Effect of clinical guidelines in nursing, midwifery, and the therapies: a systematic review of evaluations

L H Thomas, E McColl, N Cullum, N Rousseau, J Soutter, N Steen

Abstract

Background—Although nursing, midwifery, and professions allied to medicine are increasingly using clinical guidelines to reduce inappropriate variations in practice and ensure higher quality care, there have been no rigorous overviews of their effectiveness. 18 evaluations of guidelines were identified that meet Cochrane criteria for scientific rigor.

Methods—Guideline evaluations conducted since 1975 which used a randomised controlled trial, controlled before and after, or interrupted time series design were identified through a combination of database and hand searching.

Results—18 studies met the inclusion criteria. Three studies evaluated guideline dissemination or implementation strategies, nine compared use of a guideline with a no guideline state; six studies examined skill substitution: performance of nurses operating according to a guideline were compared with standard care, generally provided by a physician. Significant changes in the process of care were found in six out of eight studies measuring process and in which guidelines were expected to have a positive impact on performance. In seven of the nine studies measuring outcomes of care, significant differences in favour of the intervention group were found. Skill substitution studies generally supported the hypothesis of no difference between protocol driven by nurses and care by a physician. Only one study included a formal economic evaluation, with equivocal findings.

Conclusions—Findings from the review provide some evidence that care driven by a guideline can be effective in changing the process and outcome of care. However, many studies fell short of the criteria of the Cochrane Effective Practice and Organisation of Care Group (EPOC) for methodological quality. (Quality in Health Care 1998;7:183–191)

Keywords: clinical guidelines; nursing; midwifery; systematic review; clinical effectiveness

Introduction

Clinical guidelines in medicine have been shown to improve the performance of doctors and to promote the delivery of high quality, evidence-based health care. The benefits of developing and implementing guidelines may not, however, be unique to this profession. In nursing, midwifery, and professions allied to medicine, there is increasing interest in the use of guidelines to reduce inappropriate variations in practice and ensure higher quality care.

Preliminary searches indicated little formal evaluation of the effectiveness of the implementation strategies adopted by nurses, midwives, and professions allied to medicine. In the United States, guideline dissemination seems to be through massive distribution to professional and lay groups, although there is evidence from the medical profession to suggest that dissemination by publication or direct mailing are the least successful strategies. However, it cannot be assumed that those strategies proved to work for doctors will transfer directly to other groups of healthcare professionals with a dissimilar educational preparation and a different management structure.

We therefore undertook a systematic review of evaluations of clinical guidelines in nursing, midwifery, and professions allied to medicine. This was registered with the Cochrane Effective Practice and Organisation of Care Group (EPOC; formerly the Cochrane Collaboration on Effective Professional Practice).

Our review had two main objectives.

1. To identify evaluations (in any setting) of clinical practice guidelines and of related dissemination and implementation strategies in nursing, midwifery, health visiting, and other professions allied to medicine—podiatry, speech and language therapy, physiotherapy, occupational therapy, dietetics, clinical psychology, pharmacy, and radiography.

2. To derive estimates of the effectiveness and cost effectiveness of clinical practice guidelines as strategies for promoting improved professional practice and enhanced patient outcomes in nursing and professions allied to medicine.

We sought, as far as possible, to replicate the medically oriented work of Grimshaw et al. We wished to describe the nature of guidelines for the target professions. (Who developed them; were they derived by consensus methods; were they evidence-based?) We also wanted to investigate whether guideline dissemination and implementation strategies found to be effective for doctors were transferable to nursing and other professions allied to medicine. However, our review was intended to be primarily descriptive, rather than driven by a hypothesis.
Methods

INCLUSION AND EXCLUSION CRITERIA

The definition of clinical guidelines by Field and Lohr[2] as "systematically developed statements to assist practitioner decisions about appropriate health care for specific clinical circumstances" was used. We were interested in all types of guideline, not simply those based on evidence.

To be eligible for inclusion in our review, a study had to be an evaluation of the effect of a clinical guideline on the behaviour of nurses or other professionals allied to medicine or on patient outcomes. We included studies which evaluated an entire, or identified component, of a guideline aimed at these professions and which specified which part or parts of the guideline should be implemented by the professional in question. We considered not only studies of guidelines developed by the target profession for their own use, but also those in which the guideline was developed or disseminated by another professional group but targeted the behaviour of nurses or other professionals allied to medicine.

We included both studies which evaluated the effectiveness of guidelines against a non-intervention control (routine care) and those comparing different dissemination and implementation strategies of a guideline—for example, dissemination strategy X versus dissemination strategy Y. Studies which simply described the development, dissemination, or implementation of a guideline, without an evaluation of the impact, were excluded, as were studies in which <50% of recipients of the guideline were nurses, midwives, or other professionals allied to medicine and studies in which it was not possible to separate the effects of guidelines on the behaviour of professionals from the impact on other health professionals. A start date of 1975 was imposed, firstly to make our review comparable with that of Grimshaw and Russell,[12] and secondly because MEDLINE does not contain abstracts before that year.

