[8]	<0.0001

[1]Not available due to participant deidentification

[2]people with Child Pugh C liver cirrhosis were excluded from the SAFO trial

[3]peripheral vascular disease, myocardial infarction or stroke

[4]implanted cardiac devices, including pacemakers and implantable automatic cardioverter-defibrillator and Left Ventricular Assist Devices, but not including prosthetic heart valves

[5]vertebral osteomyelitis, epidural abscess, native joint septic arthritis, prosthetic joint infection, deep tissue abscess.

[6]injection drug use and bone classified as 'other'

[7]'other' sources not specified

[8]data shown for trial control arms

Continuous values are shown as median (interquartile range). Categorical variables are shown as count (%). Variables not available in the ARREST dataset are represented as NA. Vital signs and laboratory measurements were recorded at the time of the index blood culture in the Edinburgh and SAFO cohorts. In the ARREST trial, baseline laboratory measurements were those closest to randomisation (preceding 4 days or 1 day post randomisation) and for vital signs, the highest value within 24h of randomisation was taken.

MRSA: methicillin-resistant *Staphylococcus aureus*

Table 2: Model fit statistics

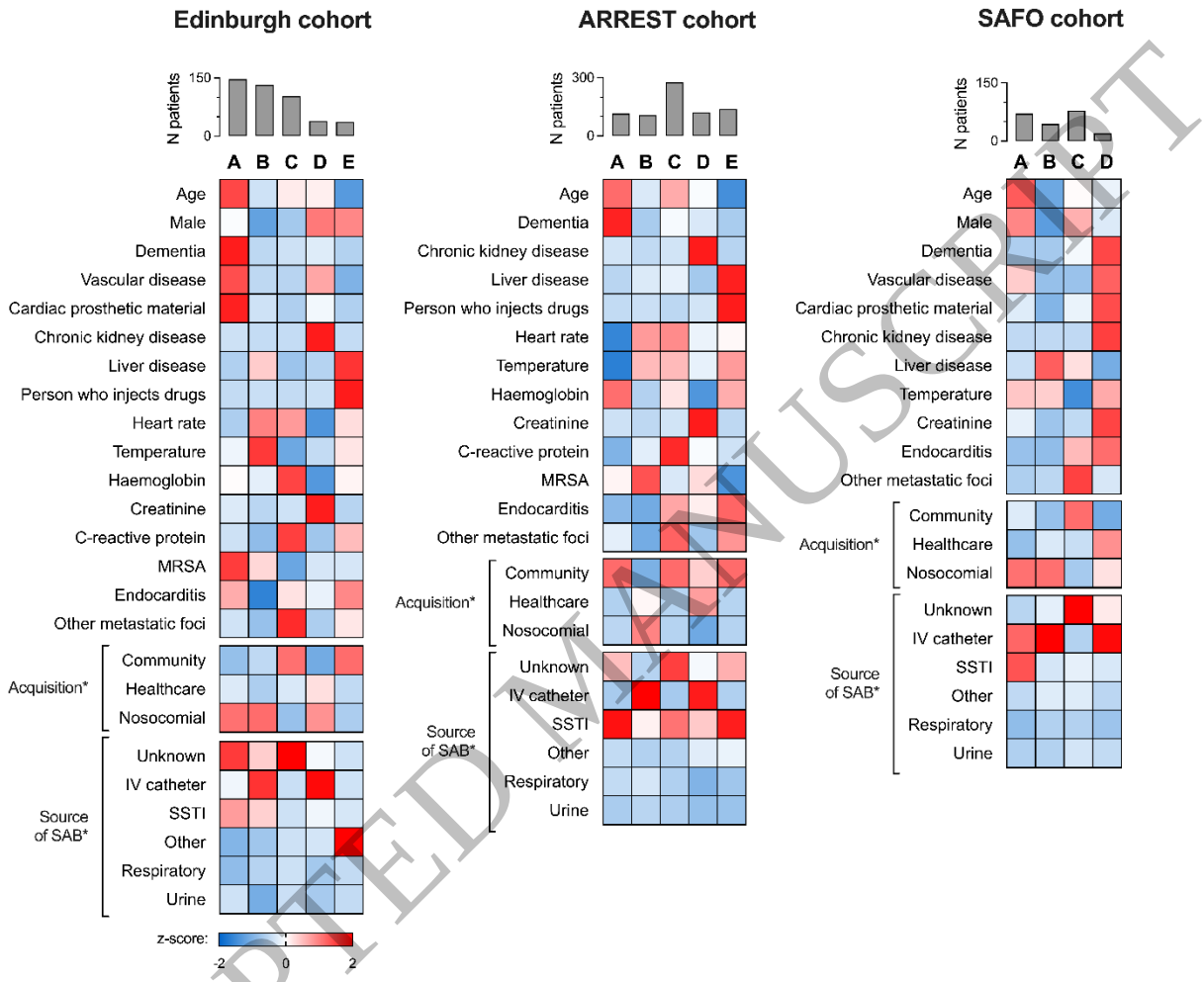
Model	BIC	LL	Npar	Entropy	Patients per class						
Edinburgh cohort											
					1	**2**	**3**	**4**	**5**	**6**	**7**
1-class	14578·4	-7200·3	29	1	458	–	–	–	–	–	–
2-class	14167·1	-6902·8	59	0·8	307	151	–	–	–	–	–
3-class	13968·5	-6711·6	89	0·8	173	152	133	–	–	–	–
4-class	13901·7	-6586·3	119	0·8	162	128	127	41	–	–	–
5-class	13829·1	-6458·1	149	0·9	147	132	103	39	37	–	–
6-class	13869·1	-6383·2	149	0·9	130	123	83	42	42	38	–
7-class	13980·5	-6350·0	209	0·9	113	103	65	57	42	41	37
ARREST cohort											
					1	**2**	**3**	**4**	**5**	**6**	**7**
1-class	19279·6	-10157·3	26	1	758	–	–	–	–	–	–
2-class	18636·4	-9741·5	53	0·8	576	182	–	–	–	–	–
3-class	18360·2	-9515·8	80	0·8	410	232	116	–	–	–	–
4-class	18284·9	-9374·1	107	0·8	284	197	162	115	–	–	–
5-class	18269·8	-9259·3	134	0·8	276	139	121	115	107	–	–

