
Technology Assessment Reports

Foreword

The following abstracts are from publications produced by the Health Technology Assessment Programme in the United Kingdom. This is a national program of government-funded research that publishes over 40 reports every year. The overall aim of the program is to ensure that high-quality research information on the costs, effectiveness, and broader impact of health technologies is produced in the most efficient way for those who use, manage, and work in the National Health Service in England and Wales. Access to the full reports can be made free of charge on the HTA website (<http://www.ncchta.org>) A fully searchable CD-ROM containing the full text of all HTA reports is also available via the HTA website free of charge worldwide. Printed versions of the reports cost £20 (plus post and packing; free in the UK; £2 in Europe; £3 for Rest of the world). See website for details.

Cost–benefit evaluation of routine influenza immunization in people 65–74 years of age

S. Allsup, M. Gosney, A. Haycox, *University of Liverpool*; and M. Regan, *Liverpool Health Authority*

Objectives: This study aimed to determine the cost-effectiveness of influenza vaccination in people 65–74 years of age in the absence of comorbidity.

Design: Primary research: randomized controlled trial.

Setting: Primary care.

Participants: People without risk factors for influenza or contraindications to vaccination were identified from twenty general practitioner (GP) practices in Liverpool in September 1999 and invited to participate in the study. There were 5,875 of 9,727 (60.4 percent) people 65–74 years of age identified as potentially eligible, and of these, 729 (12 percent) were randomized.

Intervention: Participants were randomized to receive either influenza vaccine or placebo (ratio, 3:1), with all individuals receiving pneumococcal vaccine unless administered in the previous 10 years. Of the 729 people randomized, 552 received vaccine and 177 received placebo; 726 individuals were administered pneumococcal vaccine.

Main outcome measures and methodology of economic evaluation: GP attendance with influenza-like illness (ILI) or pneumonia (primary outcome measure); or any respiratory symptoms; hospitalization with a respiratory illness; death; participant self-reported ILI; quality of life (QoL) measures at 2, 4, and 6 months poststudy vaccination; adverse reactions 3 days after vaccination. A cost-effectiveness analysis was undertaken to identify the incremental cost associated with the avoidance of episodes of influenza in the vaccination

population, and an impact model was used to extrapolate the cost-effectiveness results obtained from the trial to assess their generalizability throughout the National Health Service (NHS).

Results: In England and Wales, weekly consultations for influenza and ILI remained at baseline levels (less than 50 per 100,000 population) until week 50/1999 and then increased rapidly, peaking during week 2/2000 with a rate of 231/100,000. This rate fell within the range of “higher than expected seasonal activity” of 200–400/100,000. Rates then quickly declined, returning to baseline levels by week 5/2000. The predominant circulating strain during this period was influenza A (H3N2). Five (0.9 percent) people in the vaccine group were diagnosed by their GP with an ILI compared with two (1.1 percent) in the placebo group [relative risk (RR), 0.8; 95 percent confidence interval (CI) = 0.16 to 4.1]. No participants were diagnosed with pneumonia by their GP, and there were no hospitalizations for respiratory illness in either group. Significantly fewer vaccinated individuals self-reported a single ILI (4.6 percent vs 8.9 percent, RR, 0.51; 95 percent CI for RR, 0.28 to 0.96). There was no significant difference in any of the QoL measurements over time between the two groups. Reported systemic side-effects showed no significant differences between groups. Local side-effects occurred with a significantly increased incidence in the vaccine group (11.3 percent vs 5.1 percent, $P=0.02$). Each GP consultation avoided by vaccination was estimated from trial data to generate a net NHS cost of £174.

Conclusions: No difference was seen between groups for the primary outcome measure, although the trial was underpowered to demonstrate a true difference. Vaccination had no significant effect on any of the QoL measures used, although vaccinated individuals were less likely to self-report ILI. The analysis did not suggest that influenza vaccination in healthy people 65–74 years of age would lead to lower NHS costs.