METHODOLOGICAL CRITERIA

To be compatible with Grimshaw et al.,[2][3][11] and in line with standard practice for EPOC reviews, three types of study designs were included.

(1) Randomised controlled trials, where we distinguished between simple randomised controlled trials (in which the unit of allocation is the patient) and designs in which the unit of allocation is the health professional and which are therefore less susceptible to contamination.
(2) Controlled before and after studies.
(3) Interrupted time series designs, if there was a clearly defined point in time when the intervention occurred and at least two data collection points before and two after the intervention.

SYSTEMATIC REVIEW OF LITERATURE

We searched: (a) computerised databases (MEDLINE; CINAHL (Cumulative Index of Nursing and Allied Health Literature); PsycLIT; EMBASE; NEED (NHS Economic Evaluations Database); DARE (Database of Abstracts of Reviews of Effectiveness); DHSS-DATA; SIGLE; and the National Research Register); (b) the EPOC register of studies; (c) citations in papers identified, and (d) Quality in Health Care*. We also contacted content area experts, the libraries of appropriate professional bodies—for example, the Royal College of Nursing—and various mailbases on the internet—for example, evidence-based-health@mailbase.ac.uk.

Search strategies for electronic databases were developed sequentially, starting with MEDLINE, as this was expected to yield the highest number of relevant papers. We adapted the focus section of the EPOC search strategy (designed to identify randomised trials, controlled before and after and interrupted time series studies) to expand the section on clinical guidelines. We searched for guidelines and various synonyms in titles and abstracts as well as carrying out keyword searches. MEDLINE also identifies "publication types", which allowed us to search for papers tagged as "practice guideline" or "guideline". We then worked with an information scientist to adapt the search strategy for use with the other databases.

SCREENING OF IDENTIFIED PAPERS

The full text version of each study identified as potentially relevant to our review was read by one reviewer (NC, EM, NR, JS, or LHT) and assessed against several criteria with a decision tree developed for this review. To establish consistency between raters, all reviewers initially completed checklists for the same five papers; discrepancies were discussed and resolved. All those studies identified by one reviewer as meeting all the inclusion criteria, and those in which there was an element of doubt, were referred to a second reviewer; discussion between the two reviewers took place to resolve any remaining differences.

DATA EXTRACTION

We used the checklist and data extraction template developed by EPOC to abstract data from studies meeting our inclusion criteria. Data extraction was undertaken by two reviewers working independently and included both methodological details—for example, units of allocation and data analysis; methodological quality criteria—and information about study participants, interventions, outcomes, and results. The two reviewers then met to compare their judgements and any discrepancies were discussed and resolved. A “master version” of the data extraction template was agreed and the details were put into an Idealist[11] database.

* Quality in Health Care is now included in the following databases: HEALTH (Health Planning and Administration) Database of the National Library of Medicine and American Hospital Association Resource Center from 1995 onwards, CINAHL, ISI Science Citation Index (public health, environmental and occupational), ISI Current Contents (clinical medicine), and EMBASE.
DATA ANALYSIS
As the heterogeneity of clinical areas, design of studies, source and format of interventions, outcomes measured, and participating health professionals, was substantial, an overall estimate of effect was likely to have little practical meaning. We therefore opted simply to report the effects on the process and outcome of care in the same way as they were reported in the original papers, rather than to attempt comparisons with some common unit of change.

Results
We identified 18 studies which met our inclusion criteria. In all but one, guidelines were aimed, in part at least, at nurses. Of these studies, in five cases, 16–20 five units in a single hospital participated. No data were presented a truly multi-disciplinary guideline, aimed at an entire lipid clinic team comprising nurses, pharmacists, psychologists, and dietitians.

METHODOLOGICAL QUALITY
We identified 13 randomised controlled trials. In three studies the unit of allocation was the health professional or provider unit. In all three the randomisation process was not described and it is not clear whether there was blinded assessment of outcomes. One trial 19 had a unit of analysis error: the unit of randomisation was the provider, but the unit of analysis was the patient. Such unit of analysis errors artificially increase the power of the study and may lead to misleadingly significant findings.

There were 10 randomised controlled trials in which the unit of allocation was the patient. 18 19 21 24–26 Six of these 18 19 21 24–26 examined the effect of substituting another health professional (usually a nurse) working to protocols for usual care provided by a doctor: obviously in these studies, randomisation by health professional would not be possible. There were major differences in methodological quality: all 13 studies had good follow up; however, just over half (7/13) blindly assessed outcomes and reliably measured at least some key outcomes (7/13). A few studies (4/13) reported method of randomisation.

