Supplemental Material

#### Table S1. Prisma checklist

Section and Topic	Item #	Checklist item	Page where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	
INTRODUCTION	N		
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	
METHODS	[		
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta- analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	
RESULTS	•		

Section and Topic	ltem #	Checklist item	Page where item is reported
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	
Study characteristics	17	Cite each included study and present its characteristics.	
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	
	23b	Discuss any limitations of the evidence included in the review.	
	23c	Discuss any limitations of the review processes used.	
	23d	Discuss implications of the results for practice, policy, and future research.	
OTHER INFORM	MATION		
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	
Competing interests	26	Declare any competing interests of review authors.	
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	

#1	"Pneumonia, ventilator associated" [MeSH]
#2	Ventilator associated pneumonia [tiab]
#3	Ventilator acquired pneumonia [tiab]
#4	Ventilator pneumonia [tiab]
#5	VAP [tiab]
#6	#1 OR #2 OR #3 OR #4 OR #5
#7	"Patient care bundles" [MeSH]
#8	Care bundle [tiab]
#9	Ventilator bundle [tiab]
#10	VAP bundle [tiab]
#11	VAP prevention bundle* [tiab]
#12	Prevention [tiab]
#13	#7 OR #8 OR #9 OR #10 OR 311 OR #12
#14	#6 AND #13

# **Table S3.** VAP diagnostic criteria followed in each study.

Author, year [ref]	VAP diagnostic criteria
Al-Tawfiq, 2010 [19]	VAP was diagnosed according to the CDC criteria.
Álvarez-Lerma, 2018 [20]	VAP was diagnosed according to the CDC and the Annual meeting of the National Nosocomial Infection Surveillance Study Registry in ICU (ENVIN-HELICS registry) criteria.
Arabnejad, 2011 [21]	VAP was diagnosed according to the CPIS. Early VAP is defined as development of VAP within 48 to 72 hours after intubation, in which microorganisms
	such as Staphylococcus aureus, Haemophilus influenzae and Streptococcus pneumonia have the highest prevalence. On the contrary, late-onset VAP occurs
	involved. The results of some studies have also pointed to VAP development by multiple organisms in most patients.
	VAP was diagnosed according to the CPIS system. It consists of six components of temperature, volume of respiratory secretions, changes in white blood
Atashi, 2017 [22]	cell count, presence of infiltration in chest radiograph, hypoxemia, and secretion culture results. The overall score of this scale ranges between 0 and 10.
	Scores of 6 and higher indicate the presence of VAP.
	was diagnosed according to the CDC chiefla by a start intensivist. They used startuard chincal diagnostic chiefla. New and persistent ( $^{240}$ nours) nulmonary infiltrates on x-ray: favor >38.5% or <35% without other annarent source: laucocytosis >109.1 –1 or <3 x 109.1 –1; impaired gas exchange:
Baxter, 2005 [23]	change in sputum quality ± positive sputum culture; BAL cultures in some patients.
Birds, 2010 [24]	VAP was diagnosed according to the CDC criteria and radiological evidence.
Bukhari, 2012 [25]	VAP was diagnosed as clinical factors (fever, cough with purulent sputum), in combination with radiological evidence of pulmonary infilitrate, leukocytosis, a suggestive gram stain and growth of bacteria in cultures of sputum, tracheal aspirate, pleural fluid or blood.
Burja, 2018 [26]	VAP was diagnosed as inflammatory changes on a chest radiograph >48 hours after intubation, aspiration of purulent fluid >48 hours after the intubation, or VAP as a discharge diagnosis. Early VAP (≤ 7 days after intubation) and late VAP (>8 days after intubation).
	The CMS definition of VAP was used, which includes (1) new or evolving infiltrate or consolidation on two or more serial chest radiograms, (2) temperature
Cacheco, 2012 [54]	938-C with no other cause, leukopenia (WBC G4,000/mm3), or leukocytosis (WBC Q12,000/mm3), and (3) new-onset purulent sputum or change of character of the sputum or worsening hypoxia.
DeLuca, 2017 [27]	Patients were diagnosed with VAP if they had a new, persistent infiltrate on chest x-ray after ≥48 hours of continuous MV, temperature >38°C or <36°C,
	and leukocytes >12,000 or <4,000, or microbiologic evidence of VAP (eg, growth of a predominant organism on BAL). Discharge summaries, microbiologic
	data, and antibiotic therapy were reviewed to confirm the diagnosis.
Ding, 2013 [28]	VAP was diagnosed clinically according to six different previous clinical definitions and the newly recommended VAE algorithm: loose definition, rigorous
	definition, the CPIS and the Canadian Critical Care Trials Group classification (possible type), the International Sepsis Forum Consensus definition (probable
	type), the CDC criteria, and the new VAE algorithm.

Eom, 2013 [29]	VAP was diagnosed according to the CDC criteria by training infection control professionals. It was defined by infiltrates on chest-X-ray in patients receiving MV for >48 hours in the ICU. Along with at leash two of the following: temperature >38° or <35°, leukocytosis or leukopenia, purulent ETT secretions, potentially pathogenic bateria isolated from the ETT aspirate, and increasing oxygen requirement.
Ferreira, 2016 [30]	VAP was diagnosed as MV patients whose condition has evolved to the point where a new or progressive pulmonary infiltrate in a chest X-ray. The definition also requires at least two clinical signs and/or laboratory abnormalities that suggest an infectious process such as: fever (>38°C); leukocytosis or leukopenia; presence of purulent tracheal secretion after 48 hours of ventilation.
Hawe, 2009 [31]	
Kao, 2019 [32]	VAP was diagnosed only if it occurred 48 hours after the ETT was inserted with a MV and was based on radiological evidence (new or progressive infiltration on chest radiography or computed tomography images), clinical condition (body temperature >38°C or <36°C, tachypnea, hypoxia/desaturation, respiratory distress, and purulent sputum), and laboratory data (abnormal white blood cell count, C-reactive protein, and gas exchange).
Khan, 2016 [33]	VAP was diagnosed according to CDC criteria. It was defined as pneumonia that developed >48 hours after endotracheal intubation. It was diagnosed clinically as two or more serial chest radiographs with at least 1 of the following: new, progressive, or persistent infiltrates; consolidation; or cavitation; with 2 of the following: core temperature >38.5° or <36°C, leukocytosis (>12,000/ mm3), leukopenia (white blood cell count <1500/mm3); or new-onset purulent bronchial secretions, without another cause and a significant positive culture from blood, BAL fluid, or endotracheal aspirate or culture from another relevant site of infection. Tracheal aspirates were considered purulent at a neutrophil count with Gram stain of >25 per high-power field on lightmicroscopy.
Landelle, 2018 [34]	VAP was diagnosed according to the criteria established by Hospitals in Europe Linked for Infection Control through Surveillance, and a CPIS >6. Provable VAP required the presence of Rx changes with systemic inflammation (temperature ≥38 °C, or leukocyte count >12,000 or <4000 cells/mL) with clinical pulmonary signs (i.e. purulent tracheal secretions). Definite VAP was defined by the addition of positive quantitative cultures of distal pulmonary sampling obtained by BAL (significant threshold ≥10^4 colony-forming units/mL) or mini-BAL (significant threshold ≥ 10^3 colony-forming units/mL).
Lansford, 2007 [35]	VAP was diagnosed according to the CDC criteria. Established by the National Nosocomial Infections Surveillance System criteria, including radiographic evidence of at least one of the following: new or progressive infiltrate, consolidation, or pneumatoceles. Also required is fever or leukopenia. Finally, at least two of the following conditions must be present: New onset of purulent sputum, worsening cough, rales, or worsening gas exchange.
Lim, 2015 [36]	VAP was diagnosed as a respiratory tract infection developed after 48 hours of intubation with MV or within 48 hours after disconnecting the ventilator. The respiratory tract infection follows the definition in the Nosocomial Infection Surveillance guideline from the Taiwan Centers for Disease Control, and it is determined by the clinicians according to the clinical presentations after ruling out all other cause-induced systemic inflammatory response syndrome. The ventilators were limited to the invasive types by either tracheostomy or ETT only, and other noninvasive ventilation devices were excluded.
Liu, 2020 [38]	VAP was diagnosed based on the "VAP prevention, diagnosis and treatment guideline" published by the Intensive Care Branch of the Chinese Medical Association in 2013. This guideline was revised and actualised in 2019.
Liu, 2021 [37]	-

Morris, 2011 [39]	Diagnosis of VAP was made independently by the treating clinical team. Chest radiograph interpretation was undertaken "off-line" and by clinicians who were independent of the treating team. For VAP diagnosis, Hospitals in Europe Linked for Infection Control through Surveillance has a two-stage definition: first, clinically suspected VAP based on clinical criteria; and second, microbiologically confirmed VAP based on further investigations. From 2005 to 2008, we were unable to decrease the incidence of clinically diagnosed VAP but we had shown that increasing the use of quantitative analysis of BAL fluid for microbiological diagnosis resulted in a decrease in the reported incidence of microbiologically confirmed VAP, which was explained by superior test specificity compared with analysis of tracheal aspirates.
Okgün, 2016 [40]	VAP was diagnosed according to the CDC criteria. It was defined during daily surveillance rounds by trained infection control commitee members. The infection preventionists verified all suspected cases with radiographs and microbiologic analyses.
Omrane, 2007 [41]	VAP was diagnosed as the occurrence of a first episode for each patient. It was defined as either the presence of a new and persistent (>72 hours) radiographic infiltrate with one of the following findings: positive pleural or blood cultures for the same organism as that recovered in the tracheal aspirate or sputum; radiographic cavitation; histopathologic evidence of pneumonia; or 2 of the following: (fever (>38.3°C), leukocytosis (white blood cells [WBC] >10 x10^3/µL) or leukopenia (WBC <4 x10^3/µL), purulent tracheal aspirate or sputum (>25 leukocytes/hpf determined by Gram stain). Furthermore, the patient had to have been ventilated >48 hours to be diagnosed with VAP.
Ongstad, 2013 [42]	VAP was diagnosed according to the NHSN criteria and it was identified by an experienced infection control nurse.
Parisi, 2016 [43]	VAP was diagnosed according to the guidelines by the supervising physician. Specifically, the presence of a new infiltrate on the chest radiograph and 2 of 3 clinical criteria (leukocytosis, purulent secretions, fever), together with tracheobronchial secretions, confirmed the occurrence of VAP for the physician. Also, the CPIS was calculated, and a score greater than 6 was used to verify the diagnosis.
Pérez-Granda, 2014 [44]	VAP was diagnosed according to the CDC criteria. Patients ventilated for >48 hours were diagnosed with VAP based on the presence of new and/or progressive pulmonary infiltrates on the chest radiograph plus two or more of the following criteria: fever >38.5°C or hypo 109/ thermia <36°C, leukocytosis =12 × L, purulent tra cheobronchial secretions, and a =15% reduction in PaO2/FiO2. Patients with a CPIS higher than 6 were also considered to have pneumonia. The isolation of one or more pathogenic microorganisms in significant bacterial counts was required to confirm the diagnosis of VAP.
Rello, 2012 [45]	VAP was diagnosed according to the CDC criteria by the attending physician team. An independent investigator (intensivist), who was not part of the team caring for the patient made the final diagnosis of pneumonia, using quantitative respiratory cultures, using standardized thresholds.
Rosenthal, 2012 [46]	VAP was diagnosed according to the CDC/NHSN criteria. It was diagnosed in a MV patient with a chest radiograph that shows new or progressive infiltrates, consolidation, cavitation, or pleural effusion. The patient also must meet at least one of the following criteria: new onset of purulent sputum or change in character of sputum, organism cultured from blood, or isolation of an etiologic agent from a specimen obtained by tracheal aspirate, bronchial brushing or BAL, or biopsy.
Sachetti, 2014 [47]	VAP were defined as including all of the cases for which the area intensivist physician had registered that diagnosis.