Future research should look at ways to maximize vaccine uptake in people at greatest risk from influenza and also the level of vaccine protection afforded to people from different age and socioeconomic populations.

Lowering blood pressure to prevent myocardial infarction and stroke: A new preventive strategy

M. Law, N. Wald, and J. Morris, *Wolfson Institute of Preventive Medicine, Queen Mary's School of Medicine and Dentistry*

Objectives: This study aimed to investigate the screening performance of measuring blood pressure and other variables in identifying those who will develop, or die from, ischemic heart disease and stroke and to quantify by how much drugs that lower blood pressure will reduce the risk of ischemic heart disease and stroke in those designated "screen positive."

Data sources: Used were MEDLINE, Cochrane collaboration and Web of Science databases; Stroke registries; Health Survey for England; Office of National Statistics; BUPA (British United Provident Association) study.

Review methods: Relevant cohort studies and randomized trials were identified and analyzed. Statistical analysis was used to determine drug efficacy and adverse effects.

Results: Lowering blood pressure by 5 mm Hg diastolic reduces the risk of stroke by an estimated 34 percent and ischemic heart disease by 21 percent from any pretreatment level; there is no threshold. These estimates, from cohort studies, have been corroborated by the results of randomized trials in persons with high, average, and below average levels of blood pressure. Despite its importance in causing cardiovascular disease, blood pressure is a poor predictor of cardiovascular events. Its poor screening performance is illustrated by the findings that, in the largest cohort study, persons in the top 10 percent of the distribution of systolic blood pressure experienced only 21 percent of all ischemic heart disease events and 28 percent of all strokes at a given age. Combining several reversible risk factors adds little to the screening performance of blood pressure alone; for example, the 25 percent of men 55–64 years of age at highest computed risk (≥ 1 percent) experience only 46 percent of all ischemic heart disease events. The main methods of screening should be to identify all persons with a history of cardiovascular disease events (for example, identifying patients at the time of hospital discharge after a first myocardial infarction detects 50 percent of all heart disease deaths in a population at a false-positive rate of 12 percent), and to use a person's age. Identifying everyone with a history of myocardial infarction or stroke in a population and everyone 55 years of age or older would include 98 percent of all deaths from ischemic heart disease and stroke. The five main categories of blood pressure-lowering drugs, thiazides, beta-blockers, angiotensin-converting enzyme (ACE) inhibitors, angiotensin-II receptor antagonists, and calcium channel blockers, significantly reduce blood pressure from all pretreatment levels, although the extent of the blood pressure reduction increased with pretreatment blood pressure. The reductions were similar at standard dose for the five categories: average reduction was 9.1 systolic and 5 diastolic.

The effect of combinations of two drugs on blood pressure was additive. No effect of age was apparent, given blood pressure. There were no serious metabolic consequences of using these drugs in standard dose.

Conclusions: The evidence presented indicates that three drugs in combination may reduce stroke by approximately two thirds and ischemic heart disease by half. The report suggests that the term hypertension should be avoided, because it is not a disease and it suggests another category (normotensives) who would not benefit from lowering blood pressure. Blood pressure reduction using combinations of safe, well-established drugs is effective in preventing cardiovascular events. It is, therefore, suggested that such preventive therapy be considered more widely in people who, by virtue of existing disease or simply age, are at risk of a heart attack or stroke regardless of initial blood pressure.

Randomized controlled trial to evaluate the clinical- and cost-effectiveness of Hickman line insertions in adult cancer patients by nurses

A. Boland, A. Haycox, A. Bagust, *University of Liverpool*; and L. Fitzsimmons, *Christie Hospital NHS Trust*

Objectives: This study examined the clinical- and cost-effectiveness of image-guided Hickman line insertions versus blind Hickman line insertions undertaken by nurses in adult cancer patients.

Design: A cost-effectiveness analysis was carried out alongside a randomized controlled trial.

Setting: The study was undertaken at a large acute cancer center in Manchester, United Kingdom.