We identified two controlled before and after studies where a second site acted as the control. 22 23 In the study by Larson and Hargiss 22 five units in a single hospital participated; two received the intervention and three served as controls. No data were presented about the comparability of intervention and control sites; baseline measurement showed significant differences on one variable likely to influence differences after intervention (accuracy of recording of last intravenous infusion). At least some outcomes were measured reliably but none were assessed blindly. Protection against contamination seemed adequate. In the study by Shafer and Wexler, 23 two clinics (one intervention, one control) within the same medical centre participated. No data were presented on the similarity of the clinics; patients selected for the study from each clinic were matched according to age and cholesterol concentration. Outcomes measured were reliable and were assessed blindly. As the lipid clinic team worked only in the intervention clinic, it is likely that there was successful protection against contamination between clinics.

Three studies were interrupted time series designs. 20 21 22 In the study by French et al 20 the intervention seemed to be independent of other changes during the study period; in the study of Mitchell and Jones 21 no details were provided as to whether this was or was not the case. In the study of Tilden and Shepherd, 22 in which the targeted behaviour was identification of battered women, the authors stated that media coverage and an initiative by hospital administration (aimed at all staff) to increase the level of detail in patient documentation may have intensified the effect of the intervention. No interrupted time series study had >12 data points before and after the intervention; only one study 23 did a formal test for trend. All interrupted time series studies were prone to detection bias. In two studies 17 23 data collection methods were identical before and after the intervention. However, in the study of Mitchell and Jones 22 data on rates of falls before intervention were collected from retrospective analysis of incident forms of falls, and prospective data were collected by adding a “falls reporting document” to the incident form. As this seemed to be part of the intervention, it directly influenced data collection. Whether the intervention directly influenced data collection in the two other interrupted time series studies is not clear. Outcomes were not assessed blindly in one study; in the remaining two studies it is not clear whether this was the case. In only one study 23 was information provided about the completeness of the data set: all adult female trauma patients were included. No interrupted time series study measured outcomes reliably.

STUDIES EVALUATING DISSEMINATION AND IMPLEMENTATION STRATEGIES
Three studies (table) 16 17 20 explicitly evaluated guideline dissemination or implementation strategies. All were randomised controlled trials. Within these studies, the interventions were heterogeneous, precluding combining the findings from two or more studies. Seto et al 21 in their evaluation of opinion leaders and in-service lectures as methods of disseminating a guideline on urinary catheter care to nurses, compared three groups: all received guidelines; one received both lectures and input from local opinion leaders; the remaining two received either lectures or input from opinion leaders. Some evidence was found that the combined strategy of opinion leaders and lectures led to higher compliance with correct practice than did either strategy alone. The weakest strategy was a lecture alone.

Herman et al 20 found that a “nurse protocol” intervention (primarily an implementation strategy) led to positive changes in the process of care (offer of influenza and pneumococcal vaccine). There were three comparison groups: all received guidelines, educational materials,
### Comparison tables