Samra, 2016 [48]	VAP was diagnosed according to the CDC criteria. It was diagnosed as a pneumonia that occurs in a patient who was intubated and ventilated ≥48 hours. The patient has to present new or progressive infiltrates, consolidation or cavitations on chest X-ray with one of the following: new onset purulent bronchial secretions, leucopenia (white blood cell<1500/mm3) or leukocytosis (>12,000/mm3), c- Core temperature >38 °C or <36 °C without other cause, positive culture from blood, BAL or endotracheal aspirate.
Santana, 2022 [49]	VAP was diagnosed according to the Brazilian National Regulatory Health Agency. Itis characterized by a pulmonary infection occurring after 48 hours of endotracheal MV, associated with one or more chest radiographs with the presence of a new, persistent or progressive infiltrate, fever (>38°C) or leukocytosis or leukopenia, worsening pulmonary secretions or worsening pulmonary function.
Sen, 2016 [50]	VAP was diagnosed according to the CDC/NHSN criteria by infection control staff. It was diagnosed as including clinical, microbiologic, and radiographic data. The clinical criteria used included the following: fever, presence of infiltrate on chest radiography, quantitative bacterial culture identified through mini-BAL or bronchoscopy. Organisms that are reported are based either on mini-BAL or bronchoscopy.
Talbot, 2015 [51]	VAP was defined according to the CDC criteria. It was defined by trained infection preventionists who were masked to patient-specific bundle adherence data. Every weekday the infection practitioners reviewed every respiratory culture. Patients with an identified culture specimen then underwent medical chart review with examination of chest radiographs and clinical signs and symptoms.
Tao, 2012 [52]	VAP was defined according to the CDC/NHSN criteria. It was indicated in a MV patient with a chest radiograph that shows new or progressive infiltrates, consolidation, cavitation, or pleural effusion. The patient must also meet at least 1 of the following criteria: new onset of purulent sputum or change in character of sputum, organism cultured from blood, or isolation of an etiologic agent from a specimen obtained by tracheal aspirate, bronchial brushing, or BAL, or biopsy.
Triamvisit, 2016 [53]	VAP was defined according to the CDC criteria by 10-year-experienced staff. VAP diagnosis is HAP that occurs after using a MV >48 hours. It was diagnosed including a new persistent or progression of either opacity or cavitation on serial chest films together with high fever (>38.0°C), leukopenia (<4,000 WBC/mm3) or leukocytosis (≥12,000 WBC/mm3), altered mental status with no other causes in older than 70-years-old patients, and purulent sputum or change in sputum character or increased respiratory secretions, or increased required suction.

BAL: Bronchoalveolar Lavage; CDC: Centers for Disease Control and Prevention; CPIS: Clinical Pulmonary Infection Score; ETT: Endotracheal Tube; MV: Mechanical Ventilation; NHSN: National Health Safety Network; VAE: Ventilator Associated Event; VAP: Ventilator-Associated Pneumonia.

## Table S4. Care bundles recommendations by each clinical practice guideline

	Dodek, 2004 [8]	IHI's care bundles, 2012 [4]	Torres, 2017 [2]	Alvarez-Lerma, 2019 [7]
Physical strategies				
Oral endotracheal intubation	R	_	_	-
Daily assessment of readiness to extubate	-	R	_	_
Deep venous thrombosis prophylaxis	_	R	_	-
Hand hygiene	-	_	-	R
Cuff pressure control	-	-	-	R
Heat and moisture exchanger	R	_	_	-
Closed suction system	R	_	_	-
Drainage of subglottic secretions	С	_	-	R
Search for maxillary sinusitis	NR	-	_	-
Scheduled change of ventilator circuit and humidifiers	NR	_	_	NR
Chest physiotherapy	NR	-	-	-
Early tracheostomy	NR	_	_	_

Position strategies				
Semi-recumbent positioning/ Head of bed	R	R	_	R
Kinetic beds	С	_	_	-
Prone positioning	NR	-	-	_
Pharmacologic strategies				
Selective oral decontamination	NR	-	R	R
"Sedation vacation"	-	R	-	-
Peptic ulcer prophylaxis	_	R	-	-
Oral Care	_	R	_	_
Stress ulcer prophylaxis	R	-	_	-
Narrow-spectrum antibiotics in early VAP and low risk of resistance	_	-	R	-
Broad-spectrum therapy targeting Pseudomonas and B-lactamase producing bacteria in patients with risk of antibiotic resistance	_	_	R	-
Selective digestive decontamination	NR	-	NR	R
Prophylactic antibiotics	NR	_	_	-
Other strategies				

Lower respiratory tract samples before starting antibiotic therapy	_	_	R	-
Training in appropriate airway management	_	_	_	-

Author, year [ref]	REPORTING											EXTERNAL VALIDITY			INTERNAL VALIDITY - BIAS-							INTE ONFOU	POW ER	SCO RE				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	
Liu, 2021 [37]	1	1	1	1	2	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	26
Atashi, 2017 [22]	1	1	1	1	2	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	26
Rello, 2012 [45]	1	1	1	1	2	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	24
DeLuca, 2017 [27]	1	1	1	1	2	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	23
Khan, 2016 [33]	1	1	1	1	2	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	23
Parisi, 2016 [43]	1	1	1	1	2	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	23
Triamvisit, 2016 [53]	1	1	1	1	2	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	23
Lim, 2015 [36]	1	1	1	1	2	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	23
Ding, 2013 [28]	1	1	1	1	2	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	23
Omrane, 2007 [41]	1	1	1	1	2	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	23

**Table S5.** The quality assessment for 35 included studies by The Downs and Blacks

Liu, 2020 [38]	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	22
Okgün, 2016 [40]	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	22
Álvarez-Lerma, 2018 [20]	1	1	1	1	2	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	21
Ferreira, 2016 [30]	1	1	1	1	2	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	21
Samra, 2016 [48]	1	1	1	1	2	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	21
Sen, 2016 [50]	1	1	1	1	2	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	21
Ongstad, 2013 [42]	1	1	1	1	2	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	21
Hawe, 2009 [31]	1	1	1	1	2	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	21
Santana, 2022 [49]	1	1	1	1	2	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	0	20
Burja, 2018 [26]	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	20
Landelle, 2018 [34]	1	1	0	1	2	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	20
Pérez-Granda, 2014 [44]	1	1	0	1	2	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	20
Bukhari, 2012 [25]	1	1	1	1	1	1	1	0	1	1	1	0	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	20
Tao, 2012 [52]	1	1	0	1	2	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	20

Arabnejad, 2011 [21]	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	20
Al-Tawfiq, 2010 [19]	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	20
Lansford, 2007 [35]	1	1	1	1	2	1	1	0	0	1	1	0	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	20
Baxter, 2005 [23]	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	20
Talbot, 2015 [51]	1	1	0	1	0	1	1	0	1	1	1	0	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	19
Rosenthal, 2012 [46]	1	1	0	1	1	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	19
Morris, 2011 [39]	1	1	0	1	1	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	19
Cacheco, 2012 [54]	1	1	0	1	1	1	1	1	0	0	1	1	1	0	0	1	1	1	1	1	1	0	0	0	1	0	1	18
Sachetti, 2014 [47]	1	1	1	1	0	1	1	0	0	1	1	0	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	18
Kao, 2019 [32]	1	1	0	1	0	1	1	0	0	1	1	0	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	17
Eom, 2013 [29]	1	1	0	1	0	1	1	0	0	1	1	0	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	17
Bird, 2010 [24]	1	1	0	1	0	1	1	0	0	1	1	0	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	17
Total	36	36	25	36	52	36	36	5	14	35	36	27	36	36	0	36	36	36	36	36	36	1	1	1	36	13	35	

(<14 points) - Poor quality evidence; (14-18 points) - Fair quality evidence; (19-23) - Good quality evidence; (24-28) – Excellent quality of evidence.

The 27 questions have to be graded as "Yes", "No" and "Unable to determine" as per the available information. There are 5 sections which include: study quality (10 items), external validity (3 items), study bias (7 items), confounding and selection bias (6 items), and power (1 item). Each question if answered "yes" gets a score of 1,

except for the 5th question which can get a score of 2 if answered "yes". Thus the total score is out of 28. The modified version makes a simplification of the power question, awarding only 1 point if a study had adequate power to recognize a clinically significant effect. If a study did not mention statistical power, it was deemed either "no" or "unable to determine" and given a score of 0.

**REPORTING:** 1. Is the hypothesis/aim/objective of the study clearly described?; 2. Are the main outcomes to be measured clearly described in the Introduction or Methods section?; 3. Are the characteristics of the patients included in the study clearly described?; 4. Are the interventions of interest clearly described?; 5. Are the distributions of principal confounders in each group of subjects to be compared clearly described?; 6. Are the main findings of the study clearly described?; 7. Does the study provide estimates of the random variability in the data for the main outcomes?; 8. Have all important adverse events that may be a consequence of the intervention been reported?; 9. Have the characteristics of patients lost to follow-up been described?; 10. Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?. EXTERNAL VALIDITY: 11. Were the subjects asked to participate in the study representative of the entire population from which they were recruited?; 12. Were those subjects who were prepared to participate representative of the entire population from which they were recruited?; 13. Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?. INTERNAL VALIDITY (BIAS): 14. Was an attempt made to blind study subjects to the intervention they have received?; 15. Was an attempt made to blind those measuring the main outcomes of the intervention?; 16. If any of the results of the study were based on "data dredging", was this made clear?; 17. In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?; 18. Were the statistical tests used to assess the main outcomes appropriate?; 19. Was compliance with the intervention/s reliable?; 20. Were the main outcome measures used accurate (valid and reliable)?. INTERNAL VALIDITY (CONFOUNDING SELECTION BIAS): 21. Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population?; 22. Were study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time?; 23. Were study subjects randomized to intervention groups?; 24. Was the randomized intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?; 25. Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?; 26. Were losses of patients to follow-up taken into account?. POWER: 27. Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%?

