


UPDATE

Open Access



Statistical analysis plan for the Stepped-wedge Cluster Randomized trial of Electronic Early Notification of sepsis in hospitalized ward patients (SCREEN)

Yaseen M. Arabi^{1*} , Ramesh Kumar Vishwakarma^{2,3}, Hasan M. Al-Dorzi¹, Eman Al Qasim¹, Sheryl Ann Abdukahil¹, Fawaz K. Al-Rabeah⁴, Huda Al Ghamdi⁴, Ebtisam Al Ghamdi⁴ and the SCREEN Trial Group

Abstract

Background: It is unclear whether screening for sepsis using an electronic alert in hospitalized ward patients improves outcomes. The objective of the Stepped-wedge Cluster Randomized Trial of Electronic Early Notification of Sepsis in Hospitalized Ward Patients (SCREEN) trial is to evaluate whether an electronic screening for sepsis compared to no screening among hospitalized ward patients reduces all-cause 90-day in-hospital mortality.

Methods and design: This study is designed as a stepped-wedge cluster randomized trial in which the unit of randomization or cluster is the hospital ward. An electronic alert for sepsis was developed in the electronic medical record (EMR), with the feature of being active (visible to treating team) or masked (inactive in EMR frontend for the treating team but active in the backend of the EMR). Forty-five clusters in 5 hospitals are randomized into 9 sequences of 5 clusters each to receive the intervention (active alert) over 10 periods, 2 months each, the first being the baseline period. Data are extracted from EMR and are compared between the intervention (active alert) and control group (masked alert). During the study period, some of the hospital wards were allocated to manage patients with COVID-19. The primary outcome of all-cause hospital mortality by day 90 will be compared using a generalized linear mixed model with a binary distribution and a log-link function to estimate the relative risk as a measure of effect. We will include two levels of random effects to account for nested clustering within wards and periods and two levels of fixed effects: hospitals and COVID-19 ward status in addition to the intervention. Results will be expressed as relative risk with a 95% confidence interval.

Conclusion: The SCREEN trial provides an opportunity for a novel trial design and analysis of routinely collected and entered data to evaluate the effectiveness of an intervention (alert) for a common medical problem (sepsis in ward patients). In this statistical analysis plan, we outline details of the planned analyses in advance of trial completion. Prior specification of the statistical methods and outcome analysis will facilitate unbiased analyses of these important clinical data.

* Correspondence: yaseenarabi@yahoo.com

¹College of Medicine, King Saud Bin Abdulaziz University for Health Sciences, King Abdullah International Medical Research Center, Intensive Care Department, King Abdulaziz Medical City, Ministry of National Guard Health Affairs, Riyadh, Saudi Arabia

Full list of author information is available at the end of the article



© 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.

Trial registration: ClinicalTrials.gov [NCT04078594](https://clinicaltrials.gov/ct2/show/study/NCT04078594). Registered on September 6, 2019

Keywords: Sepsis, Alert, Screening, qSOFA, Mortality, Electronic medical records

Background

Sepsis is a major cause of morbidity and mortality among hospitalized patients. Sepsis outcome is greatly dependent on the time-sensitive administration of appropriate antimicrobials, fluid resuscitation, and source control [1–3]. Screening for sepsis using an electronic alert in hospitalized patients may improve outcomes by early sepsis recognition and timely implementation of appropriate care processes. However, the evidence for such an intervention is modest [4], and a randomized trial is needed to measure its true effect.

The objective of the Stepped-wedge Cluster Randomized Trial of Electronic Early Notification of Sepsis in Hospitalized Ward Patients (SCREEN) is to evaluate whether electronic screening for sepsis compared to no screening among hospitalized ward patients reduces all-cause 90-day in-hospital mortality [5]. The electronic screening for sepsis is based on the quick Sequential Organ Failure Assessment qSOFA [6].

In this manuscript, we describe the statistical analysis plan (SAP) of the SCREEN trial. This SAP complies with the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use and both the “Statistical principles for clinical trials E9” and “Structure and content of clinical study reports E3” [7, 8]. The final study report will follow the Consolidated Standards of Reporting Trials (CONSORT) 2010 guidelines for reporting randomized controlled trials and the CONSORT extension for stepped-wedge cluster randomized trials [9–11]. This SAP identifies the procedures to be applied to the primary and secondary analyses for the trial cohort once trial data are complete. The SAP was finalized during trial implementation, and all analyses were prospectively defined.

