Study	Ward randomised trial of Critical Care Outreach introduction in a hospital trial: Priestley 2004[131]
	medical team. Training of doctors and nurses included sessions on the use of an in-house 'patient at risk' (PAR) score to identify patients who might benefit from CCOT attention. Ward staff used PAR to trigger referral to CCOT and involvement of the admitting team's consultant. Depending on circumstances, CCOT might support and advise ward staff, remain with the patient and provide individual nursing care on the ward during crisis period, or facilitate admission to ICU. Duration 32 weeks (total trial period). Concurrent medication/care: for each of the randomised pairs of wards there were 3 consecutive 4-week periods with 1 ward in control and the other in outreach. Comments: analysed (n=1456) due to stepped wedge design not missing data. Dataset 2 (matched randomised) is reported here. Although this dataset includes fewer patients and did not allow for separate consideration of the training phase of intervention, it utilised the randomisation within ward pairings. (n=3090) Intervention 2: No critical care outreach team (rapid response team, hospital at night, medical emergency team and outreach team) present in hospital - No critical care outreach team present in hospital. The control wards moved from control to intervention wards via the 4 week training period. Duration: 32 weeks (total trial period). Concurrent medication/care: for each of the randomised pairs of wards there were 3 consecutive 4-week periods with 1 ward in control and the other in outreach. Comments: analysed (n=1336) due to stepped wedge design not missing data. Dataset 2 (matched randomised) is reported here. Although this dataset includes fewer patients, and did not allow for separate consideration of the training phase of intervention, it utilised the randomisation within ward pairings.
Funding	Academic or government funding.
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: 24 HOUR/7 DAY versus NO CRITICAL CARE OUTREACH TEAM PRESENT IN WARD. Protocol outcome 1: Length of hospital stay during study period. - Actual outcome: length of stay in hospital at 32 weeks trial; HR 0.907 (95%CI 0.835 to 0.985); Comments: Hazard ratio of data of 2733 patients. data set 2 (matched randomised). Although this dataset included fewer patients, and did not allow for separate consideration of the training phase of intervention, it utilised the randomisation within ward pairings. this had the advantage of excluding potential bias due to ward characteristics and time trends, as each outreach ward month is balanced by a control in the same month for the other (randomly chosen) member of the ward pair); Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: control a bit younger and more females; but due to step wedge design all controls turned to intervention eventually; randomisation based on ward pairings; Group 1 Number missing: 33, Reason: patients excluded because of incomplete data; Group 2 Number missing: 137, Reason: patients excluded because of incomplete data Protocol outcome 2: Mortality during study period. - Actual outcome: in-hospital mortality at 32 weeks trial; OR 0.523 (95%CI 0.322 to 0.849); Comments: odds of death of data of 2733 patients. data set 2 (matched	

Study	Ward randomised trial of Critical Care Outreach introduction in a hospital trial: Priestley 2004[131]
randomised). Although this dataset included fewer patients, and did not allow for separate consideration of the training phase of intervention, it utilised the randomisation within ward pairings. this had the advantage of excluding potential bias due to ward characteristics and time trends, as each outreach ward month is balanced by a control in the same month for the other (randomly chosen) member of the ward pair); Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: control a bit younger and more females; but due to step wedge design all controls turned to intervention eventually; randomisation based on ward pairings; Group 1 Number missing: 19, Reason: patients excluded because of incomplete data; Group 2 Number missing: 92, Reason: patients excluded because of incomplete data	
Protocol outcomes not reported by the study	Quality of life at during study period; Patient and/or carer satisfaction at during study period; Avoidable adverse events at during study period; In- hospital cardiac arrest at during study period; Number of DNAR orders at during study period; Re-admission to ICU at during study period.