Participants: Cancer patients due to have a Hickman line insertion who were older than 18 years of age and were clinically and physically compliant with specified protocols participated.

Interventions: To obtain central venous access for the patient, two interventions were investigated: (i) blind insertion of a Hickman line and (ii) image-guided insertion of a Hickman line. Both interventions involved blind venipuncture of the subclavian vein. In the blind arm, the Hickman line was routinely inserted without the use of image guidance at any point in the procedure. Transfer to the interventional X-ray suite and use of image guidance were options immediately available to the operator during the procedure if required. In the image-guided arm, the position of the guidewire was checked before the Hickman line was introduced and later the Hickman line was positioned with the use of X-ray fluoroscopy.

Main outcome measures: The primary clinical outcome measure was catheter-tip misplacement, and this misplacement was expected to be higher in the blind arm. When comparing the skill level of the trainer and the trainees, pneumothorax was the primary clinical outcome measure. Other outcomes measures included arterial puncture, hematoma, infection, failed insertion, and assistance from other health-care professionals.

Results: No statistically significant difference was found between the mean cost per patient in the two arms of the trial.

The only statistically significant difference in clinical outcomes was the frequency of catheter-tip misplacement, which was higher in the blind arm of the trial. At very low costs, the image-guided approach dominates the blind approach as fewer costs and greater benefits are incurred. It is evident that nurses previously inexperienced in the procedure can be trained to insert Hickman lines successfully both at the bedside and under image guidance within a 3-month period.

Conclusions: This report indicates that nurse insertion of Hickman lines in the majority of adult cancer patients is both safe and effective. However, there are a select group of patients for whom image-guided insertion may be preferred. The results reveal that skills and expertise can be transferred from trainer to trainee through a relatively short, but intensive, training course. It is also evident that patients support nurse insertion. Further research is suggested to compare the safety and efficacy of nurse versus doctor insertions in particular subgroups of patients and also to assess the quantity and quality of current service provision to inform National Health Service decision-making in this area.

Systematic review of isolation policies in the hospital management of methicillin-resistant *Staphylococcus aureus*: A review of the literature with epidemiological and economic modeling

B. S. Cooper, S. P. Stone, C. C. Kibbler, B. D. Cookson, J. A. Roberts, *University of London*; G. F. Medley, *University of Warwick*; G. J. Duckworth, *Communicable Disease Surveillance Centre*; R. Lai, *University of London*; and S. Ebrahim, *University of Bristol*

Objectives: This study reviewed the evidence for the effectiveness of different isolation policies and screening practices in reducing the incidence of methicillin-resistant *Staphylococcus aureus* (MRSA) colonization and infection in hospital inpatients in an effort to develop transmission models to study the effectiveness and cost-effectiveness of isolation policies in controlling MRSA.

Data sources: Sources included MEDLINE, EMBASE, CINAHL, The Cochrane Library, and SIGLE (1966–2000), as well as handsearching key journals. No language restrictions were used.

Review methods: Key data were extracted from articles reporting MRSA-related outcomes and describing an isolation policy in a hospital with epidemic or endemic MRSA. No quality restrictions were imposed on studies using isolation wards (IW) or nurse cohorting (NC). Other studies were included if they were prospective or used planned comparisons of retrospective data. Stochastic and deterministic models investigated long-term transmission dynamics, studying the effect of a fixed capacity IW, producing economic evaluations using local cost data.

Results: A total of forty-six studies were accepted: eighteen IWs, nine NC, nineteen other isolation policies. Most were interrupted time series, with few planned formal prospective studies. All but one reported multiple interventions. Consideration of potential confounders, measures to prevent bias, and

appropriate statistical analysis were mostly lacking. No conclusions could be drawn in a third of the studies. Most others provided evidence consistent with reduction of MRSA acquisition. Six long interrupted time series provided the strongest evidence. Four of these studies provided evidence that intensive control measures that included patient isolation were effective in controlling MRSA. In two others, IW use failed to prevent endemic MRSA. There was no robust economic evaluation. Models showed that improving the detection rate or ensuring adequate isolation capacity reduced endemic levels, with substantial savings achievable.

