Profound effect of study design factors on ventilator-associated pneumonia incidence of prevention studies: benchmarking the literature experience

James C. Hurley1−3*

1School of Rural Health, University of Melbourne, Australia; 2Infection Control Units, Ballarat Health Services and St John of God Hospital, Ballarat, Australia; 3Division of Internal Medicine, Ballarat Health Services, PO Box 577, Ballarat, Victoria 3353, Australia

Received 30 December 2007; returned 23 January 2008; revised 6 February 2008; accepted 11 February 2008

Background: The ventilator-associated pneumonia incident proportion (VAP-IP) is highly variable among control groups of studies of methods for its prevention. The objective here is to develop and validate a literature-derived benchmark against which these groups can be profiled.

Methods: A literature search yielded 95 cohort groups and control and intervention groups of 150 studies of either non-antimicrobial or antimicrobial methods of VAP prevention. The 95 cohort groups comprise a benchmark set (30 groups), from which the reference funnel plot (RFP) was derived, and a search set (65 groups), against which the benchmark was validated. The VAP-IP data of the benchmark set were found in five published systematic reviews, whereas the VAP-IP data of the search set were abstracted directly from the literature.

Findings: Among the 95 cohort groups, the VAP-IP of groups with size >399 was significantly lower than the VAP-IP of smaller groups. Compared with the RFP, 15 of 51 (29%) control groups from studies of antimicrobial methods of VAP prevention with concurrent design were high outlier versus 2 of 110 (2%) control groups from other types of study design (P<0.001). There were only 22 (14%) outlier groups, all low outlier, among the 162 intervention groups.

Conclusions: Study design factors such as concurrency and study size have potentially greater influence on the VAP-IP than do the VAP prevention methods under study. The outlier status of control groups were inapparent in the individual studies and the meta-analyses and yet would have confounded the estimates of treatment effect.

Keywords: antimicrobial prophylaxis, cross-infection, funnel plots

Introduction

There are several published range estimates for the ventilator-associated pneumonia incident proportion (VAP-IP) among patients receiving prolonged (>48 h) mechanical ventilation (Table 1). These estimates represent expert opinion and are in turn derived from the results of over 30 observational studies.1−3 In addition, there are >140 intervention studies of various non-antimicrobial-based methods and antimicrobial-based methods of VAP prevention in this patient group. The results of these studies, which are summarized in over 20 narrative and systematic reviews,4−28 are marked by heterogeneity in treatment effect and VAP-IP data, which are highly variable, particularly so for studies of antimicrobial-based methods of VAP prevention.

The effect of widespread antimicrobial use in the ICU environment is 2-fold. It may prevent infection, but it may also alter the ecology of the ICU and increase the colonization pressure and cross-infection risk.29,30 The impact on cross-infection on the results of these studies of VAP prevention methods is unknown.31 Outbreaks and cross-infection occur in the ICU setting but both are thought to be under-recognized by conventional surveillance methods.32,33 However, using molecular typing techniques to identify cross-colonization, this accounts for up to 23% of colonization and up to 37% of cases of infection with Staphylococcus aureus in the typical ICU setting.34,35
Patterns of VAP isolates including an increase in *S. aureus* have been noted among control groups of studies of antimicrobial methods of VAP prevention suggesting that inapparent outbreaks had occurred. Partly because of these concerns, in studies of antimicrobial-based methods, different study designs had been used such as concurrent versus non-concurrent group design, and use or non-use of topical placebo to achieve study blinding where topical antimicrobial is one of the study interventions.

In the interpretation of these prevention studies, both for any one study and also in the summary result of a meta-analysis, the presumption is that the study effect occurs only in the intervention groups (i.e. a reduction in VAP-IP) and not in the control groups (i.e. an increase in VAP-IP). This presumption has never been challenged and cannot be tested without a benchmark and a method for profiling the results. Moreover, an objective benchmark against which infection rates could be profiled would also facilitate the detection of outbreaks.

Funnel plot methodology is commonly used to detect outlier results in systematic reviews. More recently, funnel plot methodology has been applied as a type of statistical process control charting, to profiling the performance of hospitals having different levels of patient risk and size to enable the identification of outliers.

The objective of this analysis is to profile the VAP-IP of intervention and control groups of these studies.

### Methods

#### Overview

The objectives of this analysis are: (i) to develop and validate a literature-derived benchmark of VAP-IP using two sets of cohort groups from observational studies; and (ii) to profile the VAP-IP among groups, whether control or intervention, from studies of VAP prevention methods against this benchmark.