# **Table S6.** Comparison between previous systematic reviews and the current study

	Current study	Pileggi <i>et al.</i> 2018 [1]
Study Design	Systematic Review & Meta-analysis	Systematic Review & Meta-analysis
Period of publication	January 1985 to July 2022	Until June 2017
Databases	Pubmed, Cochrane Library, Web of Science	Pubmed, Scopus, Cochrane Library, Web of Science
Eligible study design	RCTs and Observational studies	RCTs and Observational studies
PROSPERO register	CRD42022341780	CRD42017054268
Inclusion criteria	<ul> <li>-Adult (≥18 years) ICU patients under mechanical ventilation</li> <li>-Application of care bundles only in the intervention group, control group did not receive care bundles.</li> </ul>	<ul> <li>-Adult (≥18 years) ICU patients under mechanical ventilation</li> <li>-Make reference to a ventilator bundle</li> <li>-Assess mortality (report enough data to estimate RR or OR)</li> </ul>
Exclusion criteria	-Patients admitted with pneumonia and patients with nosocomial pneumonia other than VAP.	-Care bundles for prevention of other hospital acquired infections.
Outcomes	-Main: VAP occurrence -Other: Duration of MV, ICU and hospital length of stay, ICU and hospital mortality, VAP-related mortality and length of stay, compliance	-Main: Mortality (overall, hospital, ICU and VAP-related mortality) -Other: VAP occurrence, ICU and hospital length of stay, duration of mechanical ventilation, days of antibiotic therapy, compliance
Age, years	≥18 years	≥18 years
N studies included	29	13
N Subjects	<mark>116,873</mark>	11,664

**Figure S1.** Funnel plot (A) and Forest plots (B-G) on **VAP episodes** reported in the care bundles and standard care.

A) Funnel plot of studies reporting data about VAP episodes in the care bundles and standard care groups.



## **B)** Forest plot of VAP episodes by baseline measures subgroups.

	care bu	Indles	standard of	of care		Odds Ratio	Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI	M-H, Random, 95% CI
1.4.1 Baseline Meas	ures						
Lansford 2007	2	132	11	218	1.8%	0.29 (0.06, 1.33)	
Atashi 2018	3	38	9	38	2.0%	0.28 [0.07, 1.12]	
Triamvisit 2016	7	68	22	66	3.1%	0.23 [0.09, 0.58]	
Arabnejad 2011	9	46	28	71	3.3%	0.37 [0.16, 0.89]	
DeLuca 2017	11	192	22	195	3.7%	0.48 [0.23, 1.01]	
Burja 2018	19	74	23	55	3.8%	0.48 [0.23, 1.02]	
Landelle 2018	13	356	64	291	4.2%	0.13 [0.07, 0.25]	
Samra 2016	23	250	24	130	4.2%	0.45 [0.24, 0.83]	
Omrane 2007	22	360	23	349	4.3%	0.92 [0.50, 1.69]	
Tao 2012	27	1745	130	1999	5.0%	0.23 [0.15, 0.34]	
Lim 2015	56	14212	176	12913	5.4%	0.29 [0.21, 0.39]	
Liu 2020	90	2687	96	2029	5.4%	0.70 [0.52, 0.94]	
Rosenthal 2012	2191	51618	226	3889	5.8%	0.72 [0.62, 0.83]	
Subtotal (95% CI)		71778		22243	52.0%	0.39 [0.28, 0.56]	◆
Total events	2473		854				
Heterogeneity: Tau² =	= 0.29; Ch	i² = 82.12	2, df = 12 (P	< 0.0000	1); I² = 85%	6	
Test for overall effect:	Z = 5.24	(P < 0.00	001)				
1.4.2 No Baseline Me	easures						
Ongstad 2013	5	96	8	87	2.5%	0.54 [0.17, 1.73]	
Sen 2016	5	65	18	66	2.7%	0.22 [0.08, 0.64]	
Santana 2022	17	34	15	30	3.0%	1.00 [0.37, 2.67]	
Ferreira 2016	7	73	30	115	3.3%	0.30 [0.12, 0.73]	
Liu 2021	8	100	34	100	3.5%	0.17 [0.07, 0.39]	
Hawe 2009	10	215	49	374	3.9%	0.32 [0.16, 0.65]	
Parisi 2016	21	136	53	226	4.5%	0.60 [0.34, 1.04]	
Khan 2016	14	1453	144	2212	4.5%	0.14 [0.08, 0.24]	
Rello 2012	104	885	23	149	4.7%	0.73 [0.45, 1.19]	
Sachetti 2014	52	235	42	198	4.9%	1.06 [0.67, 1.67]	
Morris 2011	43	501	216	1460	5.3%	0.54 [0.38, 0.76]	
Baxter 2005	154	3507	47	705	5.3%	0.64 [0.46, 0.90]	
Subtotal (95% CI)		1300		5/22	48.0%	0.45 [0.31, 0.65]	-
l otal events	440	7 60 00	6/9		A. R. 700	,	
Heterogeneity: Tau*=	= 0.30; Ch	r= 50.38	5, df = 11 (P	< 0.0000	1); 1= 789	6	
i est for overall effect	2 = 4.28	(P < 0.00	01)				
Total (95% CI)		79078		27965	100.0%	0.42 [0.33, 0.54]	◆
Total events	2913		1533				
Heterogeneity: Tau <sup>2</sup> =	0.26; Ch	i <sup>2</sup> = 132.7	4, df = 24 (F	< 0.000	01); I <sup>2</sup> = 82	:%	
Test for overall effect:	Z=7.05	(P < 0.00	001)				0.1 0.2 0.5 1 2 5 10 Eavoure (CP) Eavoure (SC)
Test for subgroup dif	ferences:	Chi <sup>2</sup> = 0.	27. df = 1 (P	= 0.60), I	<sup>2</sup> =0%		Favouis [OD] Favouis [OO]

**C)** Forest plot of VAP episodes by compliance subgroups.



Test for subgroup differences: Chi<sup>2</sup> = 2.75, df = 1 (P = 0.10), l<sup>2</sup> = 63.6%

## **D)** Forest plot of VAP episodes by quality of evidence (risk of bias) subgroups.

	care bu	ndles	standard o	of care		Odds Ratio	Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% Cl	M-H, Random, 95% Cl
1.5.1 High quality of	evidence						
Atashi 2018	3	38	9	38	2.0%	0.28 [0.07, 1.12]	
Liu 2021	8	100	34	100	3.5%	0.17 [0.07, 0.39]	
Rello 2012	104	885	23	149	4.7%	0.73 [0.45, 1.19]	
Subtotal (95% CI)		1023		287	10.2%	0.34 [0.12, 1.00]	
Total events	115		66				
Heterogeneity: Tau <sup>2</sup> =	= 0.68; Chi <sup>a</sup>	<sup>2</sup> = 9.59,	df = 2 (P = 0)	0.008); <b>I</b> ₹ =	= 79%		
Test for overall effect:	Z = 1.95 (	P = 0.05	)				
1.5.2 Moderate quali	ty of evide	ence					
Arabnejad 2011	9	46	28	71	3.3%	0.37 [0.16, 0.89]	
Baxter 2005	154	3507	47	705	5.3%	0.64 [0.46, 0.90]	
Burja 2018	19	74	23	55	3.8%	0.48 [0.23, 1.02]	
DeLuca 2017	11	192	22	195	3.7%	0.48 [0.23, 1.01]	
Ferreira 2016	7	73	30	115	3.3%	0.30 [0.12, 0.73]	
Hawe 2009	10	215	49	374	3.9%	0.32 [0.16, 0.65]	
Khan 2016	14	1453	144	2212	4.5%	0.14 [0.08, 0.24]	
Landelle 2018	13	356	64	291	4.2%	0.13 [0.07, 0.25]	
Lansford 2007	2	132	11	218	1.8%	0.29 [0.06, 1.33]	
Lim 2015	56	14212	176	12913	5.4%	0.29 [0.21, 0.39]	
Liu 2020	90	2687	96	2029	5.4%	0.70 [0.52, 0.94]	
Morris 2011	43	501	216	1460	5.3%	0.54 [0.38, 0.76]	_ <b>-</b>
Omrane 2007	22	360	23	349	4.3%	0.92 [0.50, 1.69]	
Ongstad 2013	5	96	8	87	2.5%	0.54 [0.17, 1.73]	
Parisi 2016	21	136	53	226	4.5%	0.60 [0.34, 1.04]	
Rosenthal 2012	2191	51618	226	3889	5.8%	0.72 [0.62, 0.83]	-
Samra 2016	23	250	24	130	4.2%	0.45 [0.24, 0.83]	
Santana 2022	17	34	15	30	3.0%	1.00 [0.37, 2.67]	
Sen 2016	5	65	18	66	2.7%	0.22 [0.08, 0.64]	
Tao 2012	27	1745	130	1999	5.0%	0.23 [0.15, 0.34]	_ <b></b>
Triamvisit 2016	7	68	22	66	3.1%	0.23 [0.09, 0.58]	
Subtotal (95% CI)		77820		27480	85.0%	0.41 [0.31, 0.53]	◆
Total events	2746		1425				
Heterogeneity: Tau² =	= 0.25; Chř	<sup>2</sup> = 114.2	26, df = 20 (F	P < 0.000	01); I <sup>z</sup> = 82'	%	
Test for overall effect:	Z = 6.82 (	P < 0.00	001)				
1.5.3 Low quality of e	evidence						
Sachetti 2014	52	235	42	198	4.9%	1.06 [0.67, 1.67]	<u> </u>
Subtotal (95% CI)		235		198	4.9%	1.06 [0.67, 1.67]	-
Total events	52		42				
Heterogeneity: Not ap	oplicable						
Test for overall effect:	: Z = 0.23 (	P = 0.82	)				
Total (05% CI)		70070		27065	100.0%	0 42 [0 22 0 54]	
Total (95% CI)	204.0	19018	4600	21905	100.0%	0.42 [0.33, 0.34]	▼
i otal events	2913 - 0.08: 05:	8 - 400 7	1533 74 df - 24 0		043-12-02-	Ω/ -	
Toot for overall offer th	= 0.26; Chr - 7 - 7 05 4	-=132.7 D×0.00	·4, ui = ∠4 (f 004)	- < 0.000	01); 1= 82	70	0.1 0.2 0.5 1 2 5 10
Test for overall effect:	.∠=7.05() ¥aran	F 5 U.UU Olaiz - 40	001) \44 df - 24	n - 0.004	12-04-7	Y	Favours [CB] Favours [SC]
restion subdroub all	ierences: (	unit = 13	o.ii,ui=∠(	r = 0.001	7, 17 = 84.7	70	