Appendix E: Economic evidence tables

Study	Simmes 2014[146]			
Study details	**Population & interventions**	**Costs**	**Health outcomes**	**Cost effectiveness**
Economic analysis: CCA (health outcomes: cardiac arrests and/or deaths averted, severity of disease) **Study design:** before-and-after observational study **Approach to analysis:** bottom-up costing approach was used to calculate the mean cost per-patient day during the before and after study periods. **Perspective:** The Netherland healthcare perspective **Follow-up:** 1 year before and 2 year after **Treatment effect duration**[(a)]**:** 4 months **Discounting:** None	**Population:** Patients who stayed on the surgical ward for ≥ 72 hour after major general surgery. **Cohort settings:** Mean age: NR Male: NR **Intervention 1: (n=1376)** No rapid response system with consultation of doctor after observing abnormal vital signs was left to the discretion of the nurse, vital signs not routinely recorded 3 times daily and oxygen saturation and respiratory rate were not included in the standard observation protocol. **Intervention 2: (n=2410)** The introduction of a rapid response system which included the introduction of a medical emergency team (MET) and the use of a single parameter track and trigger system. The MET was doctor-led and included an intensivist and a critical care nurse and was accessible 24/7.	**Total costs (mean per patient-day):** Intervention 1: £463 Intervention 2: £484 Incremental (2−1): £21 (95% CI: NR; p=NR) **Currency & cost year:** 2009 Euros (presented here as 2009 UK pounds[(b)])] **Cost components incorporated:** Implementation, maintenance, training, nursing time, consultations, unplanned ICU admissions	**Cardiac arrests and/or deaths:** Intervention 1: 0.5% Intervention 2: 0.25% Incremental (2−1): -0.25% (95% CI: NR; p=NR) **Severity of disease (APACHEII score):** Intervention 1: 17.5 Intervention 2: 17.6 Incremental (2−1): 0.1 (95% CI: NR; p=NR) **ICU (length of stay) (Median):** Intervention 1: 3.5 Intervention 2: 3.0 Incremental (2−1): -0.5 (95% CI: NR; p=0.94) **Unplanned ICU admissions:** Intervention 1: 2.5% Intervention 2: 4.2% Incremental (2−1): 1.7% (95% CI: NR; p=NR)	**ICER (Intervention 2 versus Intervention 1):** NR **Analysis of uncertainty:** No sensitivity analysis is reported. A scenario analysis based on using lower APACHEII score (14) for identifying patients for admission to ICU showed that the mean cost per patient-day was reduced to £8.
Data sources				

Health outcomes: the health outcomes recorded included cardiac arrests and/or deaths and severity of disease measured using the APACHEII score. Data were collected for 1 year before and 2 years after the introduction of the RRS. The authors report that the RRS continued for 4 months. **Cost sources:** prices of personnel and ICU costs were retrieved from the Dutch guideline for cost analysis in health care (National unit costs).
Comments
Source of funding: NR. **Applicability and limitations :** the population is patients recovering from general surgery, not acute medical emergency. Some uncertainty regarding the applicability of resource use and costs from the Netherlands in 2009 to the current UK NHS context. QALYs were not used as an outcome. Costs and outcomes were not discounted. Longitudinal observational study with no adjustment for temporal variation or confounders. The follow-up was different in the before and after periods (1 year versus 2 years) and it is not clear whether this follow-up adequately captures all relevant costs and outcomes. Only 1 scenario analysis is reported.
Overall applicability[(c)]:partially applicable **Overall quality[(d)]:** potentially serious limitations

Abbreviations: CCA: cost–consequence analysis; 95% CI: 95% confidence interval; ICER: incremental cost-effectiveness ratio; NR: not reported; QALYs: quality-adjusted life years.

(a) For studies where the time horizon is longer than the treatment duration, an assumption needs to be made about the continuation of the study effect. For example, does a difference in utility between groups during treatment continue beyond the end of treatment and if so for how long?

(b) Converted using 2009 purchasing power parities.[127]

(c) Directly applicable/Partially applicable/Not applicable.

(d) Minor limitations/Potentially serious limitations/Very serious limitations.

Appendix F: GRADE tables

Table 6: Clinical evidence profile: Critical care outreach team versus usual care

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Critical care outreach team	Control	Relative (95% CI)	Absolute		
In-hospital mortality												
3	randomised trials	serious[1]	serious[2]	no serious indirectness	no serious imprecision	none	-	-	OR 0.95 (0.8 to 1.12)	See footnote[4]	⊕⊕OO LOW	CRITICAL
Length of stay (hazard ratio)												
1	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	-	-	HR 0.91 (0.84 to 0.99)	See footnote[4]	⊕⊕OO LOW	CRITICAL
Cardiac arrest												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	-	-	OR 0.94 (0.79 to 1.12)	See footnote[4]	⊕⊕⊕O MODERATE	CRITICAL
Cardiopulmonary resuscitation												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	none	-	-	OR 1.00 (0.69 to 1.45)	See footnote[4]	⊕⊕OO LOW	CRITICAL
Unplanned ICU admission												
1	randomised trials	Serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	-	-	OR 1.04 (0.89 to 1.22)	See footnote[4]	⊕⊕⊕O MODERATE	IMPORTANT