<table>
<thead>
<tr>
<th>Study and study design</th>
<th>Participants</th>
<th>Clinical area</th>
<th>Results comparisons</th>
<th>Main process effect</th>
<th>Main patient outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Herman et al (1994) RCT-Pr US</td>
<td>Nurses (number not specified) and 66 doctors in 3 ambulatory medical practices</td>
<td>Pneumococcal and influenza vaccinations</td>
<td>I: Guidelines; educational materials; lectures</td>
<td>Offer of influenza vaccination: I: 48% (129/271) I2: 50% (122/242) I3: 68% (166/243) p=0.006 Offer of pneumococcal vaccination: I: 5.4% (18/329) I2: 6.3% (19/292) I3: 28.3% (89/314) p=0.001</td>
<td>Influenza vaccination: I1: 5.9%; I2: 5.8%; I3 13.2% p=0.003 Pneumococcal vaccination: I1: 2.0%; I2: 1.4%; I3: 6.7% p&lt;0.001</td>
</tr>
<tr>
<td>McDonald et al (1980) RCT-Pr (Crossover trial) US</td>
<td>3 Nurses and 26 Doctors in 1 medical outpatient clinic 3-Treatment crossover design: all providers experienced all conditions in 1 of 6 possible different treatment sequences Proportion of patients followed up: not applicable</td>
<td>Urinary catheter care</td>
<td>I: Guidelines; local opinion leaders; lecture</td>
<td>Change in reported practice, 2 weeks before vs 3 weeks after intervention: I: 5.63 (50% I2: 4.96 (35%) C: 3.29 (38%) Mean change in C significantly lower than I1 and I2 p&lt;0.05 I1 and I2 not significant: insufficient data provided to calculate p values or 95% CIs Percentage of correct catheter practices, 1 month after intervention: I1 vs I2: p&lt;0.05 I2 vs C: p&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>Seto et al (1991) RCT-Pr Hong Kong</td>
<td>220 Nurses in 1 hospital Provider: 2 wards in each of 353 proportion of patients followed up: not applicable</td>
<td>Non-insulin dependent diabetes mellitus</td>
<td>I: Guidelines; local opinion leaders; lecture</td>
<td>Changes in treatment: I: 18% (17/94); C: 11% (9/85) p=0.203</td>
<td></td>
</tr>
<tr>
<td>Comparison of guideline v no guideline: Franz (1995) RCT-Pr US</td>
<td>Nurses (number not specified) in outpatient clinics Patients: 247 with non-insulin dependent diabetes mellitus Proportion of patients followed up: 72% in total</td>
<td>Fasting plasma glucose: I: 9.1 mmol/l (from 10.2); C: 9.2 mmol/l (from 9.8) not significant: insufficient data provided to calculate p value or 95% CIs Glycated haemoglobin: I: 7.4 mmol/l (from 8.3); C: 7.6 mmol/l (from 8.3) not significant: insufficient data provided to calculate p value or 95% CIs For the following variables, insufficient data provided to calculate p value or 95% CIs Cholesterol: I: 5.4 mmol/l (from 5.6); C: 5.5 mmol/l (from 5.7) LDL-cholesterol: I: 3.25 mmol/l (from 3.34); C: 3.51 mmol/l (from 3.50) HDL-cholesterol: I: 1.07 mmol/l (from 1.09); C: 1.07 mmol/l (from 1.14) Triglycerides: I: 2.40 mmol/l (from 2.57); C: 2.62 mmol/l (from 2.54) Weight: I: 92.4 kg (from 93.8 kg); C: 92.0 kg (from 93.7 kg) Body mass index: I: 32.4 kg/m² (from 32.9 kg/m²); C: 32.4 kg/m² (from 33.0 kg/m²) Quality of life: I: 5.6); C: 5.5 mmol/l (from 5.7) Resisting temptation: I: 69; C: 79 p&lt;0.05 Positive social: I: 62; C: 76 p&lt;0.01 Exercise: I: 48; C: 64 p&lt;0.01 Negative emotions: I: 62; C: 76 p&lt;0.01 Per patient cost per unit of change in outcome (fasting plasma glucose level): I: $5.84; C: $5.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frigoletto et al (1994) RCT-Pr US</td>
<td>Two maternity units in 1 hospital Patients: 1017 intervention 917 control Protocol eligible sub-group: 678 intervention 585 control</td>
<td>Protocol (midwives) C: Standard care</td>
<td>Frequency of caesarean section: I: 19.5% (197/1009); C: 19.4% (176/906) RR = 1.0 95% CI (0.8–1.2) Mean frequency of vaginal examination (hours): I: 1.6; C: 2.3 Artificial rupture of membranes: I1: 61%; C: 51% RR = 1.295% CI (1.1–1.3) Rupture within 1 hour of admission: I1: 76%; C: 15% RR = 17.8 95% CI (13.3–23.8) Administration of oxytocin: I1: 70%; C: 56% RR = 1.3 95% CI (1.2–1.4) Administration of epidural anaesthesia: I1: 54%; C: 64% RR = 0.8 95% CI (0.8–0.9) Rate of caesarean section: I: 19.5% (197/1009); C: 19.4% (176/906) for all patients RR = 1.0 95% CI (0.8–1.2) I: 10.9%; C: 11.5% for protocol eligible group RR = 0.9 95% CI (0.4–1.9) Duration of labour (protocol eligible subgroup): I: 6.2; C: 8.9 hours Insufficient data provided to calculate CIs Rate of maternal fever (protocol eligible subgroup): I: 7%; C: 11% p=0.007 Rate of other maternal outcomes: no significant differences in foetal distress, placental abruption, shoulder dystocia, vaginal lacerations Duration of labour &gt;12 hours (protocol eligible subgroup): I: 9%; C: 26% p&lt;0.001 Infant outcomes: no significant differences in frequency of jaundice, seizures, treatment for sepsis, resuscitation at birth or admission to neonatal ICU</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Comparison tables (continued)