Methods

Study design

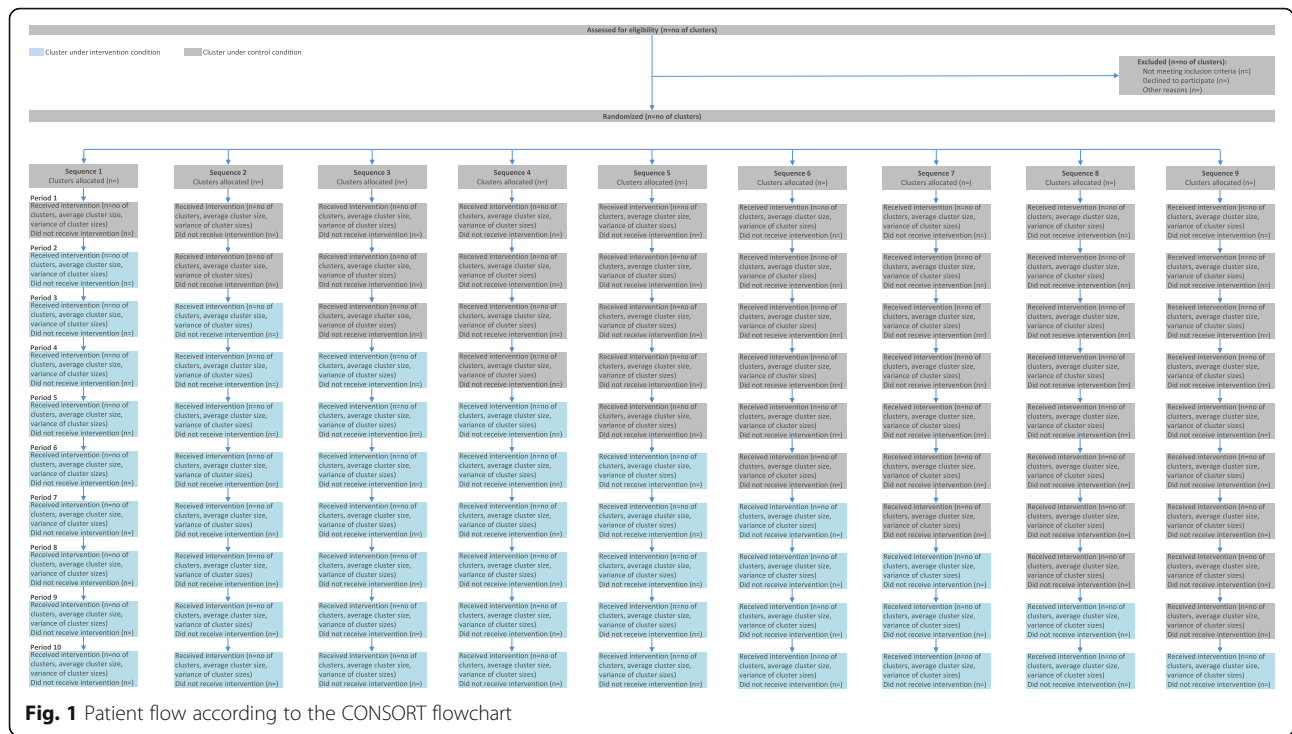
The study is conducted in the 5 Ministry of National Guard Health Affairs (MNGHA) hospitals which share the same electronic medical record (EMR) system (BESTCare, South Korea). This study is designed as a stepped-wedge cluster randomized trial, which allows to sequentially deliver the study intervention to all trial clusters over a number of periods. We present a glossary of terms in line with the CONSORT extension for stepped-wedge cluster randomized trials [11] in Table 1. The cluster refers to the unit of randomization, which is the hospital ward, and we will refer to it in the rest of the document as “ward.” A list of ward-level and patient-level eligibility criteria is outlined in the study protocol [5]. Wards are randomized into 9 sequences of 5 wards each to receive the intervention. After a baseline period of 2 months, the intervention is implemented in a new sequence of five new randomly selected wards, until it is eventually implemented in all 9 sequences (45 wards) (Fig. 1). A computer-generated non-stratified randomly allocated concealed list determines the order in which the wards receive the intervention. The randomization list was maintained with a research coordinator who was not involved in this trial, and the ward allocation remained concealed from the research and clinical teams throughout the study and was revealed for a given sequence only 1 month before the implementation of the intervention to allow training.

Intervention and control groups

The intervention included the implementation of an electronic alert system and the associated training, feedback, and audit. The alert was developed in the electronic medical record (EMR), with the feature of being

Table 1 Glossary of terms used in the current report, based on the CONSORT extension for cluster randomized trials [11]

Term	Definitions [11]	Applicability to SCREEN trial
Cluster	The unit of randomization	Hospital ward
Control condition	The comparator treatment	Masked alert
Duration of period	Time (e.g., months) between each step	10 periods, each is 2 months
Intervention condition	The treatment under evaluation	Active alert
Sequence	A sequence of codes defining the order of implementation of the treatment conditions for each cluster. More than one cluster can be allocated to each sequence	9 sequences, with 5 clusters in each sequence
Step	A planned point at which a cluster or a group of clusters crosses from control to intervention	The period between two steps is 2 months



active (visible to the treating team) or masked (inactive in EMR frontend for the treating team but active in the backend of the EMR). Once a ward is randomized to the intervention group, the alert system was activated. The intervention group constitutes of patients admitted to the wards with the active alert, and the control group constitutes of patients admitted to the wards with masked alert.