#### Data sources and inclusion criteria

The flow of studies through the analysis is detailed in Figure 1. The study groups were sourced in three ways. First, the benchmark groups are derived from the five publications that report range estimates and also the VAP-IP data on which each is based. Second, a computerized search of PUBMED (including MEDLINE) was performed, using the following keywords: ventilator-associated pneumonia, incidence rate and mechanical ventilation up to December 2007. Third, the literature has been hand searched over a period of 14 years, and systematic reviews and meta-analysis of VAP prevention methods were also searched for eligible studies.

The study inclusion criteria were as follows: a publication reporting the VAP-IP for a group of patients receiving prolonged (>48 h) mechanical ventilation. An eligible study could provide more than one group for inclusion if discrete groups could be identified meeting the entry criteria. Studies in a language other than English were included when the required data had been abstracted in an English language systematic review.

The study exclusion criteria were as follows: studies conducted in animals, studies published before 1984, publications without VAP-IP data (e.g., letters, editorials and studies limited to VAP diagnosis and therapy), studies of populations limited to either paediatric or burn-injured patients, studies of populations for which <75% of patients were ventilated for more than 48 h, duplicate publication and a study design requiring the administration of components of an antimicrobial prevention regimen routinely to control group patients (duplex control design).

#### Designation of study design and cohort, control and intervention groups

**Observational studies** [Table S1, available as Supplementary data at JAC Online (http://jac.oxfordjournals.org/)]. Observational studies are those designed to survey the VAP-IP among cohorts of patients receiving mechanical ventilation without a specific study intervention. For this analysis, two sets of cohort groups were derived from observational studies, a benchmark set and a search set.

**Prevention studies.** The prevention studies were designed to study two broad types of VAP prevention methods, non-antimicrobial-based [Table S2, available as Supplementary data at JAC Online (http://jac.oxfordjournals.org/)] or antimicrobial-based [Table S3, available as Supplementary data at JAC Online (http://jac.oxfordjournals.org/)]. The designation of studies as either antimicrobial-based methods or non-antimicrobial-based methods of VAP prevention and the designation of a study group as control or intervention were as stated in any of the overview documents.

**Prevention studies: non-antimicrobial studies (Table S2).** These had studied either gastric-based or airway-based methods of VAP prevention. The interventions in these studies are further described in Hurley.

**Prevention studies: antimicrobial studies (Table S3).** Antimicrobial-based methods of VAP prevention are those using antimicrobials (administered enterally, parenterally or both) or antiseptics (administered only parenterally) to reduce VAP.

The prevention studies were further classified in relation to those with a study design using concurrent control and intervention groups and those that had a non-concurrent, i.e. historic or geographic segregation, design.

#### Data extraction

The benchmark set represents the expected VAP-IP of published studies as abstracted by experts and used to formulate the VAP-IP range estimates in previously published overview documents. Hence, the VAP-IP data for the groups of the benchmark set were taken as abstracted in the overview documents in which they appeared rather than from the original publication. Only those

---

**Table 1. Literature-based range estimates for ventilator-associated pneumonia incident proportion (VAP-IP)**

<table>
<thead>
<tr>
<th>Authors a</th>
<th>Year</th>
<th>No. of abstracted studies b</th>
<th>VAP-IP range (%) c</th>
</tr>
</thead>
<tbody>
<tr>
<td>George 1</td>
<td>1993</td>
<td>11</td>
<td>9–30</td>
</tr>
<tr>
<td>Cook and Koller 2</td>
<td>1998</td>
<td>8</td>
<td>13–38</td>
</tr>
<tr>
<td>Chastre and Fagon 3</td>
<td>2002</td>
<td>10</td>
<td>8–28</td>
</tr>
<tr>
<td>Bergmans and 4</td>
<td>2004</td>
<td>15</td>
<td>8.6–65</td>
</tr>
<tr>
<td>Safdar et al. 5</td>
<td>2005</td>
<td>28</td>
<td>7–12.5d</td>
</tr>
<tr>
<td>This study</td>
<td>2007</td>
<td>31</td>
<td>8–46e</td>
</tr>
</tbody>
</table>

aReferences 1–5 are source documents.  
bNumber of studies abstracted in defining a VAP-IP range estimate.  
cThe VAP-IP range estimates 1–4 are maximum-minimum range intervals.  
dThe range estimate of Safdar et al. 5 is a weighted average and 95% confidence interval.  
eThe range estimate from this study corresponding to the two sigma limits of the RFP at a group size of 100 patients.
groups with group size <400 patients were used in the benchmark set as all but one of the prevention studies to be profiled have group sizes of <400 patients.

