Research priorities in critical care medicine in the UK

Abstract Objectives: To establish priorities for research in critical care medicine in the UK using survey and nominal group (NG) techniques.

Design: The senior doctor and nurse from 325 intensive care units (ICUs) in the UK were invited to contribute up to ten research questions relevant to intensive care organisation, practice or outcomes. These were then ranked twice using a Likert scale by a panel (nominal group) consisting of ten doctors (two trainees) and two nurses from university teaching and district general (community) hospitals. The first ratings were performed privately, and the second after group discussion. Thirty questions, ten each with strong, moderate or weak support, were then returned for rating by the originating ICU staff and the results compared with those of the NG.

Results: One hundred eighty-five respondents (35.6% university teaching, 62.1% district general, 2.3% not stated) provided 811 questions of which 722 were research hypotheses. The most frequently identified topics were the evaluation of high dependency care, ICU characteristics, treatments for acute lung injury and acute renal failure, nurse:patient ratios, pulmonary artery catheterisation, aspects of medical and nursing practice, protocol evaluation, and interhospital transfers. These were condensed into 100 topics for consideration by the NG. Discussion and re-rating by the group resulted in strong support being offered for 37 topics, moderate support for 48, and weak support for 21. Following circulation of ten questions from each category, nine questions achieved strong support from both ICU staff and the NG. These were the effect on outcomes from critical illness of early intervention, high dependency care, nurse:patient ratios, interhospital transfers, early enteral feeding, optimisation of perioperative care, hospital type, regionalisation of paediatric intensive care and the use of pulmonary artery catheters. The absence of any questions relating to interventions targeting mediators of the immuno-inflammatory response could be a consequence of the failure of recent studies in sepsis to demonstrate benefits in outcome.

Conclusions: The intensive care community in the UK appears to prioritise research into organisational aspects of clinical practice and practical aspects of organ-system support. Health services research and the biological sciences need to develop collaborative methods for evaluating interventions and outcomes.

Key words Intensive care · Critical care · Consensus techniques · Nominal groups · Outcomes · Monitoring techniques · Health services
Introduction

Intensive care presents particular challenges for research. Heterogeneity of case mix and clinical practice, and comparatively small patient numbers, makes clinical and health services research difficult. However, intensive care has a substantial number of methods for case mix adjustment and several large observational databases. Once case mix-adjusted data are pooled from many intensive care units (ICUs), variations in clinical practice (for example, the use of a particular treatment or monitoring device) can become an asset rather than an obstacle for exploring differences in outcomes in multicentre projects.

A practical problem with multicentre research is that it requires a degree of commitment that may be difficult to sustain unless research priorities have been agreed by the participants. There might also be value in employing ‘collective wisdom’ to identify those research topics of greatest importance, as these are likely to motivate potential participants. Consensus development techniques offer methods for achieving these aims and for managing group decision making in the presence of uncertainty.

Formal consensus development techniques have been employed before to establish research priorities in nursing [1] and occupational medicine [2], as well as for establishing the appropriate use of interventions such as angioplasty [3] and to determine standards of practice [4, 5]. There are three methods: Delphi surveys, the use of expert panels or nominal groups (NGs) and consensus development conferences [6]. We have recently investigated one of these methods for determining research priorities for intensive care [7]. In this paper we present the clinical outcomes from that process and reduction was performed by grouping into domains, and then the most representative question within the subject area was used verbatim with minimal editing and without reference to its relative frequency.

The nominal group

A NG was convened by invitation. The group was constituted to reflect a 50:50 teaching hospital:district general (community) hospital ratio, and included two nurses and two medical trainees. To allow for discussion, the size of the group was capped at 12 participants, and the number of hypotheses for review was limited to the 100 most frequently occurring. Four weeks before the meeting the members of the NG were asked to rate all 100 hypotheses privately, and without knowledge of the relative frequency with which they had been suggested, using a Likert scale from 1 (no support) to 9 (strong support). The NG then met to discuss each topic, the discussion being facilitated by one of the authors (NAB). Sufficient time was allowed for discussion of each topic, the duration depending on the degree of controversy which was aroused. Ambiguous topics were re-phrased by agreement, and six were split into two separate topics, resulting in a final total of 106. Immediately after each topic had been discussed, each participant rated it again privately. The differences between the group’s pre- and post-discussion ratings were compared subsequently, allowing calculation of the change in ratings and in degree of consensus, and estimation of both the absolute and relative importance attached to each topic.