6-class	19409·7	-9171·1	161	0·8	171	153	119	116	106	93	–
7-class	19462·4	-9107·9	188	0·8	172	131	117	112	88	78	60
SAFO cohort											
					1	**2**	**3**	**4**	**5**	**6**	**7**
1-class	2065·0	-976·1	21	1	214	–	–	–	–	–	–
2-class	1963·3	-866·3	43	0·8	125	89	–	–	–	–	–
3-class	1945·9	-798·6	65	0·9	125	67	22	–	–	–	–
4-class	1981·6	-757·4	87	0·8	79	71	44	20	–	–	–
5-class	2035·2	-725·1	109	0·9	59	54	45	36	20	–	–
6-class	2121·1	-709·1	131	0·9	58	53	38	34	18	13	–
7-class	2174·1	-676·5	153	0·9	51	40	36	37	23	20	7

Model fit statistics for latent class models from one to seven classes in the Edinburgh, ARREST, and SAFO cohorts. BIC: Bayesian Information Criteria, defined in Methods. LL: Log-likelihood, measures the fit of the model to the data. Npar: number of parameters, measure of model complexity. Entropy is a measure for class separation: it ranges from zero to one and values of ≥0·8 indicate good separation of the different classes.

FIGURE LEGENDS

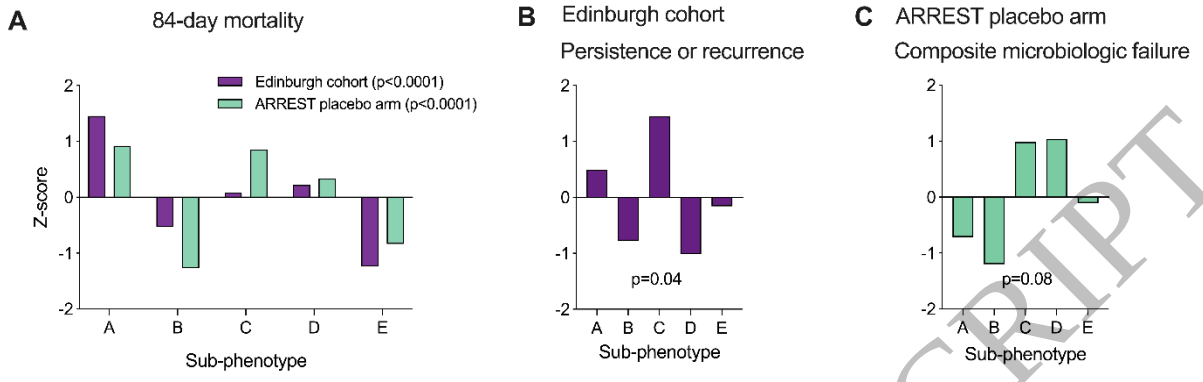
Figure 1: Comparison of class-defining variables between SAB sub-phenotypes



Vertical bars show the number of patients assigned to each sub-phenotype. Cells are shaded according to row z-score (i.e. comparing sub-phenotypes) except 'Acquisition' and 'Source of SAB', where shading is by column z-score (i.e. comparing within each sub-phenotype). Intensity of red shading reflects a more positive z-score (i.e. above the mean) and intensity of blue shading reflects a more negative z-score (i.e. below the mean). In the SAFO trial, people with Child Pugh C liver cirrhosis, MRSA infection, and people who inject drugs were not recruited.

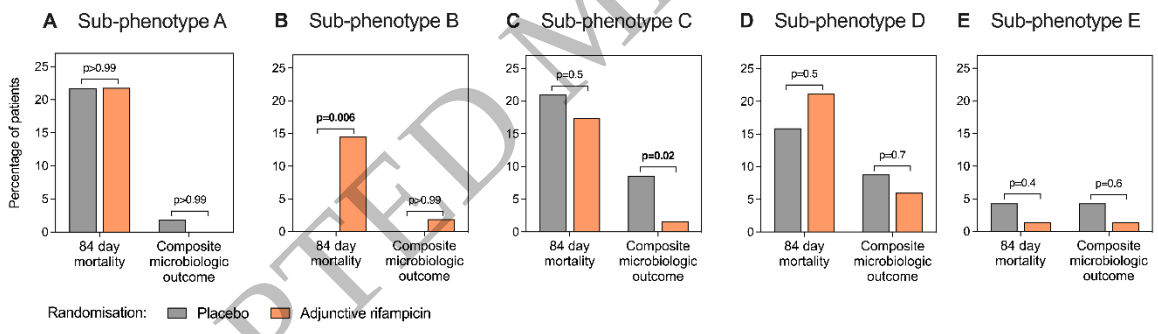
IV: intravenous; SSTI: skin or soft tissue infection; MRSA: methicillin-resistant *S. aureus*.

Figure 2: Comparison of outcomes between SAB sub-phenotypes



Comparison of **(A)** all-cause 84-day mortality, **(B)** persistent or recurrent bacteraemia in the Edinburgh retrospective observational cohort, and **(C)** composite microbiologic failure in the ARREST trial placebo arm. Bars represent z-scores comparing the outcome between sub-phenotypes within the same cohort. Differences in the proportion of patients with each outcome between sub-phenotypes were compared using Fisher's Exact test or Chi-squared test.

Figure 3: Effect of adjunctive rifampicin in SAB sub-phenotypes



Comparison of outcomes of patients randomised to placebo or adjunctive rifampicin when SAB sub-phenotypes considered separately. Treatment outcomes within each sub-phenotype were compared using Fisher's exact test. Two comparisons were made within each sub-phenotype so the significance level was set at 0·025 (α=0·05, n=2).