For data other than the benchmark set, the data extraction was performed by the author directly from the original publication. The primary outcome is the VAP-IP, which is the incidence of ventilator-associated pneumonia per 100 patients. Other information abstracted were the first author name, year of publication, broad category of intervention method under study (antimicrobial or non-antimicrobial), the type of study design (concurrent or non-concurrent) and numbers of patients studied (group size). For cohort groups, additional information abstracted were: whether the mode of VAP diagnosis required sampling using bronchoscopic methods, the proportion of patients admitted for trauma and whether the data collection was retrospective. Data were extrapolated from tables and figures if not available in the text. Care was taken to stratify patient groups appearing across more than one publication.

Reference funnel plot (RFP)

A funnel plot is a commonly used tool to visually assess the symmetry and variability among the size of study treatment effects from individual studies in a meta-analysis. It consists of a scatter plot and a funnel. The horizontal axis of the plot is a measure of effect size and the vertical axis is a measure of study precision. The scatter plot of study effects from individual studies allows assessment of symmetry in effect size and also identifying studies that disproportionately contribute to heterogeneity in the summary effect estimate. The funnel represents the upper and lower boundaries of expected variability in effect size for all levels of study precision, most commonly derived as the 95% confidence limits calculated using fixed effect methods. The funnel is centred on the average effect size derived from all studies in the scatter plot.

A funnel plot used for profiling purposes differs from the above in several ways. The effect size of interest is the group event rate. Second, the measure of study precision is the group size on a log scale. Third, the 95% confidence limits are usually calculated using a random effects model to accommodate heterogeneity (over-dispersion) and repeated measures. In profiling applications, the control limits are termed two sigma limits instead of confidence limits as they define the limits of ‘common cause variation’ for observations that are not confined to those from which the limits were calculated. Observations falling outside these limits are said to be subject to ‘special cause variation’.

Statistical analysis and profiling analysis

The VAP-IP data are summarized as medians using inter-quartile ranges (IQRs) to express variation and compared using the Kruskal–Wallis test.
Study designs and ventilator-associated pneumonia

The VAP-IP for control and intervention groups was profiled against the RFP. Those groups lying within the two sigma limits of the RFP are inliers, those above the upper two sigma limit are high outliers and those below the lower two sigma limit are low outliers. The result of profiling is presented as the distribution of outliers and inliers, and the χ² test was used to test these results. The statistical analyses were performed using STATA statistical software (release 8.0, STATA Corp., College Station, TX, USA).

Ethics

Ethics approval was not required for this study.

Results

The literature search identified 243 studies (419 groups) that had reported original VAP-IP data. There were a total of 95 cohort groups from 93 observational studies (Table S1), of which 30 groups comprise the benchmark set and 65 groups comprise the search set.

There were 162 control and 162 intervention groups from studies of non-antimicrobial (Table S2) and antimicrobial (Table S2) methods of VAP prevention. There were 11 control groups from studies with a non-concurrent design among studies of non-antimicrobial-based methods and 12 among studies of antimicrobial-based methods of VAP prevention.

Description of the cohort groups

The size of the 95 cohort groups ranged from 19 to 9080 patients. There is asymmetry in the distribution of VAP-IP for large versus small groups (Figure 2). The VAP-IP of groups with size >399 patients (14; 10–17; median IQR) is less than that of groups with size <400 (24; 16–32; P = 0.0001; Kruskal–Wallis test).

Among the cohort groups with size <400 patients, the median year of publication was 1997, the median proportion of trauma admissions was 25%, 12 were retrospective in design and bronchoscopic methods were used for VAP diagnosis in 19 groups. To assess the group level effects of year of publication, proportion of patients admitted for trauma and mode of VAP diagnosis on the VAP-IP, these cohort groups were stratified by each of these variables sequentially into those above and below the median for the two quantitative variables and also for mode of VAP diagnosis and for retrospective or prospective design. With each of these sequential stratifications, differences in median VAP-IP were <4 per 100 patients in each case and not significant (Kruskal–Wallis test). Similar results were obtained by stratifying all 95 cohort groups.