ICU admission												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	none	–	-	OR 1.15 (0.64 to 2.07)	See footnote[4]	⊕OOO VERY LOW	IMPORTANT
DNAR orders issued												
1	randomised trials	Serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	160/4161 (3.8%)	1.7%	RR 2.24 (1.61 to 3.1)	21 more per 1000 (from 10 more to 36 more)	⊕⊕⊕O MODERATE	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias.
[2] Downgraded by 1 or 2 increments because: the point estimate varies widely across studies, the confidence intervals across studies show minimal or no overlap, Heterogeneity, I2>50%, unexplained by subgroup analysis.
[3] Downgraded by 1 increment if the confidence interval crossed 1 MID or by 2 increments if the confidence interval crossed both MIDs.
[4] Absolute values could not be calculated as the papers reported adjusted analyses only without control event rates.

Appendix G: Excluded clinical studies

Table 7: Studies excluded from the clinical review

Study	Exclusion reason
Adelstein 2011[8]	Incorrect study design (not RCT, prospective cohort)
Aftyka 2014[10]	Incorrect study design (not RCT, before and after)
Aftyka 2014A[9]	Not relevant as it is not pertaining to in-hospital medical emergency teams
Al kadri 2010[11]	Incorrect population (obstetrics). not comparable to UK setting (Saudi Arabia)
Al-qahtani 2013[12]	Not comparable to UK setting (Saudi Arabia)
Aneman 2006[13]	Systematic review: literature search not sufficiently rigorous
Anon 2005C[2]	Incorrect study design (not RCT)
Anon 2005D[1]	Correction for Hillman 2005 (data not relevant for our analysis)
Anon 2006A[4]	Incorrect study design (not RCT)
Anon 2006F[3]	Incorrect study design (not RCT)
Anon 2008A[5]	Incorrect study design (not RCT, commentary)
Anon 2009B[6]	Incorrect study design (not RCT, commentary)
Anon 2013[7]	Incorrect study design (not RCT, commentary)
Anwar 2010[14]	Incorrect age group
Austin 2014[15]	Not comparable to UK setting (USA)
Ball 2003[16]	Incorrect study design (not RCT, before and after study)
Bannard-Smith 2016[17]	Incorrect study design (non-RCT; prospective observational cohort study)
Barbetti 2008[18]	Systematic review: literature search not sufficiently rigorous
Barnes 2015[19]	Incorrect study design (not RCT, before and after study)
Baxter 2008[20]	Incorrect study design (not RCT, audit)
Beckett 2009[21]	Incorrect study design (not RCT, cohort study)
Beitler 2011[22]	not comparable to UK setting (USA)
Bellomo 2003[23]	Incorrect study design (not RCT, cohort study)
Bellomo 2004[24]	Incorrect study design (not RCT, cohort study)
Blotsky 2016[25]	Non-RCT; before/after study
Bokhari 2010[26]	Incorrect study design (not RCT, cohort study n<200)
Bonafide 2014[27]	Incorrect age group
Boniatti 2014[28]	Not comparable to UK setting (Brazil)
Bosch 2008[29]	Incorrect study design (not RCT, before and after study)
Brilli 2007[30]	Incorrect age group
Bristow 2000[31]	Incorrect study design (not RCT, cohort study)
Buist 2002[33]	Incorrect study design (not RCT, cohort study)
Buist 2007[32]	Incorrect study design (not RCT, audit)
Cabrini 2009[34]	Incorrect intervention and comparison
Calzavacca 2008[37]	Incorrect study design (not RCT, prospective cohort)
Calzavacca 2009[35]	Incorrect study design (no RCT, poster of a retrospective observational study)
Calzavacca 2010[38]	Incorrect study design (not RCT, cohort study)