#### **E)** Forest plot of VAP episodes by study design subgroups.

	care bu	ndles	standard of care			Odds Ratio	Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI	M-H, Random, 95% Cl
1.6.1 RCT							
Arabnejad 2011	9	46	28	71	3.3%	0.37 [0.16, 0.89]	
Atashi 2018	3	38	9	38	2.0%	0.28 [0.07, 1.12]	
Subtotal (95% CI)		84		109	5.3%	0.34 [0.16, 0.72]	
Total events	12		37				
Heterogeneity: Tau² =	= 0.00; Chi	i <sup>z</sup> = 0.13,	df = 1 (P = 0)	0.72); I² =	0%		
Test for overall effect:	Z = 2.84 (	(P = 0.00)	5)				
1.6.2 Cohorts							
Baxter 2005	154	3507	47	705	5.3%	0.64 [0.46, 0.90]	_ <b>-</b>
Burja 2018	19	74	23	55	3.8%	0.48 [0.23, 1.02]	
DeLuca 2017	11	192	22	195	3.7%	0.48 [0.23, 1.01]	
Ferreira 2016	7	73	30	115	3.3%	0.30 [0.12, 0.73]	
Hawe 2009	10	215	49	374	3.9%	0.32 [0.16, 0.65]	
Khan 2016	14	1453	144	2212	4.5%	0.14 [0.08, 0.24]	
Landelle 2018	13	356	64	291	4.2%	0.13 [0.07, 0.25]	
Lansford 2007	2	132	11	218	1.8%	0.29 [0.06, 1.33]	
Lim 2015	56	14212	176	12913	5.4%	0.29 [0.21, 0.39]	
Liu 2020	90	2687	96	2029	5.4%	0.70 [0.52, 0.94]	_ <b></b> -
Liu 2021	8	100	34	100	3.5%	0.17 [0.07, 0.39]	
Morris 2011	43	501	216	1460	5.3%	0.54 [0.38, 0.76]	- <b>-</b>
Omrane 2007	22	360	23	349	4.3%	0.92 [0.50, 1.69]	
Ongstad 2013	5	96	8	87	2.5%	0.54 [0.17, 1.73]	
Parisi 2016	21	136	53	226	4.5%	0.60 [0.34, 1.04]	
Rello 2012	104	885	23	149	4.7%	0.73 [0.45, 1.19]	
Rosenthal 2012	2191	51618	226	3889	5.8%	0.72 [0.62, 0.83]	
Samra 2016	23	250	24	130	4.2%	0.45 [0.24, 0.83]	
Sen 2016	5	65	18	66	2.7%	0.22 [0.08, 0.64]	
Tao 2012	27	1745	130	1999	5.0%	0.23 [0.15, 0.34]	
Subtotal (95% CI)		78657		27562	83.8%	0.40 [0.31, 0.52]	•
Total events	2825		1417				
Heterogeneity: Tau² =	= 0.26; Chi	i <sup>z</sup> = 118.0	)4, df = 19 (i	P < 0.000	01); I² = 8	4%	
Test for overall effect:	Z= 6.79 (	(P < 0.00	001)				
4.0.0.00							
1.6.3 Others							
Sachetti 2014	52	235	42	198	4.9%	1.06 [0.67, 1.67]	
Santana 2022	17	34	15	30	3.0%	1.00 [0.37, 2.67]	
Triamvisit 2016	7	68	22	66	3.1%	0.23 [0.09, 0.58]	
Subtotal (95% CI)		337		294	10.9%	0.05 [0.20, 1.05]	
lotal events	/6		79				
Heterogeneity: Tau* =	= 0.51; Chi	r = 8.47,	ατ = 2 (P = Ι	J.U1); I*=	16%		
lest for overall effect:	Z = 0.90 (	P = 0.37	)				
Total (95% CI)		79078		27965	100.0%	0.42 [0.33, 0.54]	•
Total events	2913		1533				-
Heterogeneity: Tau <sup>2</sup> =	= 0.26: Chi	<b>r</b> = 132 7	74. df = 24 (i	P < 0.000	01); I <sup>z</sup> = 8	2% -	
Test for overall effect:	Z = 7.05 (	P < 0.00	001)				U.1 U.2 U.5 1 2 5 10
Test for subgroup dif	ferences:	Chi² = 1.	21. df = 2 (P	= 0.55), I	²=0%		Favours [CB] Favours [SC]

#### **F)** Forest plot of VAP episodes by country subgroups.

	care bu	ndles	standard	of care		Odds Ratio	Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% Cl	M-H, Random, 95% Cl
1.7.1 Western EU							
Buria 2018	19	74	23	55	3.8%	0.48 [0.23, 1.02]	
Hawe 2009	10	215	49	374	3.9%	0.32 [0.16, 0.65]	
Landelle 2018	13	356	64	291	4.2%	0.13 [0.07, 0.25]	
Morris 2011	43	501	216	1460	5.3%	0.54 [0.38, 0.76]	_ <b>_</b>
Parisi 2016	21	136	53	226	4.5%	0.60 [0.34, 1.04]	
Rello 2012	104	885	23	149	47%	0731045119	<b>_</b> _
Subtotal (95% CI)		2167		2555	26.4%	0.42 [0.27, 0.67]	◆
Total events	210		428				
Heterogeneity: Tau <sup>2</sup> =	0 25: Chi	₹= 21.38	$f = 5 (P = 1)^{120}$	0.0007).	I² = 77%		
Test for overall effect:	7 = 3.64 (	P = 0.00	03) 03)	0.0001,7,			
restion overall effect.	2 - 0.04 (	, = 0.00	00,				
1.7.2 North America							
Bayter 2005	154	3507	47	705	6 3 %	0.64.0046.0.901	
Del uro 2017	11	102	22	195	37%	0.48 [0.23 1.01]	
Khan 2016	1.4	1462	111	2212	1 5 %	0.40 [0.23, 1.01]	
Lapeford 2007	2	1400	144	2212	4.570		
Operano 2007	22	132	22	210	1.070		
Onnane 2007	22	000	23	349	4.370	0.92 [0.00, 1.09]	
Cop 2016	5	90	10	07	2.3%	0.04 [0.17, 1.73]	
Subtotal (95% CI)	5	5805	10	3932	2.7 %	0.22 [0.08, 0.84]	
Total quanta	24.2	5005	272	3032	24.070	0.40 [0.22, 0.74]	
Lotorogonoity: Tou? -	213 0.40:06	z	273 ) df = 670 -	0.00043	$12 - 0.10^{\circ}$		
Teet fer everell effect	- 0.49, Chi	-= 30.90 n = 0.00	ו, ui = פ (ד ⊂ יי	0.0001),	1-= 81 %		
rest for overall effect.	Z = 2.92 (	P = 0.00	3)				
173 Other regions							
Archanoiod 2011		40	20	74	2.20	0.07 10 4 6 0.001	
Arabriejau 2011	9	40	20	20	3.3%		
Alashi 2018 Ferreire 2016	3	30	9	38	2.0%	0.28 [0.07, 1.12]	
Ferreira 2016		13	30	42042	3.3%	0.30 [0.12, 0.73]	
Lim 2015	50	14212	176	12913	5.4%	0.29 [0.21, 0.39]	
Liu 2020	90	2687	96	2029	5.4%	0.70 [0.52, 0.94]	
Liu 2021	8	100	34	100	3.5%	0.17 [0.07, 0.39]	
Rosenthal 2012	2191	51618	226	3889	5.8%	0.72[0.62, 0.83]	
Sachetti 2014	52	235	42	198	4.9%	1.06[0.67, 1.67]	
Samra 2016	23	250	24	130	4.2%	0.45 [0.24, 0.83]	
Santana 2022	17	34	15	30	3.0%	1.00 [0.37, 2.67]	
Tao 2012	27	1745	130	1999	5.0%	0.23 [0.15, 0.34]	
Triamvisit 2016	7	68	22	66	3.1%	0.23 [0.09, 0.58]	
Subtotal (95% CI)		/1106		21578	48.8%	0.42 [0.30, 0.61]	-
Total events	2490		832				
Heterogeneity: Tau <sup>2</sup> =	0.28; Chi	<sup>2</sup> = 76.57	7, df = 11 (P	< 0.0000	1); I <b>²</b> = 86	%	
Test for overall effect:	Z = 4.74 (	P < 0.00	001)				
T-4-1 (0.54) - 00		70070		07005	400.00	0.40.00.00.00.00	
Total (95% CI)		79078		27965	100.0%	0.42 [0.33, 0.54]	▼
Total events	2913		1533				
Heterogeneity: Tau <sup>z</sup> =	0.26; Chi	<sup>2</sup> = 132.7	74, df = 24 (i	P < 0.000	01); I <b>²</b> = 8	2%	
Test for overall effect:	Z=7.05 (	P < 0.00	001)				Favours [CB] Favours [SC]
Test for subgroup diff	erences: •	Chi <sup>z</sup> = 0.	02, df = 2 (P	' = 0.99), I	<b>z</b> =0%		· ····································