The survey

A survey was then conducted using a single letter to the medical directors of all 325 ICUs except the 12 represented by members of the NG. Thirty questions were selected from the 106 considered by the NG: the first ten, respectively, with strong support (median rating 7–9), ten with modest support (4–6) and ten with weak support (1–3). Respondents were unaware of the ratings by the NG. They were asked to rate the importance of the questions from 1 (no support) to 9 (very strong support). Analysis of survey ratings was performed to determine the reliability of the NG.

Methods

Generation of research questions

Letters were sent during July 1998 to 325 ICUs in the UK, one copy to the clinical director/lead consultant and one to the senior nurse. We asked them to send us up to ten research questions that they considered important in the context of intensive care organisation, practice or outcomes, and to do this in discussion with their colleagues. An example of a research question was given in order to distinguish hypotheses from questions about numerical frequency. Respondents were not obliged to identify their hospital.

The responses were categorised into common domains by one of the authors (JFB), using the predominant theme of the question or hypothesis. Thus, had a question such as ‘Does pulmonary artery catheterisation affect long-term survival in the elderly?’ been posed, it would have been classified under ‘monitoring’ and not under interventions or post-ICU care; whereas had it been phrased ‘Should techniques like pulmonary artery catheterisation be offered to elderly patients with little chance of survival’, this would have been classed as an ethical issue. The process of abstraction and calculation was performed by grouping into domains, and then the most representative question within the subject area was used verbatim with minimal editing and without reference to its relative frequency.

The nominal group

A NG was convened by invitation. The group was constituted to reflect a 50:50 teaching hospital:district general (community) hospital ratio, and included two nurses and two medical trainees. To allow for discussion, the size of the group was capped at 12 participants, and the number of hypotheses for review was limited to the 100 most frequently occurring. Four weeks before the meeting the members of the NG were asked to rate all 100 hypotheses privately, and without knowledge of the relative frequency with which they had been suggested, using a Likert scale from 1 (no support) to 9 (strong support). The NG then met to discuss each topic, the discussion being facilitated by one of the authors (NAB). Sufficient time was allowed for discussion of each topic, the duration depending on the degree of controversy which was aroused. Ambiguous topics were re-phrased by agreement, and six were split into two separate topics, resulting in a final total of 106. Immediately after each topic had been discussed, each participant rated it again privately. The differences between the group’s pre- and post-discussion ratings were compared subsequently, allowing calculation of the change in ratings and in degree of consensus, and estimation of both the absolute and relative importance attached to each topic.

The survey

A survey was then conducted using a single letter to the medical directors of all 325 ICUs except the 12 represented by members of the NG. Thirty questions were selected from the 106 considered by the NG: the first ten, respectively, with strong support (median rating 7–9), ten with modest support (4–6) and ten with weak support (1–3). Respondents were unaware of the ratings by the NG. They were asked to rate the importance of the questions from 1 (no support) to 9 (very strong support). Analysis of survey ratings was performed to determine the reliability of the NG.

Data and statistical analysis

Median values were calculated for the level of support for each topic and the degree of agreement or dispersion was calculated using the mean absolute deviation from the median (MADM). A reduction in MADM thus represents an increase in agreement. The level of consensus was defined by tertiles of the MADM as low (< 1.41), moderate (1.08–1.41) or high (< 1.08). Statistical significance was determined using the Wilcoxon’s signed rank test for changes in rank order, and the chi-squared test for the degree of consensus. The level of association between the NG and the survey ratings was examined using Pearson’s correlation coefficient and the extent of agreement using the kappa statistic.