A comparison of two trials

Results

• Relative risk 0.84 (relative reduction of 16%)
• Relative risk 0.84 (relative reduction of 16%)

Interpretation

• The intervention did not reduce risk of the outcome
• The intervention did reduce the risk of the outcome
The Optimise trial

Sample size calculation
- Predicted 12.5% absolute risk reduction (RR 25%)
- 690 participants per group

Results
- Absolute risk reduction 6.8% (RR 16%)
- P = 0.07

Conclusion
- The intervention did not reduce the incidence of the outcome
Figure 2. Cumulative Incidence of Mortality Up to 180 Days After Surgery Using a Cardiac Output-Guided Hemodynamic Therapy Algorithm Intervention vs Usual Care

Cumulative Mortality, %

Time From Start of Surgery, d

Usual care
Intervention

Log-rank $P = .09$

No. at risk
Intervention 368 350 344 339 334 333 306
Usual care 365 348 331 325 321 317 286

Pearse 2014. Optimise
The POISE trial

Sample size calculation
- Predicted 1.5% absolute risk reduction (RR 25%)
- 4000 participants per group

Results
- Absolute risk reduction 1.1% (RR 16%)
- P = 0.04

Conclusion
- The intervention reduces the incidence of the outcome
Choose the correct type of study

SIGN classification
Choose the correct type of study

Box 2. SIGN classification for grading evidence

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1+++</td>
<td>High-quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias</td>
</tr>
<tr>
<td>1+</td>
<td>Well-conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias</td>
</tr>
<tr>
<td>1−</td>
<td>Meta-analyses, systematic reviews, or RCTs with a high risk of bias</td>
</tr>
<tr>
<td>2+++</td>
<td>High-quality systematic reviews of case-control or cohort studies</td>
</tr>
<tr>
<td></td>
<td>High-quality case-control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal</td>
</tr>
<tr>
<td>2+</td>
<td>Well-conducted case-control or cohort studies with a low risk of confounding or bias and a moderate probability that the relationship is causal</td>
</tr>
<tr>
<td>2−</td>
<td>Case-control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal</td>
</tr>
<tr>
<td>3</td>
<td>Non-analytic studies; for example, case reports, case series</td>
</tr>
<tr>
<td>4</td>
<td>Expert opinion</td>
</tr>
</tbody>
</table>

RCT: randomised controlled trial; SIGN: Scottish Intercollegiate Guidelines Network

Belsey 2009. What is EBM?
Figure 3. Meta-analysis of Number of Patients Developing Complications After Surgery