**G)** Forest plot of VAP episodes by VAP diagnostic criteria subgroups.

	care bu	ndles	standard of	of care		Odds Ratio	Odds Ratio					
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI	M-H, Random, 95% Cl					
1.8.1 Clinical+micro	obiological											
Baxter 2005	154	3507	47	705	5.7%	0.64 [0.46, 0.90]						
DeLuca 2017	11	192	22	195	4.0%	0.48 [0.23, 1.01]						
Khan 2016	14	1453	144	2212	4.9%	0.14 [0.08, 0.24]						
Landelle 2018	13	356	64	291	4.6%	0.13 [0.07, 0.25]						
Morris 2011	43	501	216	1460	5.7%	0.54 [0.38, 0.76]	_ <b>-</b>					
Omrane 2007	22	360	23	349	4.6%	0.92 [0.50, 1.69]						
Rello 2012	104	885	23	149	5.1%	0.73 [0.45, 1.19]						
Rosenthal 2012	2191	51618	226	3889	6.3%	0.72 [0.62, 0.83]	-					
Samra 2016	23	250	24	130	4.6%	0.45 [0.24, 0.83]						
Sen 2016	5	65	18	66	2.9%	0.22 [0.08, 0.64]						
Tao 2012	27	1745	130	1999	5.4%	0.23 [0.15, 0.34]	_ <b>-</b> _					
Subtotal (95% CI)		60932		11445	53.9%	0.41 [0.28, 0.59]	◆					
Total events	2607		937									
Heterogeneity: Tau <sup>z</sup>	= 0.32; Chi	<b>≈</b> = 85.58	, df = 10 (P	< 0.0000	1); I <sup>z</sup> = 88	%						
Test for overall effect	t: Z = 4.70 (	(P < 0.00	001)									
		-	-									
1.8.2 Clinical												
Burja 2018	19	74	23	55	4.0%	0.48 [0.23, 1.02]						
Ferreira 2016	7	73	30	115	3.5%	0.30 [0.12, 0.73]						
Lansford 2007	2	132	11	218	1.9%	0.29 [0.06, 1.33]						
Lim 2015	56	14212	176	12913	5.9%	0.29 [0.21, 0.39]	- <b>-</b> -					
Liu 2020	90	2687	96	2029	5.9%	0.70 [0.52, 0.94]	_ <b></b>					
Ongstad 2013	5	96	8	87	2.7%	0.54 [0.17, 1.73]						
Sachetti 2014	52	235	42	198	5.3%	1.06 [0.67, 1.67]	<b>-</b> _					
Santana 2022	17	34	15	30	3.2%	1.00 [0.37, 2.67]						
Triamvisit 2016	7	68	22	66	3.3%	0.23 [0.09, 0.58]						
Subtotal (95% CI)		17611		15711	35.6%	0.49 [0.32, 0.76]	◆					
Total events	255		423									
Heterogeneity: Tau²	= 0.28; Chi	<sup>2</sup> = 34.79	), df = 8 (P ≺	0.0001);	l² = 77%							
Test for overall effec	t: Z = 3.24 (	(P = 0.00)	1)									
1.8.3 CPIS criteria												
Arabnejad 2011	9	46	28	71	3.6%	0.37 [0.16, 0.89]						
Atashi 2018	3	38	9	38	2.1%	0.28 [0.07, 1.12]						
Parisi 2016	21	136	53	226	4.8%	0.60 [0.34, 1.04]						
Subtotal (95% CI)		220		335	10.5%	0.49 [0.31, 0.76]	◆					
Total events	33		90									
Heterogeneity: Tau²	= 0.00; Chi	²=1.50,	df = 2 (P = 0)	0.47); I <sup>2</sup> =	0%							
Test for overall effec	t: Z = 3.16 (	(P = 0.00)	2)									
Total (95% CI)		78763		27/01	100.0%	0 44 [0 34 0 56]						
Total evente	2005	.0105	1460	21431	.00.070	0.44 [0.04, 0.00]	•					
Lotorogonoity: Tour	2895 - 0.26: CM	<b>Z</b> = 102.0	1450 16 df = 22 /		01\\ 8=0	20%						
Telefoyeneity, Tau-	Test for overall effect: $7 = 6.51$ (P < 0.00001) Test for overall effect: $7 = 6.51$ (P < 0.00001)											
Test for overall effec	∠= 0.01 ( ifforonoco:	(F' ≦ 0.001 Chi≩ – 0.4	001) 55 df - 270	- 0.763	z - 0%		Favours [CB] Favours [SC]					
restion subdroub a	merences.	$C \Pi = 0.3$	55, ui = 2 (P	= 0.70, 1	- 0.70							

**Figure S2.** Funnel plot (A) and Forest plots (B-F) on **days on mechanical ventilation** in subjects treated with care bundles or standard care.

A) Funnel plot on days of mechanical ventilation in subjects treated with care bundles or standard care.



## B) Forest plot on days of mechanical ventilation by compliance subgroups.

	care bundles			standa	ard of o	are		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
1.8.1 More than 70%									
Landelle 2018	6.8	4.2	356	7.6	4.8	291	18.5%	-0.80 [-1.50, -0.10]	
Khan 2016	6.9	14.1	1453	6.8	9	2212	17.1%	0.10 [-0.72, 0.92]	_ <b>+</b> _
Ding 2013	6.4	5.2	137	5.7	4.5	213	14.3%	0.70 [-0.36, 1.76]	+
Ongstad 2013	4.5	5.3	96	4.7	5.2	87	10.0%	-0.20 [-1.72, 1.32]	
Cacheco, 2012	18.9	11	655	21.5	15.4	299	7.4%	-2.60 [-4.54, -0.66]	
Subtotal (95% CI)			2697			3102	67.2%	-0.37 [-1.20, 0.46]	-
Heterogeneity: Tau² =	0.55; Cł	ni <b>²</b> = 11	1.70, df	= 4 (P =	0.02);1	<b>≈</b> =66%	5		
Test for overall effect: .	Z = 0.88	(P = 0	1.38)						
1.8.2 Less than 70%									
Morris 2011	5.4	5.9	501	5.1	5.2	1460	19.9%	0.30 [-0.28, 0.88]	
Perez-Granda 2014	2.8	(.4	1534	4.1	11.9	401	12.6%	-1.30 [-2.52, -0.08]	
Santana 2022 Subtotal (95% CI)	22.1	23.7	2069	26.8	28.7	30 1891	0.2%	-4.70 [-17.70, 8.30] -0.45 [-1.92, 1.02]	
Heterogeneity: Tau <sup>2</sup> =	n 91+ Cł	u² = 5	87 df=	2(P = 0)	) 05\±₽	= 66%			
Test for overall effect:	7 = 0.60	(P = 0	01, 01- 155)	20-0		- 00 %			
reerier erefail eneor.	_ 0.00	v - 0							
Total (95% CI)			4766			4993	100.0%	-0.36 [-0.99, 0.27]	◆
Heterogeneity: Tau² =	0.43; Cł	ni <b>≥</b> = 18	8.38, df	= 7 (P =	0.01);1	<b>≈</b> =62%	6		
Test for overall effect: .	Z = 1.11	(P = 0	.27)						Favours ICB1 Favours ISC1
Test for subgroup diffe	erences	: Chi <b></b> ≇⊧	= 0.01,	df=1 (P	= 0.93)	), I <sup>z</sup> = 0%	б		

	car	e bund	les	stand	ard of c	are		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% Cl
1.13.1 High quality of	evidenc	e:							
Subtotal (95% CI)			0			0		Not estimable	
Heterogeneity: Not ap	plicable								
Test for overall effect:	Not app	licable	!						
1.13.2 Moderate qual	ity of ev	idence	•						
Santana 2022	22.1	23.7	34	26.8	28.7	30	0.1%	-4.70 [-17.70, 8.30]	
Sen 2016	16.2	27.2	65	19.7	25.3	66	0.2%	-3.50 [-12.50, 5.50]	
Ferreira 2016	22.9	28	115	20	26.3	73	0.3%	2.90 [-5.01, 10.81]	
Arabnejad 2011	12.6	9.1	46	17.8	12.8	71	1.2%	-5.20 [-9.17, -1.23]	
Burja 2018	8.4	7.3	74	9.1	7.8	55	2.4%	-0.70 [-3.35, 1.95]	-+-
Triamvisit 2016	2.9	3	68	5.8	9.3	66	2.9%	-2.90 [-5.25, -0.55]	
Lansford 2007	5.3	7.4	132	7.3	12.2	218	3.6%	-2.00 [-4.05, 0.05]	
Ongstad 2013	4.5	5.3	96	4.7	5.2	87	5.6%	-0.20 [-1.72, 1.32]	+
Pérez-Granda 2014	2.8	7.4	1534	4.1	11.9	401	7.3%	-1.30 [-2.52, -0.08]	
Ding 2013	6.4	5.2	137	5.7	4.5	213	8.5%	0.70 [-0.36, 1.76]	+-
Khan 2016	6.9	14.1	1453	6.8	9	2212	10.6%	0.10 [-0.72, 0.92]	+
Landelle 2018	6.8	4.2	356	7.6	4.8	291	11.7%	-0.80 [-1.50, -0.10]	-
Morris 2011	5.4	5.9	501	5.1	5.2	1460	13.0%	0.30 [-0.28, 0.88]	+
DeLuca 2017	2.1	2.2	192	2.7	3	195	13.5%	-0.60 [-1.12, -0.08]	•
Rosenthal 2012	6.3	10.6	51618	6.8	11.2	3889	15.0%	-0.50 [-0.86, -0.14]	-
Subtotal (95% CI)			56421			9327	96.0%	-0.49 [-0.91, -0.06]	•
Heterogeneity: Tau <sup>2</sup> =	0.25; CI	hi <sup>z</sup> = 29	9.25, df=	: 14 (P =	0.010)	; I <sup>2</sup> = 52	%		
Test for overall effect:	Z = 2.26	(P = 0	.02)						
1.13.3 Low quality of	evidenc	e							
Cacheco, 2012	18.9	11	655	21.5	15.4	299	4.0%	-2.60 [-4.54, -0.66]	
Subtotal (95% CI)			655			299	4.0%	-2.60 [-4.54, -0.66]	◆
Heterogeneity: Not ap	plicable								
Test for overall effect:	Z = 2.63	(P = 0	.009)						
Total (95% CI)			57076			9626	100.0%	-0.59 [-1.03, -0.15]	•
Heterogeneity: Tau <sup>2</sup> =	0.30; CI	hi² = 34	4.09, df=	: 15 (P =	0.003)	; I <sup>z</sup> = 56	%		
Test for overall effect:	Z = 2.61	(P = 0	.009)	-					-20 -10 0 10 2 Eavoure [CP] Eavoure [SC]
Test for subaroup diff	erences	∶ Chi <del>"</del> =	- 4.36. d	f=1 (P=	= 0.04).	$ ^{2} = 77.2$	1%		