<table>
<thead>
<tr>
<th>Study and study design</th>
<th>Participants</th>
<th>Clinical area</th>
<th>Results comparisons</th>
<th>Main process effect</th>
<th>Main patient outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naylor 1990&lt;sup&gt;16&lt;/sup&gt; RCT/Pt US</td>
<td>2 Nurses (in intervention group) in 1 hospital</td>
<td>Patients: 20 intervention 20 control</td>
<td>Discharge planning for &gt;70 year-olds in medical and surgical units</td>
<td>I: Protocol C: Standard practice</td>
<td>Mean duration of stay (days): I: 8.2 ± 4.9; C: 9.05 ± 7.66 not significant; insufficient data to calculate p value or 95% CIs. Postdischarge infection rates: I: 33.3% (6/20); C: 50% (10/20) RR = 0.43 95% CI (0.11–1.57) Postdischarge rates of readmissions to hospital: I: 16.7% (4 readmissions, 3 patients); C: 64.7% (12 readmissions to hospital, 11 patients) p &lt; 0.05 Results reported for medical and surgical diagnostic related group Medical and surgical separately. Mean duration of initial admission to hospital (days): Medical: I: 7.4 (3.8); C: 7.5 (5.2) – difference (95% CI) 0.1 (–1.6 to 1.4) Surgical: I: 15.8 (9.4); C: 14.8 (8.3) difference (95% CI) 1.0 (–2.0 to 1.0) Mean days to readmission to hospital: I: 45.6; C: 61. p = 0.12 Rate of readmission to hospital: Medical: I: 15% (11); C: 16% (11) difference (95% CI) 1% (–8% to 12%); Surgical: I: 10%; C: 7% (5% difference (95% CI) 3% (–7% to 13%) Mean duration of readmission to hospital: Medical: I: 94; C: 100 difference (95% CI) 6 (–83 to 71) Surgical: I: 52; C: 26 difference (95% CI) 26 (–8 to 60) Occurrence of IV related phlebitis: I: decrease from 33.5% to 20.0%, p &lt; 0.05 C; increase from 23.8% to 26.7% p &lt; 0.25 Frequency of bacterial colonisation of infusion devices: I: increase from 12.7% to 19.4%, p &lt; 0.1 C: decrease from 7.1% to 5.9%, p = 0.90. Patient comfort with IV insertion: not reported in appropriate format</td>
</tr>
<tr>
<td>Naylor 1994&lt;sup&gt;17&lt;/sup&gt; RCT/Pt US</td>
<td>2 Nurses and doctors (number not specified) in 1 hospital</td>
<td>Patients: 364 patients randomised; 276 analysed; 140 intervention 136 control</td>
<td>Discharge planning for &gt;70 year-olds in medical and surgical units</td>
<td>I: Protocol C: Standard practice</td>
<td></td>
</tr>
<tr>
<td>Larson and Hargiss 1984&lt;sup&gt;18&lt;/sup&gt; CBA US</td>
<td>10 nurses (in intervention group) and doctors (number not specified) in 2 intervention and 3 control wards in 1 university affiliated hospital</td>
<td>Patients: 707 (in total) Proportion of patients followed up: not clear</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shaffer and Wexler 1984&lt;sup&gt;19&lt;/sup&gt; CBA US</td>
<td>1 lipid clinic team (nurse led) and 1 general medicine clinic team in 1 medical centre</td>
<td>Provider: 1 intervention 1 control Patients: 60 intervention 60 control</td>
<td>Hyperlipidaemia</td>
<td>I: Guidelines (nurse, clinical pharmacist, dietitian, clinical psychologist) C: No intervention (doctors)</td>
<td>Percentage of patients for whom a new drug was prescribed: I: 85% (51/60); C: 18% (11/60) p &lt; 0.001 Number of patients in whom an existing drug was stopped: I: 22 patients; C: 3 patients. p &lt; 0.001 Percentage of patients receiving expert dietary counselling: I: 100% (60/60); C: 42% (25/60) p &lt; 0.001</td>
</tr>
<tr>
<td>French et al 1989&lt;sup&gt;20&lt;/sup&gt; ITS Hong Kong</td>
<td>3 Infection control nurses in 1 hospital</td>
<td>Patients: ranged from 809 to 1260 in 7 prevalence surveys</td>
<td>Infection rates</td>
<td>I: Guidelines C: No intervention</td>
<td>Total cholesterol at follow up: I: 6.08 mmol/l (from 7.65); C: 6.72 mmol/l (from 7.50) p = 0.004 LDL-cholesterol at follow up: I: 3.78 mmol/l (from 5.20); C: 4.60 mmol/l (from 5.17) p &lt; 0.001 HDL-cholesterol at follow up: I: −1.03 mmol/l (from 1.11); C: −1.14 mmol/l (from 1.19) not significant: insufficient data to calculate p value or 95% CIs Triglycerides at follow up: I: −3.39 mmol/l (from 4.13); C: −3.56 mmol/l (from 4.45) not significant: insufficient data to calculate p value or 95% CIs Body mass index: I: 28.3 kg/m² (from 27.9 kg/m²); C: 28.5 kg/m² (from 28.8 kg/m²) not significant: insufficient data to calculate p value or 95% CIs Percentage achieving national cholesterol educational programme LDL-cholesterol goal (3.36 mmol/l): I: 44% (19/43); C: 11% (3/28) RR = 4.1 95% CI (1.4–12.7) Percentage on lipid lowering drug: I: 91% (51/56); C: 35% (21/60) RR = 3.75 95% CI (2.05–6.87) Preintervention hospital acquired infection rates (crude) 8.9%, 10.5% Postintervention hospital acquired infection rates (crude) 7%, 7.8%, 6.2%, 6.5%, 5.6% p &lt; 0.001 Preintervention hospital acquired infection rates (adjusted) 8.5%, 9.9% Postintervention hospital acquired infection rates (adjusted) 6.7%, 7.5%, 6.7%, 6.0%, 6.2% p &lt; 0.001 Hospital acquired urinary tract infection: 3.1%, 3.8%, 2.8%, 3.0%, 2.1%, 1.0%, 2.0% p = 0.008 3.2%, 3.1%, 2.6%, 2.9%, 2.3%, 1.9%, 2.3% p = 0.032 Preintervention hospital acquired infection rates (crude) 7%, 7.8%, 6.2%, 6.5%, 5.6% p &lt; 0.001 Preintervention hospital acquired infection rates (adjusted) 8.5%, 9.9% Postintervention hospital acquired infection rates (adjusted) 6.7%, 7.5%, 6.7%, 6.0%, 6.2% p &lt; 0.001 Hospital acquired urinary tract infection: 3.1%, 3.8%, 2.8%, 3.0%, 2.1%, 1.0%, 2.0% p = 0.008 3.2%, 3.1%, 2.6%, 2.9%, 2.3%, 1.9%, 2.3% p = 0.032</td>
</tr>
<tr>
<td>Mitchell and Jones 1995&lt;sup&gt;21&lt;/sup&gt; ITS Australia</td>
<td>Nurses (number not specified) on 1 medical ward in an acute care hospital</td>
<td></td>
<td>Falls</td>
<td>I: Protocol C: No intervention</td>
<td></td>
</tr>
<tr>
<td>Tilden 1987&lt;sup&gt;22&lt;/sup&gt; ITS US</td>
<td>2 Nurses in a hospital emergency department</td>
<td>Female adult trauma</td>
<td>I: Guidelines; staff education programme; local consensus processes C: No intervention</td>
<td>Increase in rate of identification of positive battering from 7 cases (9.72% of total) to 17 cases (22.97% of total) p = 0.03 OR likelihood of identification 2.74 times greater post-intervention</td>
<td></td>
</tr>
</tbody>
</table>
## Comparison tables (continued)