Derivation and validation of the RFP

The RFP and the 30 cohort groups of the benchmark set together with the VAP-IP data for the 65 cohort groups of the search set are presented in Figure 2. The two sigma limits of the RFP for a hypothetical group with 100 patients correspond to a VAP-IP of 8–46 (Table 1). Eighty-two per cent of the search set cohort groups (53 of 65) were inlier (Figure 2). Only one of the outliers among the cohort groups had a group size <400 patients versus eleven for groups with a size >399 patients (P = 0.003; χ² test df = 2; Table 2). This reflects the asymmetry in the VAP-IP with group size.

Profiling analysis: groups of prevention studies

The scatter plot of groups from all studies of VAP prevention methods with either concurrent (Figures 3 and 4) or non-concurrent (Figure 5) design is presented together with the RFP (Figure 2) against which the groups were profiled.

Among the intervention groups, 22 of 162 (14%) were low outlier and none were high outlier for VAP-IP. In contrast, a triple disparity is evident between the VAP-IP of control groups versus other groups of the prevention studies in the following respects (Table 2). The distribution of VAP-IP is asymmetrical among the control groups from studies of antimicrobial-based methods of VAP prevention with concurrent design with a concentration of high outlier groups (P < 0.001; χ² test df = 2; Table 2 and Figure 4). There is significant inequality between the median VAP-IP of control groups from the four types of study design (P = 0.004, Kruskal–Wallis test). Third, the variability of VAP-IP among the four types of control groups is unequal and the IQR of the concurrent control groups of studies of antimicrobial-based methods is more than twice that of the corresponding concurrent intervention groups. The inequality of control groups remains when the profiling analysis was limited to those groups derived from publications sourced from the systematic reviews (data not shown).

Discussion

This analysis is a novel approach to the investigation of the effect of study design on the event rates of published studies. Currently available methods of benchmarking VAP rates are problematic due to issues of risk adjustment for multiunit comparisons and lack of a gold standard.41 Hence, a method of statistical process control charting is used here that is similar to methods used to profile production standards in industry and, more recently, to profile the outcome data derived from multiple healthcare units, to identify outliers.36–40 This method offers a readily interpretable plot that allows for variable sample size

![Figure 2](https://example.com/figure2.png)

Figure 2. The RFP, which is the unweighted group average surrounded by the two sigma limits derived from the benchmark set of cohort groups (circles) as described in the Methods section. Also shown is the VAP-IP data of the search set of cohort groups (plus symbols).
and can be applied to crude (not case mix or risk adjusted) data-sets.\textsuperscript{37,38} The funnel plot indicates the decreasing precision in study results associated with decreasing group size.\textsuperscript{36}

In the derivation of the benchmark, five range estimates were found as part of the literature search (Table 1). These range estimates, while representing expert opinion, were problematic. Although there was substantial overlap in the five range estimates, they were unequal, particularly for the upper limit. Moreover, the group sizes to which they corresponded were not clear and with one exception they were produced by methods that were unspecified.

The VAP-IP data for groups of size up to 400 patients as abstracted in these specified overview documents\textsuperscript{1–5} were used in the derivation of the benchmark. Data abstraction from these specified overview documents rather than abstraction from the original publications is a novel approach to data extraction. This process of data abstraction was used here for two reasons: the VAP-IP data are obtained by a process that is objective, independent and transparent, and the original studies from which the VAP-IP data were abstracted are regarded by the authors of the overviews as exemplary for deriving a range estimate of VAP-IP.

The RFP derived here is a benchmark of common cause variation for the VAP-IP of published studies having different patient populations, different diagnostic criteria for VAP and derived from studies over a 20 year interval. It accommodates event rates from groups which vary in size and the consequential variability in levels of precision. Being wider than most previous range estimates (Table 1), the benchmark developed here would be expected to be more conservative in identifying outlier groups when applied to the profiling objective of this analysis.

**Limitations of the benchmark**

The benchmark developed in this analysis is applicable only to published studies. It would be expected that the patient populations in these studies, the ICUs in which these studies might