Dovato
dolutegravir/lamivudine

EVIDENCE SUPPORTS THE HIGH BARRIER TO RESISTANCE OF DOVATO UP TO 5 YEARS[1-3]

CONFIDENCE IN DOVATO ACROSS TREATMENT SETTINGS[4–9]

Treatment-naïve resistance rates, with up to **3 years of evidence**[5–7]

0% (n=0/1,885)*[,4]
REAL-WORLD EVIDENCE

0.1% (n=1/953)**[,†,‡,§,5–7]
RANDOMISED CONTROLLED TRIALS

Treatment-experienced resistance rates, with up to **5 years of evidence**[1–3]

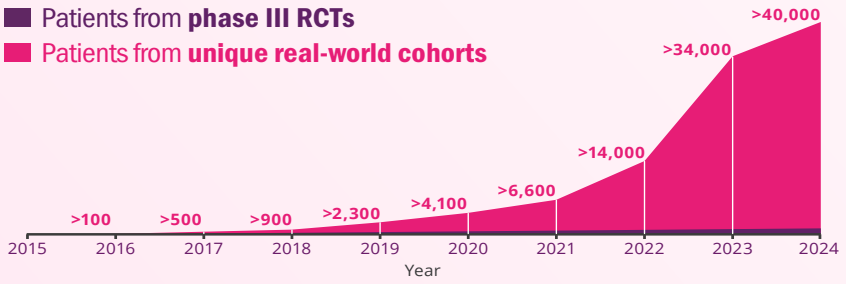
0.03% (n=10/35,888)*[,4]
REAL-WORLD EVIDENCE

0% (n=0/615)[||,¶,#,8,9]
RANDOMISED CONTROLLED TRIALS

>300,000 PEOPLE LIVING WITH HIV HAVE BEEN TREATED WITH DOVATO GLOBALLY[10]

DOVATO is supported by a wealth of evidence, with the outcomes of **>40,000** people living with HIV captured within clinical trials and real-world evidence, including those with:[4–9,11,12]

- **NO PRIOR TREATMENT EXPERIENCE**[13]
- **NO BASELINE RESISTANCE TESTING**[13]
- **HIGH BASELINE VIRAL LOAD** (>100,000 copies/mL and even >1M copies/mL)[6,13]
- **LOW CD4 + COUNT** (≤200 cells/mm³)[13]



IS IT TIME TO RECONSIDER THE VALUE OF THE 2ND NRTI?

LEARN MORE

DOVATO is indicated for the treatment of Human Immunodeficiency Virus type 1 (HIV-1) infection in adults and adolescents above 12 years of age weighing at least 40 kg, with no known or suspected resistance to the integrase inhibitor class, or lamivudine.[13]

Adverse events should be reported. Reporting forms and information can be found at **https://yellowcard.mhra.gov.uk/** or search for MHRA Yellowcard in the Google Play or Apple App store. Adverse events should also be reported to GSK on 0800 221441

REFERENCES

1. Maggiolo F et al. BMC Infect Dis 2022; 22(1): 782.
2. Taramasso L et al. AIDS Patient Care STDS 2021; 35(9): 342–353.
3. Ciccullo A et al. JAIDS 2021; 88(3): 234–237.
4. ViiV Healthcare. Data on File. REF-223795. 2024.
5. Cahn P et al. AIDS 2022; 36(1): 39–48.
6. Rolle C et al. Open Forum Infect Dis 2023; 10(3): ofad101.
7. Cordova E et al. Poster presented at 12th IAS Conference on HIV Science. 23–26 July 2023. Brisbane, Australia. TUPEB02.
8. De Wit S et al. Slides presented at HIV Glasgow. 23–26 October 2022. Virtual and Glasgow, UK. M041.
9. Llibre J et al. Clin Infect Dis 2023; 76(4): 720–729.
10. ViiV Healthcare. Data on File. REF-220949. 2024.
11. Rolle C et al. Poster presented IDWeek. 11–15 October 2023. Virtual and Boston, USA. 1603.
12. Slim J et al. Abstract presented IDWeek. 11–15 October 2023. Virtual and Boston, USA. 1593.
13. DOVATO. Summary of Product Characteristics. June 2023.