Conclusions: Major methodological weaknesses and inadequate reporting in published research mean that many plausible alternative explanations for reductions in MRSA acquisition associated with interventions cannot be excluded. No well-designed studies allow the role of isolation measures alone to be assessed. Nonetheless, there is evidence that concerted efforts that include isolation can reduce MRSA even when endemic. Little evidence was found to suggest that current isolation measures recommended in the United Kingdom are ineffective, and these practices should continue to be applied until further research establishes otherwise. The studies with the strongest evidence, together with the results of the modeling, provide testable hypotheses for future research. Guidelines to facilitate design of future research are produced.

Routine examination of the newborn: The EMREN study. Evaluation of an extension of the midwife role including a randomized controlled trial of appropriately trained midwives and pediatric senior house officers

J. Townsend, *London School of Hygiene and Tropical Medicine*; D. Wolke, J. Hayes, S. Davé, C. Rogers, L. Bloomfield, *University of Hertfordshire*; E. Quist-Therson, *Mount Vernon and Watford Hospitals NHS Trust*; M. Tomlin, and D. Messer, *University of Hertfordshire*

Objectives: This study assesses the implications and cost-effectiveness of extending the role of midwives to include the routine (24-hour) examination of the healthy newborn usually carried out by junior doctors.

Design: The study included a prospective randomized controlled trial (RCT) with mother and baby dyads randomized to either senior house officer (SHO) or midwife for the routine examination of the newborn. Midwives and SHOs were also video taped while performing the examinations, and the videos were rated by an independent consultant and senior midwife. In addition, extensive interviews, surveys, consultations, and assessments were carried out.

Setting: Several settings were involved, including a district general hospital (for the RCT), a London teaching hospital, general practices, and mothers' homes (for interviews). Questionnaires were sent to all maternity units in England (for the National Survey).

Participants: A total of 826 mother and baby dyads in a District General Hospital in southeast England participated.

Midwives and SHOs, as well as midwifery managers, pediatric consultants; general practitioners (GPs) and representatives of key organizations also participated.

Interventions: A routine examination of a newborn baby was carried out at approximately 24 hours from birth, and another examination for half the babies in each group was carried out at 10 days at home by the community midwife.

Main outcome measures: Referrals were assessed as appropriate and as major or minor by three independent consultants. Problems identified during the first year of life were assessed as identifiable at 24 hours. Other measures included quality assessment by video against an agreed written proforma, maternal satisfaction, and opinion of professionals and mothers about aspects of the examination.

Results: There was no statistical difference between SHO and midwife examinations in appropriate referral rates to hospital or community or in inappropriate referral rates to hospital. Video taped assessments were assessed to be carried out more appropriately by the midwives than by the SHOs. Overall, maternal satisfaction was high and higher when a midwife rather than an SHO was involved. Few new health problems were identified at the 10-day examination. From the National Survey, it was estimated that approximately 2 percent of babies in England are examined by a midwife. If midwives were to examine all babies where there were no complications of birth or antenatal history, there would be savings of approximately £2 per baby born, equivalent to savings of £1.2 million nationally. Were midwives to examine all babies on normal wards, savings would increase to approximately £4.30 per baby born or £2.5 million nationally. Representatives of the professional bodies were of the opinion that having trained midwives carrying out the examination would be valuable.

Conclusions: All component aspects of the study were consistent in showing benefits or at least no significant barriers to suitably qualified, trained midwives carrying out the examinations. Developing the role of the midwife to include examination of the newborn is likely to result in improved quality of examinations and higher satisfaction from mothers. It would slightly reduce overall health service costs, with some increased resources needed by midwifery departments and some decrease in resource needs of pediatric departments. There is a need for further research into the value of the examination being carried out at home rather than in a hospital, the overall unsatisfactory quality of the examination of the hips, and appropriate inclusion criteria for which babies the midwives should examine.