In the intervention group, once a patient meets the qSOFA criteria, an alert appears in EMR as a pop-message and appears also on a hand-held device carried by the charge nurse of the corresponding ward. The alert prompts the nurse to notify the physician and prompts the physician to assess the patient for possible sepsis. At the beginning of the study, we launched a hospital-wide sepsis awareness campaign in all 5 participating hospitals focusing on the importance of timely interventions for sepsis. We also provided in-service training sessions to the related medical and nursing departments at the beginning of the study and before implementation in each ward. We conducted regular webinars with the leaders of active wards. We also created an intranet page with educational resources (videos, presentations, documents, posters, and related links) that explained the project and provided clinical guidance and resources. We developed a dashboard to display data for each active ward on the number of alerts and the percentage of acknowledged alerts and time to acknowledgment by nurses and physicians. We have set 15 min as an overstretched target for

nursing acknowledgment and 30 min for physician acknowledgment, and we have tracked these indicators on the electronic dashboards. The data on the dashboards have been made available in real-time for nursing and medical managers of each ward. In addition, the same data reports are generated every 2 weeks and shared with the teams in the active wards. The hospital administration and quality management department are engaged in the process of feedback and of improving the response time to alerts. The bedside management and assessment are at the discretion of the treating team, and the project has adopted the 2016 Surviving Sepsis Campaign guidelines and the hour-1 bundle, but without specific monitoring of bundle compliance [2, 12].

Study population

The study has been conducted in the 5 Ministry of National Guard Health Affairs (MNGHA) hospitals: King Abdulaziz Medical City, Riyadh; King Abdulaziz Medical City, Jeddah; and Prince Mohammed Bin Abdul Aziz Hospital, Al Madinah; King Abdulaziz Hospital, Al Ahsa; and Imam Abdulrahman Al Faisal Hospital, Dammam. A CONSORT diagram will be generated according to the CONSORT extension for stepped-wedge cluster randomized trials (CRTs) [11]. For each sequence at each period, the number of clusters receiving the intervention, the average cluster size, and its variance, and the number of clusters not receiving the intervention will be presented (Fig. 1).

Sample size

The sample size for this stepped-wedge cluster randomized design was calculated for 45 clusters with 10 periods (including one baseline period) using Power Analysis and Sample Size (PASS) software (PASS 15 Power Analysis and Sample Size Software (2017). NCSS, LLC. Kaysville, UT, USA, [ncss.com/software/pass](https://www.ncss.com/software/pass)). After a baseline period, 5 new clusters switch from the control group to the intervention group at the beginning of each subsequent period. Using historical data obtained from the development domain of the EMR for ward patients admitted from 01 July 2018 to 30 June 2019, we calculated a baseline in-hospital mortality rate by day 90 of 3.13%. Based on the same dataset, 18.3% of eligible ward patients had an alert based on qSOFA criteria [6] with in-hospital mortality of 8.16% compared to 2% in the non-alert patients. For sample size calculations, we made the following assumptions: (A) the impact of the intervention on mortality occurs only in patients who have the alert; (B) only half of the patients with the alert have sepsis; (C) 90% of deaths among the patients with the alert occur among septic patients; (D) early intervention resulting from the alert will reduce the in-hospital mortality by 50%, i.e., from 8.16 to 4.08% in patients with sepsis, and would lead to an overall change in in-hospital mortality for the whole cohort from 3.13 to 2.46% yielding a relative risk of 0.79 and a risk difference of 0.67% which is the target effect size; (E) 80% power using two-sided Wald Z-test and significance level of 5%; and (F) an intra-cluster correlation (*ICC*, a measure of the relatedness of cluster) of 0.22 as estimated from the same retrospective electronic database. As the primary analysis would be adjusted for random effect to account for the correlation between patients within the same cluster, we used the estimation variance (σ^2), which is calculated from responses (P_1 , P_2), as the within-cluster variance (σ_w^2) as suggested by Hussey and Hughes and Hemming et al. (2015) [13, 14]. As such, a reduction of in-hospital mortality by 90 days from 3.13 to 2.46% (relative risk of 0.79 and a risk difference of 0.67%) requires a total sample size of 65,250 subjects (average of 1450 subjects per cluster with an average of 145 subjects per cluster per period). With all five hospitals combined, this is expected to require 20 months (2 months per period). There is no planned interim analysis.

Study cohorts

Intention-to-treat cohort

We will report patient flow according to the CONSORT flowchart for stepped-wedge cluster randomized trials by the allocated sequence and period [11] (Fig. 1). The intention-to-treat (ITT) cohort includes all eligible patients admitted to the eligible wards. The ITT analysis

also implies that patients in the ITT cohort in the wards belonging to a particular period will be analyzed per their planned randomization regardless of what happens during the trial. For example, if a ward was planned to have the intervention during a given period and for technical reasons the alert system was not operational, patients admitted in that ward during that period will be analyzed as receiving an active alert. Although it is not anticipated that there will be wards that crossover their allocated group (i.e., change from alert to non-alert, or vice versa), such instances will be documented. Patients who are transferred from one ward to another will be counted as part of the first ward. The primary analysis will be based on this population.