<table>
<thead>
<tr>
<th>Study and study design</th>
<th>Participants</th>
<th>Clinical area</th>
<th>Results comparisons</th>
<th>Main process effect</th>
<th>Main patient outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Skill-substitution:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greenfield et al 1975</td>
<td>5 doctors</td>
<td>Management of low back pain</td>
<td>Protocol (nurses) C: No intervention (doctors)</td>
<td>Final diagnosis: diagnosis of muscle headache: I: 62.1% (126/203); C: 43.5% (84/193) p&lt;0.001</td>
<td>Symptom relief: resolved or improved: I: 73% (162/222); C: 68.5% (135/197) p=0.33</td>
</tr>
<tr>
<td>RCT-Pt</td>
<td>US</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greenfield et al 1976</td>
<td>5 doctors</td>
<td>Management of headache</td>
<td>Protocol (nurses) C: No intervention (doctors)</td>
<td>Percentage of patients with record of: Smoking habit (once): I1: (100%); I2: (84%); 16 patients p=0.32 Urine test (once): I1: (93%); I2: (21%); 4 patients p&lt;0.00002 Blood pressure (at each visit): I1: (100%); I2: (100%); 103 visits p=1.00 Pulse (at each visit): I1: (98%); I2: (65%); 67 visits p&lt;0.01 Weight (at each visit): I1: (90%); I2: (37%); 38 visits p&lt;0.001</td>
<td>Symptom relief: resolved or improved: I: 73% (162/222); C: 68.5% (135/197) p=0.33</td>
</tr>
<tr>
<td>RCT-Pt</td>
<td>US</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jewell and Hope 1988</td>
<td>1 nurse and 1 doctor in a general practice</td>
<td>Hypertension</td>
<td>Protocol (nurses) C: No intervention (doctors)</td>
<td>Percentage of patients having s x ray films ordered: I1: 51.4% (56/108); I2: 50.0% (73/146) p=0.59</td>
<td>Symptom relief: resolved or improved: I: 73% (162/222); C: 68.5% (135/197) p=0.33</td>
</tr>
<tr>
<td>RCT-Pt</td>
<td>UK</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klassen et al 1993</td>
<td>21 nurses and doctors (number not specified) in 1 emergency room</td>
<td>Radiology referrals</td>
<td>Protocol (nurses) C: No intervention (doctors)</td>
<td>Percentage of patients with record of: Smoking habit (once): I1: (100%); I2: (84%); 16 patients p=0.32 Urine test (once): I1: (93%); I2: (21%); 4 patients p&lt;0.00002 Blood pressure (at each visit): I1: (100%); I2: (100%); 103 visits p=1.00 Pulse (at each visit): I1: (98%); I2: (65%); 67 visits p&lt;0.01 Weight (at each visit): I1: (90%); I2: (37%); 38 visits p&lt;0.001</td>
<td>Symptom relief: resolved or improved: I: 73% (162/222); C: 68.5% (135/197) p=0.33</td>
</tr>
<tr>
<td>RCT-Pt</td>
<td>Canada</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zeller et al 1992</td>
<td>Nurses and doctors (numbers not specified) working in 1 cardiothoracic intensive care unit</td>
<td>Management of postoperative bleeding after cardiac surgery</td>
<td>Protocol (nurses) C: No intervention (doctors)</td>
<td>Percentage of patients in whom coagulation tests performed: I: 13% (14/108); C: 14% (20/146) p=0.40</td>
<td>Mean postoperative blood loss: I: 817 ml; C: 976 ml p&lt;0.08 Mean activated partial thromboplastin time: I: 31.7 s; C: 47.3 s p&lt;0.01 Number requiring emergency sternotomy: I: 4% (4/108); C: 8% (12/146) p=0.32</td>
</tr>
<tr>
<td>RCT-Pt</td>
<td>Australia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RCT-Pt=randomised controlled trial by provider; RCT-Pt=randomised controlled trial by patient; CBA=controlled before and after study; ITS=interrupted time series study; I=intervention; I1, I2, I3=intervention 1, intervention 2, intervention 3; C=control; LDL-cholesterol=low density lipoprotein cholesterol; HDL-cholesterol=high density lipoprotein cholesterol; RR=risk ratio; BP=blood pressure; OR=odds ratio.