Impact of screening on future health-promoting behaviors and health beliefs: A systematic review

C. R. Bankhead, J. Brett, C. Bukach, P. Webster, *Oxford University*; S. Stewart-Brown, *The University of Warwick*; M. Munafo, *Oxford University, Radcliffe Infirmary*; and J. Austoker, *Oxford University*

Objectives: A systematic review was performed to examine the effects of cholesterol, breast cancer, and cervical cancer

screening on actual or intended health-promoting behaviors and health-related beliefs.

Methods: Eleven electronic databases (between 1980 and 2000) served as data sources. All English language studies that investigated the impact of cholesterol, breast, and cervical screening programs on health-promoting behaviors and beliefs were assessed for inclusion. The data extraction form and quality assessment criteria were developed using the National Health Service Centre for Reviews and Dissemination guidelines. Data were extracted, and a nonquantitative synthesis was conducted. Reviewers categorized the outcomes into those that could be considered beneficial or detrimental to health. This categorization was based on a value judgment that considered both statistical and clinical significance.

Results: The cholesterol studies used prospective designs more frequently, possibly as many investigators focused on observing changes in lifestyle after screening. Participants who went for breast or cervical screening were not offered advice on lifestyle changes, and most of the research into cancer screening programs investigated issues related to uptake of screening services, explanations of why people are or are not screened, and interventions to improve uptake. All three screening programs are associated with high levels of favorable health behaviors and beliefs that have been measured, although there is evidence that recommended follow-up after screening is often not adhered to. There was no literature on the cost-effectiveness regarding the wider implications of screening (only on reduction of disease-specific mortality/morbidity), possibly due to the outcomes being very broad and not easily categorized and classified.

Conclusions: The studies reviewed suggest that cholesterol screening had a positive effect on health behaviors, although participation was voluntary and those screened were possibly more motivated to make changes. These results, therefore, are not generalizable to the entire population, and other factors need to be taken into account. Reduction in blood cholesterol levels was reported in all but two of the studies that assessed this outcome, suggesting that successful lifestyle changes were made. However, as most of the studies only reported follow-up of those screened, some of the reduction can be attributable to regression to the mean. Whether breast and cervical screening affect future health behaviors and beliefs has not been directly measured in many studies, and few studies have collected baseline measures. However, evidence suggests that women who attend breast and cervical screening once are likely to reattend, and attendance is associated with several positive health behaviors, although it cannot be confirmed whether the associations observed were a result of screening or because these women have a certain set of health behaviors and beliefs irrespective of their experience of screening. Areas of further research include measuring a much wider range of behaviors and beliefs before and after screening is accepted or declined, examining the subgroup of participants who receive "desirable" results and the impact of this on health beliefs and health-promoting behavior, and qualitative research into the experiences of screening and how this finding interacts with knowledge and beliefs about other aspects of health.

What is the best imaging strategy for acute stroke?

J. M. Wardlaw, S. L. Keir, *Western General Hospital (University of Edinburgh)*; J. Seymour, *University of Aberdeen*; S. Lewis, P. A. G. Sandercock, M. S. Dennis, *Western General Hospital (University of Edinburgh)*; and J. Cairns, *University of Aberdeen*

Objectives: The cost-effectiveness of computed tomographic (CT) scanning after acute stroke was investigated to assess the contribution of brain imaging to the diagnosis and management of stroke and to estimate the costs, benefits, and risks of different imaging strategies to provide data to inform national and local policy on the use of brain imaging in stroke.

Methods: A decision-analysis model was developed to represent the pathway of care in acute stroke, using "scan all patients within 48 hours" as the comparator against twelve alternative scan strategies. Hospitals in Scotland were the setting. The participants were patients admitted to the hospital with a first stroke and those managed as outpatients. The interventions included the effect on functional outcome after ischemic or hemorrhagic stroke, tumors, or infections of correctly administered antithrombotic or other treatment; of time to scan and stroke severity on diagnosis by CT or magnetic resonance imaging (MRI); on management, including length of stay, functional outcome, and quality-adjusted life years (QALYs) of the diagnostic information provided by CT scanning; and the cost-effectiveness (cost versus QALYs) of different strategies for use of CT after acute stroke. The main outcome measures were death and functional outcome at long-term follow-up; accuracy of CT and MRI; cost of CT scanning by time of day and week; effect of CT diagnosis on change in health outcome, length of stay in hospital, and QALYs; and cost-effectiveness of various scanning strategies.