<table>
<thead>
<tr>
<th>Source</th>
<th>Intervention</th>
<th>Control</th>
<th>Risk Ratio (95% CI)</th>
<th>Favors Intervention</th>
<th>Favors Control</th>
<th>Weight, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoemaker et al, 1988</td>
<td>8/28</td>
<td>30/60</td>
<td>0.57 (0.30-1.08)</td>
<td>-</td>
<td>-</td>
<td>1.7</td>
</tr>
<tr>
<td>Berlauk et al, 1991</td>
<td>11/68</td>
<td>9/21</td>
<td>0.38 (0.18-0.79)</td>
<td>-</td>
<td>-</td>
<td>1.3</td>
</tr>
<tr>
<td>Mythen et al, 1995</td>
<td>0/30</td>
<td>6/30</td>
<td>0.08 (0.00-1.31)</td>
<td>-</td>
<td>-</td>
<td>0.1</td>
</tr>
<tr>
<td>Sinclair et al, 1997</td>
<td>1/20</td>
<td>1/20</td>
<td>1.00 (0.07-14.90)</td>
<td>-</td>
<td>-</td>
<td>0.1</td>
</tr>
<tr>
<td>Ueno et al, 1998</td>
<td>4/16</td>
<td>5/18</td>
<td>0.90 (0.29-2.78)</td>
<td>-</td>
<td>-</td>
<td>0.5</td>
</tr>
<tr>
<td>Wilson et al, 1999</td>
<td>38/92</td>
<td>28/46</td>
<td>0.68 (0.48-0.95)</td>
<td>-</td>
<td>-</td>
<td>6.2</td>
</tr>
<tr>
<td>Lobo et al, 2000</td>
<td>6/19</td>
<td>12/18</td>
<td>0.47 (0.23-0.99)</td>
<td>-</td>
<td>-</td>
<td>1.3</td>
</tr>
<tr>
<td>Jerez et al, 2001</td>
<td>53/181</td>
<td>65/209</td>
<td>0.94 (0.70-1.28)</td>
<td>-</td>
<td>-</td>
<td>7.6</td>
</tr>
<tr>
<td>Conway et al, 2002</td>
<td>5/29</td>
<td>9/28</td>
<td>0.54 (0.20-1.40)</td>
<td>-</td>
<td>-</td>
<td>0.8</td>
</tr>
<tr>
<td>Pearse et al, 2005</td>
<td>27/62</td>
<td>41/60</td>
<td>0.64 (0.46-0.89)</td>
<td>-</td>
<td>-</td>
<td>6.3</td>
</tr>
<tr>
<td>Wakeling et al, 2005</td>
<td>24/67</td>
<td>38/67</td>
<td>0.63 (0.43-0.93)</td>
<td>-</td>
<td>-</td>
<td>4.8</td>
</tr>
<tr>
<td>Noblett et al, 2006</td>
<td>1/51</td>
<td>8/52</td>
<td>0.13 (0.02-0.98)</td>
<td>-</td>
<td>-</td>
<td>0.2</td>
</tr>
<tr>
<td>Donati et al, 2007</td>
<td>8/68</td>
<td>20/67</td>
<td>0.39 (0.19-0.83)</td>
<td>-</td>
<td>-</td>
<td>1.3</td>
</tr>
<tr>
<td>Smetkin et al, 2009a</td>
<td>1/20</td>
<td>4/20</td>
<td>0.25 (0.03-2.05)</td>
<td>-</td>
<td>-</td>
<td>0.2</td>
</tr>
<tr>
<td>Jhanji et al, 2010</td>
<td>57/90</td>
<td>30/45</td>
<td>0.95 (0.73-1.23)</td>
<td>-</td>
<td>-</td>
<td>10.4</td>
</tr>
<tr>
<td>Mayer et al, 2010</td>
<td>6/30</td>
<td>15/30</td>
<td>0.40 (0.18-0.89)</td>
<td>-</td>
<td>-</td>
<td>1.1</td>
</tr>
<tr>
<td>Ceccon et al, 2011</td>
<td>16/20</td>
<td>20/20</td>
<td>0.80 (0.64-1.02)</td>
<td>-</td>
<td>-</td>
<td>12.8</td>
</tr>
<tr>
<td>Challand et al, 2012</td>
<td>10/89</td>
<td>13/90</td>
<td>0.78 (0.36-1.68)</td>
<td>-</td>
<td>-</td>
<td>1.2</td>
</tr>
<tr>
<td>Brandstrup et al, 2012a</td>
<td>23/71</td>
<td>24/79</td>
<td>1.07 (0.66-1.71)</td>
<td>-</td>
<td>-</td>
<td>3.1</td>
</tr>
<tr>
<td>Salzwedel et al, 2013a</td>
<td>21/79</td>
<td>36/81</td>
<td>0.60 (0.39-0.93)</td>
<td>-</td>
<td>-</td>
<td>3.6</td>
</tr>
<tr>
<td>Goepfert et al, 2013a</td>
<td>34/50</td>
<td>42/50</td>
<td>0.81 (0.65-1.01)</td>
<td>-</td>
<td>-</td>
<td>13.7</td>
</tr>
<tr>
<td>OPTIMISE, 2014</td>
<td>134/368</td>
<td>158/365</td>
<td>0.84 (0.70-1.01)</td>
<td>-</td>
<td>-</td>
<td>21.8</td>
</tr>
<tr>
<td>Total</td>
<td>488/1548</td>
<td>614/1476</td>
<td>0.77 (0.71-0.83)</td>
<td>-</td>
<td>-</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Heterogeneity: $\chi^2 = 30.44; P = .08; I^2 = 31\%$
Test for overall effect: $z = 6.22; P < .001$