The study was launched in October 2019. During the study period in 2020–2021, and in response to the surge in the number of hospitalized patients with coronavirus disease-19 (COVID-19), some of the wards had to be converted to intensive care units (ICUs), making them ineligible for the study intervention [15]. These wards will be excluded from the ITT cohort while used as ICUs. During the peak of COVID-19 cases, total admissions to the wards declined substantially to less than 50%; therefore, 2 consecutive periods (starting June 2020) were extended from 2 to 3 months each to account for the decline in cluster size. In addition, some wards were designated for admission of suspected or confirmed COVID-19 cases. Given the higher mortality associated with such designation [16], and because such designation would likely be over-represented in the intervention group, data on a ward designated as a COVID-19 ward will be documented as a ward-level variable, which will be used as a fixed effect term in the primary model (as discussed below). Additionally, patient-level data on COVID-19 diagnosis will be also obtained.

Alert cohort

This cohort represents the subset of ITT patients who had the alert whether in the intervention wards or the control wards.

Statistical tests and their confidence intervals (*CI*s) will be calculated with two-sided. The statistical significance will be set at the 5% level. All analyses will be performed using SAS version 9.4. Categorical variables will be summarized as counts and frequencies, and continuous variables as median and interquartile ranges or means and standard deviations, if deemed normally distributed.

Reporting baseline characteristics, physiological parameters, and treatments

Baseline characteristics will be presented for the ITT and alert cohorts (Table S1) including age, sex, source of

admission, admitting ward, comorbidities (extracted based on ICD-10-AM), Charlson comorbidity index, source of infection (pneumonia, urinary tract infection, skin and soft tissue infection, intra-abdominal infection, other infections, and no clear source, extracted based on ICD-10-AM), and dialysis. We will also report vital signs (systolic blood pressure, diastolic blood pressure, heart rate, temperature, and respiratory rate) as well as laboratory parameters (lactate, white blood cells, bilirubin, creatinine) and whether cultures of blood, respiratory, urine, or other body fluids (pleural, ascitic, cerebrospinal, joint) were obtained and whether treatments (intravenous fluids and antibiotics) were received at baseline. Definitions of these variables were outlined in the study protocol [5].

Reporting alert information

We will report the number of patients with at least one alert and the number of alerts per patient, the time to first alert, and the qSOFA criteria which led to the alert trigger (Table S2). We will report vital signs in the 12-h pre-alert (systolic blood pressure, diastolic blood pressure, heart rate, temperature, and respiratory rate) as well as laboratory parameters in the 12-h pre-alert (lactate, white blood cells, bilirubin, creatinine) and whether a culture of blood, respiratory, urine, or other body fluids (pleural, ascitic, cerebrospinal, joint) was obtained and whether treatments (intravenous fluids and antibiotics) were received in the 12-h pre-alert.

Reporting process measures and post-alert physiologic parameters

We will compare the following process measures between the two groups: (A) percentage of patients with lactate reported within 12 h of alert if not reported in the 12 h before alert and the highest lactate value reported in the 12 h after the alert; (B) the percentage of patients with blood culture ordered in 12 h if not performed in the 12 h before the alert; (C) percentage of patients with respiratory, urine, and body fluid cultures ordered in 12 h if not performed in the 12 h before the alert; (D) intravenous fluid administered in 12 h after alert (yes, no); (E) percentage of patients not on antibiotics in the 12 h before alert with a new antibiotic administered within 3 and 12 h of the alert; and (F) post-alert systolic and diastolic blood pressure (the lowest value in the 12 h after the alert) and heart and respiratory rate (highest value in the 12 h after the alert) (Table S2).

Study outcomes

The primary outcome is defined as all-cause in-hospital mortality within 90 days. Other outcomes include hospital length of stay (LOS) (ITT cohort and alert cohort),

transfer to ICU within 90 days (ITT cohort and alert cohort) and 14 days of alert (alert cohort), ICU-free days in the first 90 days (ITT cohort and alert cohort), critical care response team (CCRT) activation within 90 days (ITT cohort and alert cohort) and 14 days of alert (alert cohort), cardiac arrest within 90 days (ITT cohort and alert cohort) and 14 days of alert (alert cohort), the need for mechanical ventilation, vasopressor therapy, and incident renal replacement therapy within 90 days (all in ITT cohort and alert cohort) and 14 days of alert (alert cohort). Balancing/safety outcome measures include antibiotic-free days, acquisition of multidrug-resistant organisms within 90 days in both groups (ITT cohort and alert cohort), and acquisition of *Clostridium difficile* infection within 90 days (ITT cohort and alert cohort) (Table S3 and S4).