Results: CT is very sensitive and specific for hemorrhage within the first 8 days of stroke only. Suboptimal scanning used in epidemiology studies suggests that the frequency of primary intracerebral hemorrhage (PICH) has been underestimated. Aspirin increases the risk of PICH. There were no reliable data on functional outcome or on the effect of antithrombotic treatment given long-term after PICH. In 60 percent of patients with recurrent stroke after PICH, the cause is another PICH and mortality is high among PICH patients. A specific MR sequence (gradient echo) is required to identify prior PICH reliably. CT scanners were distributed unevenly in Scotland, 65 percent provided CT scanning within 48 hours of stroke, and 100 percent within 7 days for hospital-admitted patients, but access after hours was very variable and for outpatients was poor. The average cost of a CT brain scan for stroke was £30.23 to £89.56 in normal working hours and £55.05 to £173.46 after hours. Average length of stay was greatest for severe strokes and those who survived in a dependent state. For a cohort of 1,000 patients 70–74 years of age, the policy "scan all strokes within 48 hours," cost £10,279,728 and achieved 1982.3 QALYs. The most cost-effective strategy was "scan all immediately" (£9,993,676 and 1982.4 QALYs). The least cost-effective was to "scan patients

on anticoagulants, in a life-threatening condition immediately and the rest within 14 days."

Conclusions: In general, strategies in which most patients were scanned immediately cost least and achieved the most QALYs, as the cost of providing CT (even after hours) was less than the cost of inpatient care. Increasing independent survival by even a small proportion through early use of aspirin in the majority with ischemic stroke, avoiding aspirin in those with hemorrhagic stroke, and appropriate early management of those who have not had a stroke, reduced costs and increased QALYs.

Systematic review and modeling of the investigation of acute and chronic chest pain presenting in primary care

J. Mant, R. J. McManus, R. A. L. Oakes, B. C. Delaney, *University of Birmingham*; P. M. Barton, *University of Birmingham*; J. J. Deeks, *Institute of Health Sciences*; L. Hammersley, *University of Birmingham*; R. C. Davies, *Sandwell and West Birmingham NHS Trust, Sandwell General Hospital*; M. K. Davies, *University Hospital, Birmingham NHS Trust*; and F. D. R. Hobbs, *University of Birmingham*

Objectives: The objectives were to ascertain the value of a range of methods—including clinical features, resting and exercise electrocardiography, and rapid access chest pain clinics (RACPCs)—used in the diagnosis and early management of acute coronary syndrome (ACS), suspected acute myocardial infarction (MI), and exertional angina.

Methods: MEDLINE, EMBASE, CINAHL, the Cochrane Library, and electronic abstracts of recent cardiological conferences were used as data sources. Searches identified studies that considered patients with acute chest pain with data on the diagnostic value of clinical features or an electrocardiogram (ECG) and patients with chronic chest pain with data on the diagnostic value of resting or exercise ECG or the effect of a RACPC. Likelihood ratios (LRs) were calculated for each study, and pooled LRs were generated with 95 percent confidence intervals. A Monte Carlo simulation was performed evaluating different assessment strategies for suspected ACS, and a discrete event simulation evaluated models for the assessment of suspected exertional angina.