## C) Forest plot on days of mechanical ventilation by quality of evidence (risk of bias) subgroups.

	care	e bund	lles	stand	ard of c	are		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
1.14.1 RCT									
Arabnejad 2011 Subtotal (95% CI)	12.6	9.1	46 <mark>46</mark>	17.8	12.8	71 <b>71</b>	1.2% <b>1.2%</b>	-5.20 [-9.17, -1.23] - <b>5.20 [-9.17, -1.23]</b>	
Heterogeneity: Not ap	plicable								
Test for overall effect:	Z= 2.57	(P = 0	1.01)						
1.14.2 Cohorts									
Burja 2018	8.4	7.3	74	9.1	7.8	55	2.4%	-0.70 [-3.35, 1.95]	<u> </u>
Cacheco, 2012	18.9	11	655	21.5	15.4	299	4.0%	-2.60 [-4.54, -0.66]	_ <b>—</b>
DeLuca 2017	2.1	2.2	192	2.7	3	195	13.5%	-0.60 [-1.12, -0.08]	-
Ding 2013	6.4	5.2	137	5.7	4.5	213	8.5%	0.70 [-0.36, 1.76]	
Ferreira 2016	22.9	28	115	20	26.3	73	0.3%	2.90 [-5.01, 10.81]	
Khan 2016	6.9	14.1	1453	6.8	9	2212	10.6%	0.10 [-0.72, 0.92]	+
Landelle 2018	6.8	4.2	356	7.6	4.8	291	11.7%	-0.80 [-1.50, -0.10]	-
Lansford 2007	5.3	7.4	132	7.3	12.2	218	3.6%	-2.00 [-4.05, 0.05]	
Morris 2011	5.4	5.9	501	5.1	5.2	1460	13.0%	0.30 [-0.28, 0.88]	+
Ongstad 2013	4.5	5.3	96	4.7	5.2	87	5.6%	-0.20 [-1.72, 1.32]	-+-
Pérez-Granda 2014	2.8	7.4	1534	4.1	11.9	401	7.3%	-1.30 [-2.52, -0.08]	
Rosenthal 2012	6.3	10.6	51618	6.8	11.2	3889	15.0%	-0.50 [-0.86, -0.14]	-
Sen 2016	16.2	27.2	65	19.7	25.3	66	0.2%	-3.50 [-12.50, 5.50]	
Subtotal (95% CI)			56928			9459	95.8%	-0.43 [-0.83, -0.04]	•
Heterogeneity: Tau <sup>2</sup> =	0.19; Cl	ni <b>z</b> = 23	3.85, df=	: 12 (P =	0.02);1	<b>≈</b> = 50%			
Test for overall effect:	Z= 2.16	(P = 0	1.03)						
1.14.3 Others									
Santana 2022	22.1	23.7	34	26.8	28.7	30	0.1%	-4.70 [-17.70, 8.30]	
Triamvisit 2016	2.9	3	68	5.8	9.3	66	2.9%	-2.90 [-5.25, -0.55]	
Subtotal (95% CI)			102			96	3.0%	-2.96 [-5.27, -0.64]	◆
Heterogeneity: Tau <sup>2</sup> =	0.00; Cł	ni² = 0.	.07, df = 1	1 (P = 0.	79); I² =	0%			
Test for overall effect:	Z = 2.50	(P = 0	1.01)						
Total (95% CI)			57076			9626	100.0%	-0.59 [-1.03, -0.15]	•
Heterogeneity: Tau <sup>2</sup> =	0.30; CI	ni² = 34	4.09, df=	: 15 (P =	0.003)	l² = 56	%	-	
Test for overall effect:	Z = 2.61	(P = 0	1.009)	-					-10 -5 0 5 10 Eavoure [CR] Eavoure [SC]
Test for subgroup diff	erences	: Chi²⊧	= 9.75, di	f = 2 (P =	= 0.008)	. <b>I</b> ² = 79	9.5%		

D) Forest plot on days of mechanical ventilation by study design subgroups.

## E) Forest plot on days of mechanical ventilation by country subgroups.

	care	e bund	les	stand	ard of c	are		Mean Difference	Mean Difference					
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% Cl					
1.15.1 Western EU														
Burja 2018	8.4	7.3	74	9.1	7.8	55	2.4%	-0.70 [-3.35, 1.95]						
Landelle 2018	6.8	4.2	356	7.6	4.8	291	11.7%	-0.80 [-1.50, -0.10]	+					
Morris 2011	5.4	5.9	501	5.1	5.2	1460	13.0%	0.30 [-0.28, 0.88]	+					
Pérez-Granda 2014	2.8	7.4	1534	4.1	11.9	401	7.3%	-1.30 [-2.52, -0.08]						
Subtotal (95% CI)			2465			2207	34.4%	-0.51 [-1.35, 0.33]	•					
Heterogeneity: Tau² =	0.42; Cł	ni² = 8.0	69, df=3	3 (P = 0.	03); I² =	65%								
Test for overall effect: Z = 1.19 (P = 0.23)														
1.15.2 North America														
Cacheco, 2012	18.9	11	655	21.5	15.4	299	4.0%	-2.60 [-4.54, -0.66]						
DeLuca 2017	2.1	2.2	192	2.7	3	195	13.5%	-0.60 [-1.12, -0.08]	•					
Ding 2013	6.4	5.2	137	5.7	4.5	213	8.5%	0.70 [-0.36, 1.76]						
Lansford 2007	5.3	7.4	132	7.3	12.2	218	3.6%	-2.00 [-4.05, 0.05]						
Ongstad 2013	4.5	5.3	96	4.7	5.2	87	5.6%	-0.20 [-1.72, 1.32]						
Sen 2016 Subtetel (05% CI)	16.2	27.2	65	19.7	25.3	66	0.2%	-3.50 [-12.50, 5.50]						
Subiotal (95% CI)			12//			1078	33.3%	-0.70 [-1.05, 0.22]	•					
Heterogeneity: lauf =	0.65; Cr	11 <b>* =</b> 12	2.19, at =	:5(P=l	J.U3); I*	= 59%								
l est for overall effect:	Z=1.49	(P = 0	.14)											
1.15.3 Other regions														
Arabneiad 2011	12.6	9.1	46	17.8	12.8	71	1.2%	-5.20 [-9.171.23]						
Ferreira 2016	22.9	28	115	20	26.3	73	0.3%	2.90 [-5.01, 10.81]	<u> </u>					
Khan 2016	6.9	14.1	1453	6.8	9	2212	10.6%	0.10 [-0.72, 0.92]	+					
Rosenthal 2012	6.3	10.6	51618	6.8	11.2	3889	15.0%	-0.50 [-0.86, -0.14]	-					
Santana 2022	22.1	23.7	34	26.8	28.7	30	0.1%	-4.70 [-17.70, 8.30]						
Triamvisit 2016	2.9	3	68	5.8	9.3	66	2.9%	-2.90 [-5.25, -0.55]						
Subtotal (95% CI)			53334			6341	30.1%	-0.88 [-1.95, 0.18]	◆					
Heterogeneity: Tau² =	0.66; Cł	ni <sup>z</sup> = 12	2.54, df=	: 5 (P = 0	0.03); I <mark>ř</mark>	= 60%								
Test for overall effect:	Z=1.63	(P = 0	.10)											
Total (95% CI)			57076			9626	100.0%	-0.59 [-1.03, -0.15]	•					
Heterogeneity: Tau <sup>2</sup> =	0.30; Cł	ni <sup>z</sup> = 34	.09. df=	: 15 (P =	0.003)	: <b> </b> ² = 56	%							
Test for overall effect:	Z = 2.61	(P = 0	.009)						-20 -10 0 10 20					
Test for subgroup diff	erences	Chi <b></b> ⁼=	: 0.30, di	f= 2 (P =	= 0.86),	l² = 0%								

F) Forest plot on days of mechanical ventilation by VAP diagnostic criteria subgroups.

	care	e bund	les	stand	ard of c	are		Mean Difference	Mean Difference				
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI				
1.16.1 Clinical+Micro	biologic	al											
DeLuca 2017	2.1	2.2	192	2.7	3	195	13.5%	-0.60 [-1.12, -0.08]	*				
Khan 2016	6.9	14.1	1453	6.8	9	2212	10.6%	0.10 [-0.72, 0.92]	+				
Landelle 2018	6.8	4.2	356	7.6	4.8	291	11.7%	-0.80 [-1.50, -0.10]	+				
Morris 2011	5.4	5.9	501	5.1	5.2	1460	13.0%	0.30 [-0.28, 0.88]	+				
Rosenthal 2012	6.3	10.6	51618	6.8	11.2	3889	15.0%	-0.50 [-0.86, -0.14]	-				
Sen 2016	16.2	27.2	65	19.7	25.3	66	0.2%	-3.50 [-12.50, 5.50]					
Subtotal (95% CI)			54185			8113	64.1%	-0.34 [-0.71, 0.03]	•				
Heterogeneity: Tau² = 0.09; Chi² = 9.51, df = 5 (P = 0.09); I² = 47%													
Test for overall effect: Z = 1.78 (P = 0.08)													
1.16.2 Clinical													
Burja 2018	8.4	7.3	74	9.1	7.8	55	2.4%	-0.70 [-3.35, 1.95]					
Cacheco, 2012	18.9	11	655	21.5	15.4	299	4.0%	-2.60 [-4.54, -0.66]					
Ding 2013	6.4	5.2	137	5.7	4.5	213	8.5%	0.70 [-0.36, 1.76]					
Ferreira 2016	22.9	28	115	20	26.3	73	0.3%	2.90 [-5.01, 10.81]					
Lansford 2007	5.3	7.4	132	7.3	12.2	218	3.6%	-2.00 [-4.05, 0.05]					
Ongstad 2013	4.5	5.3	96	4.7	5.2	87	5.6%	-0.20 [-1.72, 1.32]					
Santana 2022	22.1	23.7	34	26.8	28.7	30	0.1%	-4.70 [-17.70, 8.30]					
Triamvisit 2016	2.9	3	68	5.8	9.3	66	2.9%	-2.90 [-5.25, -0.55]					
Subtotal (95% CI)			1311			1041	27.4%	-1.05 [-2.29, 0.19]	•				
Heterogeneity: Tau <sup>2</sup> =	1.55; CI	ni² = 16	6.72, df=	: 7 (P = (	0.02); I <b>²</b>	= 58%							
Test for overall effect:	Z=1.66	(P = 0	.10)										
1 16 3 CDIS critoria													
Archnolog 2014	10.0	0.4	40	17.0	12.0	74	4.000	5 20 1 0 4 7 4 2 21					
Arabriejau 2011 Dároz Orondo 2014	12.0	9.1	40	17.8	12.8	404	1.270	-0.20[-9.17,-1.23]					
Subtotal (95% CI)	2.0	7.4	1534	4.1	11.9	401	7.3% 8.5%	-1.30 [-2.32, -0.08]					
Hotorogonoity: Tou <sup>2</sup> -	5 26· CI		20 df - 1	1 /P = 0	07\-18-	70%	0.070	-2.11 [-0.40, 0.55]					
Toet for overall effect:	7 - 1 /7	(P = 0)	38, ui = 171	r (F = 0.	07),11=	7070							
restior overall ellect.	∠ = 1.47	(= 0	.14)										
Total (95% CI)			57076			9626	100.0%	-0.59 [-1.03, -0.15]	•				
Heterogeneity: Tau <sup>2</sup> =	0.30; CI	ni² = 34	4.09, df=	: 15 (P =	0.003)	; <b>I</b> ² = 56	%						
Test for overall effect:	Test for overall effect: $Z = 2.61$ (P = 0.009) Test for overall effect: $Z = 2.61$ (P = 0.009)												
Test for subgroup diff	erences	∶Chi <del>"</del> =	= 2.74, d	f = 2 (P =	= 0.25),	l² = 26.	9%		Favours [CB] Favours [SC]				

**Figure S3.** Funnel plot (A) and Forest (B-F) plots on **hospital length of stay** in subjects treated with care bundles or standard care.