Analysis of primary outcome

The primary outcome of all-cause in-hospital mortality by day 90 will be compared between the intervention group and the control group at the individual level with a generalized linear mixed model with a binary distribution using the jack-knife method to estimate standard errors to account for grouping within clusters and by incorporating a log-link function to estimate the relative risk as a measure of effect [13, 17]. We will include two levels of random effects to account for nested clustering within wards and periods and two levels of fixed effects: hospitals and COVID-19 ward status in addition to the intervention. The model will be selected as the best model with a unique covariance structure that produces the lowest Bayesian Information Criterion (BIC) value. The covariance structures that will be considered in the model are the first order of autocorrelation covariance structure, unstructured covariance structure, Toeplitz covariance structure, and variance component structure (VC). We will use the Satterthwaite method to adjust for denominator degree of freedom for tests of the fixed effects. The random coefficients will be modeled using G-side random effects, and we will obtain the subject-specific estimates by defining the appropriate variance-covariance structure. The overdispersion of parameters will also be assessed by calculating the ratio of the Pearson chi-square statistic and its degrees of freedom. If this ratio exceeds 1, it indicates that the variability in data has not been properly modeled and that there is residual overdispersion due to misspecification of the conditional distribution. In case of no event reported in a cluster for a particular period, a Firth correction will be used in the generalized linear mixed model to avoid the issue of quasi-separation. The Newton-Raphson optimization technique with ridging option might be used to help with the convergence of the model. Results will be expressed as relative risk with 95% CI. The primary

outcome will be also analyzed similarly in the alert cohort. In case any model fails to converge, we will report odds ratio from mixed-effect logistic regression.

Analysis of secondary outcomes

Categorical outcomes including ICU admission, CCRT activation, cardiac arrest, the need for mechanical ventilation, vasopressor therapy, incident renal replacement therapy, acquisition of multidrug-resistant organisms, and *Clostridium difficile* infection will be compared between the intervention and the control group, in a similar model to the one used in the analysis primary outcome. We will include two levels of random effects to account for nested clustering within wards and periods and two levels of fixed effects: hospitals and COVID-19 ward status in addition to the intervention. Results will be expressed as relative risk with 95% *CI*. In case any model fails to converge, we will report odds ratio from the mixed-effect logistic regression.

Continuous outcomes including hospital LOS, ICU-free days, and antibiotic-free days will be compared using a mixed-effect Poisson model (Table S3). We will include two levels of random effects to account for nested clustering within wards and periods and two levels of fixed effects: hospitals and COVID-19 ward status in addition to the intervention. The model will be selected as the best model with a unique covariance structure that produces the lowest Bayesian Information Criterion (BIC) value. The covariance structures that will be considered in the model are the first order of autocorrelation covariance structure, unstructured covariance structure, Toeplitz covariance structure, and variance component structure (VC). The random coefficients will be modeled using G-side random effects, and we will obtain the subject-specific estimates by defining the appropriate variance-covariance structure. The results will be expressed as beta estimates with 95% *CI*.

Sensitivity and subgroup analyses

To address the concern about contamination, we will conduct a sensitivity analysis excluding all patients in the control group in the 90 days before crossing over to the intervention group. Because there are fewer than 50 clusters, we will conduct a sensitivity analysis using a small sample correction with the Kenward-Roger method [13, 18, 19]. We will conduct a sensitivity analysis adjusting for the following covariates: type of wards (medical, surgical, oncology, and mixed), age, baseline systolic blood pressure, baseline respiratory rate, Glasgow Coma Scale, and Charlson comorbidity index. For the latter analysis, we will use imputation for missing variables as outlined below. In addition, we will conduct a complete case sensitivity analysis. We will conduct a

sensitivity analysis excluding the periods in which wards were assigned as COVID-19 wards.

We will analyze the primary outcome of all-cause in-hospital mortality by day 90 across predefined subgroups using the same model of the primary analysis. The predefined subgroups include age ≤ 65 years and >65 years; patients with documented infection source (including ICD-10-AM for pneumonia, urinary tract infection, skin and soft tissue infection, intra-abdominal infection, or other infections); patients with no documented infection or infection source; patients admitted to medical, surgical, oncology, and mixed wards, alert within 48 h of admission and after 48 h of admission; and patients admitted to COVID-19 and non-COVID-19 wards (Table S6). Results of the test of interaction will be reported.

Handling dropouts and missing data

We do not expect to have missing observations in the variables required for the primary analysis. Given the nature of the study, missing observations are expected, for example, not all patients will have all laboratory tests and therapies within the narrow time windows defined in the protocol. Variables used in the descriptive analysis will not be imputed. In some of the secondary and sensitivity analyses, imputation will be used. Because the Glasgow Coma Scale score is not documented for patients with normal neurologic status, missing observations will be assigned a normal value for purposes of the model adjustment. Other variables used in the model adjustment model will be assessed and characterized in terms of their pattern (i.e., Missing Completely at Random, Missing at Random, Missing Not at Random). For Missing Completely at Random data, all analyses will be based on a list-wise deletion approach where observations with complete values will be only considered for analysis. For variables with values Missing at Random, multiple imputation techniques will be utilized to impute the missing values as suggested by Rubin [20]. And for variables with values Missing Not at Random, a pattern-mixture model technique will be used to impute the missing values [21].