Study	Exclusion reason
Calzavacca 2010[36]	Incorrect study design (not RCT, retrospective cohort)
Campello 2009[39]	Incorrect study design (not RCT, before and after study)
Chaboyer 2004[41]	Incorrect study design (not RCT, commentary)
Chan 2008[43]	Not comparable to UK setting (USA)
Chan 2010[42]	Systematic review: literature search not sufficiently rigorous
Chen 2009[46]	No relevant outcomes reported (original study Hillman 2005 is included)
Chen 2014[44]	Incorrect study design (not RCT)
Chen 2014[48]	Incorrect study design (not RCT, population based study)
Chen 2015[45]	Incorrect comparison (delayed call versus non-delayed call). Data from Merit study (already included) analysed, no new outcomes
Chittawatanarat 2013[49]	Incorrect study design (not RCT, retrospective review)
Dacey 2007[50]	Incorrect study design (not RCT, before and after study)
De 2016[52]	Letter
Dechert 2013[53]	Not comparable to UK setting (USA)
Devita 2004[55]	Incorrect study design (not RCT, before and after study)
Downar 2013[56]	Incorrect study design (not RCT, retrospective review)
Downey 2008[57]	Incorrect study design (not RCT, cohort study n<200)
Eliott 2008[58]	Study not relevant (not pertaining to outreach service)
Esmonde 2006[59]	Systematic review: literature search not sufficiently rigorous
Findlay 2011[60]	Incorrect study population (trauma)
Flabouris 2010[61]	No outcomes relevant to our protocol (original paper Hillman 2005 fully included)
Galhotra 2010[62]	Not comparable to UK setting (USA)
Gao 2007[63]	Incorrect study design (not RCT, interrupted time-series analysis)
Garcea 2004[64]	Incorrect study design (not RCT, observational study)
Georgeto 2011[65]	Incorrect study design (not RCT, before and after study)
Gerdik 2010[66]	Incorrect population (Trauma)
Gessner 2007[67]	Not comparable to UK setting (USA)
Gilman 2014[68]	Incorrect comparison (hospitalised versus non-hospitalised patients)
Goncales 2012[69]	Not comparable to UK setting (Brazil)
Gray 2011[70]	Incorrect study design (not RCT, poster of observational study)
Haji 2004[71]	Incorrect study design (not RCT, retrospective audit)
Hanson 2009[72]	Incorrect age group
Hanson 2010[73]	Incorrect age group
Harrison 2010[74]	Incorrect study design (not RCT, cohort study)
Hatler 2009[75]	Incorrect study design (not RCT, before and after study)
Hayani 2011[76]	Incorrect study design (not RCT)
Hourihan 1995[79]	Incorrect study design (not RCT, prospective cohort)
Howell 2012[80]	Not comparable to UK setting (USA)
Jaderling 2011[83]	Incorrect study design (not RCT, retrospective cohort study)
Jaderling 2013[82]	Incorrect study design (not RCT, prospective observational study)
Jolley 2007[85]	Incorrect study design (not RCT, quasi experimental)
Jones 2005[87]	Incorrect study design (not RCT, prospective controlled study)
Jones 2007[91]	Incorrect study design (not RCT, before and after study)