A) Funnel plot on length of hospital stay in subjects treated with care bundles or standard care.



B) Forest plot on length of hospital stay by baseline measures subgroups.



C) Forest plot on length of hospital stay by compliance subgroups.



	care bundles			standa	ard of o	are		Mean Difference	Mean Difference		
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI		
1.22.1 RCT											
Arabnejad 2011	25.1	17.5	46	36.1	21.3	71	2.2%	-11.00 [-18.08, -3.92]			
Subtotal (95% CI)			46			71	2.2%	-11.00 [-18.08, -3.92]			
Heterogeneity: Not ap	plicable										
Test for overall effect:	Z = 3.05	i (P = 0	).002)								
1.22.2 Cohorts											
Burja 2018	20.8	17.4	74	17.4	12.2	55	3.8%	3.40 [-1.71, 8.51]			
DeLuca 2017	5.5	7.5	192	7.4	9	195	13.7%	-1.90 [-3.55, -0.25]			
Ding 2013	20.8	17.2	137	17.4	11.9	213	7.2%	3.40 [0.11, 6.69]	<b>⊢</b>		
Ferreira 2016	20.2	8	73	16.6	5.4	115	11.6%	3.60 [1.52, 5.68]			
Morris 2011	7.4	5.9	501	7.8	6.7	1460	18.6%	-0.40 [-1.02, 0.22]			
Ongstad 2013	13.3	13.3	96	15	17.7	87	4.5%	-1.70 [-6.27, 2.87]			
Rosenthal 2012	6.4	9.4	51618	6.9	11.4	3889	19.3%	-0.50 [-0.87, -0.13]	-		
Sen 2016	30.8	41.8	65	27.1	28.8	66	0.8%	3.70 [-8.61, 16.01]			
Tao 2012	3.7	7.8	3330	6.1	22.9	3250	17.8%	-2.40 [-3.23, -1.57]	+		
Subtotal (95% CI)			56086			9330	97.3%	-0.17 [-1.23, 0.90]	•		
Heterogeneity: Tau² =	1.39; C	hi² = 4:	5.38, df=	= 8 (P < 1	0.0000°	l); l² = 8	2%				
Test for overall effect:	Z = 0.31	(P = 0	).76)								
1.22.3 Others											
Santana 2022	28	28.7	34	30.5	29.7	30	0.6%	-2.50 [-16.85, 11.85]			
Subtotal (95% CI)			34			30	0.6%	-2.50 [-16.85, 11.85]			
Heterogeneity: Not ap	plicable										
Test for overall effect:	Z = 0.34	(P = 0	).73)								
Total (95% CI)			56166			9431	100.0%	-0.37 [-1.47, 0.74]	•		
Heterogeneity: Tau <sup>2</sup> =	1.61; C	hi² = 5	3.68, df=	= 10 (P <	0.000	01); I <sup>z</sup> =	81%				
Fest for overall effect:	Z = 0.65	i (P = 0	).51)						-20 -10 0 10 20		
									Favours [CB] Favours [SC]		

D) Forest plot on length of hospital stay by study design subgroups.

#### E) Forest plot on length of hospital stay by country subgroups.



F) Forest plot on length of hospital stay by VAP diagnostic criteria subgroups.

	care bundles			stand	ard of c	are		Mean Difference	Mean Difference			
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI			
1.24.1 Clinical+Micro	obiologia	al										
DeLuca 2017	5.5	7.5	192	7.4	9	195	13.7%	-1.90 [-3.55, -0.25]				
Morris 2011	7.4	5.9	501	7.8	6.7	1460	18.6%	-0.40 [-1.02, 0.22]	+			
Rosenthal 2012	6.4	9.4	51618	6.9	11.4	3889	19.3%	-0.50 [-0.87, -0.13]	•			
Sen 2016	30.8	41.8	65	27.1	28.8	66	0.8%	3.70 [-8.61, 16.01]				
Tao 2012	3.7	7.8	3330	6.1	22.9	3250	17.8%	-2.40 [-3.23, -1.57]	÷.			
Subtotal (95% CI)			55706			8860	70.2%	-1.15 [-2.06, -0.24]	◆			
Heterogeneity: Tau² =	Heterogeneity: Tau² = 0.67; Chi² = 20.53, df = 4 (P = 0.0004); l² = 81%											
Test for overall effect	Z = 2.48	8 (P = 0	).01)									
1.24.2 Clinical												
Burja 2018	20.8	17.4	74	17.4	12.2	55	3.8%	3.40 [-1.71, 8.51]				
Ding 2013	20.8	17.2	137	17.4	11.9	213	7.2%	3.40 [0.11, 6.69]				
Ferreira 2016	20.2	8	73	16.6	5.4	115	11.6%	3.60 [1.52, 5.68]				
Ongstad 2013	13.3	13.3	96	15	17.7	87	4.5%	-1.70 [-6.27, 2.87]				
Santana 2022	28	28.7	34	30.5	29.7	30	0.6%	-2.50 [-16.85, 11.85]				
Subtotal (95% CI)			414			500	27.7%	2.62 [0.69, 4.55]	◆			
Heterogeneity: Tau² =	= 1.00; C	hi <sup>z</sup> = 4.	.99, df =	4 (P = 0)	.29); <b>i²</b> =	= 20%						
Test for overall effect	Z= 2.68	6 (P = 0	0.008)									
1.24.3 CPIS criteria												
Arabnejad 2011	25.1	17.5	46	36.1	21.3	71	2.2%	-11.00 [-18.08, -3.92]				
Subtotal (95% CI)			46			71	2.2%	-11.00 [-18.08, -3.92]				
Heterogeneity: Not a	oplicable	9										
Test for overall effect	Z = 3.05	5 (P = 0	).002)									
Total (95% CI)			56166			9431	100.0%	-0.37 [-1.47, 0.74]	♠			
Heterogeneity: Tau <sup>2</sup> =	= 1.61; C	hi² = 53	3.68, df=	= 10 (P =	(0.000	01); I <sup>z</sup> =	81%					
Test for overall effect	Z = 0.65	5 (P = 0	).51)						Favours [CB] Eavours [SC]			
Test for subgroup dif	ferences	: Chi <b></b> ≇∘										

Figure S4. Funnel plot (A) and Forest plots (B-F) on ICU length of stay in subjects treated with care bundles or standard care.

A) Funnel plot on ICU length of stay in subjects treated with care bundles or standard care.



B) Forest plot on ICU length of stay by educational interventions subgroups.



C) Forest plot on ICU length of stay by baseline measures subgroups.



#### D) Forest plot on ICU length of stay by study design subgroups.

	care	e bund	les	stand	ard of	care		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
1.34.1 RCT									
Arabnejad 2011 Subtotal (05% CI)	18.6	16.3	46	25.3	15.5	71	0.6%	-6.70 [-12.63, -0.77]	
Listeregeneitr bist on	nlianhla		40				0.0%	-0.70 [-12.05, -0.77]	
Test for everall effect:	piicapie 7 - 2 24	/D = 0	0.23						
restior overall ellect.	2 = 2.21	(P=0	.03)						
1.34.2 Cohorts									
Baxter 2005	5.8	8.3	3507	4.7	6.8	705	13.6%	1.10 [0.53, 1.67]	+
Burja 2018	11.3	10.9	74	11.7	9.5	55	1.6%	-0.40 [-3.93, 3.13]	
Cacheco, 2012	20.6	11.5	655	22.8	15.1	299	4.4%	-2.20 [-4.12, -0.28]	
DeLuca 2017	3.4	3.7	192	4.1	5.2	195	10.5%	-0.70 [-1.60, 0.20]	
Ding 2013	11.7	9	137	10.1	6.7	213	5.0%	1.60 [-0.16, 3.36]	+
Khan 2016	9.8	11.9	2212	8.5	9.7	1453	12.3%	1.30 [0.60, 2.00]	+
Landelle 2018	9.2	6.7	356	9.3	6.3	291	9.6%	-0.10 [-1.10, 0.90]	4
Lim 2015	3.7	3	14212	3.7	3	12913	17.0%	0.00 [-0.07, 0.07]	•
Omrane 2007	7.1	9.9	360	7.7	14	349	4.9%	-0.60 [-2.39, 1.19]	
Ongstad 2013	6.2	6.2	96	8.3	9.1	87	3.4%	-2.10 [-4.38, 0.18]	
Pérez-Granda 2014	4.7	3	1534	4.4	3.7	401	15.3%	0.30 [-0.09, 0.69]	+
Sen 2016	27.3	40	65	26.2	24.9	66	0.2%	1.10 [-10.33, 12.53]	
Triamvisit 2016	14.7	10.2	68	17.2	10.9	66	1.6%	-2.50 [-6.08, 1.08]	<del></del>
Subtotal (95% CI)			23468			17093	99.4%	0.13 [-0.33, 0.58]	•
Heterogeneity: Tau <sup>2</sup> =	0.30; Cł	ni² = 45	i.13, df=	= 12 (P ≺	0.0000	01); I² = 7	'3%		
Test for overall effect:	Z = 0.55	(P = 0)	.58)						
1.34.3 Others									
Subtotal (95% CI)			0			0		Not estimable	
Heterogeneity: Not ap	plicable								
Test for overall effect:	Not app	licable							
Total (95% CI)			23514			17164	100.0%	0.07 [-0.40, 0.54]	<b>♦</b>
Heterogeneity: Tau² =	0.33; Cł	ni² = 50	.07, df=	= 13 (P <	0.0000	01); I <b>²</b> = 7	4%		
Test for overall effect:	Z = 0.30	(P = 0	.76)						Favours [CB] Favours [SC]
Test for subgroup diff	erences	: Chi <sup>z</sup> =							

.....