Graphical presentation

The subgroup analyses will be displayed as a forest plot.

Adjustment for multiplicity

To adjust for multiple testing for secondary outcomes and subgroup analyses, we will use the false discovery rate (FDR) as described by Benjamini and Hochberg [22]. In this procedure, all hypothesis tests will be sorted in descending order based on their calculated *p*-value. All hypothesis tests below an index *K* will be rejected where *K* is calculated as follows:

$$K = \max \left\{ i : p(i) \leq \frac{i}{m} \cdot q \right\}$$

where $i = 1, 2, \dots, m$; m is the total number of tested hypotheses; and $q = 0.05$. The multiplicity testing adjustment will also be done on CIs by constructing $1 - K \times q/m$ CI for each selected parameter [23].

Protocol deviations

We will report protocol deviations, if they occur, including failure to implement the intervention in a given ward, or wrong implementation of the intervention in a ward assigned to the control group. These deviations will be documented on the CONSORT flow diagram. The data from such wards will be analyzed according to the ITT principle.

Study governance

The study management and development committee is responsible for the overall management of the study, data management, and maintenance of the trial master file and statistical master file following the Good Clinical Practice principles.

Discussion

The SCREEN trial provides an opportunity for a novel trial design and analysis of routinely collected and entered data to evaluate the effectiveness of an intervention for a common medical problem (sepsis in ward patients).

Stepped-wedge cluster randomized trials involve randomization of clusters to different sequences and are increasingly used in clinical medicine [14, 24–26]. This design is suitable for evaluating interventions delivered at the level of the cluster. In the SCREEN trial, it allows the assessment of the effect of the alert system by comparing the outcomes of patients in the intervention and control cohorts as well as over time [11]. In addition, by having the masked alert, an additional comparison between patients with alerts in the intervention and control groups will be performed.

Strengths of our SAP include the analysis according to the ITT principle, the large sample size, and the multicenter nature. Our analysis plan highlights some potential trial limitations, in relation to the low event rate, which we have addressed by the large sample size. We will address the multiplicity of secondary analyses by reporting the false discovery rate. Our alert which is based on qSOFA is possibly applicable in other healthcare settings, although there may be variations across different EMRs. As per the design of the stepped-wedge trial, the proportion of clusters in the intervention group increases gradually with time, and as a result, the intervention group will, on average, have more observations at later dates of the study [14]. Because external factors may influence

outcome over time, the time is a potential confounder and was accounted in the sample size calculation and will be adjusted for in the analysis [14]. The COVID-19 pandemic is an example of an external factor that is associated with increased in-hospital mortality of hospitalized patients [27]. Because certain wards were assigned as ICUs during the pandemic, we will exclude these wards for that duration of ICU assignment from the analysis. Other wards were assigned as COVID-19 wards. To address this confounder, we will account for COVID-19 ward status as a fixed effect term for each period. In addition, we will carry out a sensitivity analysis excluding the periods in which wards were assigned as COVID-19 wards. We will also conduct a subgroup analysis by whether patients were admitted to COVID-19 or non-COVID-19 wards. We used 90-day in-hospital mortality rather than a shorter mortality measure as the primary outcome to capture the effect of the intervention on a longer outcome. However, this would likely cause some degree of within-cluster contamination, as some patients in the control wards will become exposed to the alert after crossing over to the intervention group. The impact of the contamination is likely to be modest because the average stay in the hospital is likely to be <10 days. Nevertheless, we will conduct a sensitivity analysis excluding all patients in the control group in the 90 days before crossing over to the intervention group. We did not stratify randomization according to hospital given that the 5 hospitals follow a similar healthcare delivery model; however, we used hospital as a fixed effect term in the analytical model.

The study has started in October 2019 and is anticipated to have complete implementation and follow-up data by the end of October 2021.

Conclusion

In this SAP, we outline details of the planned analyses in advance of SCREEN trial completion. Prior specification of the statistical methods and outcome analysis will facilitate unbiased analyses of these important clinical data.

Abbreviations

CCRT: Critical Care Response Team; CI: Confidence interval; EMR: Electronic medical record; ICU: Intensive care unit; ITT: Intention to treat; LOS: Length of stay; qSOFA: Quick Sequential Organ Failure Assessment; SAP: Statistical analysis plan; SOFA: Sequential Organ Failure Assessment

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13063-021-05788-3>.

Additional file 1: Table S1. Baseline data. **Table S2.** Alert information and process measures. **Table S3.** Outcomes. **Table S4.** Post-alert outcomes in the alert cohort. **Table S5.** Subgroup Analysis; ITT population.

Additional file 2. Statistical analysis plan.

Acknowledgements

We would like to thank Dr. Jessica Kasza, Associate Professor in biostatistics in the School of Public Health and Preventive Medicine at Monash University, for providing statistical input to this SAP.