Results: For acute chest pain, no clinical features in isolation were useful in ruling in or excluding an ACS, although the most helpful clinical features were pleuritic pain (LR + 0.19) and pain on palpation (LR + 0.23). ST elevation was the most-effective ECG feature for determining MI (with LR + 13.1) and a completely normal ECG was reasonably useful at ruling out this condition (LR + 0.14). Results from "black box" studies of clinical interpretation of ECGs found very high specificity but low sensitivity. In the simulation exercise of management strategies for suspected ACS, the point of care testing with troponins was cost-effective. Prehospital thrombolysis on the basis of ambulance telemetry was more effective but more costly than if performed in the hospital. In cases of chronic chest pain, resting ECG features were not found to

be very useful (presence of Q-waves had LR + 2.56). For an exercise ECG, ST depression performed only moderately well (LR + 2.79 for a 1-mm cutoff), although this performance did improve for a 2-mm cutoff (LR + 3.85). Other methods of interpreting the exercise ECG did not result in dramatic improvements in these results. Weak evidence was found to suggest that RACPCs may be associated with reduced admission to hospital of patients with noncardiac pain, better recognition of ACS, earlier specialist assessment of exertional angina and earlier diagnosis of noncardiac chest pain. In a simulation exercise of models of care for investigation of suspected exertional angina, RACPCs were predicted to result in earlier diagnosis of both confirmed coronary heart disease (CHD) and noncardiac chest pain than models of care based around open access exercise tests or routine cardiology outpatients, but they were more expensive. The benefits of RACPCs disappeared if waiting times for further investigation (e.g., angiography) were long (6 months).

Conclusions: Where an ACS is suspected, emergency referral is justified. ECG interpretation in acute chest pain can be highly specific for diagnosing MI. Point of care testing with troponins is cost-effective in the triaging of patients with suspected ACS. Resting ECG and exercise ECG are of only limited value in the diagnosis of CHD. The potential advantages of RACPCs are lost if there are long waiting times for further investigation. Recommendations for further research include the following: determining the most appropriate model of care to ensure accurate triaging of patients with suspected ACS; establishing the cost-effectiveness of prehospital thrombolysis in rural areas; determining the relative cost-effectiveness of rapid access chest pain clinics compared with other innovative models of care; investigating how rapid access chest pain clinics should be managed; and establishing the long-term outcome of patients discharged from RACPCs.

Effectiveness and efficiency of guideline dissemination and implementation strategies

J. M. Grimshaw, R. E. Thomas, G. MacLennan, C. Fraser, C.R. Ramsay, L. Vale, *University of Aberdeen*; P. Whitty, M. P. Eccles, *University of Newcastle upon Tyne*; L. Matowe, L. Shirran, *University of Aberdeen*; M. Wensing, R. Dijkstra, *University of Nijmegen*; and C. Donaldson, *University of Calgary*

Objectives: A systematic review of the effectiveness and costs of different guideline development, dissemination, and implementation strategies was undertaken. The resource implications of these strategies was estimated, and a framework for deciding when it is efficient to develop and introduce clinical guidelines was developed.

Methods: MEDLINE, Healthstar, Cochrane Controlled Trial Register, EMBASE, SIGLE, and the specialized register of the Cochrane Effective Practice and Organization of Care (EPOC) group were used as data sources. Single estimates

of dichotomous process variables were derived for each study comparison based upon the primary end point or the median measure across several reported end points. Separate analyses were undertaken for comparisons of different types of intervention. The study also explored whether the effects of multifaceted interventions increased with the number of intervention components. Studies reporting economic data were also critically appraised. A survey to estimate the feasibility and likely resource requirements of guideline dissemination and implementation strategies in United Kingdom settings was carried out with key informants from primary and secondary care.