## E) Forest plot on ICU length of stay by country subgroups.

	care bundles			standard of care				Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% Cl
1.35.1 Western EU									
Burja 2018	11.3	10.9	74	11.7	9.5	55	1.6%	-0.40 [-3.93, 3.13]	
Landelle 2018	9.2	6.7	356	9.3	6.3	291	9.6%	-0.10 [-1.10, 0.90]	
Pérez-Granda 2014	4.7	3	1534	4.4	3.7	401	15.3%	0.30 [-0.09, 0.69]	t
Subtotal (95% CI)			1964			747	26.4%	0.24 [-0.12, 0.60]	•
Heterogeneity: Tau² =	0.00; Cł	ni² = 0.1	66, df=0	2 (P = 0.	72); I²=	0%			
Test for overall effect:	Z=1.30	(P = 0	.19)						
1.35.2 North America									
Baxter 2005	5.8	8.3	3507	4.7	6.8	705	13.6%	1.10 [0.53, 1.67]	+
Cacheco, 2012	20.6	11.5	655	22.8	15.1	299	4.4%	-2.20 [-4.12, -0.28]	
DeLuca 2017	3.4	3.7	192	4.1	5.2	195	10.5%	-0.70 [-1.60, 0.20]	
Ding 2013	11.7	9	137	10.1	6.7	213	5.0%	1.60 [-0.16, 3.36]	
Omrane 2007	7.1	9.9	360	7.7	14	349	4.9%	-0.60 [-2.39, 1.19]	
Ongstad 2013	6.2	6.2	96	8.3	9.1	87	3.4%	-2.10 [-4.38, 0.18]	
Sen 2016	27.3	40	65	26.2	24.9	66	0.2%	1.10 [-10.33, 12.53]	
Subtotal (95% CI)			5012			1914	42.0%	-0.31 [-1.49, 0.87]	-
Heterogeneity: Tau <sup>2</sup> =	1.57; Cl	ni² = 28	6.08, df=	:6(P=0	).0002)	; I <sup>z</sup> = 779	6		
l est for overall effect:	Z = 0.52	(P = 0	.61)						
1.35.3 Other regions									
Arabnejad 2011	18.6	16.3	46	25.3	15.5	71	0.6%	-6.70 [-12.63, -0.77]	
Khan 2016	9.8	11.9	2212	8.5	9.7	1453	12.3%	1.30 [0.60, 2.00]	-
Lim 2015	3.7	3	14212	3.7	3	12913	17.0%	0.00 [-0.07, 0.07]	•
Triamvisit 2016	14.7	10.2	68	17.2	10.9	66	1.6%	-2.50 [-6.08, 1.08]	
Subtotal (95% CI)			16538			14503	31.5%	-0.05 [-1.39, 1.28]	•
Heterogeneity: Tau² =	1.04; Cl	ni² = 19	9.80, df=	: 3 (P = 0	).0002)	; <b>I<sup>2</sup> =</b> 859	6		
Test for overall effect:	Z = 0.08	(P = 0	.94)						
Total (95% CI)			23514			17164	100.0%	0.07 [-0.40, 0.54]	
Heterogeneity: Tau <sup>2</sup> =	0.33; CI	ni² = 50	).07, df=	:13 (P <	0.0000	)1); I <sup>2</sup> = 7	4%		
Test for overall effect:	Z = 0.30	(P = 0	.76)						-10 -5 0 5 10
Test for subgroup diff	erences	∶Chi <b></b> =							

#### F) Forest plot on ICU length of stay by VAP diagnostic criteria subgroups.



**Figure S5.** Funnel plot (A) and Forest plots (B-E) on **hospital mortality** in subjects treated with care bundles or standard care.

A) Funnel plot on hospital mortality in subjects treated with care bundles or standard care.



B) Forest plot on hospital mortality by educational interventions subgroups.



C) Forest plot on hospital mortality by compliance subgroups.



Test for subgroup differences:  $Chi^2 = 0.24$ , df = 1 (P = 0.62),  $l^2 = 0\%$ 

#### D) Forest plot on hospital mortality by country subgroups.

	care bu	ndles	standard o	f care		Odds Ratio	Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% Cl	M-H, Random, 95% CI
1.48.1 Western EU							
Burja 2018	36	74	27	55	6.5%	0.98 [0.49, 1.97]	
Hawe 2009	49	215	112	374	9.4%	0.69 [0.47, 1.02]	
Landelle 2018	73	356	68	291	9.5%	0.85 [0.58, 1.23]	
Morris 2011	101	501	367	1460	10.7%	0.75 [0.59, 0.96]	
Parisi 2016	56	136	79	226	8.9%	1.30 [0.84, 2.02]	_ <del></del>
Subtotal (95% CI)		1282		2406	45.0%	0.85 [0.69, 1.05]	◆
Total events	315		653				
Heterogeneity: Tau <sup>2</sup> :	= 0.02; Chi <sup>a</sup>	²= 5.78,	df = 4 (P = 0)	l.22); <b>l²</b> =	31%		
Test for overall effect	: Z = 1.53 (	P = 0.12	)				
1.48.2 North Americ	a						
Baxter 2005	1143	3507	175	705	11.1%	1.46 [1.22, 1.76]	+
DeLuca 2017	49	192	63	195	8.9%	0.72 [0.46, 1.12]	
Ding 2013	25	137	49	213	7.9%	0.75 [0.44, 1.28]	
Ongstad 2013	42	1189	46	1290	9.0%	0.99 [0.65, 1.52]	<u>+</u>
Subtotal (95% CI)		5025		2403	36.9%	0.98 [0.66, 1.46]	<b>•</b>
Total events	1259		333				
Heterogeneity: Tau <sup>2</sup> :	= 0.12; Chi <sup>a</sup>	²=13.38	3, df = 3 (P =	0.004); l <sup>a</sup>	²= 78%		
Test for overall effect	: Z=0.11 (	P = 0.91	)				
1.48.3 Other regions	5						
Ferreira 2016	22	73	70	115	7.1%	0.28 [0.15, 0.52]	<b>_</b>
Samra 2016	17	250	19	130	6.5%	0.43 [0.21, 0.85]	
Santana 2022	20	34	13	30	4.4%	1.87 [0.69, 5.05]	
Subtotal (95% CI)		357		275	18.1%	0.56 [0.21, 1.51]	
Total events	59		102				
Heterogeneity: Tau <sup>2</sup> :	= 0.60; Chi <sup>a</sup>	<b>²</b> = 10.21	l, df = 2 (P =	0.006); l <sup>a</sup>	²= 80%		
Test for overall effect	: Z=1.14 (I	P = 0.25	)				
Total (95% CI)		6664		5084	100.0%	0.83 [0.63, 1.08]	•
Total events	1633		1088				
Heterogeneity: Tau <sup>2</sup> :	= 0.16; Chi <sup>a</sup>	<b>²</b> = 53.53	8, df = 11 (P ·	< 0.0000	1); I² = 79	%	
Test for overall effect	: Z = 1.40 (	P = 0.16	)				Eavours ICB1 Eavours ISC1
The state of the second st	· · · · · · · · · · · · · · · · · · ·	0 L 12 - 1	10 16 0.00	0.075	7 0.04		, atomo [ob] , atomo [oo]

Test for subgroup differences: Chi<sup>2</sup> = 1.12, df = 2 (P = 0.57), l<sup>2</sup> = 0%

E) Forest plot on hospital mortality by VAP diagnostic criteria subgroups.

	care bur	Idles	standard o	f care		Odds Ratio		Odds Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% Cl		M-H, Random, 95% Cl	
1.49.1 Clinical+Micro	biological								
Baxter 2005	1143	3507	175	705	12.1%	1.46 [1.22, 1.76]		-	
DeLuca 2017	49	192	63	195	9.8%	0.72 [0.46, 1.12]			
Khan 2016	1388	1453	988	2212	0.0%	26.45 [20.35, 34.39]			
Landelle 2018	73	356	68	291	10.5%	0.85 [0.58, 1.23]			
Morris 2011	101	501	367	1460	11.6%	0.75 [0.59, 0.96]			
Samra 2016	17	250	19	130	7.3%	0.43 [0.21, 0.85]			
Subtotal (95% CI)		4806		2781	51.3%	0.82 [0.55, 1.23]		-	
Total events	1383		692						
Heterogeneity: Tau² =	: 0.17; Chi <sup>a</sup>	= 29.96	i, df = 4 (P ≺	0.00001)	); l² = 87%	1			
Test for overall effect:	Z = 0.95 (F	P = 0.34	)						
1.49.2 Clinical									
Buria 2018	36	74	27	55	7.2%	0.98 [0.49, 1.97]		<b>_</b>	
Ding 2013	25	137	49	213	8.8%	0.75 [0.44, 1.28]			
Ferreira 2016	22	73	70	115	7.9%	0.28 [0.15, 0.52]		_ <b></b>	
Ongstad 2013	42	1189	46	1290	9.9%	0.99 [0.65, 1.52]		-+-	
Santana 2022	20	34	13	30	5.0%	1.87 [0.69, 5.05]			
Subtotal (95% CI)		1507		1703	38.9%	0.78 [0.46, 1.34]			
Total events	145		205						
Heterogeneity: Tau <sup>2</sup> =	: 0.27; Chi <sup>a</sup>	= 15.15	i, df = 4 (P =	0.004); l <sup>a</sup>	'= 74%				
Test for overall effect:	Z=0.88 (F	P = 0.38)	)						
1 49 3 CPIS criteria									
Pariei 2016	66	126	70	226	0.0%	1 20 10 94 2 021			
Subtotal (95% CI)	50	136	73	220	9.8%	1.30 [0.84, 2.02]		•	
Total events	56		79		01070	100 [010 1, 2102]		•	
Heterogeneity: Not ar	nlicahle		10						
Test for overall effect:	7 = 1.18 (F	$P = 0.24^{\circ}$	)						
	2-1.10()	0.24	, ,						
Total (95% CI)		6449		4710	100.0%	0.84 [0.63, 1.12]		◆	
Total events	1584		976						
Heterogeneity: Tau <sup>2</sup> =	0.17; Chi <sup>a</sup>	= 50.41	, df = 10 (P <	< 0.0000 <sup>-</sup>	1); I <sup>2</sup> = 809	%			-
Test for overall effect:	Z=1.18 (F	<sup>o</sup> = 0.24)	)				0.05	U.Z 1 5 2U Eavoure ICP1 Eavoure ISC1	
Test for subgroup diff	, rerences: C								

#### Figure S6. Forest plot on ICU mortality in subjects treated with care bundles or standard care.

A) Forest plot on ICU mortality by baseline measures subgroups.