---

**Table 2. Funnel plot: results of profiling analysis**

<table>
<thead>
<tr>
<th>Types of groups profiled</th>
<th>No. of groups</th>
<th>VAP-IP median (IQR)</th>
<th>low outliers</th>
<th>inliers</th>
<th>high outliers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cohort groups</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>benchmark set</td>
<td>30</td>
<td>25 (21–34)</td>
<td>0</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td>search set</td>
<td>65</td>
<td>15 (11–22)</td>
<td>11</td>
<td>53</td>
<td>1</td>
</tr>
<tr>
<td><strong>Control groups\textsuperscript{a,b,c}</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-antibiotic studies (CC design)</td>
<td>88</td>
<td>24 (16–32)</td>
<td>2</td>
<td>85</td>
<td>1</td>
</tr>
<tr>
<td>non-antibiotic studies (NCC design)</td>
<td>11</td>
<td>19 (12–31)</td>
<td>0</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>antibiotic studies (CC design)</td>
<td>51</td>
<td>34 (22–51)</td>
<td>2</td>
<td>34</td>
<td>15</td>
</tr>
<tr>
<td>antibiotic studies (NCC design)</td>
<td>12</td>
<td>19 (13–41)</td>
<td>1</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td><strong>Intervention groups</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-antibiotic studies (CC design)</td>
<td>87</td>
<td>15 (9–24)</td>
<td>8</td>
<td>79</td>
<td>0</td>
</tr>
<tr>
<td>non-antibiotic studies (NCC design)</td>
<td>11</td>
<td>16 (12–32)</td>
<td>2</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>antibiotic studies (CC design)</td>
<td>56</td>
<td>16 (9–22)</td>
<td>8</td>
<td>48</td>
<td>0</td>
</tr>
<tr>
<td>antibiotic studies (NCC design)</td>
<td>8</td>
<td>7 (2–9)</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

VAP-IP, ventilator-associated pneumonia incident proportion; IQR, inter-quartile range (25th–75th percentile); CC, concurrent control; NCC, non-concurrent control.

\textsuperscript{a}Distribution of high outlier control groups; non-antibiotic study concurrent control groups versus antibiotic study concurrent control groups (\(P < 0.001\)).

\textsuperscript{b}Equality of medians; non-antibiotic study concurrent control groups versus antibiotic study concurrent control groups (\(P = 0.004\)).

\textsuperscript{c}Equality of variances (Bartlett’s test); all control groups (\(P < 0.001\)).

---

and can be applied to crude (not case mix or risk adjusted) data-sets.\textsuperscript{37,38} The funnel plot indicates the decreasing precision in study results associated with decreasing group size.\textsuperscript{36}

In the derivation of the benchmark, five range estimates were found as part of the literature search (Table 1). These range estimates, while representing expert opinion, were problematic. Although there was substantial overlap in the five range estimates, they were unequal, particularly for the upper limit. Moreover, the group sizes to which they corresponded were not clear and with one exception they were produced by methods that were unspecified.

The VAP-IP data for groups of size up to 400 patients as abstracted in these specified overview documents\textsuperscript{1–5} were used in the derivation of the benchmark. Data abstraction from these specified overview documents rather than abstraction from the original publications is a novel approach to data extraction. This process of data abstraction was used here for two reasons: the VAP-IP data are obtained by a process that is objective, independent and transparent, and the original studies from which the VAP-IP data were abstracted are regarded by the authors of the overviews as exemplary for deriving a range estimate of VAP-IP.

The RFP derived here is a benchmark of common cause variation for the VAP-IP of published studies having different patient populations, different diagnostic criteria for VAP and derived from studies over a 20 year interval. It accommodates event rates from groups which vary in size and the consequential variability in levels of precision. Being wider than most previous range estimates (Table 1), the benchmark developed here would be expected to be more conservative in identifying outlier groups when applied to the profiling objective of this analysis.

**Limitations of the benchmark**

The benchmark developed in this analysis is applicable only to published studies. It would be expected that the patient populations in these studies, the ICUs in which these studies might
have been undertaken and also the studies themselves would each have been subject to recruitment, publication and possible citation biases which would limit the projection to the broader population of patients receiving prolonged mechanical ventilation outside of a published study. For example, ICUs that might be prepared to undertake a study may not be representative of the broader population of ICUs.

Pertinent to the objectives of this study, all the prevention studies except one had group sizes of <400 patients. In this respect, study size is itself a study design factor. There is asymmetry in the VAP-IP of cohort groups of observational studies, in which the VAP-IP of very large groups (>399 patients) for all but three groups were lower than the average VAP-IP of studies with <400 patients (Figure 2). There are several possible explanations for this finding such as volume effect, recruitment bias resulting in lower risk patients being over represented in very large (registry) studies and ascertainment bias with closer scrutiny for VAP in smaller studies. This VAP-IP asymmetry would account for the low range estimate for VAP-IP derived using a method weighted for study size. Indeed, a funnel plot constructed using all 95 cohort groups (including those with size >400 patients) yields two sigma limits corresponding to a group size of 100 patients of 2.3 to 44 (data not shown). Using a funnel plot derived using all 95 cohort groups to profile the groups of the intervention studies would have slightly increased the numbers of high outlier groups but eliminated the low outlier groups.