SCREEN Trial Group**King Abdulaziz Medical City – Riyadh**

Yaseen M Arabi
Abdulmohsen Alsaawi
Ali M Al Khathaami
Raed H AlHazme
Abdullah Al Mutrafy
Eman Al Qasim
Sheryl Ann Abdukahil
Fawaz K Al-Rabeah
Huda Al Ghamdi
Ebtisam Al Ghamdi
Khadega A Abuelgasim
Abdulaleem Alatassi
Georges Nammour
Angela Caswell
John Alchin
Hasan M Al-Dorzi
Mufareh Edah AlKatheri
Ahmad Alharbi
Raed Al Almoody
Joan Jones
Saad Al-Qahtani
Salih Bin Salih
Nahar Alselaim
Nabiha Tashkandi
Ramesh Kumar Vishwakarma
Zeyad Alyousef
Mohamed S. Al Moammary
Soud Rasheed
Amal Matroud
Rasha Ebeid Al Anazi
Haifa Al Shammari
Majid M Alshamrani
Saleh Qasim
Saeed Obbed
Mohammed Al Qarni
Abdul Rahman Jazieh

King Abdulaziz Medical City – Jeddah

Mohammed Al Zahrani
Fahad Al-Hameed
Adnan A Munshi
Wasil Jastaniah
Hassan AlMarhabi
Emad AlWafi
Ali H Alyami
Arwa O Yamani
Hadia Al Tabsh
Basem R Banat
Omar Abuskout
Anna Liza Marcelo
Mayadah M Alhabshi
Abdullah Al Hamdan

King Abdulaziz Hospital – Al Ahsa

Ali Al Qarni
Ahmed Al Arfaj
Hani T Mustafa
Mohammad Shahin
Abdulaziz A Ghamdi
Jamielah Yaakob
Jamal Chalabi
Johanna E Greyvenstein
Abdulaziz Al Qasem
Hattan Esilan
Sami Musalam Aliyyen
Christa Myumi Sian

Prince Mohammad Bin Abdulaziz Hospital – Madinah

Turki Alwasaidi
Amar M Alhasani
Ibrahim J. Jaber
Noha Omaish
Mohammad Abdrabo
Ahmad S Qureshi
Azura Abdrahim
Naif Almughamisi
Azurahazri Abd Rahim

Imam Abdulrahman Al Faisal Hospital – Dammam

Ahmed Al Shouabi
Shaher Qahtani
Mariam Alansari
Hanan Al Somali
Clara Masala
Gaber Madram
Wafa Nasser
Fatimah Talaqof
Maryam Almulhim

Authors' contributions

YA is the chief investigator and contributed to the conception and design, the analytical plan, the drafting of the manuscript, and the critical revision of the manuscript for important and intellectual content. RV, HD, EQ, SA, FR, HG, and EG contributed to the development of the protocol and critical revision of the manuscript for important intellectual content. All authors agree to be accountable for all aspects of the work and have read and approved the final manuscript.

Funding

This is a non-funded KAIMRC-sponsored study.

Availability of data and materials

The datasets will be available from the corresponding author as per the regulations of King Abdullah International Medical Research Center (KAIMRC).

Declarations**Ethics approval and consent to participate**

The study was approved by the Institutional Review Board of the Ministry of National Guard-Health Affairs. Informed consent was waived because of the nature of the study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹College of Medicine, King Saud Bin Abdulaziz University for Health Sciences, King Abdullah International Medical Research Center, Intensive Care Department, King Abdulaziz Medical City, Ministry of National Guard Health Affairs, Riyadh, Saudi Arabia. ²Biostatistics and Bioinformatics Department, King Saud Bin Abdulaziz University for Health Sciences, King Abdullah International Medical Research Center, King Abdulaziz Medical City, Ministry of National Guard Health Affairs, Riyadh, Saudi Arabia. ³Statistics Department, European Organisation for Research and Treatment of Cancer (EORTC) Headquarters, Brussels, Belgium. ⁴College of Health Informatics, King Saud Bin Abdulaziz University for Health Sciences, King Abdullah International Medical Research Center, Information Systems and Informatics Division, King Abdulaziz Medical City, Ministry of National Guard Health Affairs, Riyadh, Saudi Arabia.

Received: 24 August 2021 Accepted: 29 October 2021

Published online: 22 November 2021

References

1. Evans L, Rhodes A, Alhazzani W, Antonelli M, Coopersmith CM, French C, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Intensive Care Med.* 2021;47(11):1181-247.