Study	Exclusion reason
Jones 2007[93]	Incorrect study design (not RCT, cohort study n<200)
Jones 2007[88]	Incorrect study population (surgical patients)
Jones 2008[89]	Incorrect study design (not RCT, retrospective cohort)
Jones 2012[92]	Incorrect study design (not RCT, prospective observational study)
Jones 2013[90]	Systematic review: literature search not sufficiently rigorous
Jones 2013[86]	Incorrect study design (not RCT, retrospective cohort study)
Karpman 2013[95]	Not comparable to UK setting (USA)
Karvellas 2012[96]	Not comparable to UK setting (Brazil)
Kenward 2004[97]	Incorrect study design (not RCT, cohort study)
Kim 2012[98]	Incorrect study design (not RCT, prospective observational study)
King 2006[99]	Incorrect study design (not RCT, before and after study)
Knott 2011[100]	Incorrect study design (not RCT, Retrospective cohort) Not relevant (pertains to effect of outreach teams on documentation of advance care directives)
Konrad 2010[101]	Incorrect study design (not RCT, prospective before and after trial)
Kotsakis 2011[102]	Incorrect age group
Kwak 2014[103]	Incorrect study design (not RCT, observational study)
Laurens 2010[105]	Systematic review: literature search not sufficiently rigorous
Laurens 2011[104]	Incorrect study design (not RCT, before and after)
Leary 2003[106]	Incorrect study design (not RCT, before and after study)
Lee 1995[107]	Incorrect study design (not RCT, observational study)
Lighthall 2010[108]	Not comparable to UK setting (USA)
Lim 2011[109]	Incorrect study design (not RCT, before and after study)
Maharaj 2015[111]	Systematic review (study designs are inappropriate)
Mailey 2006[112]	Not comparable to UK setting (USA)
Massey 2010[113]	Systematic review: literature search not sufficiently rigorous
Mcarthur-rouse 2001[114]	Systematic review: literature search not sufficiently rigorous
Mcfarlan 2007[115]	Not comparable to UK setting (USA)
Mcneill 2013[117]	Systematic review: literature search not sufficiently rigorous
Medina-rivera 2010[118]	Not comparable to UK setting (Puerto Rico)
Meredith 2005[119]	Incorrect study design (not RCT, before and after study)
Moriarty 2014[121]	Not comparable to UK setting (USA)
Moroseos 2014[122]	Not comparable to UK setting (USA). Incorrect study population (surgery patients)
Morris 2013[123]	Incorrect study design (not RCT, retrospective cohort study n<200)
Muchoki 2015[124]	Poster presentation of an observational study
Niven 2014[125]	Systematic review: literature search not sufficiently rigorous
Offner 2007[126]	Incorrect population (Trauma)
Orosz 2014[128]	Incorrect study design (not RCT, retrospective cohort)
Pirret 2008[129]	Incorrect study design (not RCT)
Pittard 2003[130]	Incorrect study design (not RCT, before and after study)
Ranji 2007[132]	Systematic review: literature search not sufficiently rigorous
Rashid 2014[133]	Not comparable to UK setting (India)
Reza 2015[134]	Incorrect study design (report on the implementation of a pulmonary

Study	Exclusion reason
	embolism response team)
Rothschild 2008[135]	Incorrect study design (not RCT)
Sabahi 2012[137]	Not comparable to UK setting (Dubai)
Salamonson 2001[138]	Incorrect study design (not RCT, retrospective review of hospital data)
Salvatierra 2014[139]	Not comparable to UK setting (USA)
Santamaria 2010[140]	Incorrect study design (not RCT, prospective cohort study)
Sarani 2011[141]	Incorrect study design (not RCT, retrospective review)
Sebat 2007[142]	Not comparable to UK setting (USA)
Segon 2014[143]	Not comparable to UK setting (USA)
Shah 2011[144]	Not comparable to UK setting (USA)
Sharek 2007[145]	Incorrect age group
Simmes 2012[148]	Incorrect study population (surgical patients)
Simmes 2013[147]	Incorrect study population (surgical patients)
Smith 2014[149]	Incorrect study design (not RCT, retrospective cohort)
Solomon 2016[150]	Systematic review (references screened)
Story 2004[152]	Incorrect study design (not RCT, cohort study)
Story 2013[151]	Incorrect study design (not RCT, audit)
Subbe 2003[153]	Conference abstract of RCT but looking at effect of physiological scoring system rather than outreach team
Tam 2014[154]	Incorrect study design (not RCT, retrospective chart review)
Tan 2014[155]	Systematic review: literature search not sufficiently rigorous
Tibballs 2005[156]	Incorrect age group
Tibballs 2009[157]	Incorrect age group
Tobin 2012[158]	Incorrect study design (not RCT, retrospective cohort study)
Vazquez 2009[159]	Not comparable to UK setting (USA)
Williams 2010[160]	Incorrect study design (not RCT, before and after study)
Winters 2007[161]	Systematic review: literature search not sufficiently rigorous
Winters 2013[162]	Systematic review: literature search not sufficiently rigorous
Young 2002[163]	Incorrect study design (not RCT, abstract of a before and after study)
Young 2008[164]	Incorrect study design (not RCT, retrospective analysis of audit forms)
Zorko 2013[165]	Incorrect age group

Appendix H: Excluded economic studies

No studies were excluded.