The differences between the results of large studies versus small studies are as occasionally seen with the summary findings of meta-analyses of several small studies being contrary to the results of a single mega-trial of the same topic.

**Profiling results**

Over 90% of control and intervention groups from non-antimicrobial-based methods of VAP prevention are inlier. By striking contrast, 30% of control groups from studies of antimicrobial-based methods of VAP prevention with a concurrent design are high outlier, whereas only 2.5% would be expected. Alternatively, this finding could be explained as a deficit of inlier control groups from studies of antimicrobial-based methods. This interpretation implies an unlikely deficit of inlier control groups from at least 550 studies of antimicrobial prevention methods that were ‘missing’ because they had not been published due to a null result (i.e. no difference in VAP-IP between control groups and intervention groups) or published in journals that are more difficult to access (e.g. foreign language or abstracts).

This difference in VAP-IP of different types of control groups is also apparent among the abstracted VAP-IP data as appearing in the largest meta-analyses of VAP prevention methods. For example, among the largest meta-analyses of antibiotic methods of VAP prevention, there are 16 of 32 studies with control groups VAP-IP’s >40% versus only 2 of a total of 42 studies among the three largest meta-analyses of non-antimicrobial methods.

There is a smaller excess of low outlier groups (21 of 160; 13%) among the intervention groups. Profiling is not designed to address the size of treatment effect in the various studies. However, an excess of low outlier groups would be an expected finding among intervention groups receiving a therapy intended to prevent VAP.
A further limitation is that the mechanism for special cause variation remains speculative. An increased proportion of *S. aureus* among the VAP isolates has previously been noted among the outlier control groups of groups of the studies of antimicrobial-based methods, suggesting that inapparent outbreaks had occurred in these studies. However, there is a pattern of discrepancies which suggest that study design factors account for the striking concentration of special cause variation among the concurrent control groups of the studies of antimicrobial-based methods of VAP prevention. This pattern of discrepancies is between the intervention and concurrent control groups in respect of the proportion of inlier groups and the degree of variability of VAP-IP (with respect to both median VAP-IP and IQR; Table 2): in both respects, the concurrent intervention groups of the studies of antimicrobial-based methods more closely resemble the cohort groups of the benchmark than do the corresponding concurrent control groups. This pattern of discrepancies is surprising given the broad variety of antimicrobial interventions under study. In contrast, the control groups of both the non-concurrent design antimicrobial studies and the control groups of the studies of either design of the non-antimicrobial-based methods of VAP prevention more closely resemble the cohort groups than do their corresponding intervention groups, as might be expected.

The difference between the median of the benchmark cohort groups and, on the one hand, the median of the concurrent intervention groups of studies of antimicrobial-based methods and, on the other hand, the median of the concurrent control groups of studies of antimicrobial-based methods, is 9 per 100 patients for both. This implies that in studies of antimicrobial-based prevention methods the effect of concurrency as a study design factor on VAP-IP is at least as great as the effect of the antimicrobial interventions under study. The effect is also greater than the group level effect of proportion of patients admitted for trauma and mode of VAP diagnosis on VAP-IP, neither of which was found to be significant in this analysis although significant group level effects for each has previously been demonstrated in an ecological analysis.

Conclusions

There is a concentration of special cause variation among the concurrent control groups of studies of antimicrobial-based methods for the prevention of VAP as reflected in the large number of high outlier control groups. This concentration of special cause variation exceeds that for the intervention groups, for which low outlier status is explainable on the basis of use of prevention therapies. The likely mechanism of this special cause variation is enhanced cross-infection resulting from study design factors such as the concurrency of control and intervention groups and the use of topical placebo and antimicrobial as study interventions, respectively. The apparent efficacy of antimicrobial-based methods for the prevention of VAP would have been inflated as a consequence.

Acknowledgements

This manuscript was accepted as a minor thesis for the degree Master in Epidemiology in the School of Population Health, University of Melbourne.

Funding

Paulson Rural Medical Research Fellowship.

Transparency declarations

None to declare. The funder had no role in the design, analysis or writing of this study or in its submission for publication. The author is independent from the funder.

Supplementary data

Tables S1, S2 and S3 are available as Supplementary data at *JAC* Online (http://jac.oxfordjournals.org/).

References

Study designs and ventilator-associated pneumonia

42. Dersimonian R, Levine R. Resolving discrepancies between a meta-analysis and a subsequent large controlled trial. JAMA 1999; 282: 664–70.