and lectures; group two received additional provider and patient education at each clinic visit; and group three received an additional protocol. This was a complex intervention, comprising a flowchart record incorporating the recommendations of the guideline—that is, a patient specific prompt—and a redifinition of the nurse’s role. Herman et al.27 found that an...
active dissemination strategy (involve
provider education) had a minor effect on profes-
sional behaviour in the group receiving only
guidelines, educational materials, and lectures.
The impact of the guidelines on outcomes of
care was also assessed: a significant difference
between intervention and control groups was
found, but in the opposite direction to that
which was expected: the percentage of patients
decreasing immunisation increased in the inter-
vention period to 13.2% in the prevention group,
5.8% in the patient education group,
and 5.9% in the control group. However, there
was an error in the unit of analysis in this study:
hence findings were more likely to be signifi-
cant.

McDonald et al., compared the effect of
computerised prompts specific to the patient
at the time of consultation, computerised
prompts plus bibliographic citations, and
standard care on the use and follow up of
medications. However, the results for nurses
were presented with the two intervention
groups combined, so it is not possible to com-
pare the two methods of implementation. With
the computerised prompts and computerised
prompts plus citations groups combined, the
intervention did not seem to lead to a
significant change in nurses’ behaviour; this
finding was contrary to that for physicians in
the same study. It is possible that this negative
finding was due to a lack of power to detect a
difference (n=five nurses).

GUIDE VERSUS NO GUIDE STUDIES
Nine studies (table) compared the introd-
tuction of guidelines with a no guideline
control. These studies aimed to improve
the quality of care delivered by the target
profession. Of these, four were randomised
controlled trials. Three were interrupted
time series studies, and two were control-
led before and after studies. Targeted
behaviours were nutrition therapy for patients
with non-insulin dependent diabetes, man-
agement of labour, discharge planning,
management of intravenous therapy, lipid
lowering, identification of battered women,
infusion control, and reduction in rates of
falls.

Five studies assessed the effect of guidelines
in terms of the process of care. In four of these,
significant changes in at least some of the
assessed processes of care were identified.
For example, the study of Frigoletto et al.,
the management of labour by nurse midwives
found significant differ-
ences in practices relating to the management
of labour between the active management pro-
tocol and usual care groups. In the active man-
agement group, the frequency of vaginal
examinations was significantly higher (mean
interval 1.6 hours v 2.5 hours for standard
care group); significantly more women received
oxytocin to hasten labour (70% v 56%) and
significantly fewer women received epidural
anaesthesia (45% v 64%). The
analysis reported by Larson and Hargiss
is almost certainly inappropriate, but insufficient
data were provided to be certain: there is no

explanation as to how the dependent variable
was calculated for the analysis of variance
(ANOVA), and clustering of the data has prob-
ably not been taken into account.

Eight studies measured outcomes of care. In
six of these, significant differences in favour
of the intervention group were found for at least
some outcomes of care. However, the
findings from several of these studies were
equivocal, with no significant change being
noted in other outcome variables assessed.
For some studies, the observed change in what
seemed to be the primary outcome measure
was non-significant.