PRESCRIBING INFORMATION

Dovato Prescribing Information

Legal Notices

Privacy Policy

Contact Us

ABBREVIATIONS

3TC, lamivudine; **CD4,** cluster of differentiation 4; **DTG,** dolutegravir; **FDA,** United States Food and Drug Administration; **FTC,** emtricitabine; **HIV,** human immunodeficiency virus; **ITT-E,** intention-to-treat exposed; **NRTI,** nucleoside/nucleotide reverse transcriptase inhibitor; **RCT,** randomised controlled trial; **RNA,** ribonucleic acid; **TAF,** tenofovir alafenamide fumarate; **TDF,** tenofovir disoproxil fumarate; **XTC,** emtricitabine.

FOOTNOTES

*Data extracted from a systematic literature review of DTG+3TC real-world evidence. Overlap between cohorts cannot be fully excluded.

**The reported rate reflects the sum-total of resistance cases calculated from GEMINI I and II (n=1/716, through 144 weeks), STAT (n=0/131, through 52 weeks), and D2ARLING (n=0/106, through 24 weeks).[5–7]

†GEMINI I and II are two identical 148-week, phase III, randomised, double-blind, multicentre, parallel-group, non-inferiority, controlled clinical trials testing the efficacy of DTG/3TC in treatment-naïve patients. Participants with screening HIV-1 RNA ≤500,000 copies/mL were randomised 1:1 to once-daily DTG/3TC (n=716, pooled) or DTG + TDF/FTC (n=717, pooled). The primary endpoint of each GEMINI study was the proportion of participants with plasma HIV-1 RNA <50 copies/mL at Week 48 (ITT-E population, snapshot algorithm).[13]

‡STAT is a phase IIIb, open-label, 48-week, single-arm pilot study evaluating the feasibility, efficacy, and safety of DTG/3TC in 131 newly diagnosed HIV-1 infected adults as a first line regimen. The primary endpoint was the proportion of participants with plasma HIV-1 RNA <50 copies/mL at Week 24.[6]

§D2ARLING is a randomised, open-label, phase IV study designed to assess the efficacy and safety of DTG/3TC in treatment-naïve people with HIV with no available baseline HIV-1 resistance testing. Participants were randomised in a 1:1 ratio to receive DTG/3TC (n=106) or DTG + TDF/XTC (n=108). The primary endpoint was the proportion of participants with plasma HIV-1 RNA <50 copies/mL at Week 48.[7] Results at week 24 of the study.

||The reported rate reflects the sum-total of resistance cases calculated from TANGO (n=0/369, through 196 weeks) and SALSA (n=0/246, through 48 weeks).[8,9]

¶TANGO is a randomised, open-label, trial testing the efficacy of DOVATO in virologically suppressed patients. Participants were randomised in a 1:1 ratio to receive DOVATO (n=369) or continue with TAF-containing regimens (n=372) for up to 200 weeks. At Week 148, 298 of those on TAF-based regimens switched to DOVATO. The primary efficacy endpoint was the proportion of subjects with plasma HIV-1 RNA ≥50 copies/mL (virologic non-response) as per the FDA Snapshot category at Week 48 (adjusted for randomisation stratification factor).[8,13]

#SALSA is a phase III, randomised, open-label, non-inferiority clinical trial evaluating the efficacy and safety of switching to DTG/3TC compared with continuing current antiretroviral regimens in virologically suppressed adults with HIV. Eligible participants were randomised 1:1 to switch to once-daily DTG/3TC (n=246) or continue current antiretroviral regimens (n=247). The primary endpoint was the proportion of subjects with plasma HIV-1 RNA ≥50 copies/mL at Week 48 (ITT-E population, snapshot algorithm).[9]

ViiV Healthcare, 980 Great West Road, Brentford, Middlesex, London, UK.

ViiV Healthcare

PM-GBL-DLL-BNNR-240004 May 2024