Results: In total, 235 studies reporting 309 comparisons met the inclusion criteria; of these studies, 73 percent of comparisons evaluated multifaceted interventions, although the maximum number of replications of a specific multifaceted intervention was eleven comparisons. Overall, the majority of comparisons reporting dichotomous process data observed improvements in care; however, there was considerable variation in the observed effects both within and across interventions. Commonly evaluated single interventions were reminders, dissemination of educational materials, and audit and feedback. There were twenty-three comparisons of multifaceted interventions involving educational outreach. The majority of interventions observed modest to moderate improvements in care. No relationship was found between the number of component interventions and the effects of multifaceted interventions. Only 29.4 percent of comparisons reported any economic data. The majority of studies only reported costs of treatment; only twenty-five studies reported data on the costs of guideline development or guideline dissemination and implementation. The majority of studies used process measures for their primary end point, despite that only three guidelines were explicitly evidence-based (and may not have been efficient). Respondents to the key informant survey rarely identified existing budgets to support guideline dissemination and implementation strategies. In general, the respondents thought that only dissemination of educational materials and short (lunchtime) educational meetings were generally feasible within current resources.

Conclusions: There is an imperfect evidence base to support decisions about which guideline dissemination and implementation strategies are likely to be efficient under different circumstances. Decision-makers need to use considerable judgment about how best to use the limited resources they have for clinical governance and related activities to maximize population benefits. They need to consider the potential clinical areas for clinical effectiveness activities, the likely benefits and costs required to introduce guidelines and the likely benefits and costs as a result of any changes in provider behavior. Further research is required to develop and validate a coherent theoretical framework of health professional and organizational behavior and behavior change to inform better the choice of interventions in research and service settings and to estimate the efficiency of dissemination and implementation strategies in the presence of different barriers and effect modifiers.

Clinical-effectiveness and cost-effectiveness of neonatal screening for inborn errors of metabolism using tandem mass spectrometry: A systematic review

A. Pandor, *The University of Sheffield*; J. Eastham, *University of Leeds*; C. Beverley, J. Chilcott, S. Paisley, *The University of Sheffield*

Objectives: The clinical- and cost-effectiveness of tandem mass spectrometry (MS)-based neonatal screening for inborn errors of metabolism (IEM) were evaluated.

Methods: Fourteen electronic bibliographic databases covering biomedical, science, economic, and gray literature; the reference lists of relevant articles and abstracts of conference proceedings; and eighteen health services research-related resources served as data sources. This review is an update of two previous health technology assessment reports of neonatal screening for IEM. These reports have been updated by a systematic review of published research (between 1995 and January 2002) on neonatal screening of inherited metabolic disorders using tandem MS. This update was supplemented by a search for economic literature and the application of a modeling exercise to investigate the economics of using tandem MS within a neonatal screening program in the United Kingdom.

Results: Evidence from the reviews of IEM found that the UK screening program for phenylketonuria (PKU) was well established and that there was universal agreement that neonatal screening for PKU was justified. Of the many other disorders that can be detected by tandem MS, the best candidate condition for a new screening program was

medium-chain acyl-coenzyme A dehydrogenase (MCAD) deficiency. For many other IEM that can be detected by tandem MS, robust clinical evidence was limited. Cost-effectiveness analysis using economic modeling indicated that substituted the use of tandem MS for existing technologies for the screening of PKU alone could not be justified. However, results from the economic modeling indicated that the addition of screening for MCAD deficiency as part of a neonatal screening program for PKU using tandem MS would be economically attractive. Using an operational range of 50,000–60,000 specimens per system per year, the mean incremental cost for PKU and MCAD deficiency screening combined using tandem MS from the model was –£23,312 for each cohort of 100,000 neonates screened. This cost saving is associated with a mean incremental gain of 59 life-years. Additional economic modeling using the available evidence does not support including other inherited metabolic diseases within a neonatal screening program at present.

Conclusions: The evidence appears to support the introduction of tandem MS into a UK neonatal screening program for PKU and MCAD deficiency combined. Tandem MS has the potential for simultaneous multi-disease screening using a single analytical technique. Although the marginal cost of extending the program to include other conditions may be relatively small, the application of this new technology to PKU and MCAD deficiency screening does not suggest the wholesale inclusion of all disorders detectable by tandem MS. It is suggested that the primary focus of further research should be on the long-term effectiveness of treatment strategies on adverse outcomes (disabilities and impairments) under conventional management and the potential impact of early diagnosis using tandem MS.