Only one study included a formal eco-
omic evaluation. This study assessed the
impact of medical nutrition therapy for patients
with non-insulin dependent diabetes admin-
istered according to practice guidelines with
medical nutrition therapy administered in line
with standard care. Findings were suggestive
of a cost efficiency advantage in favour of the
control group when one outcome measure
(glycated haemoglobin) was considered, but no
difference when an alternative outcome meas-
ure (fasting plasma glucose concentration) was
examined. The studies described by Naylor et
al., evaluating the effects of a protocol for
discharging patients compared with standard
discharge procedures, and Larson and
Hargiss, evaluating the effectiveness of spe-
cially trained nurses who used guidelines in the
maintenance of intravenous treatment com-
pared with standard care, suggested that direct
healthcare costs are decreased by care driven
by a guideline, but these authors did not
provide sufficient data to ascertain whether
these cost savings were offset by the cost of
guideline dissemination and implementation.

SKILL SUBSTITUTION STUDIES
In the remaining six studies (table), the
performances of nurses operating in accord-
ance with a guideline were compared with
standard care, generally provided by a physi-
cian. All six studies were randomised control-
led trials. The aim was skill substitution,
generally in the interests of greater efficiency.
Targeted behaviours were: management of
dysuria, frequency and vaginal discharge,
management of low back pain, management
of headache, management of hypertension,
referral for x-ray examination, management
of postoperative bleeding after cardiac
surgery. Study settings were: outpatients, in-
patients, the emergency room, and primary
care.

In these studies, a finding of no significant
difference, or of a positive effect in favour of the
nurse or other professional allied to medicine,
was appropriate. Findings from all six studies
which examined the processes and the five
which examined outcomes of care generally
supported the hypothesis of no difference.
However, appropriate sample size calculations
were not reported in any study; thus we cannot
be sure that there was sufficient power to detect
a clinically relevant difference. The finding of
no significant difference may therefore have
been due to small sample size. Clearly this is
weak evidence, as there was no comparison with patients cared for by nurses not using guidelines.

Discussion

Studies which compared different dissemination and implementation strategies suggest educational interventions are of value in the dissemination of guidelines, and confer a benefit over passive dissemination approaches. This finding is in line with that of Grimshaw and Russell for physicians. However, from the limited evidence presented in the two studies in which there were positive changes in process or outcomes, we cannot tell which forms of education are most effective in bringing about changes in behaviour. Grimshaw and Russell suggested that educational strategies requiring more active participation, such as educational outreach visits and targeted seminars, are more likely to lead to changes in behaviour among physicians; there is an urgent need to investigate whether this assertion holds true for nurses, midwives, and other professions allied to medicine. By contrast, a computerised prompt specific to the patient at the time of consultation did not seem to alter nurses' behaviour, despite evidence that this is an effective way of changing the behaviour of physicians. However, this finding was based on a small sample of nurses: further research is required before this implementation strategy can be rejected.

Interventions evaluated by Herman et al and by Larson and Hargiss, were complex; complex interventions were also described by French et al and by Mitchell and Jones. Wensing and Gro 

Many of the studies fall short of the quality criteria for randomised controlled trials, controlled before and after studies, and interrupted time series studies set down by EPOC. Important methodological details were often not reported—for example, few studies specified initially the size of difference in process or outcome that would be considered to be “clinically significant” and reported whether sample sizes were calculated accordingly. Thus we cannot be sure that the studies were adequately powered to detect a clinically important difference. Similarly, in the studies seeking to show no significant difference between care provided by a nurse and driven by a guideline and standard care provided by a physician, appropriate sample size calculations were not reported. In these studies, the finding of no significant difference may have been due to small sample size. Pocock advises that, in general, larger sample sizes are required when the aim is to infer that there is no difference.

Potential sources of bias include a lack of blinding to allocation. If data collectors were aware of which patients were in the intervention and control groups, data collection procedures may have been biased in favour of the intervention group. Few studies reported an intention to treat analysis, although there was often substantial (and, in some cases, differential) attrition from intervention and control samples.
It is mostly impossible to tell whether the guidelines evaluated were based on evidence. Although many were based on a literature review, the extent to which these reviews were systematic was not described, nor were the quality criteria by which any evidence was assessed. The information contained in the papers does not allow us to state whether this lack of an evidence base was due to failure to search for the evidence systematically, or whether there was virtually none. This calls into question the validity of the guidelines.

More research is clearly required in the development, dissemination, implementation, evaluation, and cost effectiveness of clinical guidelines as a strategy for improving professional practice in nursing and professions allied to medicine. Only then will a decision be possible on whether their potential for improving the practice of nurses and professions allied to medicine is as great as it is for doctors.1

We thank the NHS research programme “evaluating methods to promote the implementation of research findings” for funding this project. We acknowledge the helpful comments and advice of members of the EPOC Advisory Board and our three external peer reviewers, Philip Ayres, Ellen Hodnett, and Mary Ann Thompson. We also thank Carol Riccallon, a freelance information scientist, and Julie Glanville, from the NHS Centre for Reviews and Dissemination at York, for help with searching. A big thank you also goes to our superb project secretary, Catherine Brown. We thank Trevor Sheldon and two anonymous reviewers for helpful comments on an earlier draft.1