Pearse 2014. Optimise
Primary outcome
Core Outcome Measures for Perioperative and Anaesthesics Care (COMPAC)

Standards for definitions and use of outcome measures for clinical effectiveness research in perioperative medicine: European Perioperative Clinical Outcome (EPCO) definitions

A statement from the ESA-ESICM joint taskforce on perioperative outcome measures

Ib Jammer, Nadine Wickboldt, Michael Sander, Andrew Smith, Marcus J. Schultz, Paolo Pelosi, Brigitte Leva, Andrew Rhodes, Andreas Hoeft, Bernhard Walder, Michelle S. Chew and Rupert M. Pearse

Sample size

- Larger samples ➔ more precise estimates
- Greater precision = narrow confidence intervals
- Two approaches:
  1. Estimation of a measure with a specified precision
  2. Based on hypothesis test or demonstrating a specified significance level
Error

Type I error

- Probability of rejecting null hypothesis when it is true

Type II error

- Probability of not rejecting the null hypothesis when it is false
Error

Type I error
(false positive)

You’re pregnant

Type II error
(false negative)

You’re not pregnant
Power

- Probability that a hypothesis test will give a statistically significant result if an effect truly exists
  
- $1 - \text{type II error rate}$

- Chosen \textit{a priori}

- \textit{Usually 80%}
Sample size for a hypothesis test

- Base level of incidence in one group (usually control)
- Minimum between group difference that you want to detect e.g. RRR of >25%
- Power (usually 80%)
- Significance level or type I error rate (usually 5%)
Sample size for a hypothesis test
Sample size for a hypothesis test

<table>
<thead>
<tr>
<th>What confidence level do you need?</th>
<th>95 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical choices are 90%, 95% or 99%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>What power do you need?</th>
<th>80 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>A common choice is 80%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>What do you believe the likely sample proportion in group 1 to be?</th>
<th>60 %</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>What do you believe the likely sample proportion in group 2 to be?</th>
<th>40 %</th>
</tr>
</thead>
</table>

Your recommended sample size is 95

This is the minimum sample size you need for each group to detect whether the stated difference exists between the two proportions (with the required confidence level and power).

http://select-statistics.co.uk/sample-size-calculator-two-proportions
Primary outcome

Composite outcome?
We aimed to compare NEWS to the Patient at Risk Score (PARS), the existing early warning score in use at our hospital (Table 1). We chose two outcome measures: a composite of mortality and critical care unit (level two or three care) escalation within the first 48 h of the admission, and hospital length of stay. This composite outcome measure has previously been used in similar studies and will capture all patients suffering cardiac arrest at our institution.⁹,¹⁰,¹³
Trial design

Day of surgery
- Consent
- Randomisation
- Elective Surgery
- Intervention

Before surgery
- Screening
- Patient information sheet

In-hospital follow-up

1-year outcomes

30-day outcomes

PRISM trial (unpublished) | www.prismtrial.org
Trial registration

Browse Studies
- Cancer
- Circulatory System
- Digestive System
- Ear, Nose and Throat
- Eye Diseases
- Genetic Diseases
- Haematological Disorders
- Infections and Infestations

ISRCTN registry


Queen Mary
University of London
www.qmul.ac.uk
P I C O

Population

Intervention

Comparator

Outcome measure(s)
Top Tips

• Consider likely treatment effect and power appropriately
• Involve a clinical trials unit
• Approach a statistician early (and include in the grant)
• Use the Research Design Service (RDS)
• Many stats MSc require projects for their students (!)
Any questions?
Further reading

Essential medical statistics (2003). Kirkwood and Sterne

Pearse et al. Effect of a perioperative cardiac output guided hemodynamic therapy algorithm on outcomes follow major gastrointestinal surgery. *JAMA* 2014