2. Levy MM, Evans LE, Rhodes A. The surviving sepsis campaign bundle: 2018 update. *Intensive Care Med.* 2018;44(6):925–8. <https://doi.org/10.1007/s00134-018-5085-0>.
3. Rivers E, Nguyen B, Havstad S, Ressler J, Muzzin A, Knoblich B, et al. Early goal-directed therapy in the treatment of severe sepsis and septic shock. *New Engl J Med.* 2001;345(19):1368–77. <https://doi.org/10.1056/NEJMoa010307>.
4. Makam AN, Nguyen OK, Auerbach AD. Diagnostic accuracy and effectiveness of automated electronic sepsis alert systems: a systematic review. *J Hosp Med.* 2015;10(6):396–402. <https://doi.org/10.1002/jhm.2347>.
5. Arabi YM, Alsaawi A, Al Zahrani M, Al Khathaami AM, AlHazme RH, Al Mutrafy A, et al. Electronic early notification of sepsis in hospitalized ward patients: a study protocol for a stepped-wedge cluster randomized controlled trial. *Trials.* 2021;22(1):695. <https://doi.org/10.1186/s13063-021-05562-5>.
6. 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(8):801–10. <https://doi.org/10.1001/jama.2016.0287>.
7. ICH. The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH): statistical principles for clinical trials 5 February 1998 [Available from: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf].
8. ICH. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use: E3 - structure and content of clinical study reports 30 November 1995 [Available from: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf].
9. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ.* 2010;340(mar23 1):c869. <https://doi.org/10.1136/bmj.c869>.
10. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ.* 2010;340(mar23 1):c332. <https://doi.org/10.1136/bmj.c332>.
11. Hemming K, Taljaard M, McKenzie JE, Hooper R, Copas A, Thompson JA, et al. Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration. *BMJ.* 2018; 363:k1614. <https://doi.org/10.1136/bmj.k1614>.
12. Rhodes A, Evans LE, Alhazzani W, Levy MM, Antonelli M, Ferrer R, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock: 2016. *Intensive Care Med.* 2017;43(3):304–77. <https://doi.org/10.1007/s00134-017-4683-6>.
13. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials.* 2007;28(2):182–91. <https://doi.org/10.1016/j.cct.2006.05.007>.
14. Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ.* 2015; 350(feb06 1):h391. <https://doi.org/10.1136/bmj.h391>.
15. Al-Dorzi HM, Aldawood AS, Almatrood A, Burrows V, Naidu B, Alchin JD, et al. Managing critical care during COVID-19 pandemic: the experience of an ICU of a tertiary care hospital. *J Infect Public Health.* 2021;14(11):1635–41. <https://doi.org/10.1016/j.jiph.2021.09.018>.
16. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. *Jama.* 2020;323(20):2052–9. <https://doi.org/10.1001/jama.2020.6775>.
17. Shakeshaft A, Doran C, Petrie D, Breen C, Havard A, Abudeen A, et al. The effectiveness of community action in reducing risky alcohol consumption and harm: a cluster randomised controlled trial. *PLoS Med.* 2014;11(3): e1001617. <https://doi.org/10.1371/journal.pmed.1001617>.
18. Thompson JA, Hemming K, Forbes A, Fielding K, Hayes R. Comparison of small-sample standard-error corrections for generalised estimating equations in stepped wedge cluster randomised trials with a binary outcome: a simulation study. *Stat Methods Med Res.* 2021;30(2):425–39. <https://doi.org/10.1177/0962280220958735>.
19. Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics.* 1997;53(3):983–97. <https://doi.org/10.2307/2533558>.
20. LITTLE, RJA, and D B RUBIN. Statistical analysis with missing data. New York: John Wiley & Sons; 1987.
21. Thijs H, Molenberghs G, Michiels B, Verbeke G, Curran D. Strategies to fit pattern-mixture models. *Biostatistics.* 2002;3(2):245–65. <https://doi.org/10.1093/biostatistics/3.2.245>.
22. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B (Methodological).* 1995;57(1):289–300.
23. Benjamini Y, Yekutieli D, Edwards D, Shaffer JP, Tamhane AC, Westfall PH, et al. False discovery rate: adjusted multiple confidence intervals for selected parameters [with comments, Rejoinder]. *J Am Stat Assoc.* 2005; 100(469):71–93. <https://doi.org/10.1198/016214504000001907>.
24. Beard E, Lewis JJ, Copas A, Davey C, Osrin D, Baio G, et al. Stepped wedge randomised controlled trials: systematic review of studies published between 2010 and 2014. *Trials.* 2015;16(1):353. <https://doi.org/10.1186/s13063-015-0839-2>.
25. Lloyd M, Karahalios A, Janus E, Skinner EH, Haines T, De Silva A, et al. Effectiveness of a bundled intervention including adjunctive corticosteroids on outcomes of hospitalized patients with community-acquired pneumonia: a stepped-wedge randomized clinical trial. *JAMA Intern Med.* 2019;179(8):1052–60. <https://doi.org/10.1001/jamainternmed.2019.1438>.
26. Mdege ND, Man MS, Taylor Nee Brown CA, Torgerson DJ. Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. *J Clin Epidemiol.* 2011;64(9):936–48. <https://doi.org/10.1016/j.jclinepi.2010.12.003>.
27. Kadri SS, Sun J, Lawandi A, Strich JR, Busch LM, Keller M, et al. Association Between Caseload Surge and COVID-19 Survival in 558 U.S. Hospitals, March to August 2020. *Annals of Internal Medicine.* 2021;174(9):1240–51.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

